Quarterly Reviews of Biophysics, Page 1 of 132. © Cambridge University Press 2013
doi:10.1017/S0033583512000157 Printed in the United States of America

Why nature really chose phosphate

Shina C. L. Kamerlin[1], Pankaz K. Sharma[2], Ram B. Prasad[2] and Arieh Warshel[2]*

[1] Department of Cell and Molecular Biology (ICM), Uppsala Biomedical Centre, Uppsala University, Box 596, S-751 24 Uppsala, Sweden
[2] Department of Chemistry (SGM 418), University of Southern California, 3620 McClintock Avenue, Los Angeles, CA 90089, USA

Abstract. Phosphoryl transfer plays key roles in signaling, energy transduction, protein synthesis, and maintaining the integrity of the genetic material. On the surface, it would appear to be a simple nucleophile displacement reaction. However, this simplicity is deceptive, as, even in aqueous solution, the low-lying d-orbitals on the phosphorus atom allow for eight distinct mechanistic possibilities, before even introducing the complexities of the enzyme catalyzed reactions. To further complicate matters, while powerful, traditional experimental techniques such as the use of linear free-energy relationships (LFER) or measuring isotope effects cannot make unique distinctions between different potential mechanisms. A quarter of a century has passed since Westheimer wrote his seminal review, 'Why Nature Chose Phosphate' (Science 235 (1987), 1173), and a lot has changed in the field since then. The present review revisits this biologically crucial issue, exploring both relevant enzymatic systems as well as the corresponding chemistry in aqueous solution, and demonstrating that the only way key questions in this field are likely to be resolved is through careful theoretical studies (which of course should be able to reproduce all relevant experimental data). Finally, we demonstrate that the reason that nature *really* chose phosphate is due to interplay between two counteracting effects: on the one hand, phosphates are negatively charged and the resulting charge-charge repulsion with the attacking nucleophile contributes to the very high barrier for hydrolysis, making phosphate esters among the most inert compounds known. However, biology is not only about reducing the barrier to unfavorable chemical reactions. That is, the same charge-charge repulsion that makes phosphate ester hydrolysis so unfavorable also makes it possible to regulate, by exploiting the electrostatics. This means that phosphate ester hydrolysis can not only be turned on, but also be turned off, by fine tuning the electrostatic environment and the present review demonstrates numerous examples where this is the case. Without this capacity for regulation, it would be impossible to have for instance a signaling or metabolic cascade, where the action of each participant is determined by the fine-tuned activity of the previous piece in the production line. This makes phosphate esters the ideal compounds to facilitate life as we know it.

* Author for correspondence: A. Warshel.
Tel: (213) 740 4114; Email: warshel@usc.edu

1. Introduction

Chemical reactions involving the formation or cleavage of P–O bonds in phosphate esters are ubiquitous in biological systems. Any undergraduate chemistry or biology student is familiar with the hydrolysis of adenosine triphosphate (ATP) to yield ADP and inorganic phosphate (P_i), a reaction that is one of the principal sources of biochemical energy. Additionally, the transfer of phosphoryl groups from one entity to another plays a central role in biosynthesis, the control of secondary messengers, the regulation of protein function and the maintenance of the integrity of the genetic material. Both DNA and RNA are phosphate esters, as are many intermediate metabolites. Some examples of biochemically relevant phosphates are compiled in Table 1.

As shown in the Table 1, a wide range of enzymes has evolved to catalyze reactions involving phosphates, operating via a range of different mechanisms, and under a wide range of different conditions. Thus, there exist enzymes catalyzing phopshoryl transfer that can function at both high and low pH, enzymes that utilize direct attack by water or alternately employ an enzyme-derived nucleophile, as well as enzymes that employ metal ions in catalysis, and others that do not. Phosphoryl transfer has also become the focus of much effort in the design of artificial catalysts, many of which draw their inspiration from known enzyme structures (Aguilar-Pérez *et al.* 2006; Feng *et al.* 2009; Williams, 2004a) as well as in artificial enzyme design (e.g. Alcolombri *et al.* 2011). The phosphorylation and dephosphorylation of a protein can affect the function of a protein in many ways, namely by: (i) decreasing or increasing the biological activity of the protein, (ii) either stabilizing the protein or marking it for breakdown, (iii) facilitating or inhibiting movement between subcellular compartments, or (iv) initiating or disrupting protein–protein interactions (de Grauuw *et al.* 2006). The importance of protein phosphorylation in cellular signaling in particular can be emphasized by the fact that one-third of all proteins in the cell are phosphorylated at any given time (Manning *et al.* 2002a). Also, early estimates have

Table 1. *Some examples of biologically relevant phosphates* (*Westheimer, 1987*)

Phosphate	Biological role
DNA/RNA	Genetic material
Adenosine di/triphosphate	Intracellular energy transfer
Creatine phosphate	Intracellular energy transfer
Phosphoenolpyruvate	Metabolism
Pyridoxal phosphate	Coenzyme
Nicotine adenine dinucleotide	Calcium signaling
Fructose 1,6-diphosphate	Metabolism
Glucose-6-phosphate	Metabolism
Isopentenyl pyrophosphate	Metabolism
Ribose-6-phosphate	Metabolism
Glycerol 3-phosphate	Metabolism
Dihydroxyacetone phosphate	Calvin cycle, metabolism
Inositol phosphates	Cellular signaling

suggested that the human proteome possesses approximately 100 000 potential phosphorylation sites (Zhang *et al.* 2002). Therefore, it is critical that protein phosphorylation is tightly controlled in cells, as abnormal changes in enzymes responsible for protein phosphorylation and dephosphorylation can have devastating consequences, resulting in many human diseases, such as diabetes, type II obesity, and, perhaps most importantly, many forms of cancer. Due to the fact that these enzymes are so important in cellular transformation, and that their deregulation is frequently implicated in the onset of tumor formation, they have also become the focus of increasing pharmaceutical interest for the development of targeted therapeutics (e.g. Fardilha *et al.* 2010).

A significant biological constraint on living systems is the stability of the system. While at least some metabolites can be short lived, the genetic material has to be reasonably stable (when it is not manipulated by specialized enzymes). Even a simple gene will possess over a thousand ester bonds. If a single of these is cleaved during the lifetime of the organism, it may fail to reproduce. Therefore, natural selection will favor a genetic material whose stability is comparable with that of the living organism itself (Westheimer, 1987). Some phosphate esters have exceptional kinetic stability. For instance, the half-life for the uncatalyzed attack of water on alkyl phosphate dianions is approximately $1{\cdot}1 \times 10^{12}$ s^{-1} at 25 °C (Lad *et al.* 2003b), whereas the upper limit for the hydrolysis of the dimethyl phosphate monoanion at the same temperature is expected to be approximately 1×10^{15} s^{-1} (Wolfenden *et al.* 1998). Similarly, the hydrolysis of double-stranded DNA is estimated to take well over 100 000 years (Radzicka & Wolfenden, 1995). Following from this, the stability of the phosphate ester bond makes it an ideal switch in information and energy transfer processes, where specialized enzymes can change the hydrolysis barrier in quite a drastic way, allowing for it to be fast and exquisitely controlled. While the extreme stability makes phosphate esters excellent systems for information storage, experimental chemists require compounds that are reactive at room (or only moderately elevated) temperature. Therefore, the stability of phosphates makes them highly undesirable for synthetic chemistry, and synthetic organic chemists preferentially use other groups for linking hydroxyl, carboxyl and amino groups, and for activating them for chemical reactions. Similarly, while leaving groups such as chlorides, bromides, iodides, tosylates, and triflates are routinely used for nucleophilic displacement reactions by synthetic chemists, they rarely use phosphate (this despite the fact that in metabolic reactions, the leaving group is usually phosphate or pyrophosphate (Flaks *et al.* 1957; Kornberg *et al.* 1955; Preiss & Handler, 1958; Westheimer, 1987).

Biological molecules containing phosphate esters and anhydrides dominate the living world. Yet often, the lion's share of attention is given to the chemistry of carbon or sulfur, and, in comparison, little attention is paid to phosphorus chemistry. Nevertheless, the importance of phosphate esters is well known, and phosphate-containing compounds are the focus of a great body of both academic and pharmaceutical interest. Despite this, there remain many unanswered questions in this important field, particularly with regard to the preferred mechanism for the hydrolysis of phosphate esters. This review will discuss the general aspects of phosphate chemistry, and the mechanism of phosphate ester hydrolysis will be examined in great detail, from both an experimental and a computational perspective. This will include discussions of the problems with oversimplified qualitative interpretations of the experimental data, as well as with the clarification of our view that, at present, only theoretical findings can establish detailed reaction pathways and quantify catalytic effects. This requires, however, that the specific studies reach a level of sufficient reliability and ability to reproduce relevant experimental observations. We will start with an extensive review of studies of phosphate chemistry in solution. We will then move to key examples of enzymatic reactions that involve phosphate ester hydrolysis, including G-proteins, ATPase, DNA polymerases, and other systems. A significant part of this review will focus on computational studies, highlighting their great power as the way to resolve controversies in the field, while at the same time pointing out the problems and traps associated with the superficial use of computational tools. We will repeatedly emphasize the difference between the interpretation of experimental facts, and the actual experimental findings, trying to encourage the reader to be critical, without losing the fact that the detailed understanding of this field is rapidly growing. Overall, we will keep in mind the question put forward in the 1987 review by Westheimer (1987), 'Why nature chose phosphates?'. We will answer this question by concluding the electrostatic basis for this choice.

2. The fundamentals of phosphorus chemistry and relevant mechanistic issues

2.1 Basic phosphate chemistry

2.1.1 Definitions: what *is* a phosphate?

By definition, a phosphate is a salt with an anionic entity, which is built of either a single PO_4 tetrahedron, or by the condensation of multiple PO_4 anions (Averbuch-Pouchot & Durif, 1996). The main exception to this rule are compounds where one or more of the oxygen atoms have been substituted by other atoms such as H (giving rise to the phosphorous(III) $[PO_3H]^{2-}$ anion), S (giving rise to thiophosphates such as PO_3S, PO_2S_2, POS_3, and PS_4) or F (giving fluorophosphates such as PO_3F or PO_2F_2). The most common form of phosphates are compounds based on an anionic phosphorous(V) entity ($[PO_4]^{3-}$), that is composed of a nearly regular tetrahedral arrangement of four oxygen atoms, surrounding a central phosphorus atoms. Such compounds (and their corresponding salts) are referred to as 'monophosphates', and these compounds are made particular interesting by the fact that they are not only the most stable phosphates, but also the *only* phosphates to be found in the natural world.

2.1.2 Elemental phosphorus: an overview

Originally discovered by Brand in 1669 (Beatty, 2001; Emsley, 2000), elemental phosphorus exists in three generalized forms: red, white, and black (although each of these has subdivisions

with different mechanism of transformation and unique crystalline forms). Additionally, each of these types of elemental phosphorus has its own unique physical properties, although white phosphorus can be transformed into both red and black phosphorus. Historically, one of the main industrial uses of white phosphorus in the past was matchstick production (though of course, this usage dropped quite dramatically with the advent of the safety match containing only red phosphorous (Beatty, 2001). Elemental phosphorus is still widely used for synthesis of various compounds, including, herbicides, hydraulic fluids, water purification systems, and even some food products.

Due to its fundamental importance, the chemistry of phosphorus has been described in detail in the literature (e.g. Averbuch-Pouchot & Durif, 1996; Beatty, 2001; Shriver & Atkins, 2006). Here, however, we will be providing a brief précis to set the tone of the review. Phosphorus has an atomic number of 15, and atomic mass of 30·974 amu, and a covalent radius of 106 ppm. There exist seven different isotopes of phosphorus, each of which has its own unique properties. However, the ^{31}P isotope is the most common, and the only isotope found in nature. Elemental phosphorus is found as a P_4 tetrahedron (in white form and planar P_3 in black form). However, the electronic configuration of phosphorus is [Ne] $3s^2 3p^3$, leaving five empty low-lying 2-d orbitals, which are used in p–d bonding in a fashion not too dissimilar from conventional p–p (π) bonding. Therefore, common phosphorus compounds have oxidation states that range from -3 to $+5$, allowing phosphorus to readily expand its coordination number to accommodate five bonds, resulting in the associative transition states (TSs) and intermediates that have been a topic of such controversy (see Section 3). Additionally, most organophosphorus compounds (with the exception of P(V)) are strong reducing agents. Following from this, a well-known commercial application of phosphorus compounds is for 'electrodeless plating', which involves the use of $H_2PO_4^-$ in order to reduce Ni^{2+} (aq) ions, coating surfaces with metallic nickel in the process. More interestingly, elemental phosphorus can also be easily oxidized by combustion. When this combustion is complete, it yields P(V) oxide (P_4O_{10}), and, when there is only a limited supply of oxygen, P(III) oxide (P_4O_6) is obtained instead. Both of these oxides are easily hydrated – the P(III) form to give phosphorous acid, H_3PO_4, and the P(V) form to give phosphoric acid, H_3PO_4. Esters of this latter acid lie at the core of phosphate hydrolysis.

2.1.3 Phosphoric acids

Although there exists a wide range of phosphoric acids, the main focus of this review is biological phosphoryl transfer, and hence the discussion will be limited to the most basic phosphoric acid, namely orthophosphoric acid (H_3PO_4, henceforth simply referred to as phosphoric acid) and its esters. The crystal structure of H_3PO_4 studied by Smith *et al.* (1955), exist in only one crystalline form a monoclinic crystal, with lattice constants: $a = 5{\cdot}80$ Å, $b = 4{\cdot}85$ Å, $c = 11{\cdot}62$ Å, and $\beta = 95{\cdot}2^\circ$ (Smith *et al.* 1955). The space group of this crystal is $C_{2h}^5 - P2_1/c$, and there are four formula weights of $2H_3PO_4{\cdot}2H_2O$ per unit cell (Smith *et al.* 1955). Within the phosphate tetrahedron, one short P–O distance of 1·52 Å (corresponding to P=O) was observed, as well as three slightly longer P–O distances of 1·55, 1·57, and 1·59 Å (corresponding to the three P–OH bonds), as shown in Fig. 1.

As can be seen from Fig. 1, phosphoric acid is a triprotic acid, which can dissociate three times to give dihydrogen phosphate, ($H_2PO_4^-$), hydrogen phosphate (HPO_4^{2-}), and the phosphate anion (PO_4^{3-}). The pK_a values for each successive deprotonation are shown in Sections 2.1–2.3. Note also that the neutral form is only present in very acidic conditions, and has therefore so far

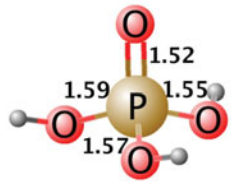

Fig. 1. Representative bond distances in phosphoric acid, based on (Smith *et al.* 1955).

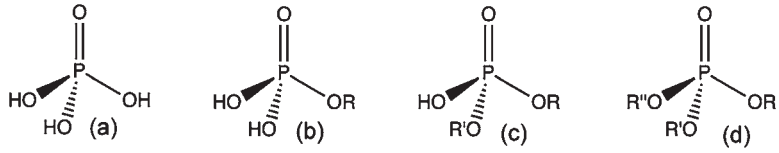

Fig. 2. A comparison of (*a*) phosphoric acid, (*b*) a phosphate monoester, (*c*) a phosphate diester, and (*d*) a phosphate triester.

not been subjected to extensive study (Cleland & Hengge, 2006).

$$H_3PO_4(s) + H_2O(l) \rightleftharpoons H_3O^+(aq) + H_2PO_4^-(aq)\ K_{a,1} = 7{\cdot}5 \times 10^{-3}, pK_{a,1} = 2{\cdot}12, \quad (1)$$

$$H_2PO_4^-(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + HPO_4^{2-}(aq)\ K_{a,1} = 6{\cdot}2 \times 10^{-8}, pK_{a,2} = 7{\cdot}21, \quad (2)$$

$$HPO_4^{2-}(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + PO_4^{3-}(aq)\ K_{a,3} = 2{\cdot}14 \times 10^{-13}, pK_{a,3} = 12{\cdot}67. \quad (3)$$

In addition, phosphoric acid can be esterified in any of one, two, or three positions, in order to form either phosphate monoesters (Fig. 2*b*), diesters (Fig. 2*c*), or triesters (Fig. 2*d*), each of which having different chemical properties and reactivities. Additionally, while phosphate triesters do not occur naturally, phosphate mono- and diesters have well-known roles as intermediates in biochemical transformations, and, additionally, monoesters formed by protein phosphorylation have key roles in regulating a wide range of processes (Cleland & Hengge, 2006).

Of additional interest is the behavior of phosphate anions with regard to metal ions that is, phosphate is able to coordinate with both mono- and divalent cations, to yield complexes like $M^I_3PO_4$, $M^I_2HPO_4$, and $M^IH_2PO_4$ in the case of the monovalent cations, and $M^{II}_3(PO_4)_2$, $M^{II}_2(HPO_4)_2$, and $M^{II}(H_2PO_4)_2$ in the case of the divalent cations. These are merely the simplest entities in these classes of compound; however, different combinations of the various anionic entities in the same compounds can also frequently be found in monophosphates. The solubility of such complexes is highly dependent on the nature of the associated cation, as well as on the degree of activity. Thus, salts such as Li_3PO_4, Ag_3PO_4, Tl_3PO_4, and $M^{II}_3(PO_4)_2$ are all highly insoluble, whereas the other complexes are moderately to highly soluble (Averbuch-Pouchot & Durif, 1996). Finally, as will be seen in subsequent sections, metalloenzymes frequently exploit the ability of phosphate esters to coordinate to mono- and divalent cations such as Ca^{2+}, Mg^{2+}, and Zn^{2+}.

2.1.4 Phosphorylation

Phosphorylation is ubiquitous in biology (Todd, 1959), where the fact that it is a reversible reaction that is effective in altering the conformation of phosphorylated molecules, which means

that it is often employed in regulatory mechanisms. Essentially, phosphorylation involves the substitution of a hydrogen atom for a phosphoryl group (PO_3) of an $-OH$ or $-NH$ group of an organic compound which may be either a small molecule or a side chain of a protein. This could introduce a conformational change in the structure of the protein *via* interaction with other hydrophobic and hydrophilic residues. For example, phosphorylation of a polar amino acid side chain can render an otherwise hydrophobic part of a protein polar and extremely hydrophilic. Another important example is the phosphorylation of ADP in order to yield ATP, which occurs in mitochondria *via* oxidative phosphorylation.

Generally speaking, phosphorylation can be grouped into two broad categories: reactions yielding phosphate triesters, and reactions yielding phosphate diesters as their initial products. In the first case, the phosphorylating entity derives from a phosphate diester, such as the mixed anhydride of a phosphate diester with a strong acid, for instance, the diester of phosphorochloric acid, or the tetraester of pyrophosphoric acid, often in the presence of a base. In the latter case, the phosphorylating agents are based on mixed anhydrides of phosphate monoesters, yielding predominantly phosphodiesters (Todd, 1959). In proteins, phosphorylation can in principle occur on a range of residues, most commonly of serine and threonine but the phosphorylation of histidine, tyrosine, and aspartate does also occur, with biological relevance. There are thus a large number of possible phosphorylation sites in proteins, and the impact of phosphorylation varies from system to system. Protein phosphorylation is overall a complicated procedure, and the enzymes regulating it will be discussed in detail in Section 4.

2.1.5 Phosphate hydrolysis

At its core, phosphate hydrolysis is the opposite of phosphorylation, and the reaction involved is quite simple: it merely involves the displacement of an oxygen-based ligand (usually an alcohol or alkoxide), by either hydroxide or water attack at the phosphorus atom:

$$RPO_4^{2-} + OH^- \rightleftharpoons [RHPO_5]^{3-} \rightleftharpoons HPO_4^{2-} + OR^-. \tag{4}$$

This reaction can essentially proceed in one of two possibly ways, that is, the reaction can either proceed as outlined in (4), with the nucleophile attacking the phosphorus atom, displacing the alcohol or alkoxide *via* phosphorus oxygen bond fission. However, an alternate competing possibility is that the C-1 carbon of the leaving group serves as an electrophile, in such a way that the entire phosphate group is replaced by the O-atom of the attacking nucleophile, as shown in Fig. 3, complicating the measurement of rate constants for P–O cleavage. As an example of this, the uncatalyzed hydrolysis of dimethyl phosphate has been found to proceed by at least 99% bond cleavage at 25° (Wolfenden *et al.* 1998), leading to an estimated upper limit of $1 \times 10^{-15}\ s^{-1}$ for the rate of spontaneous P–O cleavage. Clearly, this can be quite problematic when performing experimental studies of phosphate hydrolysis, and it is often necessary to choose systems where C–O cleavage is sterically precluded, so that the reaction cannot proceed through elimination. Fortunately, however, the two potential pathways shown in Fig. 3 give different isotopically labeled products, and can therefore be easily distinguished between by hydrolysis in ^{18}O–water. However, while it would appear from Fig. 3 that hydrolysis *via* P–O cleavage is straightforward, the reality is actually quite the opposite, and, as will be seen in Section 3, the reaction can proceed through a variety of different mechanisms even in aqueous solution.

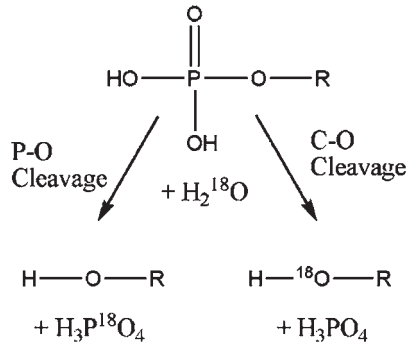

Fig. 3. Fundamental mechanisms for phosphate ester hydrolysis (see also Gani & Wilkie, 1995).

2.2 Stereochemistry and mechanism

2.2.1 Basics of chemical reaction mechanisms

While the analysis of small molecules in the gas phase can be complicated by the requirements of matching vibrational energy levels and exact collision conditions (Atkins, 1998), the situation becomes, in fact, simpler in the case of condensed phase reactions. In such cases, energy transfer from the surroundings leads to a situation where the system moves back and forth in the reactant state, until it reaches a random fluctuation that takes it to the TS, and it can then move back and forth in the product state (Warshel, 1991). The probability of having this rare fluctuation can be given exactly by the Boltzmann factor, $\exp(-\Delta g^{\ddagger}/RT)$, which reflects the probability of the system being found at the TS. The rate constant, k, can therefore be written using a modified version of the Eyring–Polanyi equation:

$$k = \frac{k_B T}{h} \kappa \exp\left(-\Delta g^{\ddagger}/RT\right), \tag{5}$$

where $\Delta g^{\ddagger}$ is the molar activation free energy, T is the temperature, k_B is Boltzmann's constant, h is Planck's constant and R is the ideal gas constant. Here, κ is a 'transmission factor', which the probability of a trajectory that has reached the TS moving to the product state. Clearly, this discussion is making the assumption that the reaction proceeds directly from reactant(s) to product(s) *via* a single TS. However, the reaction can also proceed with the intervention of in intermediate. While some intermediates can be considered to be real chemical entities with finite lifetimes, that undergo chemical reactions, only few intermediates have sufficiently long lifetimes to be directly observed experimentally, and many are so short-lived that one can only assume their existence through circumstantial evidence, muddying mechanistic discussion. Additionally, the *kinetic stability* of an intermediate is determined by the barrier heights flanking it (as this determines the ease with which it can form a product or return to its precursor), whereas the *thermodynamic stability* of an intermediate is determined by the free-energy difference between the intermediate and the reactant (or product) species, and depends on the relative energies of the reacting species. Therefore, the situation can be further complicated by the fact that it is fully possible to have a thermodynamically unstable yet kinetically stable intermediate, and, in fact the nature (and even very existence of) such intermediates in the case of phosphoryl transfer reactions has been a topic of very heated dispute (see particularly the discussion in

Sections 3 and 4). However, we would like to emphasize that the issue of the existence of intermediates, or the use of equilibrium considerations for the TS, as well as referencing the irrelevant issue of the vibrational time of the TS (Jencks, 1980), is a reflection of misunderstandings from the time when the enormous insight provided by TS theory (TST) was not understood or fully appreciated. In fact, as was very eloquently clarified by Bennett (Bennett, 1977) the issue that is encompassed by the $\Delta g^{\ddagger}$ of Eq. (5) is simply the probability of the system being found at the TS. Therefore, invoking the TS vibrational time in mechanistic analyses is risky.

Clearly, reaction mechanisms are complex, and it is most often very hard to definitively prove what path they follow (i.e. the nature of the relevant TS (s) and any intermediates etc.) by experimental techniques alone. Rather, such techniques are more useful for the identification and elimination of impossible pathways for a system. However, advances in computational techniques have helped to alleviate this problem. That is, ever increasing computational speed now allows for the detailed mapping of reaction surfaces, which in turn allows for the identification and characterization of key stationary points (particularly TSs), which are hard to conclusively characterize experimentally. This can be achieved while simultaneously reproducing relevant experimental observables, and is particularly facile when studying gas-phase reactions of small molecules. However this concept can also be extended to the condensed phase, and, increasingly, enzymatic reaction mechanisms as well. Sections 2.2.2 and 2.2.3 will address potential pathways for the hydrolysis of phosphate ester bonds, as well as introducing common and popular methods for characterizing viable reaction mechanisms, and the remaining sections will discuss current progress in distinguishing between different viable mechanisms using computational approaches.

2.2.2 Conventional inline mechanisms

Section 2.1.5 introduced phosphate ester hydrolysis as the symmetric displacement of an oxygen-based ligand (usually alcohol or alkoxide), by hydroxide or water. However, this representation is deceptively simple, as, not only can phosphate esters have various protonation states and be esterified in any or all of three positions, but also, as will be seen in Section 3, the preferred reaction mechanism changes, depending not only on the esterification level of the phosphate ester, but also its protonation state. Therefore, characterization of all the possible ways in which phosphate monoester cleavage can occur is far from trivial, as there are eight distinct mechanisms through which this reaction can potentially proceed. However, it is conventionally common to generalize this reaction into two different pathways (Fig. 4). The first of these (Fig. 4*a*) is a dissociative, S_N1-type mechanism, formally termed a $D_N + A_N$ mechanism in the IUPAC nomenclature. This is a stepwise pathway, in which leaving group departure precedes nucleophilic attack. The reaction proceeds *via* a metaphosphate intermediate, the stability of which drives a mixture of retention and inversion of configuration (note that a highly unstable intermediate will probably lead to inversion of configuration at the P-atom, due to the fact that one face of the phosphate can still be blocked by the leaving group as the nucleophile attacks).

An additional complication, however, lies in the availability of low-lying d-orbitals on the phosphorus atom, which can be utilized in bonding interactions. As a result, the phosphorus atom is readily able to expand its coordination number, and attacking nucleophiles can easily add to the four-coordinate phosphorus to give pentavalent phosphorus species (phosphoranes),

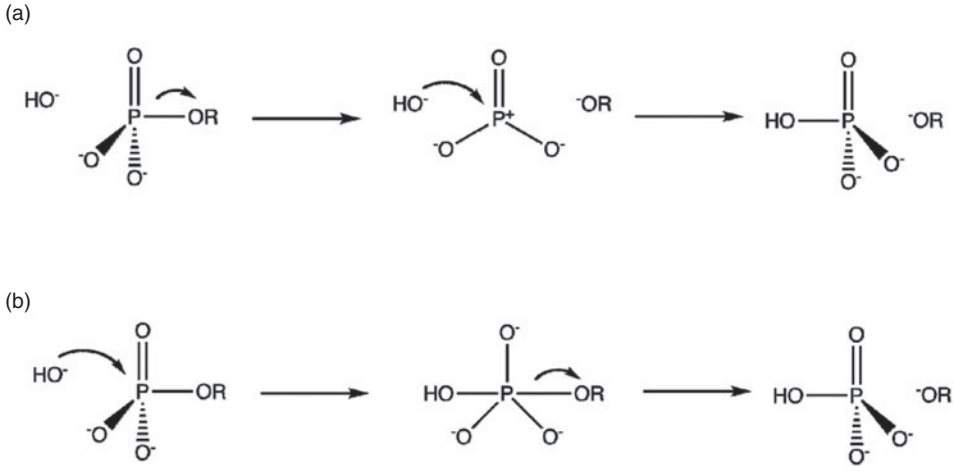

Fig. 4. Generalized pathways for phosphate monoester hydrolysis, using the illustrative example of hydroxide attack on a phosphate dianion. Shown here are stepwise (*a*) dissociative and (*b*) associative mechanisms.

which will normally have trigonal bipyramidal geometry. As a result of this, the reaction can also proceed *via* an alternative associative pathway (Fig. 4 *b*), in which nucleophilic attack occurs prior to the departure of the leaving group, and the reaction proceeds *via* a pentavalent phosphorane intermediate (formally termed an $A_N + D_N$ mechanism). This reaction generally proceeds with inversion of configuration at the phosphorus atom. Note that both pathways discussed so far have been stepwise processes. However, there is an alternative possibility that the reaction proceeds *via* a concerted (A_ND_N) mechanism, in which bond-breaking and making occur in a single reaction step, eliminating the intermediates shown in Fig. 4. The concerted pathways follow S_N2-like mechanisms, in which the leaving group departs as the nucleophile approaches. However, there exist again two possibilities, the first being that the reaction proceeds *via* a concerted *associative* pathway, with partial bond formation to the nucleophile and partial bond cleavage to the leaving group. The converse of this is a concerted *dissociative* pathway, in which, once again, bond cleavage to the leaving group and bond formation are simultaneous; however, the reaction proceeds *via* an expansive concerted TS with slightly more bond cleavage to the leaving group than bond formation to the nucleophile. It is important, however, to note that despite the perception that might be obtained from the above discussion, the associative and dissociative *concerted* pathways are not separate mechanisms, but rather lie on a spectrum of mechanistic possibilities, without a clear cutoff. Therefore, describing a mechanism as associative or dissociative based on the structure of isolated TS alone is problematic, as the entire path needs to be taken into account, as discussed in (e.g. Kamerlin *et al.* 2008b).

2.2.3 Non-inline mechanisms with pseudo-rotation

Up to this point, four distinct reaction pathways have been discussed. However, the focus has only been on inline pathways, in which the leaving group departs from the opposite face as the attacking nucleophile. There is, however, an alternate possibility, involving a non-inline pathway, in which the leaving group departs from the same face as the attacking nucleophile. Such a

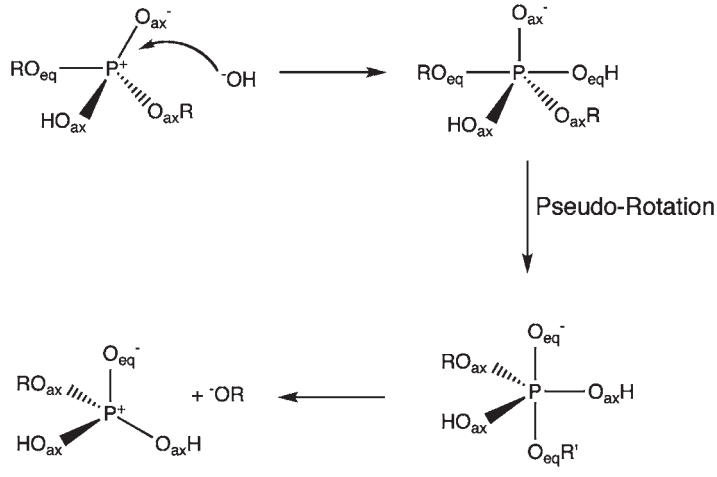

Fig. 5. The pseudo-rotation mechanism for the hydrolysis of phosphate diesters.

non-inline pathway is very rare, but has been experimentally observed e.g. during the transfer of a chirally labeled phosphoryl group from the 2-O to the 1-O atom of 2-phosphoproane-1,2-diol under acidic conditions (Buchwald *et al.* 1984). This reaction proceeds *via* a pseudo-rotation mechanism, as shown in Fig. 5, which illustrates a hypothetical pseudo-rotation mechanism for neutral phosphate with a hydroxyl anion acting as the nucleophile. Nucleophilic attack results in the formation of a pentacoordinated phosphorane intermediate, in which there are two axial and three equatorial positions. In the pseudo-rotation mechanism, the oxygen farthest away from the phosphorous atom is the incoming nucleophile, which is in an axial position. Therefore (as this is a non-inline pathway), the leaving group (OR′) is on the same face of the phosphate, in an equatorial position. In the pseudo-rotation step, internal bond angles in the intermediate are deformed in such a way that the groups change their status from equatorial to axial. This change is predominantly a movement of the phosphorus atom, with simultaneous distortion of internal bond angles. However, there is no change in the *absolute* positions of the groups relative to each other. Thus, after the pseudo-rotation, the nucleophile is in an equatorial position, while the leaving group departs from an axial position, making the leaving group the group that is now farthest from the phosphorus atom.

Note that this non-inline pathway with pseudo-rotation is in keeping with the principle of microscopic reversibility, which requires that in a reversible reaction, the mechanism in one direction has to be the exact reverse of the mechanism in the opposite direction. In this case, leaving group (OR′) departure is the exact opposite of the addition of a nucleophile ($^{-}$OH), and must follow the exact reverse pathway (Carpenter, 2005; Tolman, 1938). Since the nucleophile attacks from an axial position, the leaving group has to depart from an axial position also, which is facilitated by the pseudo-rotation. Finally, while this pseudorotation mechanism is quite rare and specialized (Westheimer, 1968), semi-empirical studies of phosphate monoester hydrolysis at different levels of phosphate and nucleophile protonation have demonstrated that not only is the non-inline associative mechanism with pseudo-rotation a reasonable reaction pathway, but there are certain conditions under which it can even be the lowest energy pathway (Wilkie & Gani, 1996). The barrier for the pseudo-rotation step is so low, however, that it is difficult to observe it either computationally or experimentally.

2.3 Characterization of different pathways

As discussed in Section 2.2, once one takes into account all possibilities of stepwise or concerted, associative or dissociative and inline or non-inline pathways, the deceptively simple nucleophilic displacement reaction that is phosphate hydrolysis can suddenly potentially proceed through eight distinct mechanistic possibilities (although non-inline phosphate hydrolysis has yet to be observed in an enzymatic system). Clearly, therefore, prior to any detailed discussion about phosphoryl transfer, it is essential to make a distinction between these different mechanisms, and there are various ways of achieving this.

2.3.1 Free-energy surfaces provide the most unique mechanistic definition

In light of the large number of mechanistic possibilities for a simple phosphoryl transfer reaction, even in solution, it is essential to have an approach that can provide a unique mechanistic definition that allows for distinction between different options in a way that they are directly comparable. The best way around this problem is by describing the full free-energy surface by means of a More O'Ferrall–Jencks (MFJ) diagram (Jencks, 1972; More O'Ferrall, 1970). By definition, this involves the visualization of the free-energy surface of a reacting system as a function of two given coordinates. These are incrementally increased, and the energy is determined for each point, allowing for the generation of a two-dimensional (2D) projection of the full multidimensional free-energy surface.

There are multiple advantages to using MFJ diagrams in order to study reacting systems. First of all, such representations are versatile – the surface can be defined in terms of any two coordinates that are relevant to the system being studied, be they bond orders, bond distances, angles or torsion angles (or any combination thereof). Also, reaction coordinate mapping makes it possible to directly compare multiple reaction pathways, as identical reactant and product states will be involved for all pathways. As a result, it is possible to obtain relative barriers for associative and dissociative pathways in the presence of both the leaving group and the nucleophile, and, in doing so, to determine the preferred reaction pathway. Therefore, since the entire free-energy system is taken into account, MFJ plots can provide a thorough definition of different mechanistic options for any given system.

Figure 6 presents an example of an MFJ diagram for examining phosphate ester hydrolysis. In this particular example, the P–O distances to the leaving group (R1, *x*-axis) and nucleophile (R2, *y*-axis) oxygen atoms have been used to define the reaction coordinate. An associative pathway will proceed through the top left corner of the plot, with more bond formation than bond cleavage. A dissociative pathway will proceed through the bottom right corner of the plot, with more bond cleavage than bond formation. A concerted pathway would lie somewhere between these two mechanistic extremes, depending upon the precise degrees of bond cleavage and formation. However, despite the power of this approach, a principle shortcoming to bear in mind is that it is depending on the viability of representing the reacting system in terms of two reaction coordinates. This is clearly not always the case, a relevant example being phosphoryl transfer in the presence of either an acid or base catalyst (Fig. 7). In such a system, there are three coordinates that need to be taken into account: bond breaking/formation to the leaving group and nucleophile, respectively, as well as the relevant proton transfer (PT). It has been suggested that there are many reacting systems, such as the example of Fig. 7, where rather than using a 2D MFJ plot, the system is best represented as a reaction cube (Grunwald, 1985; Scudder, 1990; Trushkov *et al.* 1990).

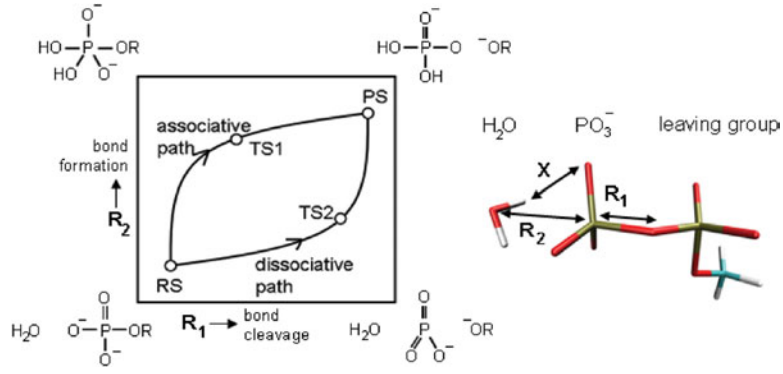

Fig. 6. A sample MFJ diagram (Jencks, 1972; More O'Ferrall, 1970) for examining phosphate ester hydrolysis. The particular example in this figure illustrates a hypothetic surface for water attack on phosphate monoester dianions. The surface is defined as a function of the distance between the phosphorus atom and the leaving (*x*-axis) and entering (*y*-axis) oxygen atoms, respectively. This figure was originally presented in (Klähn *et al.* 2006).

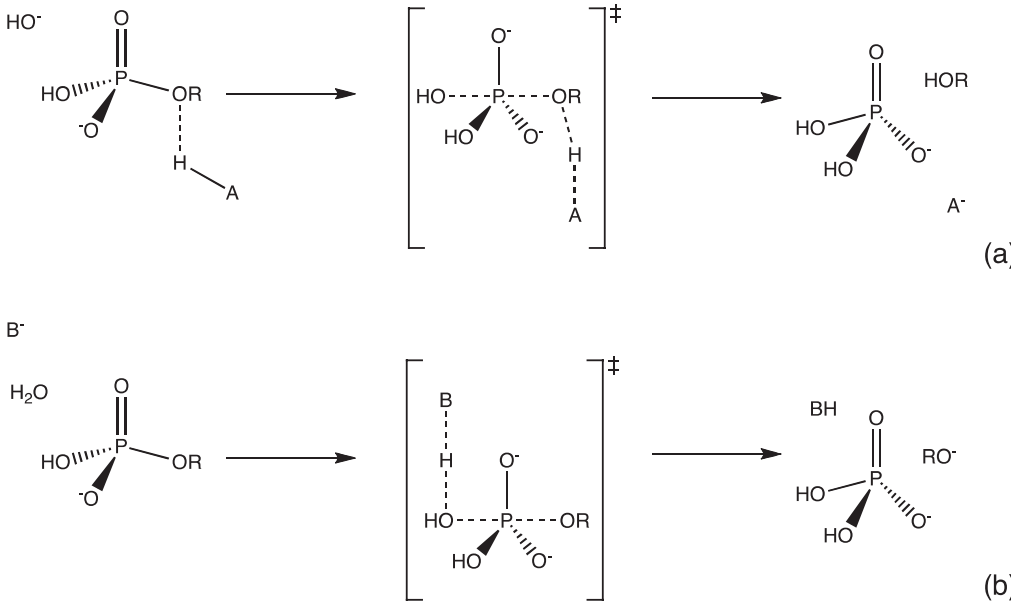

Fig. 7. Phosphate ester hydrolysis in the presence of an (*a*) acid or (*b*) base catalyst.

Figure 8 shows an example of a cubic reaction coordinate diagram (Guthrie, 1996). Here, the cube is depicted in terms of three distinct reaction coordinates, *x*, *y*, and *z*, which could for the example shown in Fig. 7 correspond to P–O distances to the leaving group and nucleophile, with the distance to the proton being transferred forming a third reaction coordinate. These three coordinates, *x*, *y*, and *z*, represent the edges of the cube. The free energy is subsequently expressed as a function of the position along a bond-order coordinate, which runs from 0 (starting material) to 1 (product) (Kurz, 1978). For cubic reaction coordinates, (0,0,0) and (1,1,1) are the starting point and product respectively. The corner intermediates that correspond to the reaction along only one edge coordinate are found at (1,0,0), (0,1,0), and (0,0,1), whereas (1,1,0), (1,0,1),

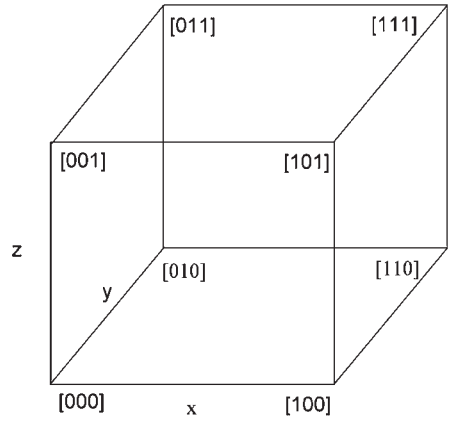

Fig. 8. A cubic reaction coordinate diagram. The three edge coordinates are represented by the x, y and z-axes. (0,0,0) and (1,1,1) are the starting point and product, respectively. All other points represent the corner intermediates (Guthrie, 1996).

and (0,1,1) are corner intermediates corresponding to a reaction along two edge coordinates (Guthrie, 1996). Multidimensional reaction coordinate mapping can also be extended to more complex mechanistic problems, involving more than three reaction coordinates (Guthrie, 1996). However, fortunately (and conveniently), in the majority of cases, it is sufficient to consider projections on limited dimensions to obtain a very informative surface. Nevertheless, there also exist more qualitative approaches to achieve this, which will be considered in Sections 2.3.2–3.3.4.

2.3.2 Bond orders and fractional associativity

One possibility that has been proposed for characterizing the different mechanisms through which phosphate ester hydrolysis can proceed is by taking the axial distances from the entering and leaving groups to the phosphorus reaction center in either the trigonal bipyramidal intermediate or TS into account, which measures the extend of the bonding of the leaving group and nucleophile to the phosphorus atom (Mildvan, 1997). Pauling (Pauling, 1960) has defined the relationship between the bond distance ($D(n)$) and the fractional bond number (n) as:

$$D(n) = D(1) - 0{\cdot}60\ln(n). \tag{6}$$

Here, $D(1)$ is the single bond distance for the P–O bond. A dissociative intermediate has an axial bond number of 0 to both the nucleophile and leaving groups; therefore, axial bond distances cannot be calculated. On the other extreme, the associative intermediate has an axial bond number of 1 to both the entering and leaving groups. Therefore, it is assumed that associative TS will have a combined axial bond order of greater than one for the P–O_{lg} and the P–O_{nuc} bonds, whereas a dissociative TS will have a combined axial bond order of less than one.

While this hypothesis sounds logical enough on paper, there are some practical problems with this approach. Firstly, the accepted bond distances are 1·63 Å for P–O and 1·5 Å for P=O (Benson, 1965; Cottrel, 1958; Huheey *et al.* 1993). However, in reality, the presence of a

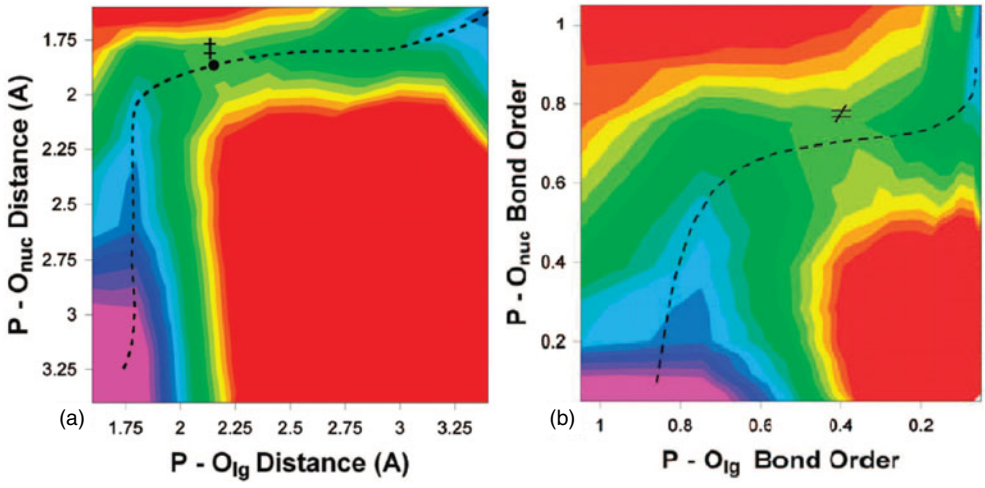

Fig. 9. Free-energy surfaces for hydroxide attack on dineopentyl phosphate, as functions of (*a*) bond order and (*b*) bond distance. This figure is adapted from (Kamerlin *et al.* 2008b).

substituent on the oxygen atom will vary the P–O bond length according to the nature of the substituent. For instance, a P–O bond in the presence of a metal ion would be expected to be longer than a simple P–O bond. Furthermore, even assuming a bond length of 1·63 Å, the scaling factor of 0·6 causes further complications, in that it scales the bond order down too fast with increasing P–O distances. An example of this can be seen in Fig. 9, which shows a comparison of the calculated MFJ plots for hydroxide attack on dineopentyl phosphate (Kamerlin *et al.* 2008b) as functions of (A) bond order and (B) bond distance. From this figure, it can be seen that the TS for this reaction would appear significantly more expansive when representing the surface in terms of bond order than when representing the surface in terms of bond distance. By Pauling's equation a TS with P–O distances of 2·2 and 1·8 Å to the leaving group and nucleophile, respectively (Fig. 9 *a*), will have bond orders of 0·4 and 0·8 to the leaving group and nucleophile, respectively (Fig. 9 *b*), rendering what would be a compact associative TS using a bond distance definition effectively a dissociative TS using a bond order definition. For a more extreme example, an even more compact TS with P–O distances of 2·1 Å to the leaving group and nucleophile, respectively, will have a bond order of approximately 0·45 to each, whereas logically, such a TS should be considered associative, as 2·1 Å are still relatively short bond lengths (particularly if either the leaving group or nucleophile is coordinated to a metal ion). However, the combined axial bond order of 0·9 would categorize these TS as effectively dissociative according to the proposal above, and this is even before one starts taking bond distances of approximately 2·4–2·5 Å into account, which, using a bond distance definition would lie on the boundary between associative and dissociative. Yet, according to Pauling's equation, a concerted TS with P–$O_{lg/nuc}$ distances in that range would only have a combined bond order of approximately 0·6.

Now an alternative to this is to use the axial bond number (n) to the entering group in the first TS to define the fractional associativity of the mechanism, such that a fully associative mechanism would be described as 100% associative, and, conversely, a fully dissociative mechanism would be described as 0% associative (Mildvan, 1997). While this is clearly an improvement on just considering combined axial bond orders calculated using Pauling's equation, the problem

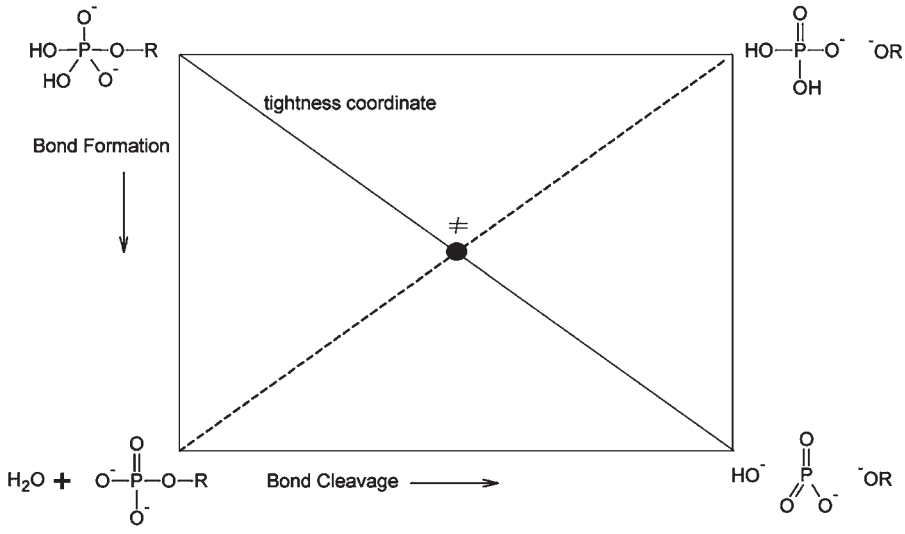

Fig. 10. Reaction coordinate diagram, defined in terms of bond order, for a reaction involving a symmetrical concerted TS.

introduced by the scaling factor still exists, such that any TS with a distance of greater than approximately 2 Å between the phosphorus atom and the oxygen of the incoming nucleophile would have a bond order of $<0{\cdot}53$, and thus be less than 50% associative.

Finally, it is possible to plot the relationship between bond order and reaction mechanism on a reaction coordinate diagram, as has previously been done by e.g. Williams and co-workers (Ba-Saif *et al.* 1989a). An example of such a diagram is shown in Fig. 10, which has been defined in terms of bond order to the incoming nucleophile and the departing leaving group. For simplicity, this figure depicts the relevant situation for a symmetrical concerted TS ('symmetrical' here refers to the P–O distances to the incoming nucleophile and departing leaving group, rather than the nature of the substituents). Here, the dissociative and associative limits are at the bottom right and top left corners of the plot, respectively, and the line connecting these two limits is referred to as the 'tightness coordinate'. A symmetrical reaction from reactants to products will proceed orthogonally to the tightness coordinate, and will cross this coordinate once and *only* once. At the point of intersection, these reaction pathways will appear as minima along the tightness coordinate. However, changing the pK_a values, atom identities or charges of the leaving group or nucleophile can result in reactions with asymmetric TSs (Jencks, 1985), for which the surface is more complex. Additionally, while such reaction coordinate diagrams can be informative, they suffer from the fact that they neither provide information about the energetics of the reaction, nor do they provide any information about charge distributions (Barnes *et al.* 1994). In light of this, and the problems presented by the scaling factor of Eq. (6), therefore, we believe it is far more reliable to use a bond distance definition than a bond order definition in defining mechanism, as a bond distance definition does not suffer from a scaling factor that scales too fast, or from problems defining the correct equilibrium bond distance. Additionally, rather than determining mechanism based on the geometry of an isolated TS, we believe it is important to take into account the topology of the full free-energy surface (see Section 2.3.1), and how the system arrived at that TS. Using this perspective, a pathway in which bond formation to the nucleophile precedes exceeds bond formation to the leaving group on the approach to the TS

would be considered 'associative', and the converse would constitute a dissociative pathway. Such a definition would also remove a lot of the semantic problems that arise from the discrepancy between using a bond order *versus* a bond distance representation.

2.3.3 Qualitative Brønsted-type linear free-energy relationships (LFER)

There are several experimental techniques available for characterizing TSs, one of which is the use of Brønsted-type LFER (Jencks, 1987), which will be discussed in greater detail in Section 3. Here, the idea is to examine the rate constants of the cleavage reaction for a range of leaving groups with different pK_as, in order to measure the extent of bond breaking to the leaving group and bond formation to the nucleophile in the TS, based on Eqs. (7) and (8) (see e.g. (Davis *et al.* 1988; Herschlag & Jencks, 1989a; Hollfelder & Herschlag, 1995a; Jencks, 1987), among others):

$$\log(k) = \beta_{lg} \mathrm{p}K_a + \mathrm{C}, \tag{7}$$

$$\log(k) = \beta_{nuc} \mathrm{p}K_a + \mathrm{C}. \tag{8}$$

Here, β_{nuc} and β_{lg} are 'Brønsted coefficients' for bond formation and bond cleavage, respectively, and k is the measured rate constant for the reaction. In the case of an associative reaction, bond formation dominates over bond cleavage, and one would traditionally expect a large value of β_{nuc}. Conversely, in the case of a dissociative reaction, bond cleavage is more important than bond formation, and one would expect a large value of β_{lg}. The difference between these coefficients can be used in order to determine the nature of the TS, by taking into account the value of β, in the identity reaction β_i:

$$\beta_i = \beta_{nuc} + \beta_{lg} (\text{where } \beta_i = \delta \log k_i / \delta p\mathrm{K}_a) \tag{9}$$

Ideally, as discussed in Sections 2.3.1 and 2.3.2, one would want to represent the system as a map in terms of two reaction coordinates: R1 and R2. However, an oversimplified interpretation of this relationship suggests that a dissociative reaction would be expected to have a negative value of β_i (due to a large β_{lg}), whereas, conversely, an associative reaction would have a positive value of β_i, as β_{nuc} is larger than β_{lg} (Herschlag & Jencks, 1989b). From Eqs. (7)–(9), it should be possible (though somewhat difficult) to calculate the value of β_i directly by observation of the rate constant and pK_a of the reaction, and, in particular, the rate constant for the uncatalyzed reaction of phosphate ester dianions has been found to be highly sensitive to the dissociation constant of the leaving group (Kirby & Varvoglis, 1968b):

$$\log(k) = 0{\cdot}86 - 1{\cdot}23 \mathrm{p}K_a, \tag{10}$$

while at the same time, being found to be practically independent of the pK_a of the nucleophile (R′OH) (Kirby & Varvoglis, 1968b):

$$\log(k) = 0{\cdot}13 \Delta \mathrm{p}K_a(\mathrm{R'OH}). \tag{11}$$

Based on Brønsted coefficients of $-1{\cdot}23$ and $0{\cdot}13$ for the nucleophile and leaving group, respectively, the reaction was assumed to proceed *via* a dissociative mechanism, with the involvement of a monomeric metaphosphate anion (Kirby & Varvoglis, 1968b).

However, in contrast to the seemingly convincing picture that emerged from the above studies, a careful computational study of the experimental LFER for this system (Florián *et al.* 1998)

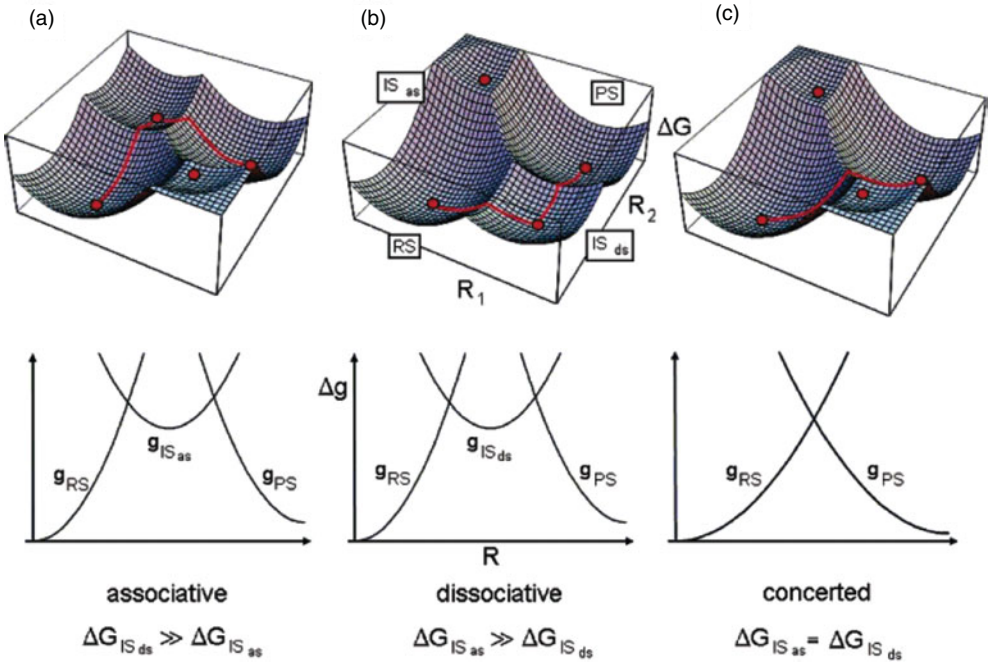

Fig. 11. A VB description of the energy surfaces for phosphate hydrolysis in three limiting cases for associative (*a*), dissociative (*b*), and concerted (*c*) mechanisms. The surfaces for the hydrolysis reaction were generated by mixing four states using an EVB formulation (the effect of the off-diagonal term is in included for simplicity). RS, IS, and PS represent, respectively, reactant, intermediate, and product states and the indices as and ds stand for associative and dissociative, respectively. The dots on the surfaces designate the minima of the corresponding diabatic surfaces (Klähn *et al.* 2006). The change of the highest TS with $\Delta G_{RS \rightarrow PS}$ defines the corresponding LFER

has demonstrated that, while the LFER for this reaction can indeed correspond to a dissociative mechanism, the predicted LFER for the associative mechanism is also consistent with the experimental LFER. This is feasible as both a late TS in the associative mechanism and early TS in the dissociative mechanism involve a largely broken bond between the phosphorus atom and the leaving group oxygen. Subsequent computational studies that have compared calculated and observed LFER for the hydrolysis of phosphate monoester dianions with different leaving groups (Klähn *et al.* 2006) demonstrated that despite the linear relationship between the leaving group pK_a and the rate constant for the reaction, the nature of the TS is highly dependent on the pK_a of the leaving group. Leaving groups with high pK_a appear to favor a more associative reaction pathway, which becomes progressively more dissociative with increasing acidity of the leaving group. A similar trend has also been observed in the case of hydroxide attack on a substituted phosphate diesters (Alkherraz *et al.* 2010; Rosta *et al.* 2008), suggesting that obtaining a unique mechanistic assignment from experimentally obtained LFER is far from trivial, and the challenges with this will be discussed in greater detail in Section 3. Note, however, that the above LFER considerations are entirely empirical, and can therefore be problematic.

A much deeper analysis is conceptually based on the empirical valence bond (EVB) description (Kamerlin & Warshel, 2011a; Warshel, 1991; Warshel & Florian, 2004). In this description, the four corners of the MFJ plot illustrated in Fig. 6 are represented as centers of parabolic surfaces, corresponding to different zero-order diabatic states (Fig. 11). This is related to

Marcus-type LFER treatments, which correlate TS energetics with the crossing of the relevant parabolas. The actual *adiabatic* reaction surface is generated by mixing these parabolic surfaces, using the EVB off-diagonal term. When the reaction follows one of the extreme paths (either associative or dissociative), the nature of the LFER can be easily quantified by shifting the parabolas in the lower part of Fig. 11 (Klähn *et al.* 2006) (see also Åqvist *et al.* 1999). The converse, however, is not true, that is, shifting the parabolas for the different mechanisms can give rise to similar LFER (Åqvist *et al.* 1999). The molecular origin for this will be discussed further in Section 3, however, as mentioned above, there exists very strong computational evidence to suggest that the interpretation of observed LFER is, despite assertions otherwise (Lassila *et al.* 2011), actually rather complicated, and that LFER therefore do not need to have a unique mechanistic interpretation. Here we again emphasize (as we will do in other sections) that our arguments about LFER and mechanism cannot be resolved experimentally in any unique way, and the final judgment must involve reliable computations.

2.3.4 Isotope effects

Section 2.3.3 examined the utilization of kinetic information in order to study reaction mechanisms, by means of LFER. An alternative is to study reactions by means of isotope effects. Isotopic substitution does not alter the chemical identity of a compound, and therefore, the potential energy for a system is independent of the patterns of isotopic substitution (with the nature of the reactions of different compounds remaining unchanged). However, the substitution of an atom of the reactive species by an isotope will affect the mass of that part of the molecule, and hence it will alter the corresponding vibrational frequencies. In general terms, this means that while the reaction path remains essentially the same, substitution of an atom by a heavier isotope will change the value of the rate constant for the reaction, k. This is called a 'kinetic isotope effect' (KIE), and it can be defined as the ratio between the rate constants for the light and heavy atoms (Maskill, 1999), i.e.

$$\mathrm{KIE} = k_{\mathrm{L}}/k_{\mathrm{H}}. \tag{12}$$

Here, k_{L} and k_{H} are the rate constants for the light and heavy atoms, respectively. However, since $\Delta G = -\mathrm{RTln}(k)$, we have:

$$\mathrm{KIE} = \exp((\Delta g^{\ddagger}_{\mathrm{H}} - \Delta g^{\ddagger}_{\mathrm{L}})/\mathrm{RT}), \tag{13}$$

where $\Delta g^{\ddagger}_{\mathrm{L}}$ and $\Delta g^{\ddagger}_{\mathrm{H}}$ are the activation free energies of the light and heavy atoms, respectively. Note that the treatment of Eq. (13) includes the quantum mechanical nuclear tunneling effects, which are included in $\Delta G^{\ddagger}$ (Warshel *et al.* 2006b). Now since equilibrium is a balance between forward and reverse kinetic processes, isotopic substitution can also alter the value of an equilibrium constant, without altering the nature of the equilibrium (Maskill, 1999). This is called an equilibrium isotope effect (EIE), and can be defined as follows:

$$\mathrm{EIE} = K_{\mathrm{L}}/K_{\mathrm{H}}. \tag{14}$$

Here, K_{L} and K_{H} are equilibrium constants for the compounds with the lighter and heavier isotopes, respectively. If this ratio is greater than unity that is, when the parameter for the lighter isotope is greater than the heavier isotope, this is referred to as a 'normal' isotope effect. If the converse is true, however, and the value is less than unity, this is referred to as an 'inverse' isotope effect.

Isotope effects can, in turn, be classified into two categories. A 'primary isotope effect' involves the cleavage of a bond to the isotopically labeled atom, whereas, if this bond is not broken in the chemical reaction, we have a 'secondary isotope effect'. Secondary isotope effects are invariably smaller than primary isotope effects, and their magnitude decreases the further the isotopically labeled atom is from the reaction site. Also, since isotopic substitution alters the zero-point energy levels of the vibrational modes of the reactant as it passes to a TS, it has been argued (e.g. Hengge, 2002) that the magnitude of a KIE is also related to bond order changes at the TS. Qualitatively, one could assume that a dissociative pathway, which would result in a *decrease* in bond order at the TS, would result in a normal isotope effect, whereas associative TS, which would result in an *increase* in bond order at the TS, would result in an inverse isotope effect. Therefore, it is assumed that secondary isotope effects can provide information about the nature of TSs, and several studies have been performed on the TSs for non-enzymatic and enzymatic hydrolyses of phosphate mono-, di-, and triesters (see for instance (Cleland & Hengge, 2006; Hengge, 1999, 2002), to name just a few examples). However, since the magnitude of the KIE is related to changes in bond order (Northrop, 1975) (which occur in the chemical step of an enzymatic reaction), KIE is not so effective in stuides of systems for which a non-chemical step (e.g. substrate binding or conformational changes) is rate limiting. More importantly from the perspective of this work, even in the case of a chemical step, further problems arise from the fact that isotope effects for phosphoryl transfer reactions do not necessarily have a unique and unambiguous mechanistic interpretation. This crucial issue will be discussed in greater detail in Section 3.

2.3.5 Overview of different factors affecting reaction mechanisms and energetics

Up to this point, we have examined the different mechanistic possibilities for phosphate hydrolysis, pointing out that the protonation state of the phosphate ester, and whether it is a phosphate mono-, di-, or triester will all affect the preferred pathway. There are, however, a number of external factors that will also affect the reaction mechanism, a few of which will be touched on in this section.

The first of these is the presence of metal ions,that is, many phosphatases possess at least one metal ion in their active sites, the roles of which vary, depending on the specific enzyme. The role of metal ions has usually been described in a very qualitative fashion, for instance by describing such factors as binding to a conjugated base and the dissipation of its charge, the lowering of the pK_a of the generated nucleophile (Glusker *et al.* 1999), or the correct positioning of the nucleophile for attack on the phosphorus center (Bruice *et al.* 1996; Gani & Wilkie, 1995; Wilkie & Gani, 1996). However, although such terms have been useful as guiding principles, they lack clear definition, and are frequently ill-defined. In actual fact, elucidating the role of metal ions is far from simple (Åqvist & Warshel, 1989, 1990; Luo *et al.* 2012a; Xiang *et al.* 2006) even in solution (Kamerlin & Warshel, 2009), and requires clear definitions of concepts which can distinguish the electrostatic effects of the metal charge from charge transfer effects and covalent effects. It is also critical to explore the possible entropic effects of the interaction of the interaction between metal ions and the reacting systems, and this can only be achieved by exploiting the power of computational modeling.

In additional to metal ions, there are a number of additional factors that can affect reaction pathways. The most obvious of these is the presence of some form of catalyst (e.g. an enzyme or a biomimetic catalyst), where the interaction of the substrate with specific functional groups or

the active site environment is capable of changing the preferred mechanism. Also, such catalysts (particularly enzymes) can act on a wide-range of leaving groups with different chemical properties, and, finally, the presence of a solvent and the nature of that solvent will have an effect on the reaction pathway. Therefore, while at first glance, phosphate ester hydrolysis would appear to be a very simple reaction that is tempting to write off for being 'too trivial', this seemingly uncomplicated reaction can in fact proceed through multiple potential reaction pathways, with the preferred pathway being dependent not only on factors directly related to the reacting system (such as the protonation state or the specific leaving group/nucleophile), but also on external factors such as the presence of a catalyst or solvent effects. Additionally, as will be seen repeatedly in subsequent sections, the interpretation of experimental studies on the mechanism by which phosphoryl transfer occurs is not without problems. Therefore, it is necessary to perform a thorough comparative analysis of both experimental and computational studies of non-enzymatic phosphoryl transfer before examining enzymatic systems, and how they relate to their non-enzymatic counterparts. As will be seen even after decades of intensive research, there remain many issues that are highly controversial in this field, and sensitive to a careful examination of both computational and experimental evidence.

3. Non-enzymatic phosphoryl transfer

As outlined in Section 2, studies of phosphoryl transfer are complicated by the fact that this reaction can proceed through multiple pathways. Not only are the energetics of the key TSs in this process relatively high, but also, the characterization of such TSs by key intermediates is quite difficult, making it very hard to deduce the nature of the TS in a unique way by experimental means. Additionally, several studies have demonstrated that traditional experimental markers such as LFER, isotope effects and activation entropies cannot be used to unambiguously determine the reaction mechanism. Finally, factors such as protonation states, esterification level and catalysis by either metal cofactors or enzymes can all potentially alter reaction mechanisms. Therefore, at present, despite great progress in this area with the advent of computers powerful enough to allow for thorough mapping of free-energy surfaces and examination of different viable mechanistic options (that can reproduce the relevant experimental markers), our knowledge of the detailed reaction pathways for even uncatalyzed phosphoryl transfer remains limited. Therefore, in this section, we will discuss what is currently known about the non-enzymatic hydrolysis of phosphate mono, di- and triesters, based on both experimental and theoretical studies, as well as presenting a detailed discussion on the advantages and shortcomings of popular experimental techniques for elucidating reaction mechanisms. From this discussion, it will be seen that despite decades of intensive research, and the frequent problematic implications (e.g. Admiraal & Herschlag, 1995, 2000; Lassila *et al.* 2011) that many of the key mechanistic issues have been experimentally resolved, there nevertheless still remain huge gaps in our knowledge that can only be bridged by combining experimental work with thorough computational studies that are capable of reproducing all relevant experimental markers.

3.1 The mechanism(s) of phosphate monoester hydrolysis

3.1.1 Phosphate monoester dianions

Traditionally, the kinetic investigation of the hydrolysis of phosphate monoester dianions have been hindered by the fact that, with the exception of esters with highly activated leaving groups

(such as 2,4-dinitrophenol (Kirby & Varvoglis, 1967a)), these compounds are generally quite unreactive. Thus, most mechanistic conclusions drawn on experimental data have been based on work with either aryl phosphate monoesters with generally good leaving groups (see e.g. Cleland & Hengge, 2006; Hengge, 1999; Kirby & Jencks, 1965a; Wolfenden, 2006), or with the methyl phosphate dianion (Lad *et al.* 2003b; Wolfenden, 2006; Wolfenden *et al.* 1998), which has been estimated to have a hydrolysis rate just below the threshold of detectability ($2 \times 10^{-20}\ s^{-1}$ at 25 °C) (Lad *et al.* 2003b). Experimental studies have demonstrated that these compounds exhibit large β_{lg} values (approximately $-1{\cdot}2$) (Kirby & Varvoglis, 1967a), small β_{nuc} values (Kirby & Jencks, 1965a), an inversion of configuration at the phosphorus atom (Friedman *et al.* 1988), small inverse $^{18}k_{nonbridge}$ isotope effects (Hengge, 2002) and a very small activation entropy ($\Delta S^{\ddagger}$) (Hoff & Hengge, 1998b; Kirby & Jencks, 1965a; Kirby & Varvoglis, 1967a, 1968b; Wolfenden *et al.* 1998). Additionally, there is no clear experimental evidence for the formation of free metaphosphate (Cleland & Hengge, 2006) (with the exception of a recent study that examined the kinetics and product distribution of the solvolysis of the *p*-nitrophenyl phosphate dianion in ternary DMSO/alcohol/water mixtures and argued for a dissociative mechanism with the full on intervention of a metaphosphate intermediate (Grigorenko *et al.* 2006; Hu & Brinck, 1999; Lima *et al.* 2012). On this basis, it was generally concluded that the reaction proceeds through a concerted A_ND_N pathway with a very loose TS.

There have been several theoretical studies of the hydrolysis of phosphate monoesters in solution (e.g. Florián *et al.* 1998; Florián & Warshel, 1997, 1998; Grigorenko *et al.* 2006; Hu & Brinck, 1999; Iche-Tarrat *et al.* 2005, 2007; Kamerlin, 2011; Kamerlin *et al.* 2008a; Kamerlin & Wilkie, 2007, 2011; Klähn *et al.* 2006). However, some of these studies only considered a single possible mechanism for the system being studied, ignoring the fact that there exist multiple mechanistic possibilities for the same system. Additionally, one of the few computational studies that has considered both associative and dissociative mechanisms is a gas phase study, where dissociative TSs have been determined in the absence of a nucleophile (Arantes & Chaimovich, 2005). Therefore, the two pathways (associative and dissociative) are not directly comparable. Alternately, a systematic study by Wang *et al.* (2003) has taken the associative hydrolysis of the methyl phosphate dianion into account as a model for phosphate hydrolysis by GTPases (like Ras), but subsequently mainly focused on the dissociative hydrolysis of the monoanion, which is then presented as general evidence for a dissociative mechanism (even though the Ras reaction in reality most likely involves a dianion).

There have, however, been a number of computational studies that have considered both the associative and dissociative hydrolyses of phosphate monoester dianions for the same substrate and nucleophile (Klähn *et al.* 2006), by means of computationally mapping MFJ plots for the systems of interest (Kamerlin, 2011; Kamerlin *et al.* 2008a). In the case of the *p*-nitrophenyl phosphate dianion, it was established that contrary to arguments that the reaction proceeds through a loose dissociative TS (e.g. Hengge, 2002) the TS for this reaction is in fact quite compact, and involves a concerted intramolecular PT that effectively results in hydroxide attack on a phosphate monoanion (Kamerlin, 2011). A similar compact associative pathway was obtained for the hydrolysis of the less activated phenyl phosphate dianion in (Klähn *et al.* 2006); however, in the case of methyl phosphate and pyrophosphate hydrolysis, both associative and dissociative pathways were observed, with a slight preference for a more dissociative pathway (Kamerlin *et al.* 2008a; Klähn *et al.* 2006). Additionally, contrary to arguments that a small activation entropy is suggestive that the hydrolysis proceeds through a TS that is dissociative in character (Hoff & Hengge, 1998b), a small activation entropy can also be fully consistent with

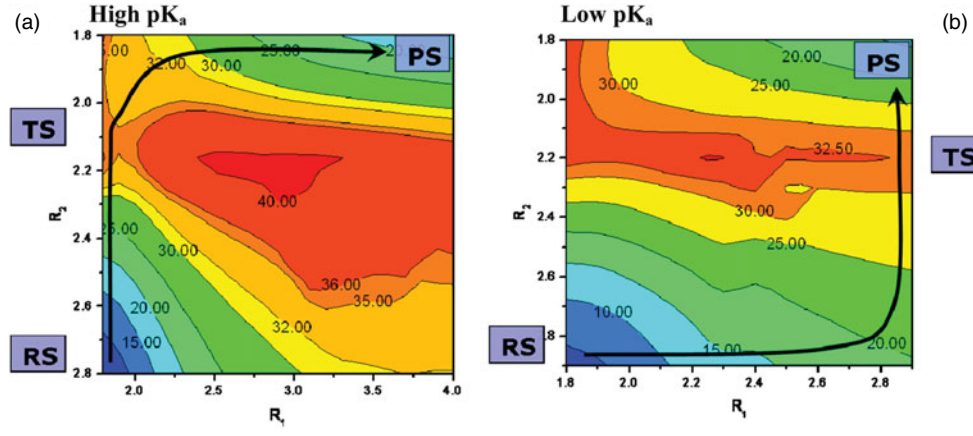

Fig. 12. Sample free-energy surfaces for the hydrolysis of phosphate monoester dianions, adapted from (Klähn *et al.* 2006). Systems with leaving groups with high pK_a prefer a concerted associative pathway (A, left), whereas those with more acidic leaving groups prefer a dissociative pathway (B, right).

an associative TS (Kamerlin, 2011; Kamerlin *et al.* 2008a), and, in fact, both associative and dissociative pathways can have similar activation entropies (Åqvist *et al.* 1999; Kamerlin *et al.* 2008a), complicating their interpretation. Finally, to complicate matters even further, mapping out free-energy surfaces for the hydrolyses of a range of phosphate and pyrophosphate dianions with increasingly acidic leaving groups (Klähn *et al.* 2006) demonstrated that, while the reaction does indeed in all cases proceed through concerted A_ND_N pathways, as was also the case with the examples mentioned above, the precise nature of the TS for the hydrolysis reaction is highly sensitive to the pK_a of the leaving group. For systems with leaving groups with high pK_a, the preferred pathway appears to be associative in nature. However, this gradually changes to an increasingly dissociative pathway, as the leaving group becomes more acidic (see Fig. 12). Nevertheless, despite the progressive switch in mechanism, this systematic study managed to accurately reproduce the observed LFER for this series of compounds. Note that a similar trend was also observed for both aryl (Rosta *et al.* 2008) and fluorophosphate diesters (Alkherraz *et al.* 2010), though in these cases, the leaving groups examined spanned a much narrower pK_a range and thus the mechanistic switch was not as drastic as that shown in Fig. 12 for the phosphate monoester dianions.

The challenges with interpreting experimental LFER will be discussed in greater detail in Section 3.4. However, here it is worth mentioning that it has been traditionally believed that LFER provide a measure of charge buildup at the position of bond cleavage, and thus that the partial charges at the TS can act as a potential mechanistic marker (e.g. Davis *et al.* 1988; Leffler & Grunwald, 1963; Williams, 1984). However, systematically evaluating partial charges at the TSs for the hydrolyses of phosphate monoester dianions has demonstrated that there is, in fact, no clear correlation between the calculated charges and the nature of the TS (Klähn *et al.* 2006). Thus, partial charges can, at best, be considered a bookkeeping description of the observed LFER. Therefore, while computational studies appear to agree with the experimental interpretation of a concerted, A_ND_N TS, the precise nature of the TS is far more complicated, as the reaction can proceed through pathways that are both associative and dissociative in character depending on factors such as the acidity of the leaving group or the precise electrostatic environment. Additionally, it is possible for both pathways to be viable for the same system

(as was the case in e.g. Kamerlin *et al.* 2008a), and that both mechanistic possibilities can be virtually indistinguishable in solution. This highlights the necessity of combining experimental work with careful theoretical studies that are capable of reproducing the available experimental data, in order to elucidate reaction mechanisms.

Again we note that the mechanism of the hydrolysis of phosphate monoester dianions is likely to be much more complex than that usually envisioned by the experimental community (see Section 3.5).

3.1.2 Phosphate monoester monoanions

While the hydrolysis of phosphate monoester dianions is, as discussed in Section 3.1.1, generally considered to proceed through a concerted (A_ND_N) mechanism, metaphosphate was first identified as a viable intermediate for the hydrolysis of phosphate monoester monoanions as early as 1954 (Butcher & Westheimer, 1955). However, although metaphosphate can exist as a stable entity in the gas phase (Grigorenko *et al.* 2006; Henchman *et al.* 1985), its existence in solution can only be inferred by indirect experimental evidence, due to its inherent instability in solution (see e.g. Guthrie, 1977; Henchman *et al.* 1985; Hengge *et al.* 1994; Jankowski *et al.* 1994 among others). Therefore, despite the tendency to interpret the experimental data for the hydrolysis of phosphate monoester monoanions as being indicative of a dissociative pathway, the possibility that this pathway may involve a metaphosphate intermediate has not been ruled out *a priori*.

While there have been extensive computational studies on this class of reaction, they have often given contradictory results (Florián & Warshel, 1997, 1998; Grigorenko *et al.* 2006; Hu & Brinck, 1999; Kamerlin & Wilkie, 2007, 2011; Liu *et al.* 2006; Mercero *et al.* 2000; Patel *et al.* 2011). For instance, both stepwise associative ($A_N + D_N$) (Liu *et al.* 2006; Mercero *et al.* 2000) and dissociative ($D_N + A_N$) (Mercero *et al.* 2000) mechanisms have been proposed for the solution reaction on the basis of computational results. However, firstly, some of these are gas-phase studies that are not relevant to solving mechanistic problems in condensed phases. Secondly (and more critically), very few of these works directly compare associative and dissociative pathways for the same system. Interestingly, a detailed study on the hydrolysis of the methyl phosphate monoanion in the presence of one or two water molecules has suggested that, in the gas phase and in aqueous solution, the dissociative pathway is slightly more favorable than the associative pathway (Hu & Brinck, 1999). However, while elegant, this study overlooks the entropic cost that is associated with the extra water molecules, which can be significant (Kamerlin *et al.* 2009a). Considering the likely entropic effect, it seems to us that, at least in the case of methyl phosphate, both associative and dissociative pathways are equally viable (though this could to some extent be affected by the nature of the leaving group (Kamerlin & Wilkie, 2011), much like the phosphate dianions).

An additional issue that has been a cause of much controversy in the case of phosphate monoester hydrolysis is that of substrate-assisted catalysis, or, more explicitly, whether a 'substrate-as-base' mechanism where the attacking water molecule transfers a proton to the phosphate is a viable mechanism. This mechanism was proposed as a viable pathway for phosphate monoester hydrolysis in 1994 (Schweins *et al.* 1994) and quantified in 1997 (Florián & Warshel, 1997), although it was initially considered highly controversial (Admiraal & Herschlag, 2000). In the context of enzymatic reaction mechanisms, this is a particularly important issue for GTP hydrolyzing enzymes, where the issue of the correct base for activating a nucleophilic water

molecule has been a question of great debate (this is discussed further in Section 4). At present, however, we will be focusing specifically on the hydrolysis in solution. Specifically, water attack on a phosphate monoester can occur in one of two ways:

$$H_2O + ROPO_3H^- \overset{k_1}{\rightleftharpoons} PO_4H_2^- + ROH, \tag{15}$$

$$H_2O + ROPO_3H^- \overset{K}{\rightleftharpoons} OH^- + ROPO_3H \overset{k_2}{\rightleftharpoons} PO_4H_2^- + CH_3OH. \tag{16}$$

The hydrolysis of the monoester can proceed *via* either water attack on a phosphate monoanion (Eq. (15)), with concerted PT to the leaving group, or, alternately, *via* a stepwise process (Eq. (16)), in which the PT from the attacking nucleophile occurs in a pre-equilibrium step, and is subsequently followed by hydroxide attack on a neutral phosphate ester. The two possibilities are shown schematically in Fig. 13, where the concerted pathway is shown in purple, and the stepwise pathway are shown in green. $\Delta g^{\ddagger}_{CON}$ and $\Delta g^{\ddagger}_{SW}$ denote $\Delta g^{\ddagger}$ for the concerted and stepwise pathways, respectively, and ΔG_{PT} represents the free-energy cost associated with transferring a proton from the attacking nucleophile to the phosphate through a pre-equilibrium step. Mathematically, this can be represented as:

$$\Delta g^{\ddagger}_{CON} = \Delta g^{\ddagger}((H_2O + ROPO_3H^-) \rightarrow (PO_4H_2^- + ROH)) \tag{17}$$

for the concerted pathway, and:

$$\begin{aligned}\Delta g^{\ddagger}_{SW} &= 2{\cdot}3RT(pK_a(H_2O) - pK_a(ROPO_3H^-)) + \Delta g^{\ddagger}((HO^- + ROPO_3H_2) \rightarrow (PO_4H_2^- + ROH)) \\ &= \Delta G_{PT} + \Delta g^{\ddagger}_{OH^-}\end{aligned} \tag{18}$$

for the stepwise pathway.

In Eq. (18), the first term (ΔG_{PT}) represents the energetics of the PT step to the phosphate, such that:

$$\Delta G_{PT} = 2{\cdot}3RT(pK_a(H_2O) - pK_a(ROPO_3H^-)). \tag{19}$$

The second term of Eq. (18) represents the barrier for hydroxide attack on the neutral phosphate ($\Delta g^{\ddagger}_{OH^-}$). Thus, the energetics of the stepwise process is simply determined by the pK_a difference between the attacking water molecule and the phosphate. Unfortunately, it is not possible to distinguish between the concerted and stepwise possibilities experimentally, as they have the same kinetic form. Additionally, it is not possible to observe hydroxide and neutral phosphate together simultaneously, as the concentration of neutral phosphate decreases as the concentration of hydroxide ion increases in solution with increasing pH. Therefore, some researchers have rejected the idea of pre-equilibrium PT to the phosphate not by direct experimental evidence of the rate of hydroxide attack on neutral phosphate, but rather, by comparing the rates of the hydrolysis of methyl phosphate (Bunton *et al.* 1958) and trimethyl phosphate (Barnard *et al.* 1961; Bunton *et al.* 1958). Here, the argument is that in order for the stepwise mechanism to be viable, the rate of hydroxide attack on neutral phosphate has to be very fast to compensate for the unfavorable pre-equilibrium step. In the specific case of the methyl phosphate monoanion, for instance, $K = 10^{-14}$ (based on pK_as of 1·65 and 15·7 for methyl phosphate and water, respectively, Florián & Warshel, 1997). Thus, to account for the observed overall reaction rate of $8 \times 10^{-6}\ s^{-1}$ (at 100 °C and pH 4·2, Bunton *et al.* 1958) (i.e. a $\Delta g^{\ddagger}$ of 31 kcal mol^{-1}), k_2, which corresponds to hydroxide attack on the corresponding neutral phosphate monoester, has to be

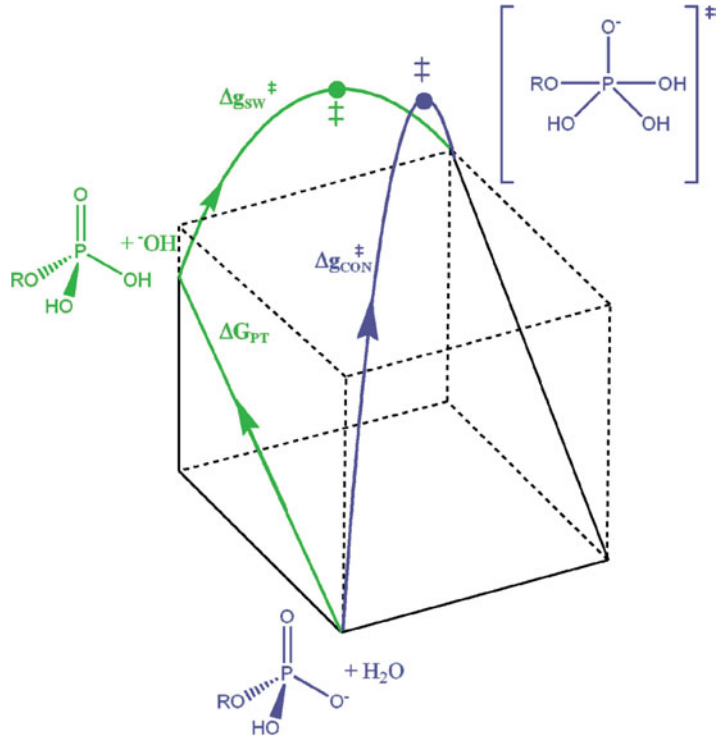

Fig. 13. Schematic 3D representation for stepwise *versus* concerted PT. Here, $\Delta g^{\ddagger}_{\mathrm{CON}}$ and $\Delta g^{\ddagger}_{\mathrm{SW}}$ represent the $\Delta g^{\ddagger}$ for concerted and stepwise pathways respectively, and ΔG_{PT} denotes the free energy required for PT to the phosphate through a pre-equilibrium step. The concerted pathway is depicted in purple, and the stepwise pathway in green. Finally, the positions of the TSs on this figure are only provided for illustration, and the actual TS can occur at any point along the pathway. This figure was originally presented in (Kamerlin *et al.* 2008b).

very fast (about $10^7\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$; see Bunton *et al.* 1958) to account for the added cost of the initial PT step in a stepwise mechanism. This in turn corresponds to an activation free energy of approximately 10 kcal mol^{-1} from TS theory. As discussed above, it is not possible to measure k_2 directly. However, the rate for OH^- attack on the corresponding phosphate triester, trimethyl phosphate, can be experimentally measured, and is equivalent to $3 \times 10^{-2}\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$ at 100 °C (Barnard *et al.* 1961; Bunton *et al.* 1958) (which corresponds to a barrier of about 25 kcal mol^{-1}). Since this rate constant is more than eight orders of magnitude smaller than the value needed to compensate for the above K that corresponds to ΔG_{PT} in Eq. (18) that is, $K^*k_2 = 10^{-7}$, or, alternatively, the sum of ΔG_{PT} and the presumed $\Delta g^{\ddagger}_{\mathrm{OH}^-}$ is about $19{\cdot}3 + 25 = 44{\cdot}3$, which is much larger than the expected 31 kcal mol^{-1}), it has been considered justified to rule out the possibility of hydroxide attack in the phosphate monoester hydrolysis reaction. A similar argument was also made based on comparing the reactions of the monoanions of 2,4-dinitrophenyl phosphate and its diester analog, methyl-2,4-dinitrophenyl phosphate (Admiraal & Herschlag, 2000). However, the overwhelming problem with this interpretation is that it is based on the assumption that a methyl group is a valid substitute for a proton. This assumption is not necessarily true, and its validity can *only* be resolved computationally (in order to measure the rate of OH^- attack one needs to use a basic environment, where the phosphate is simply unprotonated). In fact, a careful computational study has demonstrated that unlike hydroxide attack

on the methyl triester, the rate of hydroxide attack on the corresponding neutral phosphate monoester is in fact very fast ($\Delta g^{\ddagger}_{sw} = 11{\cdot}7$ kcal mol^{-1} Florián & Warshel, 1997), and significantly more so than the reaction of the corresponding methyl triester. Therefore, this is actually a viable pathway for the hydrolysis of phosphate monoesters. As shall be seen in Section 3.2, similar results were obtained for the hydrolysis of phosphate diesters. Resolving this controversy is central to understanding, for instance, the mechanism of action of G-proteins (discussed in Section 4). However, a general consensus is yet to be reached, and despite overwhelming evidence to support it, arguments are still brought against it (Lassila *et al.* 2011) (see again discussion in Section 4). Nevertheless, signs that this may be changing have started to appear in the literature, for instance, in a study proposing that while the stepwise pathway is unlikely to apply to aryl phosphates, it is a feasible possibility for the hydrolysis of alkyl phosphate esters (Iché-Tarrat *et al.* 2007).

3.1.3 Neutral phosphate monoesters

As seen in Section 2, the protonation state of a phosphate monoester is pH dependent. While the dianion dominates at high pH, and the monoanionic species at neutral pH, the fully protonated species only exists under very acidic conditions. However, unfortunately, outside of the studies that addressed the substrate-as-base mechanism, virtually no work has been done on the reactivity of these compounds.

3.2 The mechanism(s) of phosphate diester hydrolysis

While phosphoryl transfer in general is ubiquitous in biology, phosphate diesters are particularly important, as they are used to hold together the genetic code in DNA. Also, they form the functional group that is acted on by enzymes such as DNA polymerases (a family of enzymes which are key for maintaining the integrity of the genome, by regulating a wide variety of activities, including DNA recombination, repair and damage bypassing (Hübscher *et al.* 2002), as well as having been found in a high percentage of human carcinomas, suggesting a link between e.g. DNA polymerase β-activity and carcinogenesis (Lang *et al.* 2004; Wang *et al.* 1992). Understanding the precise mechanism of phosphate diester hydrolysis in solution is key to being able to fully understand the precise catalytic mechanism of diesterases such as DNA polymerase, which, at this point, remains unclear (Arndt *et al.* 2001; Florián *et al.* 2003a, 2005).

Experimental studies on phosphate diester hydrolysis have been complicated, however, by the fact that the rate of the solution reaction is extremely slow (Bunton *et al.* 1960; Kirby & Younas, 1970c; Kumamoto *et al.* 1956). As an example, the rate for the uncatalyzed hydrolysis of dimethyl phosphate at 25 °C (Radzicka & Wolfenden, 1995) has an estimated rate constant of approximately 2×10^{-13} s^{-1}. Further complicating studies of phosphate hydrolysis is the fact that the reaction can not only proceed via P–O cleavage but also via C–O cleavage, which, in this case has been shown to be responsible for up to 99% of the reactivity (and is not the site of cleavage by the relevant enzymes). This suggests an upper limit of approximately 1×10^{-15} s^{-1} (again at 25 °C) on the rate constant for spontaneous P–O cleavage (Wolfenden *et al.* 1998).

Values of β_{lg} for phosphate diester hydrolysis that have been cited in the literature range from $-0{\cdot}64$ to $-0{\cdot}97$ (Ba-Saif *et al.* 1989b; Chin *et al.* 1989a; Kirby & Younas, 1970c; Liao *et al.* 2001a; Williams *et al.* 1998a; Zalatan & Herschlag, 2006). This, combined with the presence of normal ^{18}O KIEs in the scissile oxygen atom of such reactions (Cleland & Hengge,

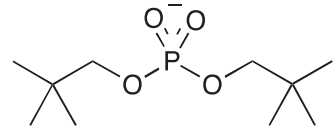

Fig. 14. Dineopentyl phosphate (Np_2P).

2006) has led to the suggestion that the reaction proceeds *via* a concerted mechanism with a TS that is 'tighter than that for monoesters' (Cleland & Hengge, 2006; Hengge & Cleland, 1991; Hengge *et al.* 1995; Kirby & Younas, 1970b, 1970c; Zalatan & Herschlag, 2006). However, this assumption does not take into account the fact that both associative and dissociative pathways are viable for phosphate monoester hydrolysis, and thus that this comparison is a moot point. Additionally, while there are several theoretical studies of phosphate diester hydrolysis (many of which are reviewed in Zhou & Taira, 1998), these are fraught with the same difficulties as studies on phosphate monoesters. Most of these works are gas-phase studies that do not take the possibility of multiple mechanisms into account. Firstly, as was also the case with the monoesters, it is highly questionable if gas-phase simulations of phosphate diester hydrolysis actually have any relevance to solving biological problems, most of which occur in solution. In fact, such gas-phase works often yield very large reaction barriers of >80 kcal mol^{-1}, at which point the significance of a difference between a local minimum (intermediate) of a few kcal mol^{-1} and its absence becomes negligible. Also, the barrier for phosphate diester hydrolysis in solution tends to be much lower than that obtained by such gas-phase studies. For instance, computational studies on the hydrolysis of dimethyl phosphate have obtained an activation free energy in the gas phase of 96 kcal mol^{-1} even though the experimental value in solution is significantly lower at 32 kcal mol^{-1} (Dejaegere *et al.* 1994). Much more relevant information has been provided by a computational study (Rosta *et al.* 2008) that mapped the full free-energy surface in solution and also took the activation entropy of the reaction into account for the hydrolysis of a series of homologous substituted methyl phenyl phosphate diesters. That is, it was shown that the TS character is again dependent on the pK_a of the leaving group, with a compact associative A_ND_N TS at high pK_a that becomes progressively more expansive as the leaving group becomes more acidic (despite reproducing the LFER with chemical accuracy). Interestingly, this study covered a much smaller pK_a range than that of the monoester dianions (Klähn *et al.* 2006); however, the overall correlation between leaving group pK_a and mechanism appears to be similar, and qualitatively similar TSs have recently been obtained using QM/MM studies (Hou & Cui, 2012).

It is important to note that with the phosphate monoesters, it is possible that the mechanism of phosphate diester hydrolysis is also dependent on pH (i.e. protonation state). Dineopentyl phosphate (Np_2P, Fig. 14) is an interesting model system for addressing this question by means of kinetic studies, as phosphate diester hydrolysis through C–O cleavage is strerically prevented, and the reaction proceeds almost exclusively through P–O cleavage (Schroeder *et al.* 2006).

More interestingly, kinetic studies on dineopentyl phosphate hydrolysis (Schroeder *et al.* 2006) have obtained virtually identical experimental parameters for the base-catalyzed (OH^- attack on Np_2P^-) and pH independent (H_2O attack on Np_2P^-) hydrolyses of this compound, giving rise to the possibility that the pH independent reaction does not involve water attack on the Np_2P^- monoanion but rather proceeds through the kinetically equivalent mechanism of hydroxide attack on the neutral phosphate diester (Fig. 15). Mapping the full free-energy surfaces for this

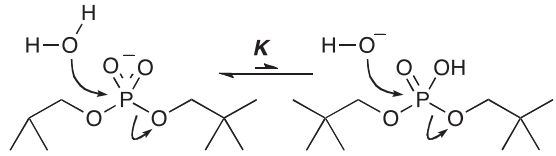

Fig. 15. Alternative kinetically equivalent mechanisms for the pH-independent hydrolysis of the Np_2P anion (Schroeder *et al.* 2006).

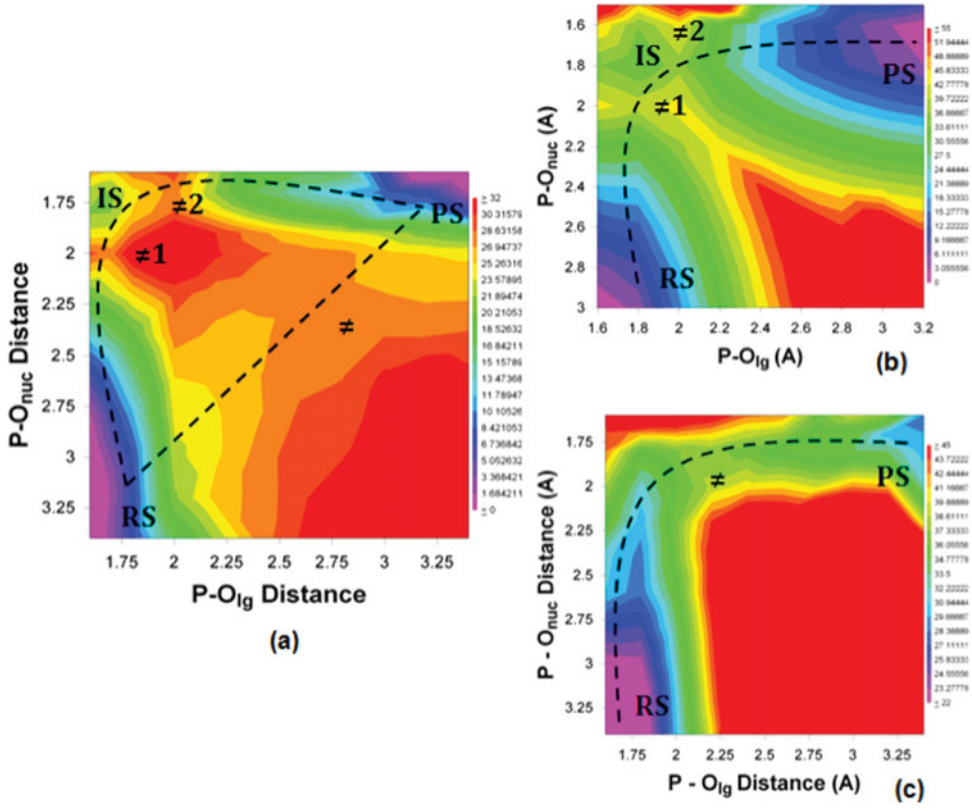

Fig. 16. Calculated free-energy surfaces for (*a*) the acid-catalyzed, (*b*) the pH-independent, and (*c*) the base catalyzed hydrolysis of dineopentyl phosphate. RS, IS, and PS denote reactant, intermediate and product states, respectively. ≠1 and ≠2 denote TSs for P–O_{nuc} formation and P–O_{lg} cleavage in an $A_N + D_N$ pathway, respectively, and ≠denotes a single concerted A_ND_N TS. This figure is adapted from (Kamerlin *et al.* 2008b).

reaction for the acid-catalyzed, pH-independent and base catalyzed reactions has demonstrated, however, that the precise mechanism is dependent on pH (Kamerlin *et al.* 2008b) (Fig. 16). At high pH, the reaction proceeds exclusively through a compact A_ND_N pathway, which involves hydroxide attack on monoanionic N_2P^-. The pH-independent reaction, on the other hand, proceeds exclusively via an $A_N + D_N$ mechanism with stepwise PT from the attacking nucleophile onto the phosphate prior to hydroxide attack on a neutral phosphate monoester. Finally, the acid catalyzed reaction at low pH can proceed through, either a stepwise $A_N + D_N$ pathway or a concerted A_ND_N pathway, with virtually identical activation parameters for both pathways.

Conceptually, it is quite simple to account for the mechanism change with pH (Kamerlin *et al.* 2008b). In the case of the base-catalyzed reaction, both incoming nucleophile and departing leaving group are anionic, with no PT between nucleophile and substrate. In this case, the reaction follows a simple A_ND_N mechanism that is in good agreement with previous experimental and computational studies of phosphate diester hydrolysis (Ba-Saif *et al.* 1989a; Chin *et al.* 1989b; Kirby & Younas, 1970c; Liao *et al.* 2001b; Rosta *et al.* 2008; Williams *et al.* 1998b; Zalatan & Herschlag, 2006). However, the pH independent and acid-catalyzed reactions are slightly more complicated. Here, we clearly observe a 'substrate-as-base' mechanism with PT from an attacking water molecule to the substrate, which is either a monoanionic (pH-independent) or neutral (acid catalyzed) species. However, the leaving group is also protonated (i.e. an alcohol) upon departure. Therefore, there is also a PT reaction to the leaving group, as expected from KIEs (Cassano *et al.* 2002; Hengge, 2002). Considering the fact that in all obtained TSs for Np_2P hydrolysis (Kamerlin *et al.* 2008b), the distance between O_{nuc} and O_{lg} is greater than approximately 4 Å, it is highly unlikely that this PT occurs directly from the nucleophile onto the leaving group, which is also the only proton source in the pH-independent reaction. For this system, this issue is easily circumvented by proceeding through a stepwise addition–elimination ($A_N + D_N$) mechanism in which the first step (bond formation between the phosphate and the incoming nucleophile) is accompanied by PT to the phosphate to form a neutral phosphane intermediate, from which the leaving group can then extract a proton in the second step (leaving group departure). As with the phosphate monoesters, it is *impossible* to experimentally distinguish between concerted water attack on a monoanionic diester and hydroxide attack on neutral phosphate (with a pre-equilibrium PT, as shown in Fig. 15). The energy cost of PT from the attacking water molecule to the phosphate is approximately 18 kcal mol^{-1} (as calculated from Eq. (19) based on a pK_a of approximately 3·6 for N_2P). However, it was demonstrated in this work that as with monomethyl phosphate (Florián & Warshel, 1997), the rate for hydroxide attack on the neutral species is actually quite fast, compensating for the cost of ΔG_{PT} (Kamerlin *et al.* 2008b). In the case of the neutral species, the reaction is also observed to proceed via a stepwise mechanism involving a fully protonated phosphorane intermediate. However, the availability of an extra proton on the phosphate now allows for a concerted A_ND_N mechanism, in which the leaving group extracts this proton upon departure to yield P_i and alcohol. Both pathways have identical activation parameters, suggesting the when there is a proton available for facile extraction by the departing leaving group, both stepwise and concerted pathways are equally viable with the system showing no preference for either. However, the key take home point is the fact that, as with the monoesters, the mechanism for diester hydrolysis is sensitive to factors such as leaving group pK_a and pH, resulting in multiple mechanistic possibilities, one of which is a substrate as base mechanism, which is discussed in much greater length in Section 4.

3.3 The mechanism of phosphate triester hydrolysis

Several mechanistic studies have been performed on both acyclic triesters with aryl leaving groups, as well as six-membered ring cyclic triesters (Khan & Kirby, 1970), which suggest that the hydrolysis of phosphate triesters is faster than that of either phosphate mono- or diesters (though their five-membered counterparts have even faster reaction rates, Cleland & Hengge, 2006). In the case of the simple aryl esters, the reaction has been shown to proceed preferentially (although not exclusively) through a stepwise $A_N + D_N$ mechanism involving a phosphorane intermediate (Khan & Kirby, 1970). However, studies on similar structures that vary both

nucleophile and leaving groups have supported the proposition that both stepwise and concerted mechanisms are viable, and that the stereochemistry of the reaction (i.e. retention *versus* inversion of configuration) depends not only on the leaving group and nucleophile but also on the solvent, any hetereoatoms in the six-membered ring, as well as any counterions that may be present (Hall & Inch, 1980; Rowell & Gorenstein, 1981). Acyclic triesters with aryl leaving groups, however, have been suggested to proceed through exclusively concerted TSs that become tighter with leaving group basicity (Ba-Saif *et al.* 1990, 1991) (much like their mono- (Klähn *et al.* 2006) and diester (Rosta *et al.* 2008) counterparts).

3.4 General considerations

3.4.1 Interpretation of LFER

LFER are a widely used experimental marker for determining the mechanism of both solution and enzymatic phosphate ester hydrolysis (Ba-Saif *et al.* 1989b; Chin *et al.* 1989a; Hollfelder & Herschlag, 1995b; Holtz *et al.* 2000; Kirby & Younas, 1970a; Lad *et al.* 2003a; Liao *et al.* 2001a; O'Brien & Herschlag, 2001; Williams *et al.* 1998a; Zalatan & Herschlag, 2006), and are considered by experimentalists to be a powerful tool for the characterization of TS structure (Jencks, 1969; Williams, 1992). The principle here is simple: kinetic measurements are performed on a series of homologous compounds, and the logarithm of the rate constant is correlated to the pK_a of the leaving group or nucleophile. If these compounds are hydrolyzed by identical mechanisms, a linear relationship is expected to arise, and the mechanism can be deduced based on the gradient of this correlation, which can be represented as the Brønsted coefficient, β, where β_{lg} and β_{nuc} denote dependence on the leaving group and nucleophile, respectively. The corresponding Brønsted coefficients can then be defined by Eq. (20):

$$\beta_{nuc/lg} = \frac{d \log k}{dpK_a^{nuc/lg}} = -\frac{\Delta\Delta g^{\ddagger}_{tot}}{1{\cdot}36\Delta pK_a^{nuc/lg}}, \tag{20}$$

where $\Delta\Delta g^{\ddagger}_{tot}$ denotes the overall (rate-limiting) activation barrier for the reaction.

From the discussion in this Sections 3.1–3.3, it was seen that phosphoryl transfer reactions can proceed through multiple distinct TSs, which is contrary to the classical interpretation of LFER, where it is assumed that each of the potential mechanisms (whether associative or dissociative) will proceed through merely a single TS. Thus, in order to accurately examine the LFER, it is important to consider the effect of the free-energy differences along each step of the pathway on the corresponding activation barriers (Åqvist & Warshel, 1993; Warshel *et al.* 1992a, 1994). Changing the pK_a of either the nucleophile or leaving group not only affects the energetics of any steps involving PT with that particular group, but also affects the pK_a of the attached phosphate group and thus the overall free energy. If one examines a hypothetical reaction:

$$XO-PO_3^{2-} + H_2O \rightleftharpoons HO-PO_3^{2-} + XOH \tag{21}$$

the effect of the leaving group on the overall free energy of the reaction can be given by the following empirical equation presented by Bourne and Williams (Bourne & Williams, 1984):

$$\Delta G^0_{rxn} \cong 0{\cdot}5pK_a^{XOH} - 7{\cdot}8 \tag{22}$$

(it should be noted that here, ΔG is obtained in kcal mol^{-1}, after a correction for the 55 M concentration of water). Finally, changing the pK_a of XOH when it is attached to the phosphate

group causes an overall shift in the pK_a of the phosphate, which can be represented as (Bourne & Williams, 1984; Guthrie, 1977):

$$\Delta pK_a(\text{phosphate})/\Delta pK_a\ (\text{lg}) \sim 2{\cdot}3 \tag{23}$$

Figure 17 shows a more detailed energetic picture, where shifts in both nucleophile and leaving group pK_a values can be related to changes in activation barriers. This hypothetical scenario has been discussed in great detail by Åqvist *et al.* (1999). This figure depicts three different sections of the overall energy surface: one along the associative corners (i→ii→iv, pre-equilibrium PT, Fig. 17 *a* and *b*), one along the dissociative corners (i→iii→iv, unimolecular elimination, Fig. 17 *c* and *d*) and finally one along a concerted pathway (i→iv, Fig. 17 *e* and *f*). These have all been drawn as the intersection of relevant parabola (including parabola representing PTs), in order to clarify the relationship between this and Marcus-type LFER treatments (Åqvist & Warshel, 1993; Warshel *et al.* 1992a, 1994) (in which TS energies are correlated with the crossing of the relevant parabola). In all cases, PTs have been depicted as stepwise processes with distinct resonance structures in order to more clearly illustrate the effects of pK_a shifts on the energy of each state. In actual fact, these PTs may be either stepwise or concerted, however, where they to occur concomitantly with bond making or breaking to the phosphorus atom, this would simply reflect the mixing of the relevant resonance structures. Finally, it is not clear how the PT from the nucleophile to the leaving group would proceed in the dissociative case; however, the pK_a difference between the two groups will nevertheless be involved in the energetics, as illustrated in Figs 17 *c* and *d*.

These relationships can in turn be analyzed by applying Marcus' relationship to each successive step that is, for steps involving crossing between two adjacent resonance structures $i \rightarrow j$, we get:

$$\beta_{i\rightarrow j} = \frac{\mathrm{d}(\Delta g)^{\ddagger}_{i\rightarrow j}}{\mathrm{d}\left(\Delta G^0_{i\rightarrow j}\right)} = \frac{\mathrm{d}\left[\left(\Delta G^0_{i\rightarrow j} + \lambda_{i\rightarrow j}\right)^2\right] =}{\mathrm{d}\left(\Delta G^0_{i\rightarrow j}\right)} = \frac{\Delta G^0_{i\rightarrow j}}{2\lambda} + \frac{1}{2}. \tag{24}$$

Here, λ is simply the reorganization free energy. If we assume that this value is significantly larger than ΔG^0, we obtain the relationship:

$$\beta_{i\rightarrow j} = \frac{\Delta\Delta g^{\ddagger}}{\Delta\Delta G^0} \approx \frac{1}{2}. \tag{25}$$

Note that in actual chemical reaction, one has to use the modified Marcus relationship (see e.g. Warshel *et al.* 1992b), but the change in the barrier, rather than the absolute barrier, still follows the Marcus' relationship.

In order to obtain the correct LFER for each possible mechanism considered, it is necessary to express the $\Delta\Delta G^0$ of all resonance structures in terms of either ΔpK_a^{lg} or ΔpK_a^{nuc}.

Figures 17(*a*) and (*b*) show the associative reaction, which is rate determining in terms of P–O cleavage to the leaving group (shown as a late TS in Fig. 17 *a*). In both cases, ΔG^0 changes due to the lowering of the pK_a of the nucleophile (Fig. 17 *a*) and leaving group (Fig. 17 *b*) are drawn relative to the product energy. If the nucleophile pK_a is reduced, the reactant state is also reduced by $0{\cdot}5\Delta pK_a^{nuc}$ kcal mol^{-1}. As the free-energy difference between the first two states is determined by the difference in pK_a between the phosphate group and the nucleophile, the energy of the second state is then reduced by $(1{\cdot}36 + 0{\cdot}5)\Delta pK_a^{nuc}$. Finally, the energy of state (4) will also drop

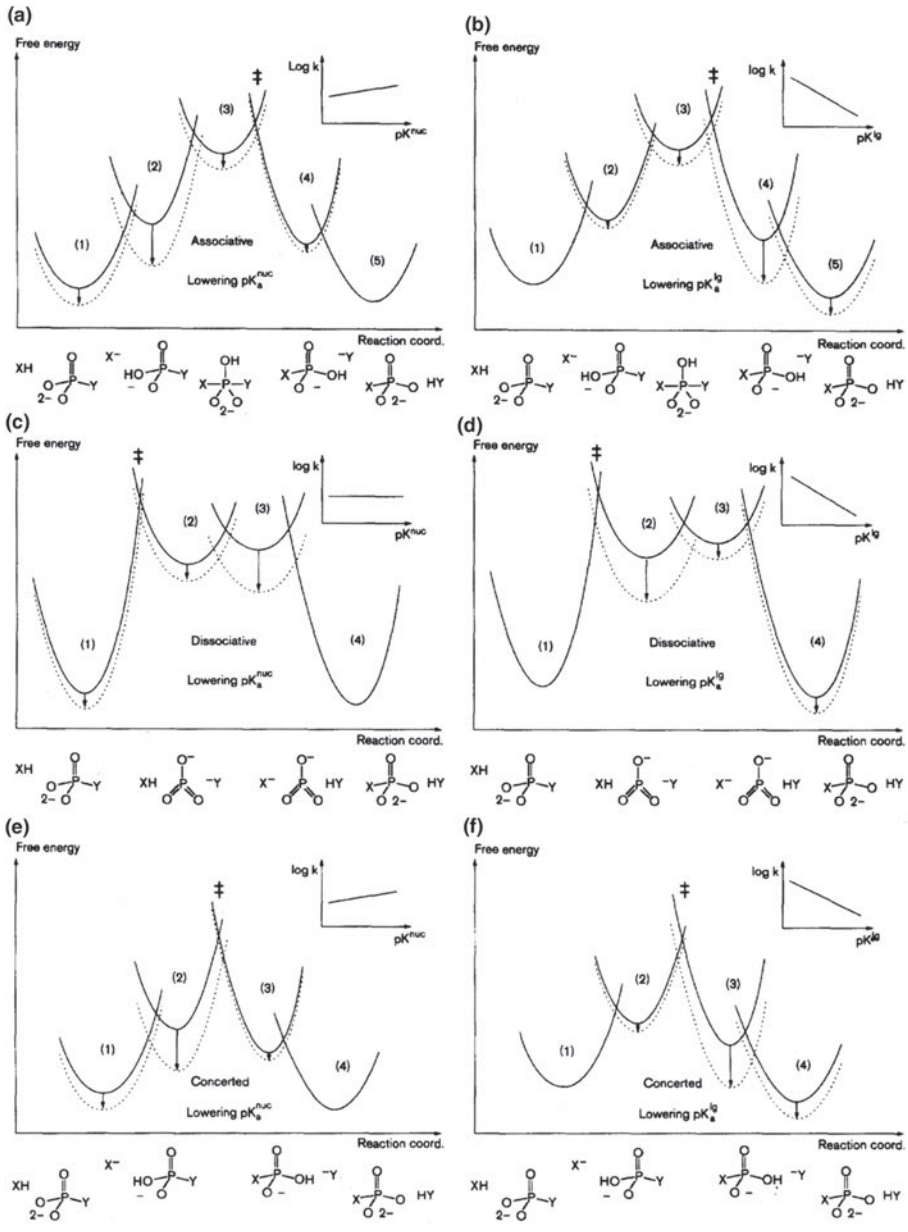

Fig. 17. Free-energy diagrams depicting different diabatic states along (*a*, *b*) a stepwise associative pathway involving pre-equilibrium PT to the phosphate group, (*c*, *d*) a stepwise dissociative pathway involving unimolecular elimination of metaphosphate from the dianion, and (*e*, *f*) a concerted pathway also involving PT. Reactions *a*–*d* are rate limiting by leaving group bond fission. The dashed curves show the shift in free energy upon reducing the pK_a of the nucleophile (*a*, *c*, *e*) or leaving group (*b*, *d*, *f*). Taken with minor corrections from (Aqvist *et al.* 1999).

relative to the product, because, in this particular case, pK_a^{phs} changes by $0{\cdot}23\Delta pK_a^{nuc}$ (Bourne & Williams, 1984; Guthrie, 1977), yielding a free-energy shift of $1{\cdot}36 \times 0{\cdot}23\Delta pK_a^{nuc}$ or basically $0{\cdot}3\Delta pK_a^{nuc}$. As the main effect of lowering the nucleophile pK_a is exerted on the second state (which is anyhow not rate determining), this results in an overall small positive value of β_{nuc}. In

contrast, we have the situation in Fig. 17*b*, where instead of the nucleophile pK_a, we are now lowering the leaving group pK_a. Now, the main effect is found in the fourth state, which is also directly involved in the curve crossing determining the barrier for late TS. As in the previous situation, the free energy of this state is lowered by $(1{\cdot}36+0{\cdot}5)\Delta pK_a^{lg}$ relative to the reactant state. Here, there are two possible extreme cases: in the first, the high-energy state (3) is lowered by the same amount as the fourth state. Thus, the rate limiting TS, which is determined by the crossing of these two states, will be subjected to the same shift. This will yield a maximum negative value of β_{lg} that corresponds to $\beta_{lg}=(-1{\cdot}36+0{\cdot}5)/1{\cdot}36$, or $-1{\cdot}4$. This is because the entire lowering of the third state can be expressed as a reduction of the activation barrier relative to the reactants. In the other extreme case, however, were the third high-energy state is entirely unaffected by lowing the pK_a of the leaving group, we would expect a value of $-0{\cdot}7$ for β_{lg}. Thus, in both cases, we observe a large negative value of β_{lg}.

The dissociative mechanism can in turn be seen in Figs 17*c* and *d*, which represent the effects of modifying the nucleophile and leaving-group pK_a values on the overall free energy, respectively. It should be noted that for a completely dissociative mechanism with an early TS, both the first and second state will be shifted by exactly the same amount, thus resulting in a value of zero for β_{nuc}. On the other hand, as the early TS is determined only by the crossing of the first and second states, once again, the relationship between $\Delta G^{\ddagger}$ and ΔG^0 can be elucidated by the Marcus-type relationship shown in Eq. (24). As with the associative mechanism, both the product state 4 and state 3 will be lowered by $0{\cdot}5\Delta pK_a^{lg}$ kcal/mol relative to the reactant state. Also, the free-energy difference between the second and third states is given by $1{\cdot}36\Delta pK_a^{lg}$, thus the second state is shifted by $(1{\cdot}36+0{\cdot}5)\Delta pK_a^{lg}$ kcal/mol. By applying Eq. (25) to the crossing of the first and second states, we thus get a value of $-0{\cdot}7$ (i.e. $0{\cdot}9/1{\cdot}36$) for β_{lg}, which is in the same range expected for associative mechanisms.

It should be noted that the analysis presented here is not related to any form of presumption with regard to the nature of state 3 in Fig. 17*a* and *b*, but rather, the same analysis could be applied to a transient dissociative-like intermediate (i.e. one that has low axial bond orders between the incoming nucleophile/departing leaving group and the phosphorus atom), as long as pre-equilibrium PT to the phosphate is considered. Here, β_{lg} would still be large and negative, but for a TS that is associated predominantly with the incoming nucleophile rather than with the departing leaving group. Thus, the protonation state of the phosphate group in the TS is as important as the bond order it has to its axial ligands. Hence, the question is, what information can one actually gain from LFER? From this analysis, it can be seen that it is not possible to easily distinguish between associative and dissociative mechanisms based on LFER – that is, if the reaction follows an extreme mechanism, the LFER can easily be quantified by shifting relevant parabola. However, in contrast, shifting the parabolas for different mechanistic possibilities can lead to similar LFER. This is best illustrated in recent computational studies of both phosphate monoester (Klähn *et al.* 2006) (Fig. 18); and diester (Rosta *et al.* 2008) hydrolysis, mentioned in Section 3.2, which demonstrated that the mechanism is highly dependent on the pK_a of the leaving group: that is, in both cases, leaving groups with a high pK_a tend to prefer a more associative mechanism, which becomes progressively more dissociative as the leaving group becomes more acidic. Overall, as shown in Fig. 18 (and Fig. 11 of Rosta *et al.* 2008), we were able to computationally reproduce the observed trend over a very wide range of rate constants. Our point is that the rate determining TS are very different for the different compounds considered in our LFER studies. Thus, it seems clear that the experimental assumption that, in the case of LFER with no breaks, a given compound follows a given LFER cannot be

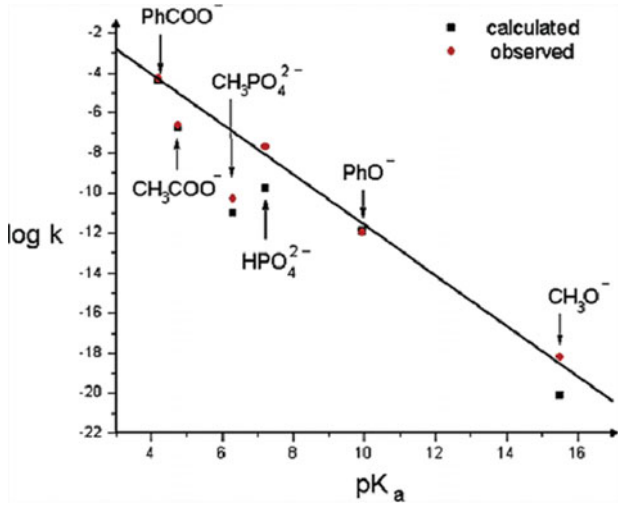

Fig. 18. A comparison of experimental (circles) and calculated (squares) LFER for the hydrolysis of a series of phosphate and pyrophosphate monoesters in aqueous solution. This figure was originally presented in (Rosta *et al.* 2008).

used to deduce the nature of the corresponding TS, as, provided that the change in TS is sufficiently gradual and correlates with the change in rate constant, this will be masked by the LFER (which could, in principle, even mask a full switch from a concerted to stepwise pathway and vice versa, as long as the rate of the rate determining step is following the trend of the previous compounds. It is worth touching on the opposite scenario, which has recently been observed as well, and which will be expanded on more in Section 4.3: where a clear split is observed in LFER for homologous mono- and dihalogenated substituents, which would superficially imply a change in mechanism, but which, upon further examination, appears to arise out of the difference in solvation effects for mono- and dihalogenated substrates, with all TSs being qualitatively similar (Kamerlin *et al.* 2009c). This provides an example of the opposite scenario, where drastic changes in LFER do not correspond to large changes in mechanism. In other words the interpretation of LFER as mechanistic markers is at best ambiguous, and, in the absence of careful theoretical studies, should be approached with care. This is particularly significant in lights of attempts, for instance, to rationalize catalytic promiscuity based on experimentally observed LFER, which will be discussed at greater length in Section 4.5.

3.5 Interpretation of effective charges

Traditionally, it has been assumed that the LFER observed in phosphate hydrolysis reactions can be used to define the effective charges on the TS (Davis *et al.* 1988) (Leffler & Grunwald, 1963; Williams, 1984). It has been assumed that the slope of the LFER (i.e. the Brønsted $\alpha-$value) or the ratio between β_{lg} and β_{nuc} can be used to measure the amount of charge build up at the position of bond cleavage. Additionally, Brønsted parameters and Leffler indices (Ba-Saif *et al.* 1989a) have been used by experimentalists to characterize reaction progress as if they report on bond order. More strictly, this analysis is linked to the fractional effective charge change that is reached in the TS, relative to complete reaction, although it should be noted that the value of β_{lg} is affected by such external factors like hydrogen bonding, general acid/base

catalysis, metal ion interactions, or solvation effects, and thus it might not be used as a reliable measure of bond order.

Such an argument correlating LFER to the effective charge on the TS may seem reasonable, particularly in the case of an intersection of two VB states. However, a detailed computational study of a series of phosphate dianions with different leaving groups (Klähn *et al.* 2006) demonstrated that the partial charges in the TS do not follow the conventional beliefs of physical organic chemistry. For instance, for the methyl phosphate dianion, the leaving group charges were found to be −0·22 and −0·68 for the associative and dissociative TSs, respectively, whereas for the methyl pyrophosphate dianion, the leaving group charges were found to be −0·30 and −0·45 for the associative and dissociative TSs, respectively. These are just examples and several systems were studied leading to the same conclusions. Based on this detailed analysis, it was demonstrated that the effective charges in the TS cannot be used to describe the different mechanisms in a unique way as the effective charges are quite similar. Thus, while the effective charges can provide a useful way of describing the change in charge of the leaving group for simple systems which can be uniquely represented by two intersecting parabola, in more complex cases they can at best act as a bookkeeping tool that provides a rough guideline as to the type of mechanism that can be expected from the LFER.

3.6 Interpretation of experimentally observed activation entropies

As ambiguous as the interpretation of LFER is the interpretation of activation entropies in terms of the elucidation of reaction mechanisms. On the one hand, it can generally be assumed that bimolecular nucleophilic substitution reactions involve a loss of translational entropy for the reactants, and thus, if the entropy were only dependent on this, $\Delta S^{\ddagger}$ could be used to distinguish between purely uni- and bimolecular TSs. However, apart from the fact that this is an oversimplified analysis (as shall be seen, entropy is not only dependent on bond order), while this argument could help distinguish between effectively stepwise ($A_N + D_N$) and concerted ($A_N D_N$) TSs, it is of little help in a situation where the reaction proceeds through either of two associative or dissociative $A_N D_N$ TSs, as was seen in previous studies of phosphate monoesters (Kamerlin *et al.* 2008a; Klähn *et al.* 2006).

Despite this, however, it has been claimed that a dissociative pathway will have a very small $\Delta S^{\ddagger}$, whereas an associative pathway will have a larger, more positive $\Delta S^{\ddagger}$ (Hoff & Hengge, 1998a; Kirby & Varvoglis, 1967b, 1968a; Kirby & Jencks, 1965b; Wolfenden *et al.* 1998), and thus, that a near-zero activation entropy is a marker for a dissociative TS. However, particularly in a case where both viable reaction pathways are concerted, this is clearly a problematic assumption. A recent computational study $-T\Delta S^{\ddagger}_{\mathrm{conf}}$ (Kamerlin *et al.* 2008a) examined precisely this issue, by generating full free-energy surfaces for the hydrolyses of the methyl phosphate dianion as well as the methyl pyrophosphate trianion by use of *ab initio* approaches, and demonstrated that in both cases, hydrolysis can proceed through either a tight associative TS or a more expansive dissociative TS, and that in each case, both pathways are virtually indistinguishable in solution merely on the basis of $\Delta g^{\ddagger}$. The activation entropy for each TS was subsequently evaluated such that the configurational entropy (solute contribution) was simulated microscopically using a combination of the quasiharmonic and restraint release approaches (Sharma *et al.* 2005; Strajbl *et al.* 2000), respectively, which evaluate the entropy using sampling from molecular dynamics runs in the presence of explicit water molecules, and the solvation entropy (solvent contribution) was evaluated using the Langevin dipoles solvation model

(Florián & Warshel, 1999). In doing so, it was demonstrated that in both cases, not only do both the associative and dissociative pathways have very similar overall barriers, but they also have very similar activation entropies (that are in good agreement with the experimental value). Most importantly, even the associative TS has a very small activation entropy, i.e. a near-zero activation entropy can equally be indicative of an associative pathway. Therefore, when considering activation entropies, bond order is not the only factor that needs to be taken into account. The overall activation entropy is also determined by issues such as steric factors that determine the configurational volumes available to the reactants during the reaction, solvation and desolvation effects that may be associated with charge redistribution upon approaching the TS, entropy changes associated with intramolecular degrees of freedom as the TS is approached, and, if PT accompanies nucleophilic substitution, this will also contain a significant entropy term contributing to the overall activation entropy (Åqvist *et al.* 1999). Thus, while entropic contributions are clearly very important when examining reaction mechanisms and not to be overlooked the use of entropic effects to elucidate mechanism is necessarily effective. The entropy cannot be interpreted solely in terms of bond order at the TS, and therefore, a near-zero activation entropy is not necessarily indicative of a dissociative mechanism, despite the fact that this is what the qualitative interpretation of the experimental data would suggest (Hoff & Hengge, 1998a; Kirby & Varvoglis, 1967b, 1968a; Kirby & Jencks, 1965b; Wolfenden *et al.* 1998).

3.7 Interpretation of isotope effects

KIEs were discussed in detail in Section 2. It has traditionally been assumed that the ratio between ^{16}O and ^{18}O isotope reaction rates (i.e. $^{16}k/^{18}k$) can be used to give information on the change in P–O bonding between the ground and TSs, and, thus, $^{16}O/^{18}O$ isotope effects have been extensively used in attempts to determine the nature of the TSs for phosphoryl transfer (Cleland & Hengge, 2006; Du *et al.* 2004; Hengge *et al.* 1994; Knight *et al.* 1986; Weiss *et al.* 1986). It has been assumed that if there is a decrease in P–O bond order in the TS, then the $^{16}k/^{18}k$ will be greater than 1, leading to a normal isotope effect. In contrast, if the bond order increases in the TS, then an inverse isotope effect would be expected.

Relating this to phosphate hydrolysis, in the case of ^{18}O non-bridge isotope effects (which have been used by e.g. Cleland and co-workers to study glucose-6-phosphate hydrolysis in water (Knight *et al.* 1986; Weiss *et al.* 1986), one would expect that a dissociative (metaphosphate-like) TS would have an increased bond order between the three non-bridge oxygens and the phosphorus atom, and thus that one would observe an inverse isotope effect. In contrast, a pentacovalent associative TS would have decreased bond order to the phosphorus atom, giving rise to a normal isotope effect. The observed ^{18}O isotope effect for a single non-bridge oxygen of the glucose-6-phosphate monoanion was proposed to be 1·0046 (suggesting a dissociative TS) (Weiss *et al.* 1986), although this value was later adjusted to take into account the possibility of a pre-equilibrium PT from a non-bridging oxygen of the phosphate to one bridging the leaving group (based on EIEs) (Knight *et al.* 1986). Thus, in this case, it is not possible to use the ^{18}O isotope to distinguish between associative and dissociative pathways.

Overall, the arguments about the relationship between isotope effects and mechanism have been circular, as experimental approaches cannot directly establish the nature of a TS. The uniqueness of the interpretation of isotope effects has already been challenged (Åqvist *et al.* 1999; Rosta *et al.* 2008), and this idea is also beginning to gain acceptance within the experimental community (Anderson *et al.* 2006). For instance, a recent theoretical study on transesterification

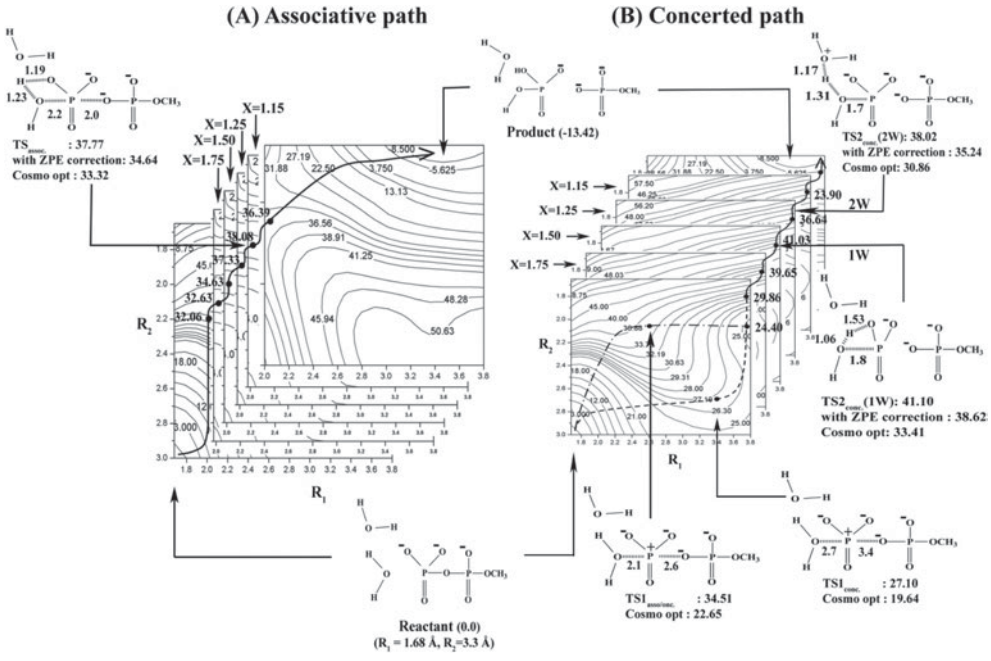

Fig. 19. The calculated free energy surfaces along the associative (A) and concerted/Late PT (B) pathways for the hydrolysis of monomethyl pyrophosphate trianion in solution utilizing both the 1 W and 2 W mechanisms. As seen from the figure the rate determining TSs for the 2 W and 1 W have similar energies. The surfaces are drawn as functions of R_1, and R_2 for several values of proton transfer coordinate X (R_1, R_2, and X are defined in Fig. 1A). TS_{assoc} describes the transition state along the associative pathway, $TS1_{conc}$ and $TS2_{conc}$ describes the transition states along the concerted pathway. The values of the TS energies evaluated by different approaches are given near the corresponding TSs figures.

(Lin *et al.* 2006) found large isotope effects that would traditionally be interpreted as evidence for a dissociative mechanism for what is essentially an associative TS (see also discussion in Section 4.5 on Wong *et al.* 2012). Similarly, a careful study of isotope effects for the hydrolysis of the magnesium monomethyl pyrophosphate trianion (Klähn *et al.* 2006) obtained very similar isotope effects for associative and dissociative pathways. Thus, the traditional interpretation of isotope effects cannot be used to distinguish between associative and dissociative pathways.

3.8 Overview and concluding comments

In this section, we have examined the mechanism of the non-enzymatic hydrolysis of phosphate mono-, di- and triesters. We have seen that the mechanism is greatly dependent not only on esterification level but also on protonation state, and that even in solution, the mechanisms of phosphate hydrolysis remain hugely controversial. Additionally, we have examined various popular experimental techniques for elucidating the mechanism of phosphate hydrolysis (such as LFER, isotope effects and entropies), and demonstrated that contrary to the traditional views, the interpretation of these experimental markers is ambiguous, and none can uniquely distinguish between associative and dissociative pathways, but rather, conventional experimental markers for determining mechanism are merely guidelines that can at best be used to eliminate potential

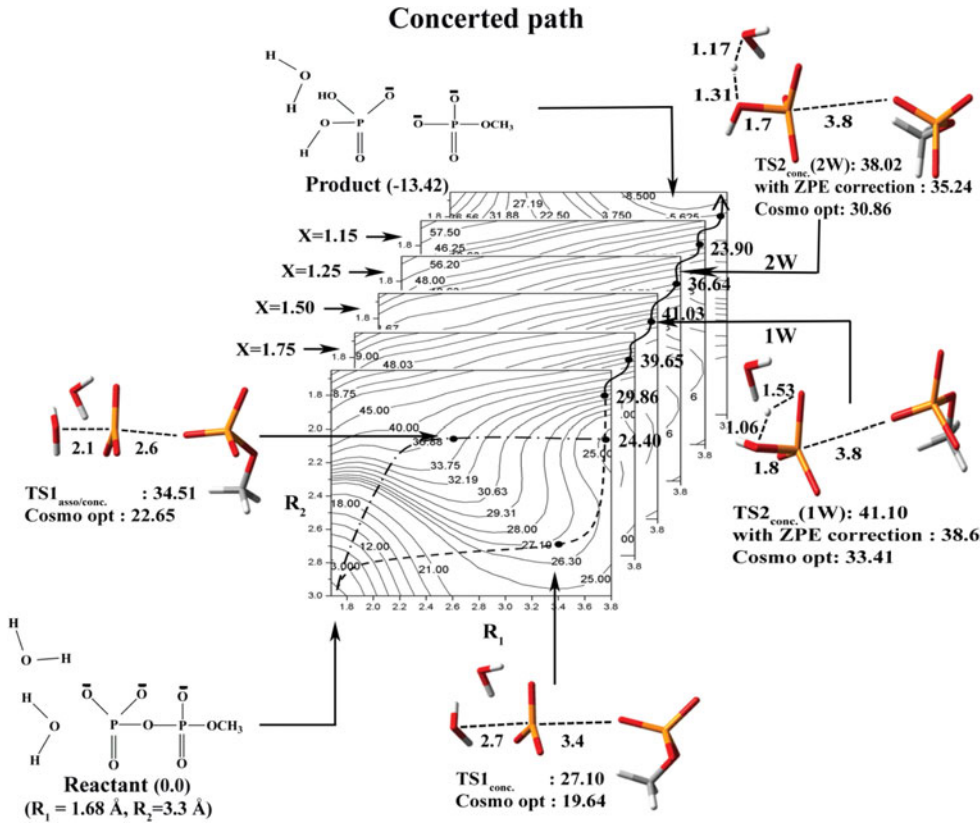

Fig. 20. Focusing on the calculated free energy surfaces for the reaction paths that starts with an associative or dissociative paths that follows by a late PT to the phosphate oxygen. As seen from the figure the rate determining TSs for the 2 W and 1 W paths have similar energies. Note that the activation entropy of the 2 W path (which is not included) should be larger than that of the 1 W barrier. The surfaces are drawn as functions of R_1, and R_2 for several values of the proton transfer coordinate X, and focuses on comparing the PT barriers that starts from the intermediate where R_2 is around 2 Å and R_1 is partially broken.

pathways. However, as was seen throughout this section, there exist a variety of computational techniques that allow us to resolve mechanistic issues that would otherwise be impossible to address experimentally (although, when done improperly, theoretical studies can also be problematic, as was seen here and will be highlighted at greater length in subsequent sections). Thus, at present, the only way to conclusively determine mechanism is by the combination of observed experimental data with careful theoretical studies that can accurately reproduce existing experimental data.

We conclude this section by again emphasizing that the mechanism of phosphate hydrolysis is much more complex than what has been considered traditionally, and that the corresponding free-energy surfaces involve a complex motion of the proton in the MFJ surface in a way that demands description of the type given in Fig. 8. This indicates that the final interpretation must require much more than a simplified experimentally based studies without extremely careful computational analysis. A case in point is given by considering the complexity of the free-energy surface and the need for a careful mapping in the surface defined by both the R_1 and R_2 of Fig. 6 and the PT coordinate. This issue is particularly important in view of recent suggestions

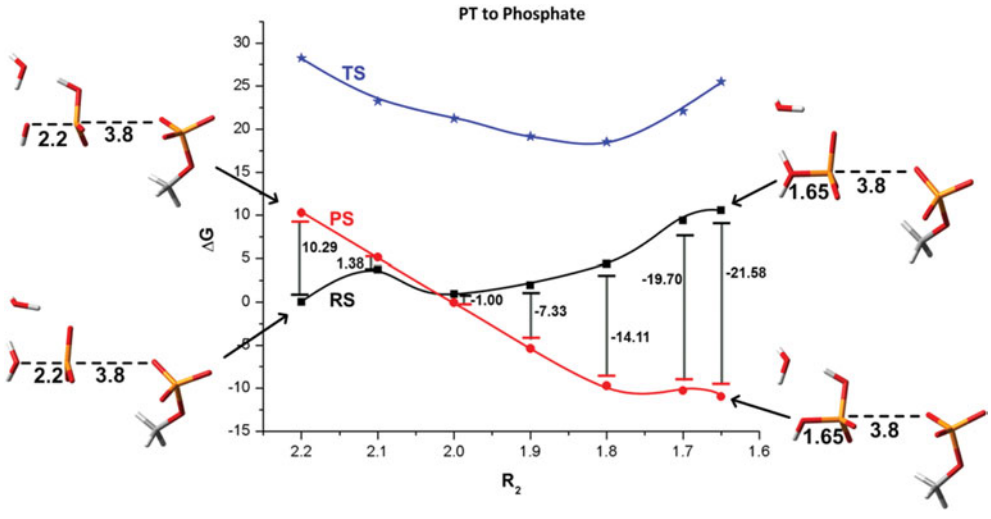

Fig. 21. The effect of the change in pK_a of the attacking water on the PT energetics in the 1 W mechanism. The figure displays the R_2 dependence of the energy of the "reactant" (where the proton is bound to the nucleophilic water and the water oxygen is in a short distance of 2.2 Å), the "product" (where the proton is attached to the acceptor phosphate oxygen) and the corresponding TSs for the PT process. As seen from the figure, the attacking water becomes much more acidic upon formation of the O-P and the PT occur. However, the PT does not become spontaneous because of the extra energy required for bending the O-P-O angel for an optimal PT.

(e.g. Grigorenko *et al.* 2006) of a PT between several water molecules instead of directly between the proton of the attacking water and the phosphate oxygen considered in Fig. 6 and in (Klähn *et al.* 2006). More specifically, in addition to the reaction paths it is also important to consider a mechanism were the associative attack of a water molecule does not involve a PT but a formation of an intermediate where the leaving group starts to break and then a late PT (Grigorenko *et al.* 2006). This PT can occur directly to the phosphate oxygen (referred to as 1 W) or through a second water molecule (referred to as 2 W) or through more water molecules. A mapping that considers consistently both options (Prasad *et al.* 2012) is given in Figs 19 and 20.

The complexity of the above surface has led to confusion in the theoretical community, which include calling a perfect concerted path a 'dissociative' path (e.g. Glaves *et al.* 2012; Grigorenko *et al.* 2006). More importantly recent theoretical studies that advocated a PT trough one or more water molecules never explored carefully the 1 W path that must be examine by careful mapping and not only by energy minimization or automated evaluation of reaction coordinates (this point is also be discussed in other sections). When careful examination of the 1 W and 2 W cases is performed (Figs. 19 and 20), it is found that both mechanisms are equally likely.

One of the interesting insights that emerged very recently form of examination of the two water (2 W) *versus* 1 W mechanism (Prasad *et al.* 2012) is the finding (see Figs 21 and 22) that the compression of R2 (the distance between the oxygen of the attacking water and the phosphate) drastically changes the pK_a of the attacking water (from about 16 to a negative value). Obviously, at some stage the attacking water becomes very acidic and transfers a proton either to another water or to the phosphate oxygen. The problem is the compression costs energy, which must be taken into account. Now adding effect of the change in pK_a between the donor and acceptor to the compression energy gives the free energy of the 'reactant' and 'product' in the PT process.

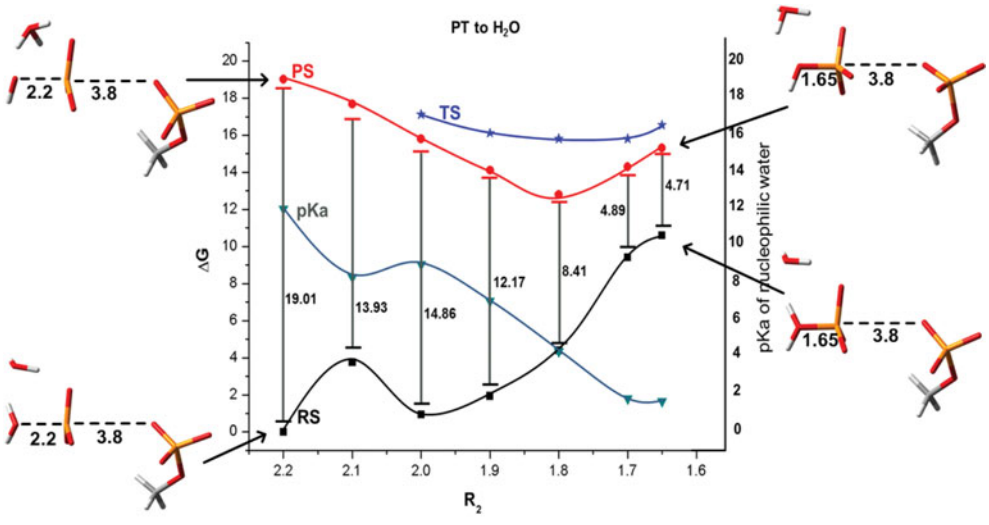

Fig. 22. The effect of the change in pK_a of the attacking water on the PT energetics in the 2 W mechanism. The figure displays the R_2 dependence of the energies of the "reactant" (where the proton is attached to nucleophilic water) and "product" (where the proton is attached to the second water) and the corresponding TSs for the PT process. As seen from the figure, the attacking water becomes much more acidic upon formation of the O-P and the PT becomes basically spontaneous.

However, what we actually need is the free energy of reaching the TS of the PT process. Here we have the main difference between the 2 W and 1 W. In the case of 2 W the exothermicity and the TS energy is almost equal since the acceptor water molecule can be arranged at an optimal distance. This means that change R_2 in an automated energy minimization or metadynamics (MTD) will reach eventually a point with a spontaneous PT. This, however, is not the case for the 1 W path. In this case, the energy will always go up along the PT coordinate, since it is necessary to bend the formed P–O bond for an optimal PT. The barrier for the PT in the case of 1 W is almost certainly the reason why this path has been overlooked in energy minimization studies.

Obviously more careful studies are needed to determine the exact PT path. However, theoretical studies will eventually resolve the nature of the complex surface, but doing it by experimental analysis is close to impossible. In fact, the concerted path that may lead to a dissociated phosphate with covalently bound water (point $TS1_{con}$ in Fig. 20) has never been considered in the traditional experimental analysis.

4. Enzymatic phosphoryl transfer

4.1 G-Proteins

4.1.1 Introduction

Nucleoside triphosphates serve as crucial mediators of life. For instance, the regulation and control of signal transduction and transport processes relies almost exclusively on GTP. ATP, in turn, is used to drive unfavorable chemical reactions, to fuel biological machines, and to regulate key biological processes (*via* the phosphorylation of proteins). Other nucleotides also play important biological roles. Their action is controlled by proteins that tune their binding affinity and

the rate of their relevant chemical transformation (see, e.g. Coleman & Sprang, 1999; Scheffzek & Ahmadian, 2005; Sprang, 1997b; Vetter & Wittinghofer, 1999; Wittinghofer, 2006). This section focuses on GTP-binding proteins, and also considers some aspects of proteins that use ATP in energy transduction.

G-proteins are a superfamily of regulatory GTP hydrolases that form a relatively stable complex with their substrate (GTP) and product [guanosine diphosphate (GDP)]. The binding to and hydrolysis of GTP by these proteins triggers conformational changes, which are used in order to communicate with other proteins. The GTP- and GDP-bound conformations generate the active (ON) and inactive (OFF) states of the protein, where the transformation between theses states allows the system to serve as a regulatory machine. The transition between the ON and OFF states of these proteins can be controlled in a variety of ways, and the G-protein superfamily includes switches (Ras and its homologues), clocks (heterotrimeric G-proteins and subunits) and sensors (elongation factors such as EF-Tu and EF-G). The exchange of GDP and GTP is also regulated by special proteins [guanine-nucleotide-exchange factors (GEF)].

Despite great progress in structural and biochemical studies of G-proteins and related systems (for reviews, see Coleman & Sprang, 1999; Geyer *et al.* 1997; Scheffzek & Ahmadian, 2005; Sprang, 1997b; Vetter & Wittinghofer, 1999; Wittinghofer, 2006), we still do not have a complete understanding of the detailed molecular mechanism of these proteins. For example, the reasons for the oncogenic effect of some Ras mutations are not clear. Further progress in this field requires quantitative structure-function correlation, which would allow for the conversion of the available structural information into activation energies, and, therefore, into the relevant catalytic effects.

4.1.2 The Ras/GAP system and the mechanism of the corresponding phosphoryl transfer

One of the best examples of this class of proteins is $p21^{ras}$ (Ras). This protein plays a central role in the signal transduction pathway that controls cell proliferation, and can be considered as a switch, where the signal is ON in the GTP-bound form, and OFF in the GDP-bound form (Barbacid, 1987; Bos, 1989; Bourne *et al.* 1991). The switch is activated by the binding of GTPase-activating proteins (GAPs), which accelerate the relatively slow GTP hydrolysis in Ras by up to 5 orders of magnitude (Gideon *et al.* 1992; Schweins *et al.* 1995). Mutations at either of the 12, 13, or 61 positions drastically slow down the GTPase reaction of the Ras–RasGAP (RasGAP) complex, leaving the signal in the ON state, and thus frequently leading to cancer (Bos, 1989; Lowy & Willumsen, 1993; Spandidos, 1989). The elucidation of the structure of Ras (Milburn *et al.* 1990; Pai *et al.* 1990) and the RasGAP complex (Scheffzek *et al.* 1997) has provided what is perhaps the most detailed link between protein structure and oncogenic mutations. A spectacular clue about the control of the Ras switch was the finding that Arg^{789} of GAP, which is also referred to as the 'arginine finger', provides major TS stabilization during the GTPase reaction (Ahmadian *et al.* 1997; Scheffzek *et al.* 1997). However, despite this finding (and other major advances), it is still unclear what the origin of the anticatalytic effect of the oncogenic mutations of Ras actually is, with the most puzzling problem being the role of Gln^{61}. Mutations of this residue in the RasGAP complex reduce the k_{cat} of the GTPase reaction by factors of 10^{-3}–10^{-7}, depending on the specific mutation. This large effect seems to be ubiquitous to different GAPs, including the Ran–RanGAP system, which does not operate through an arginine finger, but still contains a catalytically crucial glutamine (Seewald *et al.* 2002). A related glutamine is also found in other G proteins (e.g., Q204 and Q227 in $G_{i\alpha1}$ and $G_{s\alpha1}$, respectively, Sprang,

1997a). It is therefore clear that the proper positioning of Gln^{61} is crucial for effective catalysis. However, the nature of this positioning and its role in stabilizing the TS are entirely unclear. This means that one of the most crucial links between structure and catalysis in Ras (and other signal transduction proteins) is still not understood.

In principle, there are several direct ways in which Gln^{61} could be contributing to the catalytic effect. It could act 'chemically' as a general base, it could provide direct electrostatic stabilization to the TS, and it could also perhaps act by applying entropic or steric effects (see below). However, it is also equally possible that this residue is acting in an *indirect* way, the precise nature of which is not clear. Structural studies, mutational experiments and the use of TS analogs have proven unable to elucidate the origin of the catalytic role of Gln^{61}. In fact, all early works (e.g., Krengel *et al.* 1990; Pai *et al.* 1990) concluded that glutamine must be crucial as a general base that extracts a proton from the nucleophilic water molecule. This view was drastically changed by our study (Langen *et al.* 1992), in which we demonstrated that Gln^{61} cannot be the general base in Ras (since the energy of the proposed PT step in Ras is higher than the actual observed barrier for the GTPase reaction in solution). This (and subsequent) studies led to the conclusion that the GTP itself must be the elusive base, and therefore to the proposal of the GTP-as-base (or 'phosphate-as-base') mechanism (Schweins *et al.* 1994, 1995).

At this point, it is useful to clarify the misunderstandings surrounding the concept of the GTP itself being the base. There have been recent attempts to trivialize this central issue (including but not limited to Lassila *et al.* 2011). These have started with hiding the fact that the concept of Gln^{61} as a base was generally accepted by the workers in the field, and the fact that the demonstration that this concept is invalid emerged from a theoretical work, rather than from the experimental community. It continues with attempts to imply that it is experimentally well known that the reactions of phosphate monoesters, acyl phosphates and phosphorylated amines are dissociative in solution (Maegley *et al.* 1996), thus presenting a misleading view that apparently succeeds to confuse key workers (to a certain degree) about the validity of this mechanistic argument (Du *et al.* 2004; Li & Zhang, 2004; Wittinghofer, 2006). The problem continues with related arguments that it has been experimentally established that the attacking water molecule cannot be activated by direct PT to one of the non-bridging oxygens of the phosphate itself (Admiraal & Herschlag, 2000) despite the absolute fact that this is a point that cannot be tested experimentally (see discussion in Section 3 and Åqvist *et al.* 1999; Florián *et al.* 1998; Kamerlin *et al.* 2008b). This is compounded by asking (Lassila *et al.* 2011) if the general base is even necessary. Such a question would have been legitimate if it were presented without the misleading preface that 'structural studies did not find a good candidate for a general base'. Unfortunately, we have here a highly problematic statement. The ground-breaking structural elucidation of Ras (Krengel *et al.* 1990; Milburn *et al.* 1990; Pai *et al.* 1990) led to a belief in the structural community that Gln^{61} must be the general base. It was our 1992 study (Langen *et al.* 1992) and the subsequent study of Schultz and co-workers (Chung *et al.* 1993b) that established that Gln^{61} *cannot* be the general base. Next, we established the fact that the phosphate is the actual and eventual proton acceptor (Schweins *et al.* 1994) and this finding is now widely accepted (see e.g. Dall'Acqua & Carter, 2000; Kosloff *et al.* 2000; Pasqualato & Cherfils, 2005; Wittinghofer, 2006, among others). The basic question here is what the fate of the proton in the reaction of Ras is, and, since it now agreed that the proton moves to the phosphate and not to Glu^{61}, the question is reduced to the secondary question about the details of the exact pathway, and *not* to whether or not the phosphate accept this proton. This mechanistic question cannot be addressed by any of the current experiments and thus, following our initial study (Langen *et al.* 1992), we

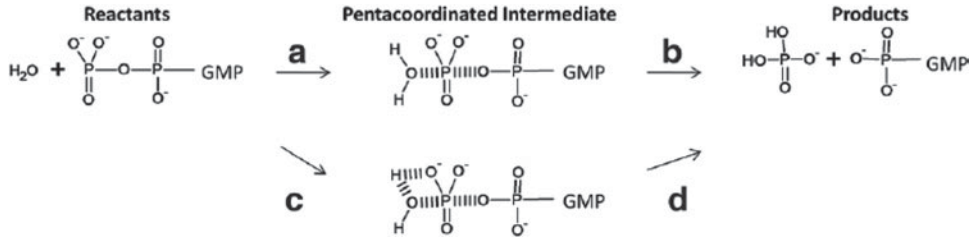

Fig. 23. Possible mechanistic pathways for the GTPase reaction mechanism. The reaction is considered as a two-step mechanism where the GTP serves as a general base. In the first step, **a**, the water molecule attacks the GTP and a penta-coordinated intermediate is formed. In the second step, **b**, the P_β–O bond is broken and the GDP and a phosphate group are formed. Note that the attack of the water molecule (step **a**) may be either concerted or stepwise mechanism. In principle, the steps b and d could also proceed with the assistance of another water or a base.

explored this issue repeatedly, eventually supplementing our earlier studies and detailed QM studies of the solution reaction by a detailed QM/MM study of the reaction surface for Ras (Klähn *et al.* 2006), obtaining a TS that is highly associative in character (P–O distance of 2·0 and 2·1 Å to the nucleophile and leaving group, respectively), with concomitant PT to a non-bridging oxygen on the phosphate (see Fig. 23 for the general mechanism (and more complex options in Figs 19 and 20, Prasad *et al.* 2012) and Section 4.1.2 for the calculated TS, as well as the discussion below).

The discussion of the phosphate as a base mechanism (Lassila *et al.* 2011) also involves the argument that the expected catalytic advantage of deprotonating water for subsequent OH^- attack is rather small, relative to the actual catalytic effect of Ras. Here we have a problematic assertion that overlooks the fact that this is not the relevant issue. The problem (which originated in, Maegley *et al.* 1996) is associated with the confusion between the uncatalyzed reaction for a given path (for instance concerted PT to the non-bridging oxygen on the phosphate) and the catalysis of the given mechanism (pathway) by the change in environment, where it is basically argued that this path with a PT provides no catalytic advantage, However, the catalysis is *not* due to the PT (this is a *mechanism* with its intrinsic energy is the energy of the reference reaction), but rather to the change of the energy of this mechanism due to the environment. The actual catalysis comes from the enzyme preorganization, including the electrostatic effect of the Mg^{2+} ion, and not from the specific mechanism used (the same mechanism used by the enzyme is much slower than in water; see Warshel *et al.* 2006b). Furthermore, the above assertion also overlooks the fact that the catalytic effect of OH^- cannot be determined experimentally (since we cannot have a protonated phosphate at the pH needed to study the OH^- effect (see discussion in Florián & Warshel, 1997; Kamerlin *et al.* 2008b, and Section 3). Moreover, the comparison of the rates k_0 and k_{OH^-} in (Lassila *et al.* 2011) as an argument against the impetus for deprotonating the attacking water molecule is not necessarily a meaningful one, as it overlooks the fact that the most likely scenario is that the observed k_0 corresponds to hydroxide attack on a protonated species in a substrate-as-base mechanism and not to water attack on the phosphate (see, e.g. Kamerlin, 2011; Kamerlin *et al.* 2008a, 2008b; Klähn *et al.* 2006). Of course, we have shown earlier (Florian *et al.* 1998) that the proton transfer from water is concerted and no OH^- is being formed.

On a related note, we would like to clarify that the additional argument against the substrate-as-base mechanism (Maegley *et al.* 1996) involves the misleading assertion that 'protonation of the γ-phosphoryl group is expected to destabilize a dissociative TS, not stabilize it'. This is based on the underlying automatic assumption that the TS is dissociative, which we have shown not to

be the case (see below). Of course, it also involves the aforementioned confusion between the energy of a given path and the concept of stabilization (i.e. as will be clarified below, even in the dissociative pathway we have to move the proton trough the γ-phosphate).

In any case, the main point in the distinction between the Gln-as-base or phosphate-as-base mechanisms has been the contrast between having the substrate as the proton acceptor, and having a protein group serving as a base. Furthermore, apart from the problem of hiding this key issue, we also may have here an interesting evolutionary issue (Luo *et al.* 2012b). If an enzyme specifically requires enzymatic residues for acid/base catalysis, then a single mutation at these points will have dramatic impact on the catalytic activity of the enzyme and affect its subsequent evolution. In contrast, if the substrate itself can act as a proton-sink, this then gives the enzyme more flexibility in how it evolves without *a priori* destroying its catalytic activity toward that substrate by 'accidentally' removing the base. Of course, mutations that change the pre-organization can have an overwhelming effect on the catalytic power of the enzyme (as is in fact the case with Gln^{61} in the Ras/Gap complex), but using a single residue as a base might cause a problem from an evolutionary perspective. Regardless of the question of evolution, and despite the fact that in solution, we frequently do not have an alternative base and the water molecule must be combined with the phosphate (as is frequently the case, see e.g., Kamerlin, 2011; Kamerlin *et al.* 2008a; Klähn *et al.* 2006), it is not useful to detach the discussion from the issue of having a protein group as the base. As hinted above, and as will be shown below, all mechanisms that involve water attack on the phosphate can be formally considered phosphate-as-base mechanisms, since the proton on the water molecule ends up being transferred to the phosphate in either a concerted or stepwise fashion (see related discussion in the Supporting Information of Adamczyk & Warshel, 2011). In fact, the realization that Gln cannot be the base led us to examine alternative mechanistic options, where the GTP accepts the proton in either a stepwise or concerted fashion (relative to the actual phosphoryl transfer), and is then hydrolyzed in either an associative or dissociative pathway (see e.g. Glennon *et al.* 2000). The relevant mechanistic considerations ranged from an examination of LFER (Schweins *et al.* 1996a) to considerations of the effect of the arginine finger on the associative and dissociative mechanisms (Glennon *et al.* 2000). In fact, the realization that the central mechanistic questions can only be resolved by theoretical studies (which must of course reproduce the relevant experimental observables) formed the basis for most of our mechanistic studies.

The mechanistic picture that subsequently emerged from our studies is largely described in Section 3 and below. We note, however, that the Ras project has forced us to also explore the nature of the surface in the protein by *ab initio* QM/MM calculations. After validating our approach by studies of the solution reaction, and, in particular, the reaction in the presence of the Mg^{2+} ion, we also evaluated the surface for the GTPase reaction in the RasGap system (Klähn *et al.* 2006). It was found that the reaction path preferentially proceeds through a concerted associative pathway, and that it is hard to experimentally establish whether this path is associative or dissociative as both TSs give rise to similar experimental observables. Interestingly, the new surfaces for the reaction in solution and in Ras involved an associative TS, where the PT from the attacking water is rather late (see Klähn *et al.* 2006). The formation of an early, short bond with the attacking water molecule was first observed by (Grigorenko *et al.* 2005, 2006), who considered it, however, to be a dissociative mechanism, and described the resulting $H_2O \ldots PO_3^-$ (or $H_2PO_4^-$) fragment incorrectly as a metaphosphate, overlooking the fact that the nucleophilic attack occurs before the cleavage of the P–O bond (the importance of examining the full surface rather than just isolated TSs is discussed in several places in this review). It is also worth pointing

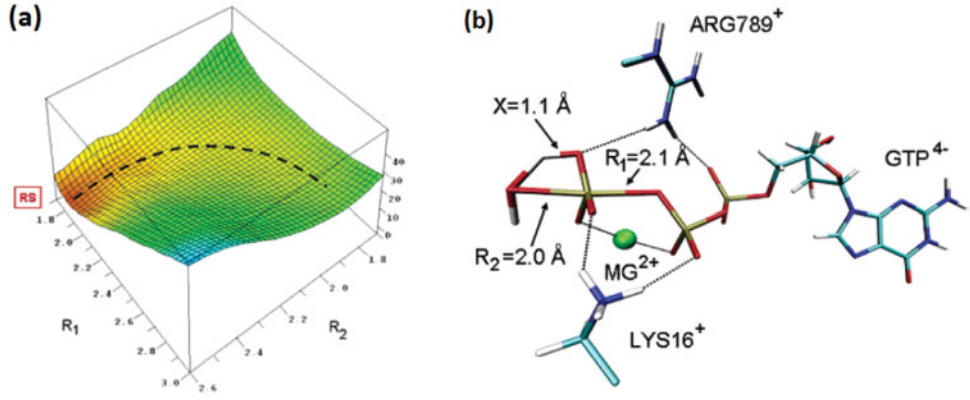

Fig. 24. (*a*) The calculated energy surface for the hydrolysis of GTP in RasGAP in the space defined by the P–O distance to the leaving group and nucleophile (R_1 and R_2, respectively). (*b*) The calculated TS structure for the hydrolysis of the GTP substrate in RasGAP. This figure was originally presented in (Klähn *et al.* 2006).

out that although the studies of (Grigorenko *et al.* 2005, 2006) were instructive (in particular for providing new insight into the initial orientation of the nucleophile), they are also problematic for a number of reasons. These studies never examined the full surface, and have not used any sampling (the corresponding problems with which have been specifically demonstrated in Klähn *et al.* 2006) and now in Figs 19 and 20 (Prasad *et al.* 2012). Thus, the authors arrived at a questionable mechanism in the Ras–GAP complex, for which the first step (P–O bond cleavage) had a barrier of only about 4 kcal mol^{-1} (Grigorenko *et al.* 2007a), where a complex of water and metaphosphate is being formed, with a subsequent rate-determining barrier of 10·5 kcal mol^{-1} relative to the reactant state, leading to Gln-assisted conversion of the nucleophilic water molecule plus metaphosphate to Pi + GDP with a protonated Gln sidechain in the enzyme. Furthermore, this work obtained very different mechanisms in enzyme and in solution (where, in the solution case, the rate-limiting step involved the formation of an associative intermediate (see Fig. 25), whereas in the enzyme the first barrier in Fig. 25 is drastically lowered and the second step becomes rate limiting). This problem become even more apparent in a study of ATP hydrolysis (Grigorenko *et al.* 2007b), (discussed in Section 4.2) and in the EF–Tu study discussed below. Incidentally, the solution study that produced a barrier of 20 kcal mol^{-1} (Grigorenko *et al.* 2006) instead of about 28 kcal mol^{-1} was done by energy minimization, which is unacceptable for the enormous dimensionality of the cluster used, and, in fact, assuming that the proposed mechanism of PT along a chain of three water molecules in water can be explored by energy minimization in somewhat unrealistic. In contrast with these works our study (Klähn *et al.* 2006) (which led to the TS shown in Fig. 24) invested extensive effort into actually validating and calibrating the calculations against a solution LFER obtained from experimental data, as well as reproducing our results with several different methods, and, of course, in evaluating the full surface (which is essential for detecting problems with the assumed paths).

One of the serious problems in attempts of some other works such as e.g. (Topol *et al.* 2004) to explore the substrate-as-base mechanism has been the misunderstanding of the nature of the corresponding surfaces reported and analyzed in many of our works. None of our works 'assumed' a two-step hydrolysis pathway, but rather, went into great length to describe and analyze the corresponding concerted PT surface (starting with Florián & Warshel, 1998), and more

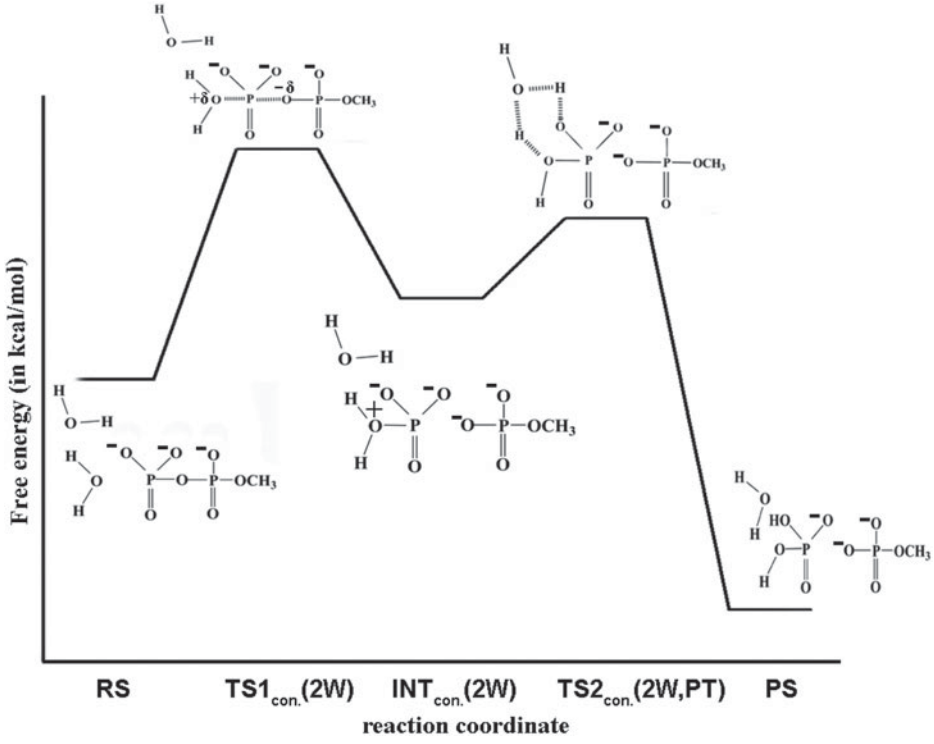

Fig. 25. A schematic diagram of the free-energy surface and the representative geometries for the hydrolysis of monomethyl pyrophosphate trianion in solution following the catalytic mechanism explored in (Grogorenko *et al.* JPCB, 2006, 110, 4407). This mechanism corresponds to the associative (2 W) pathway depicted in Fig. 20.

systematically in subsequent works). The idea that such a path can be explored by an automated TS search rather than by systematic 2D (or 3D as in Fig. 19, Prasad *et al.* 2012) mapping is particularly problematic. In fact, searching only for the PT after the bond to the leaving group has been broken (Figure 7 in Grigorenko *et al.* 2005), rather than by also considering concerted PT already in the associative and other pathways, as was done by (Klähn *et al.* 2006) and in Figs 19 and 20, Prasad *et al.* 2012), demonstrates the problem in trying to explore the substrate-as-base mechanism, without actually moving along the corresponding calculated path; this path cannot be generated by an automated reaction path.

Interestingly, as an aside, the problems seem to be compounded in the analysis presented of otherwise reasonable QM calculations of the reaction in the RasGAP complex using a cluster model (Topol *et al.* 2004), where the authors state that 'the P–O bond is already broken, but the lytic water molecule is still in a pre-reactive state', and we are also told about the existence of a stable intermediate after the $P–O_{lg}$ bond is extended to 3 Å. However, the very clear and reasonable reaction path shown in this work is in fact a stepwise *associative* pathway, with a barrier of 20 kcal mol^{-1}, whereas the dissociative pathway is significantly higher in energy. In fact, the actual path explored by (Grigorenko *et al.* 2005; Topol *et al.* 2004) is a fully associative/concerted path described in Fig. 20.

It must be mentioned here that the danger lies in the fact that it is very tempting to assume that the same *ab initio* optimization approaches that are used so effectively in studies of small molecules in the gas phase would work also in proteins. However, a very simple attempt to evaluate

the energy profile (or, say, the energy of a charge in a protein) from different starting points should convince even a major skeptic that the problem is not the basis set used, but rather the multidimensionality of the system, and thus even a low-level quantum chemical model calibrated on solution experiments will be much more informative about catalytic effects than the highest level *ab initio* QM/MM model with simple energy minimization. Similar checks will show that the energy of protonating different oxygens of the GTP γ-phosphate in Ras is drastically different with approaches that do not use very extensive sampling and relaxation (e.g. Grigorenko *et al.* 2005). The problems with any related conclusions about the phosphate-as-base mechanism are exactly as serious as the difficulties involved in calculating pK_as by QM/MM approaches (Kamerlin *et al.* 2009b).

Despite the above criticism of the specific energetics obtained by Grigenko *et al.* we must point out (see also Section 3.5) that our finding of an increase in acidity of the attacking water along the concerted path establishes that this water becomes eventually a good proton donor. Here, it is possible that the energy of PT to glutamine is lower than that of transfer to the phosphate (in analogy to the 1 W and 2 W mechanism) and this issue should be explored by careful mapping (avoiding the tarps mentioned in Section 3.5). In particular, it may be possible that the protein stabilizes more the PT to Gln at the TS because it has larger charge separation than that of TS for a direct PT to the phosphate oxygen. However, if this will be found to be a realistic description it cannot be classified as Gln as a base (which corresponds to PT before the nucleophilic attack), but rather as a Gln assisted associative or concerted attack.

In discussing our systematic computational studies, it is useful to contrast them with alternative qualitative considerations of e.g. (Lassila *et al.* 2011). The Supporting Information of that work suggests that '*KIE studies have observed leaving group KIEs for the Ras-catalyzed reaction of approximately 1·02, as expected for a loose TS that is similar to that observed for solution reactions*' (note that the authors have now changed the term dissociative to 'loose'). Unfortunately, as shown in (Klähn *et al.* 2006), in contrast with frequent implications to the contrary (e.g. Hengge, 2002; Hoff *et al.* 2001 to name a few examples), it is extremely hard to obtain unique information about the nature of the TS from the KIE in solution (see also Section 3). The problem is compounded by saying that that this proves that enzyme does not change the nature of the TS, and thus we cannot have a substrate-as-base mechanism (even though, as discussed above, the reaction in solution also most likely proceeds *via* a substrate-as-base mechanism, so there is no apparent contradiction in this interpretation of the KIE). However, here again we have a very problematic argument. First, the associative or dissociative nature of the TS cannot be deduced from KIE, and, second, it would not change any of the key issues, since a careful simulation study (Glennon *et al.* 2000) has shown that the same catalytic effect would be obtained in both associative and dissociative mechanisms.

Related problematic arguments were also used in the claim (Maegley *et al.* 1996) that the presumed charge buildup on the γ-phosphate in a dissociative TS means that there is no stabilizing interaction between the external charge and the non-bridging oxygens of γ-phosphate (i.e. that the catalytic effect of the arginine is limited to interactions with the leaving group oxygen), and that interactions with the γ-phosphate would only become important if the enzyme caused a switch to an associative TS. This argument is problematic on two fronts. The first is that stabilizing an associative TS is, as discussed above, exactly what the enzyme is doing, and therefore this argument is based entirely on the incorrect assumption that the reaction proceeds through a dissociative TS. Secondly, the poor formulation and inconsistency of the above argument resulted in a confusion in the experimental community (e.g. Chakrabarti *et al.* 2007; Wittinghofer,

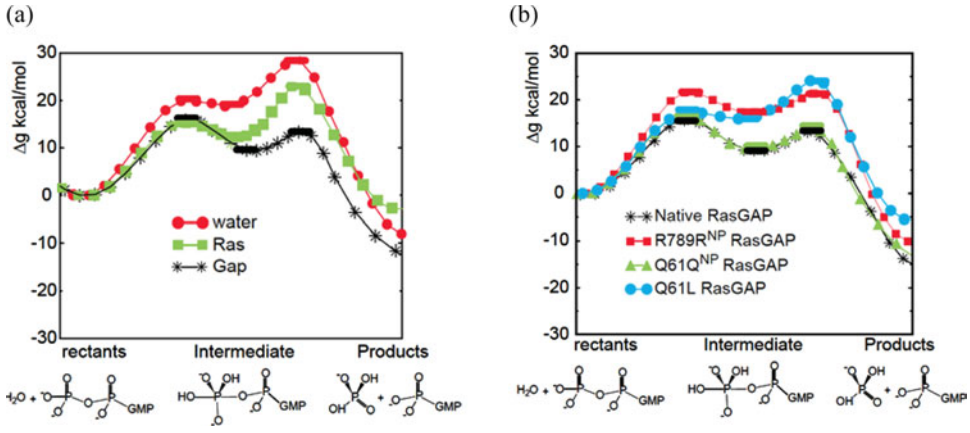

Fig. 26. (*a*) The GTPase reaction profile. The free-energy profiles for the GTPase reaction in water (circles), Ras (squares) and in the RasGAP complex (stars) obtained from an average over five different profiles, calculated using the EVB FEP/Umbrella sampling procedure.(*b*) Mutation effects on the GTPase reaction profile. The free-energy profiles of the GTPase reaction in the native RasGAP complex (stars) and its different mutations, $R789R^{NP}$ (squares), $Q61Q^{NP}$ (triangles) and Q61L (circles). The profiles were obtained from an average over five different profiles, calculated using the EVB FEP/Umbrella sampling procedure. NP stands for mutation to the non-polar form of the indicated residue. This figure was originally presented in (Shurki & Warshel, 2004).

2006), who might have been inclined to accept the presumed dissociative TS of (Maegley *et al.* 1996), but were then also willing to assume that the arginine finger can only stabilize an associative TS (e.g Allin *et al.* 2001). Obviously, the argument above seems convoluted, and some might even presume that we must be building a straw man here, but it reflects the exact logic described above. The main problem, of course, is that none of the underlying assumptions for this argument has been either established or correct. More precisely, it is unjustified to, even in 2011 (Lassila *et al.* 2011), continue to claim that the LFER are showing dissociative TSs, despite the evidence to the contrary (see Section 3). Furthermore, the implication that external charge can only stabilize a particular TS (at least when it is near the γ-phosphate), or that the qualitative effective charge considerations (which have been shown to be unjustified (Klähn *et al.* 2006) can tell us about the difference in the catalytic effect for different mechanisms is problematic, since our actual study of the interaction of the protein field with the change in the substrate charge upon the transition from the RS to TS (Glennon *et al.* 2000) (or even a related study of the GTP to GDP transitions, Muegge *et al.* 1996) has clearly demonstrated the same electrostatic catalytic effect for associative and dissociative mechanisms. Furthermore, the most profound difference is the fact that we were able to predict the quantitative catalytic effects see Fig. 26*a*) as well as the effect of different mutations (Glennon *et al.* 2000; Shurki & Warshel, 2004), (see Fig. 26*b*), whereas treatments of the type suggested in (Lassila *et al.* 2011) can, at best, tell us the sign of the mutational effect.

It seems to us that the above misunderstanding is compounded by the recent argument (Lassila *et al.* 2011) that (Maegley *et al.* 1996) actually predicted the effect of the Arg finger (Fig. 27) and its relationship to the dissociative mechanism (now suddenly called 'loose' TS). Apparently (see above) it was suggested that a dissociative TS cannot be stabilize by positive charge, stating that '*Although electrostatic interactions with the* γ*-phosphoryl oxygens are not predicted to be strengthened in a dissociative* TS *and may even be somewhat weakened, such interactions may be present because they are important for binding and positioning.*' In fact (Maegley *et al.* 1996) completely overlooked the crucial

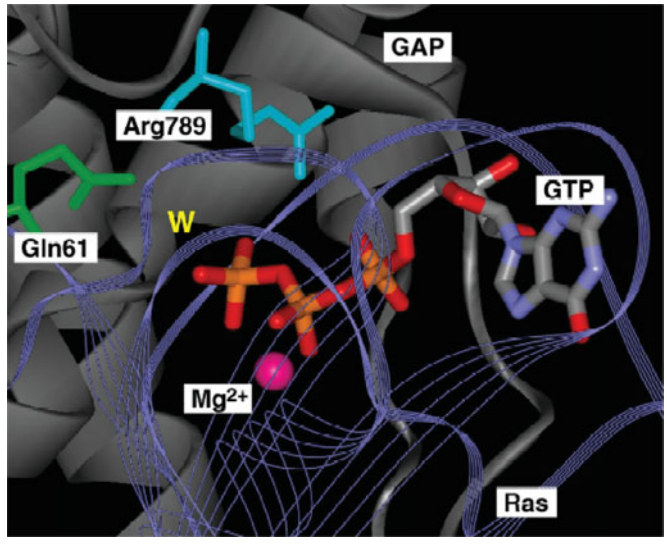

Fig. 27. An illustration of the 'Arginine Finger'.

importance of electrostatic TS stabilization identified in our studies of all significantly catalytic enzymes and of course in the case of Ras and the quantitative effects that were reproduced by (Shurki & Warshel, 2004). Furthermore they overlooked the key point already predicted in (Muegge *et al.* 1996) that the key is the stabilization of the TS is formed by breaking GTP to GDP and not the exact path for this reaction. Thus, the fact that the Arg effect is mainly electrostatic has not been recognized or analyzed correctly by the qualitative considerations of (Maegley *et al.* 1996). Another related problem that emerges from the reluctant to accept the role of electrostatic effects in catalysis has been associated with the statement (Maegley *et al.* 1996) that '*Mg^{2+} is also bound in solution, so that no large catalytic enhancement is expected relative to the solution reaction in the absence of significant perturbations within the active site*'. This overlooks the key fact that the catalytic effect of metal ion is mainly electrostatic and is entirely determined by the corresponding dielectric constant and preorganization and that, in solution, these effects are in fact small, because of the larger dielectric constant.

4.1.3 What is the true oncogenic role of Gln^{61}?

After the above rather long clarification and analysis of the underlying mechanistic issues (a problem that turned out to be less critical than anticipated, as discussed below), we can turn to the exploration of the true role of GAP and of Gln^{61}. The first theoretical progress on this front emerged after examining the reliability of our approach (first in Glennon *et al.* 2000 and then in Shurki & Warshel, 2004) by calculating the catalytic effect of Ras and the RasGAP system.

Next, we examined the effect of mutations of Arg^{789} (the so-called 'arginine finger' depicted in Fig. 27) and Gln^{61} (the alleged general base). Since electrostatic effects are usually the leading contribution in catalysis (Warshel *et al.* 2006b), we first considered the simple approach of mutating the two residues to their nonpolar analogs, where all the side-chain atoms have zero residual charges. This type of analysis allows us to quantitatively determine what the electrostatic effect of the residues under consideration actually is. The results of the corresponding simulations are presented in Fig. 26*b*. It is seen from the figure that the observed effect of Arg^{789} is

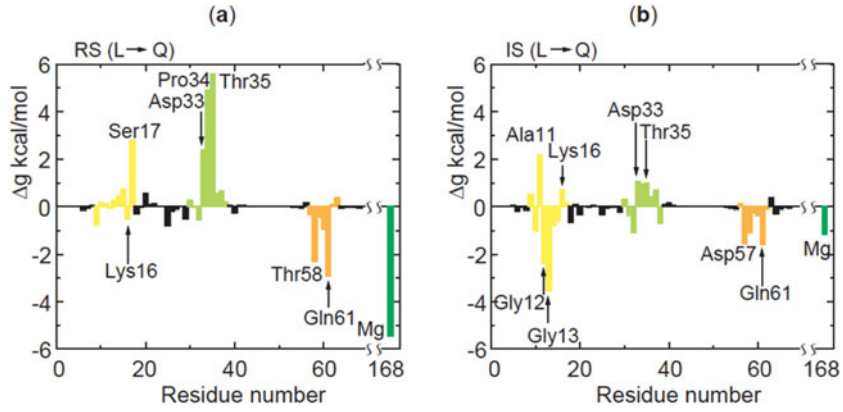

Fig. 28. Electrostatic effects due to the Q61L mutation. The main changes in the electrostatic contributions from the Ras residues due to the Q61L mutation. The changes in the RS are compared with those in the TS in (*a*) and (*b*), respectively. Negative and positive values correspond, respectively, to stabilization and destabilization due to the mutation. The changes in the key areas are highlighted as in Fig. 29, P-loop (yellow), Switch I (light green), parts of β_3 and L4 of the switch II (orange) and the Mg^{2+} (dark green). The calculated changes were obtained by the LRA procedure and then scaled by a dielectric constant of 20 and 4 for the ionized and polar residues, respectively. This figure was originally presented in (Shurki & Warshel, 2004).

reproduced using the nonpolar analog mutation technique. This provides a clear demonstration of the fact that the main role of the arginine finger is to provide direct electrostatic stabilization of the TS (Shurki & Warshel, 2004).

After our calculations were properly validated (by reproducing the overall observed catalytic effect, as well as the effects of different mutations), we could turn our attention to the examination of the origin of these effects. We started by exploring the overall catalytic effect of GAP, focusing on its indirect effect through its interaction with Ras. A comparison of the X-ray structure of the isolated Ras to the structure of Ras in the RasGAP (henceforth referred to as Ras' for simplicity) reveals some structural change induced by the binding of GAP to Ras (such as the fact that Gln^{61} moves closer to Pγ, Scheffzek *et al.* 1997). However, although this comparison is quite instructive, it cannot tell us what the exact *functional* implications of these changes are. In particular, since the enzyme is flexible, changes in the positions of various residues do not necessarily reflect changes in the overall activation barrier. Furthermore, although TS analogs (TSAs) can provide useful information, it is hard to deduce the exact relationship between TSAs and the corresponding TS (Barbany *et al.* 2003). Therefore, we used the linear response approximation (LRA) (Lee *et al.* 1992) to determine the individual contributions of Ras in the RasGAP complex (Shurki & Warshel, 2004), and compared the catalytic effect of the isolated Ras with the contribution of Ras to the RasGAP complex. This was achieved by first evaluating the protein–substrate electrostatic interactions in the RS and TS, and then subsequently decomposing them to the individual contributions from each residue (Fig. 28). This study demonstrated that the structural reorganization between the free Ras and Ras' is indeed functionally important. Significant changes occur in the P-loop, Switch I residues (in particular Asp^{33} and Thr^{35}), residues in the β_3 region (particularly Asp^{57}), L4 in Switch II (particularly Gln^{62}) and the Mg^{2+} ion (see Fig. 29 for definitions of the key regions). Therefore, identifying specific structural (rather than energetic) changes is not simple, since we obtain a similar trend with different starting configurations. Furthermore, our focus here is on average energies rather than average *structures*.

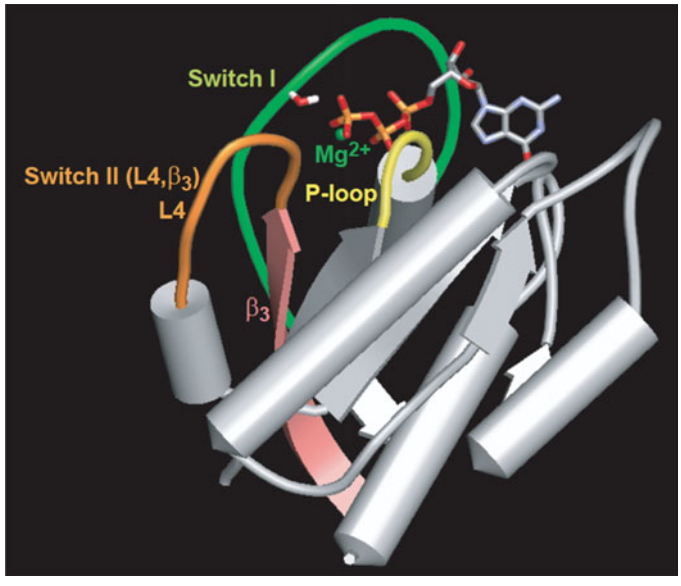

Fig. 29. An overview of the key regions in the RasGAP complex. Shown here are the P-loop (yellow), Switch I (light green), parts of Switch II (orange) and the Mg^{2+} ion (dark green). This figure was originally presented in (Shurki & Warshel, 2004).

Apparently, as was found in our study (Shurki & Warshel, 2004), the interaction between GAP and Ras upon complexation moves Ras (which is part of the complex) from its stable RS arrangement toward a configuration that stabilizes the TS and the product. The corresponding catalytic effect is therefore indirect, and thus quite different from the direct effect of the arginine finger. It is important to take into account in this respect previous studies (e.g. Muegge *et al.* 1996), which proposed that the GAP-induced transition pushes the GTP-bound form of Ras (Ras(GTP)) toward the structure of the GDP-bound form of Ras (Ras(GDP)).

Another source of support for our finding that Gln^{61} is acting in an indirect way comes from studies of the effect of mutations of $G_{s\alpha}$. The Q227L mutation (which is thought to have a similar role to that of the Q61L mutation in RasGAP) only leads to a reduction in k_{cat} of approximately 10^2 (see Graziano & Gilman, 1989; Schweins, 1991) as compared with the 10^6 effect of the Q61L mutation. If the catalytic glutamine were involved in a special direct interaction with the TS, one would have expected similar contributions in the two related systems. Further support has been provided by a mutational experiment with an unnatural amino acid (Chung *et al.* 1993a) substitution of Gln^{61} (Q61NGln) also leads to a very small anticatalytic effect in the RasGAP complex. This finding was also reproduced by our simulation studies.

In light of the above findings, we consider the indirect effect of GAP to be a major factor in the activation of the signal transduction switch. Apparently, relatively small structural rearrangements in Switch I, II, β_3, and the P-loop are sufficient to drastically change the TS stabilization by the RasGAP complex. The system can be considered to be metastable with regard to its native sequence, where mutations of some residues change the energy landscape, leading to the destruction of the catalytic effect. Mutations of Gln^{61} are particularly effective at destroying the catalytic configuration, leaving the system in its ON state, and thus causing cancer.

At this point it might be useful to comment about experimental analysis of the IR absorption bands of the different phosphate groups (measured using Fouriertransform infrared (FTIR)

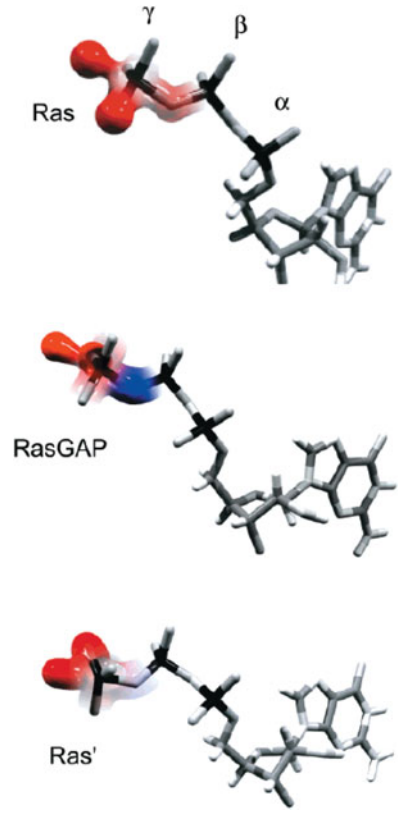

Fig. 30. An illustration of the change in electrostatic interactions between the protein residues and the γ- and β-phosphates of GTP, for the reaction catalyzed by (*a*) Ras, (*b*) Ras–GAP, and (*c*) Ras'. Blue indicates a stabilizing shift and red indicates a destabilizing shift in the interactions, with the degree of stabilization or destabilization being proportional to the intensity of the color. Note that what is presented here is *not* the change in the electrostatic potential, but rather, the actual change in electrostatic energy. This figure was originally presented in (Glennon *et al.* 2000).

spectroscopy) in the isolated GTP and in Ras-bound GTP, pointing toward a large decrease in the P–O bond order, and the fact that the oxygen atoms of the β-phosphate interact more strongly with the protein environment than those of the γ-phosphate (Allin & Gerwert, 2001). This effect becomes larger in the RasGAP complex (Allin *et al.* 2001). Remarkably, this trend validates our early prediction (Glennon *et al.* 2000; Muegge *et al.* 1996) that there is a large shift of positive electrostatic potential toward the β-phosphate upon the binding of GAP (Glennon *et al.* 2000; Klähn *et al.* 2006) (Fig. 30). However, the mechanistic interpretation of the observed effect is not immediately obvious and has been misinterpreted. Although this effect was taken as evidence for the existence of either a more dissociative TS (Allin *et al.* 2001) or a more concerted TS (Du *et al.* 2000), our detailed calculations (Glennon *et al.* 2000) demonstrated that the protein charge distribution that leads to the shift of positive charge toward the β-phosphate catalyzes the associative and dissociative processes equally (thus the charge shift induced in the substrate is not a unique mechanistic marker, in a clear contrast to attempt to argue so (see also the discussion above of the confusion caused in e.g. Chakrabarti *et al.* 2007) because of the acceptance of the unjustified assertions of Maegley *et al.* 1996. Furthermore, the FTIR reflects the charge distribution of the reactant state configuration, rather than the TS. Nevertheless, the spectral shifts are

indicative of a large change in the polarization of the substrate by its surrounding environment, which is the signature of the activation by GAP.

In conclusion, we note that it is hard to see why even now, when there is a general acceptance of the validity of *ab initio* based charge distributions, and when calculations of electrostatic effects in proteins are widely accepted as reasonable treatments, there is nevertheless still a tendency to (at times uncritically) accept arguments based on completely qualitative empirical LFER-based charge distributions, and on electrostatic considerations (such as the assertions about the presumed stabilization pattern of the Arg finger) that are not based on any valid quantitative analysis, and cannot predict the magnitude of any observed effect.

4.1.4 EFs and ribosomal activation

The elongation cycle of protein synthesis uses the thermo unstable elongation factor (EF–Tu) complex with GTP in order to deliver aminoactyl-tRNA (aa-tRNA) to the mRNA-programmed ribosome (Rodnina & Wintermeyer, 2009). More specifically, the GTP-bound state of EF–Tu forms a high-affinity ternary complex with the aa-tRNA. Upon binding of this ternary complex to the ribosome, the aa-tRNA occupies the A-site, and, when the codon–anticodon interaction is cognate, the GTPase activity of EF–Tu is significantly increased. The conformational change following GTP hydrolysis to generate GDP and P_i leads to the dissociation of EF–Tu from the ribosome and the accommodation of the aa-tRNA on the A-site for peptidyl transfer.

Breakthroughs in the elucidation of the structure of the ribosome (Agirrezabala & Frank, 2009; Ehrenberg, 2009; Liljas, 2004; Ogle *et al.* 2001; Ramakrishnan, 2010; Schmeing & Ramakrishnan, 2009; Schmeing *et al.* 2009; Schuette *et al.* 2009; Steitz, 2010; Villa *et al.* 2009; Voorhees *et al.* 2010; Yonath, 2010) as well as careful biochemical studies (Agirrezabala & Frank, 2009; Bilgin & Ehrenberg, 1994; Cool & Parmeggiani, 1991; Daviter *et al.* 2003; Hausner *et al.* 1987; Jacquet & Parmeggiani, 1988; Knudsen *et al.* 2001; Krab & Parmeggiani, 1999; Lancaster *et al.* 2008; Liljas, 2004; Moazed *et al.* 1988; Mohr *et al.* 2002; Scarano *et al.* 1995; Tinoco & Wen, 2009; Zeidler *et al.* 1995) allow one to begin thinking about the nature of the mechanism of the elongation process (see e.g. Moore, 2012). In particular, the recent elucidation of structure of the EF–Tu/ribosome complex (Voorhees *et al.* 2010) which we referred to in our study as EF–Tu', (Adamczyk & Warshel, 2011) paved the way for an exploration of the activation process at the molecular level. However, despite major biochemical and structural breakthroughs, a detailed explanation of *how* codon recognition in the 30S subunit leads to GTP hydrolysis still remains elusive. Although it is known that the precise positioning of H84 is critical for efficient catalysis (Cool & Parmeggiani, 1991; Daviter *et al.* 2003; Ramakrishnan, 2008; Scarano *et al.* 1995; Schmeing & Ramakrishnan, 2009; Schmeing *et al.* 2009; Schuette *et al.* 2009; Voorhees *et al.* 2010; Zeidler *et al.* 1995), the energetics of this positioning and its ultimate role in stabilizing the TS are unclear. More precisely, it is frequently assumed that the conserved H84 is moved to a catalytic configuration, and then serves as a general base, but this assumption is problematic (see discussion in Daviter *et al.* 2003) and has similar pitfalls as in the highly related case of the RasGAP system. As discussed above, while it was originally assumed that Gln^{61} in the RasGAP system serves as a general base, this assumption has been shown to be unlikely (e.g. Shurki & Warshel, 2004). In fact, as discussed above, it has been concluded that the activation process is due to allosteric changes in the P-loop, which are disrupted by mutations of Gln^{61}. However, it has been difficult to experimentally characterize the structural changes associated with the Gln^{61} mutation in the RasGAP complex, or to even clearly identify structural changes in the P-loop upon

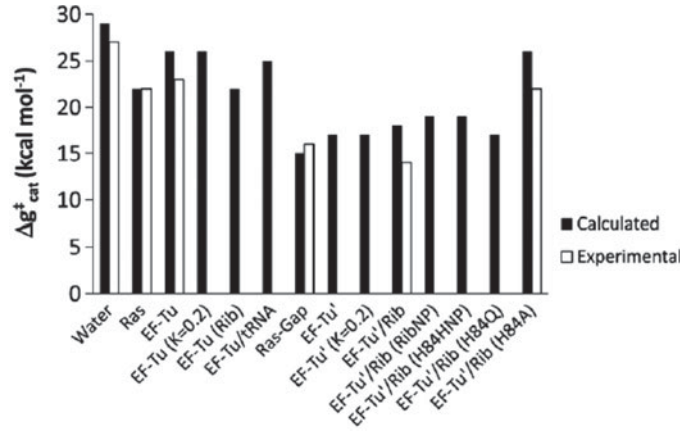

Fig. 31. Calculated and observed activation barriers (in kcal mol^{-1}) for different reacting systems. The notation NP (nonpolar) indicates that all the residual charges of the given residue or set of residues are set to zero. The notation 'rib' indicates the presence of the ribosome in the simulations. This figure was originally presented in (Adamczyk & Warshel, 2011).

complexation. Fortunately, there exists direct structural evidence (Voorhees *et al.* 2010) that, in the EF–Tu'/ribosome system, the structure of the P-loop changes. This would suggest that allosteric structural changes are far more likely candidates for the activation process than the effect of positioning the H84 side-chain in an orientation that presumably makes it a good general base.

As in the case with the RasGAP system, we were able to reproduce the experimental trend for the reaction in key systems (both EF–Tu and EF–Tu', Adamczyk & Warshel, 2011). It was found that the height of the activation barrier is already reduced in EF–Tu, and, to a much greater extent in the EF–Tu'/ribosome complex upon stabilization of the TS. We were additionally able to reproduce the effect of the H84A mutation (see Fig. 31), and to capture the important fact that the H84A mutation leads to a loss of the catalytic effect. After demonstrating our ability to reproduce the observed effect of the H84A mutation, we could begin to explore the origin of the GTPase activity, as well as its relationship to the structural orientation of H84. As a first step in this direction, we mutated H84 to its nonpolar analog (H84H^{NP}), by forcing all charges on the side chain to be zero (i.e., this 'mutation' was done for the model where H84 was explicitly ionized, to a model where H84 was completely nonpolar – see the related discussion in Section 4.1.2 on the arginine finger in Ras). Interestingly, this mutation did not reproduce the large observed anticatalytic effect of the H84A mutation.

To gain even more detailed insight into this issue, we extended our study (Adamczyk & Warshel, 2011), and simulated the catalytic free-energy landscape in the space defined by the chemical and conformational coordinate. The corresponding results for the native and H48A mutants are depicted in Fig. 32, and it shows that the H48A mutation shifts the landscape in such a way that it reduces the population of the EF–Tu' configuration. This change is a clear allosteric consequence of the binding of EF–Tu (or the ternary complex) with the ribosome. The allosteric effect is driven by a key interaction, i.e. that of H84 and the surrounding residues, including the sarcin-ricin loop (SRL). Therefore, these interactions are ultimately coupled to a conformational change that moves the system from the EF–Tu conformation to the EF–Tu' conformation. In other words, the H84A mutation disrupts the preorganization of the active site groups of the EF–Tu' system. It should be noted that a part of the effect of the H84A mutation already occurs at the EF–Tu' configuration, which reflects some local structural effects.

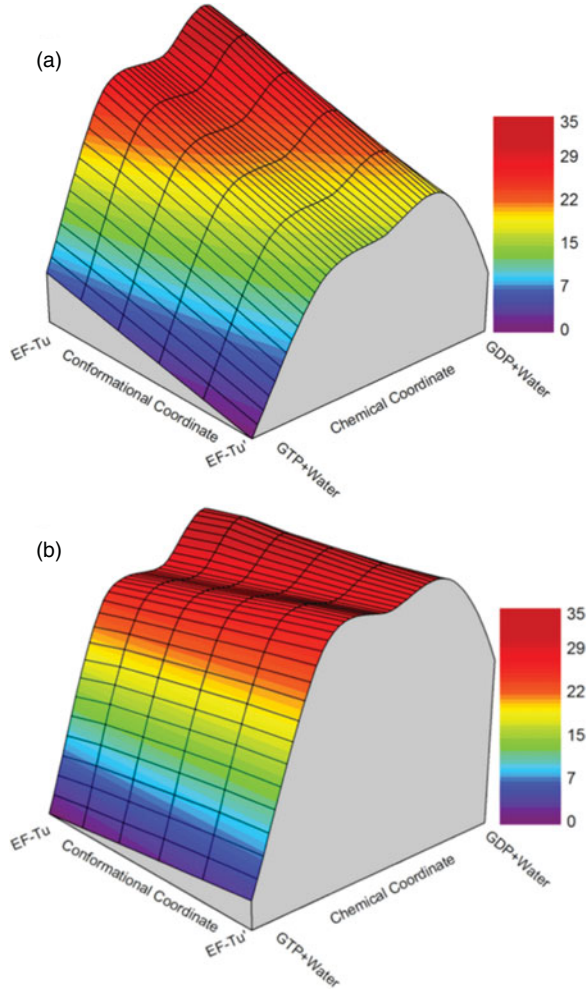

Fig. 32. An approximated catalytic free-energy landscape (kcal mol^{-1}) for the coupling between the chemical coordinate (i.e., the movement from the RS to the PS) and the conformational space that connects the EF-Tu and EF-Tu' conformations in the (*a*) native and (*b*) H84A mutant systems. This figure was originally presented in (Adamczyk & Warshel, 2011).

Overall, we succeeded in establishing that the catalytic effect of H84 is due to an indirect allosteric effect (i.e. by keeping other residues in a correct catalytic arrangement, for relative orientation see Fig. 33). Another significant point that emerged from our calculations was the finding that H84 has a pK_a above 10, and is therefore very likely to be protonated in the complex to its interaction with the phosphate group of residue A2662 of the 23S ribosomal RNA (rRNA) (see discussion in the supporting information of Adamczyk & Warshel, 2011). This finding indicates that H84 is unlikely to serve as a general base, as it is extremely unlikely to be deprotonated at physiological pH. Furthermore, the finding that the rate constant is pH-independent between pH 6 and 9 (Daviter *et al.* 2003) also provides strong evidence against the general base idea, as otherwise one would expect a decline in rate as the histidine becomes protonated (see also the discussion of this issue in Liljas *et al.* 2011; Voorhees *et al.* 2011). Nevertheless, we must point out that H84 may play somewhat more chemical role at the TS,where it can transfer its

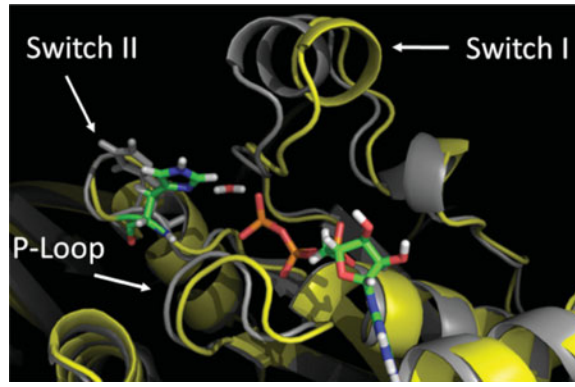

Fig. 33. The structures of the active site in both the EF-Tu (gray, from PDB code: 1EFT) and EF–Tu'/ribosome (yellow, from PDB code: 2XQD) from our simulations are compared, respectively. The changes in the backbone of this region do not occur in the EF-Tu protein alone. Critical regions (P-Loop, Switch I, and Switch II) are labeled. GTP and water are included in the RS configuration of EF–Tu'/ribosome. Mg^{2+}, aa-tRNA, and ribosome are not shown for the sake of clarity in this diagram. This figure was originally presented in (Adamczyk & Warshel, 2011). Note that even though His 84 is not presented as explicitly protonated here, the actual simulations were carried out considering His 84 as protonated.

proton directly to the phosphate oxygen and accept a proton from the nucleophilic water, whose pK_a is drastically reduced upon forming the bond to the phosphate (see Fig. 22). However, even if the energetics of such a mechanism will be found to be lower or similar to that of the 1 W, we believe that most of the catalytic effect is coming from the allosteric structural change.

In order to gain a deeper understanding of the nature of the catalytic effect in the EF–Tu'/ribosome complex, it is important to clarify the structural role of the ribosome during the activation process. Here, it is important to consider the effect of the SRL of the 23S rRNA subunit, and, in particular the role of residue A2662 (see Fig. 33 for relative positions), whose critical role has been already pointed out elsewhere (Hausner *et al.* 1987; Lancaster *et al.* 2008; Moazed *et al.* 1988). Significantly, our simulations of the EF–Tu'/ribosome complex found that the presence of the SRL loop (with its A2662 residue) was essential for the proper positioning of H84 (without any additional artificial force constraints). The remaining ribosomal components that compose the GTPase binding site play lesser (but nevertheless important) roles in the preorganization of the active site required for GTP hydrolysis. These included the rRNA and the L11 and L12 proteins (Mohr *et al.* 2002). Interestingly, mutation of the SRL residues near A2662 (Hausner *et al.* 1987) or cleavage of the SRL between residues A2662 and G2661 with the toxin a-sarcin (Bilgin & Ehrenberg, 1994) leads to significantly reduced levels of GTP hydrolysis (Voorhees *et al.* 2010). Our computational studies confirm that the SRL is necessary for moving the system to the EF–Tu' conformation, as the presence of residue A2662 (or a related external constraint on H84 that indirectly pushes EF–Tu to the EF–Tu' structure) is required in order to be able to reproduce by the computations the experimentally observed catalytic effect (see Adamczyk & Warshel, 2011).

As discussed above our exploration of the catalytic effect, indicated that the activation is not directly due to the repositioning of H84, but rather to an overall allosteric transition that forces EF–Tu (and, in particular, the P-loop) into a preorganized catalytic configuration (EF–Tu'). In light of this finding, we consider the indirect effect of the ribosome to be a major factor in the activation of GTPase hydrolysis. Specifically, relatively small structural rearrangements in Switch I, Switch II, and the P-loop (due to the formation of the EF–Tu'/ribosome complex)

drastically alter TS stabilization (see Section 4.1.2 for discussion of the related situation with RasGAP). Overall, the system can be considered metastable with regard to its native sequence, where the mutations of some residues change the free-energy landscape, leading to either the destruction (or enhancement) of the catalytic effect. Mutations of H84 are particularly effective at destroying the catalytic configuration, leaving the system in its inactivated state, and thus preventing protein synthesis.

Our analysis of the nature of EF–Tu activation by the ribosome provides significant support to our previous proposal on Ras activation. Even in the RasGAP complex (where it is harder to obtain structural information about Gln^{61} mutants), careful simulation studies (Shurki & Warshel, 2004) deduced a similar mode of activation. Consequently, these consistent results for the EF–Tu and Ras systems shed light on the general activation mechanism of GTPases, where binding to another protein or subunit greatly accelerates the intrinsic reaction. Our work on EF–Tu (Adamczyk & Warshel, 2011) provides further evidence that this activation process involves major 'allosteric' effects.

Finally, it is useful to comment on the fact that a QM/MM energy minimization study (Grigorenko *et al.* 2008) of the same type that led to support for the Gln-as-base mechanism in Ras (see Section 4.1.3) has also supported the histidine-as-base mechanism in EF–Tu. In this particular case, the problem is even more apparent; however, since, as highlighted above, His^{84} is actually protonated in this system and therefore *a priori* cannot be the base.

4.1.5 Related issues and systems

Other very important (and potentially instructive) benchmarks for the understanding of G-proteins are provided by Transducin-α (Bohm *et al.* 1997; Lambright *et al.* 1994; Sondek *et al.* 1994) and $G_{i\alpha 1}$ and $G_{s\alpha}$ (Sprang, 1997a; Sunahara *et al.* 1997). Here, it is important to understand what the main catalytic factors are, and how they are similar to or differ from those that are in operation in RasGAP. Both of these important signal transduction proteins have their catalytic glutamines in the same orientation and position as in the RasGAP complex, and both glutamines are thought to play a similar role as they do in RasGAP. Additionally, as will be also touched on in Section 4.4, there exist other G-protein complexes such as the RapGAP (Scrima *et al.* 2008), which are capably of catalyzing GTP hydrolysis much like the RasGAP complex, yet both the catalytic glutamic acid and Arg residues of the RasGAP complex are missing. This highlights the problem of oversimplifying the system to look at individual residues rather than the totality of the catalytic and allosteric effect.

4.1.6 Mechanistic insights, controversies, and misunderstandings

Since the Ras system has been the focus of the broader controversy with regard to the nature of phosphoryl transfer reactions in proteins, it may be useful to summarize here the key points considered above in a way that clarifies to the readers what has been said, the rationale beyond different statements, and the resulting confusion.

As outlined above, and as will be also summarized in Section 5, the most extreme recent argument has been about the phosphate as a base proposal of the general base. In fact, our substrate-as-base mechanism (Langen *et al.* 1992) was the paradigm shift that established that Gln cannot be the base (as was almost uniformly accepted at that time), and that the proton end up on one of the non-bridging oxygen atoms of the phosphate in the TS.

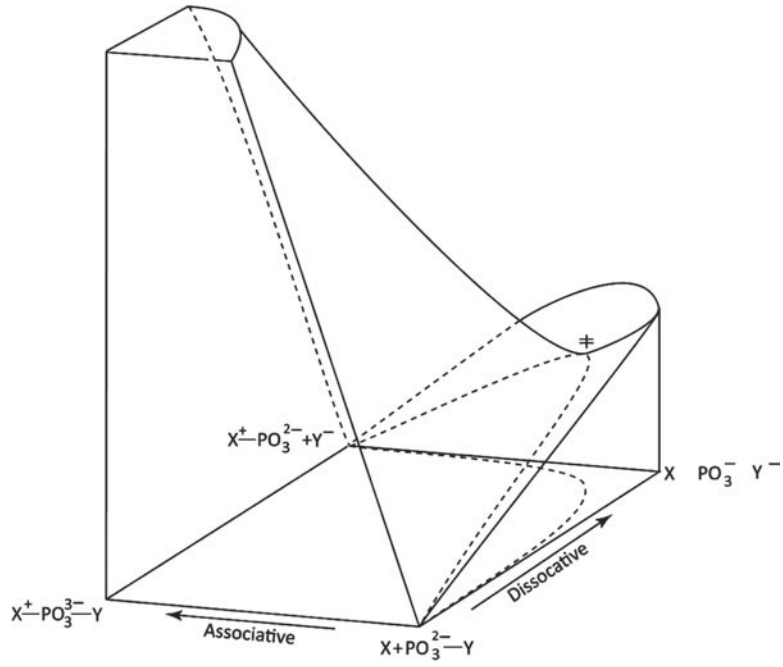

Fig. 34. The general features of a proposed energy surface for phosphate hydrolysis, adapted from (Admiraal & Herschlag, 1995), where it was suggested (without quantitative considerations) that an associative pathway will have an essentially insurmountably large energy barrier compared with a dissociative pathway (which theoretical studies have subsequently proven to be incorrect, see discussion in main text).

In this respect it might be instructive to point out that even if the PT occurs through a two-water mechanism, with the path described in Fig. 20 (Prasad *et al.* 2012) having a slightly lower energy that that of PT from the nucleophile water, it has very little to do with an alternative water as a base mechanism. That is, water as base mechanism requires that the water molecule actually accepts the proton rather than serve as a shuttle and the phosphate remain the base where the barrier is correlated with the pK_a of the phosphate oxygen. Thus, the most likely situation is that the proton is accepted by one of the non-bridging oxygens of the attacked phosphate.

The other issue is whether the TS is associative or dissociative that is, it has been stated (Admiraal & Herschlag, 2000; Zalatan & Herschlag, 2006) that '*A substantial amount of data from the field of physical organic chemistry supports a dissociative, metaphosphate-like* TS *for the direct reaction of water and other nucleophiles with phosphate monoesters*', despite theoretical evidence to the contrary (e.g. Florián & Warshel, 1998; Kamerlin *et al.* 2008a; Klähn *et al.* 2006) among others). This was most notable in (Admiraal & Herschlag, 1995), which presented an energy diagram that clearly indicated that any associative pathway is likely to be significantly higher in energy than a dissociative pathway (the key features of this diagram are given in Fig. 34). This view was (even if one ignores the unrealistic energetics) clearly demonstrated to be wrong by subsequent theoretical studies (Florián & Warshel, 1998; Kamerlin *et al.* 2008a; Klähn *et al.* 2006).

Similarly, the interpretation of experimental markers for examining charge shift during the reaction as support for a dissociative mechanism that presumably explains the action of GAP is extremely problematic. The elegant time-resolved FTIR spectroscopic studies, which examined the molecular GTPase-binding mechanism of Ras, basically corroborated out earlier theoretical

finding that Ras catalyzes GTP hydrolysis by shifting negative charge from the γ- to the β-phosphates. Unfortunately, this charge shift has incorrectly been attributed as being evidence for a dissociative pathway, by Allin *et al.* 2001, who state that '*the GAP that stimulates Ras GTPase activity can further enhance this negative charge shift from* γ*- to* β*-phosphate*) *These results support but not prove a dissociative mechanism*'. In fact, our earlier computational study (Glennon *et al.* 2000) demonstrated that the overall calculated effect of GAP and the relevant mutations examined are similar for both associative and dissociative pathways, and thus that the conclusions are in fact *independent* of the actual mechanism (and, anyhow, as discussed in this section, if anything, the most likely mechanism for GTP hydrolysis by Ras is in fact an *associative* or *concerted* and not a *dissociative* mechanism). In any case, the above argument was then further promoted by QM/MM studies which explored the charge shift upon binding of GAP to Ras in terms of the change in vibrational frequencies (Heesen *et al.* 2007), and came up with similar conclusions (missing, however, the point that the calculated frequencies are observed in the ground state rather than the TS).

We emphasize once again that the biggest root of all the resulting problems is the incorrect assumption that the TS must be dissociative, and this highlights the danger of taking a qualitative and unverified suggestion as fact without quantitative backing. In fact, as discussed extensively in this section, we have clearly demonstrated that the substrate-as-base mechanism operates through far more associative or concerted TS. However, it has been argued (most recently in Lassila *et al.* 2011) that there is no impetus for deprotonating the attacking water molecule in the TS (which as we know from the discussion in Section 3.5 is incorrect since the pK_a of the attacking water changes drastically). But, if this were true, where would the proton ultimately end up? It is trivial to show that a phosphate species (intermediate or product state) with both protons attached to the water molecule is higher in energy than the corresponding species in which one of the protons has been transferred to a non-bridging oxygen of the phosphate. So, clearly, the proton has to be transferred somewhere, and will not remain on the water molecule for the entire duration of the reaction (as this results in an energetically unfavorable species). Maegley *et al.* 1996 argued that, in the TS, the proton will have almost fully been transferred to the leaving group. However, the proton cannot fly over the phosphate from the water molecule to the leaving group, so, if leaving group protonation is even at all required, some form of proton relay is required. Here, simple pK_a considerations suggest that the non-bridging oxygens of the phosphate are far more favorable proton acceptors than a water molecule with a pK_a of $-1{\cdot}7$.

The next issue is that of what 'dissociative' TS really is. Conventionally, TSs are defined as associative or dissociative based on bond orders at the TS to the incoming nucleophile and departing leaving group. However, this is not very informative about the reaction progress, and also, using a bond order definition without looking at the full surface would make almost any TS seem dissociative. The scaling of bond orders is quite extreme such that an elongation of 0·5 Å would make the P–O_{lg} bond become defined as largely broken, whereas, for the nucleophilic oxygen, which typically is 3·0–3·5 Å from the phosphorus oxygen in the ground state complex, it would have to come at least 0·9–1·0 Å closer to the phosphorus atom for a bond to the nucleophile to be considered largely formed (bond order greater or less than 0·5), despite the fact that such a TS would clearly be associative as bond formation to the nucleophile greatly outweighs bond cleavage to the leaving group. In true dissociative TS, per definition, the bond to the leaving group would have to be essentially fully cleaved before any bond formation to the nucleophile occurs such that the reaction is driven by the bond cleavage and corresponding charge build up on the central metaphosphate moiety. However, while our extensive calculations have often shown quite expansive TSs (Florián *et al.* 1998; Florián & Warshel, 1997;

Kamerlin *et al.* 2008b; Klähn *et al.* 2006) (as one of usually multiple mechanistic possibilities), we have yet to obtain a *true* dissociative TS, where the reaction is being dominated by bond cleavage to the leaving group. Which gives rise to the question: does such a species really exist, or will the ability of phosphate to act as a proton sink always drive toward a more associative mechanism, even in aqueous solution?

The final issue comes from the fact that despite their prominence in the field, Herschlag and co-workers have been overlooking a main catalytic factor in enzymes, trying to prove that electrostatic catalysis is not important, and have used similarly problematic arguments as in the case of ketosteroid isomerase (Åqvist & Feierberg, 2002; Kamerlin *et al.* 2010; Kraut *et al.* 2006; Schwans *et al.* 2009; Sigala *et al.* 2008; Warshel *et al.* 2007). In the case of phosphoryl transfer reactions, this starts already in Fig. 1 of Maegley *et al.* 1996, which shows a hypothetical change of charge in a dissociative TS. In such a situation, one would assume that the departure of the leaving group results in a build up of positive charge on the phosphate itself, and led to the argument that, in the case of an external charge such as the arginine finger in the RasGAP system, '*Although electrostatic interactions with the* γ*-phosphoryl oxygens are not predicted to be strengthened in a dissociative* TS and may even be somewhat weakened (Maegley *et al.* 1996). It is then argued that for a catalytic interaction between the arginine and the γ-phosphate oxygens to be catalytically relevant, the enzyme has to change the nature of the TS from associative to dissociative. Therefore, (Maegley *et al.* 1996) claims that in RasGAP, the effect of the arginine finger is mainly to provide a 'strengthening hydrogen bond' to the leaving group oxygen, and the interaction with the γ-phosphate is mainly positioning, which '*cannot be taken as evidence for an associative* TS' (Maegley *et al.* 1996). Unfortunately, these logically ill-defined arguments cannot be formulated in a verifiable way, nor do they provide a quantitative prediction of any catalytic effect. Of course, the main problem is that the TS *is* in fact associative or concerted, and this problematic argument, which is based on superficial qualitative observations, led many workers to completely overlook the key point that the electrostatic effect of the arginine finger is not due to having a dissociative TS.

4.1.7 Concluding points

This section presents a detailed overview of a number of representative G-proteins, specifically the signaling protein $p21^{Ras}$ (Ras) and the EF–Tu, which plays a central role in the termination of the elongation phase of protein synthesis. In both cases, the actual chemical step involves the attack of a nucleophilic water molecule on GTP to yield GDP and P_i, and the relevant enzyme provides major electrostatic stabilization to the TS of the GTP hydrolysis. Additionally, as discussed in the text, both Ras and EF–Tu are activated by binding to other systems (e.g., protein ribosome). Our detailed theoretical studies demonstrate that, while the primary effect of this complexation is allosteric, even this is based on pure electrostatics. For both cases, the protein-protein interactions causes a conformational shift that brings the system to a conformation in which it provides even greater electrostatic stabilization of the TS, as well as the corresponding product state (i.e. GDP and P_i). As outlined in Section 4.4, due to the similarities between the different G-proteins, we believe that these conclusions are general and extendable to related systems, and theoretical studies are currently underway to verify this fact. Finally, as a concluding note to this section, we would like to again refer to our analysis above of the experimental 'evidence' for a dissociative mechanism, and the dangers of confusing the oversimplified *analysis* of elegant experiments with the actual experimental facts.

4.2 ATPase and energy transduction

4.2.1 Introduction

Phosphate hydrolysis provides a crucial source of biological energy, and, in particular the biological conversion of ATP to ADP plays a crucial role in energy transduction in life processes (as well as for the practical understanding of the action of molecular motors, Boyer, 1997; Weber & Senior, 1997). Structural, biochemical, and conceptual studies have shed major light on the details of this biological energy conversion process, but major issues remain have remained unresolved (Abrahams *et al.* 1994; Boyer, 1993, 1997; Cherepanov *et al.* 1999; Fersht, 1999; Wang & Oster, 1998; Weber & Senior, 1997). For example, despite the fact that the structures of several ATPases have been elucidated (e.g. Abrahams *et al.* 1994; Menz *et al.* 2001), and the description of the different conformational states is known, as well as recent advances in detailed elucidation of some of the key steps in the process (Adachi *et al.* 2007; Okuno *et al.* 2008; Shimo-Kon *et al.* 2010), it is still not clear at what stage in the reaction energy is actually released and, in turn, what the nature of the energy conversion process actually is. In fact, even the origin of the energy difference between ATP and ADP + P_i is far from being clear (although it is very likely that a significant part of this energy is associated with the electrostatic repulsion between the negatively charged phosphates, see Section 4.2.2).

Attempts to obtain relevant structure-function correlations for these systems include QM/MM studies of the reaction profile of the chemical steps in different subunits (Beke-Somfai *et al.* 2011; Dittrich *et al.* 2003; Štrajbl *et al.* 2003). However, with the exception of an EVB study (Štrajbl *et al.* 2003), there are (presently) no actual free-energy calculations available in the literature. Additionally, studies of the energetics of the reactant and products in different states of the F1-ATPase subunits included microscopic free-energy calculations using MM approaches (Yang *et al.* 2003), that did not actually simulate the chemical process (or the crucial separation of the ADP and P_i fragments). At present, the semi-macroscopic LRA calculations of the chemical step in different subunit states (Štrajbl *et al.* 2003) seems to yield more reliable results when we also consider a careful analysis of the reference reaction in solution (Kamerlin & Warshel, 2009). Although it is possible to obtain reasonable insight into the energetics of the chemical steps within each subunit (Štrajbl *et al.* 2003; Yang *et al.* 2003), the main challenge here is to obtain insight into the overall energy change in the overall energy conversion process of the complete system (see below).

The F_1-ATPase is composed of a stator ring and an elongated α-helical rotor, which is part of the complete F_0F_1complex. The rotor is known as the γ-subunit, and it connects the ring with the membrane-bound F_0 portion of the enzyme (Fig. 35). The stator ring is, in turn, composed of three functional dimers, each consisting of a pair of α/β-subunits, where the β-subunits bind the nucleotide involved in the hydrolysis/synthesis process. The dimers can adopt three distinct conformations, designated here as open (E), loose (D), or tight (T) forms. During the synthesis cycle, rotation of the rotor utilizes the electrochemical gradient established through the transport of H^+ ions across the membrane, while the three dimers of the stator ring show conformational alterations between the E, D, and T states, resulting in the production of three ATP molecules in one complete rotation of the γ-subunit. Fig. 35 schematically describes the ATP hydrolysis, and the associated conformational changes in the a/b- dimers, through one complete (120°) rotation of the γ-subunit, where we designate ADP + P_i, ATP, ADP, and P_i as *d*, *t*, *d′*, and *p* in the subsequent discussion. The figure reflects two alternatives mechanistic views, based on the available structural data and kinetic results, and, to a significant extent, on innovative single-molecule

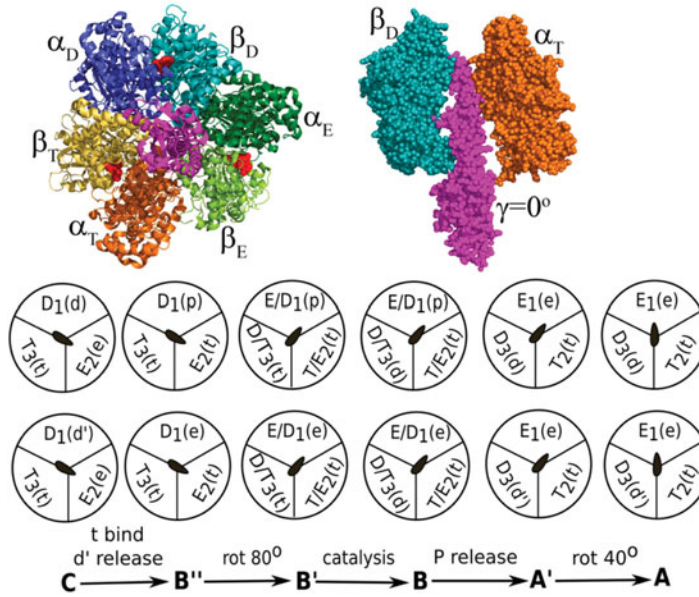

Fig. 35. Schematic representation of the ATP hydrolysis cycle of F1-ATPase, during the 120° rotation of the γ-subunit. Scheme I and II describes two of the most likely mechanisms. An atomistic picture of the system is shown on top. The stage between C and B″ may involve three nucleotides supporting a tri-site mechanism (not shown explicitly). This figure was originally presented in (Mukherjee & Warshel, 2011).

experiments that have visualized the rotation of the γ-subunit and have been able to identify key details about the overall cycle (Adachi *et al.* 2007; Okuno *et al.* 2008; Shimo-Kon *et al.* 2010). The rotation of the γ-subunit was observed to be a stepwise process involving three 120°steps, where each step involved one hydrolysis/synthesis event. The steps are intercepted by a time dwell (known as the 'catalytic dwell'), that was found to consist of 2 ms timescale events, one of which is the actual chemical step, and the other is most likely the phosphate release step.

4.2.2 The energetics of the chemical step

Our initial analysis of the energetics of some key steps in the mechanochemical process involved the computational results of (Štrajbl *et al.* 2003). This study explored different thermodynamic cycles, including the use of the LRA for capturing the conformational free energy, and obtained relatively stable and consistent results result for the binding energies in the different subunits. In addition we also performed preliminary EVB calculations of the activation barriers in the different subunits. The most interesting conceptual result was the finding that a significant part of the chemical energy might be stored in the dissociation of the ADP and P_i in water and is not related to the bond breaking energy. This indicated a key role for the release of the electrostatic repulsion between the phosphate units. This can be best illustrated by the analogy between F_1-ATPase and the action of the RasGAP system discussed in Section 4.1. In the latter case, the binding of GAP to $p21^{ras}$ greatly accelerates GTP hydrolysis, due to a shift of positive charge from P_γ to P_β (Glennon *et al.* 2000; Muegge *et al.* 1996; Schweins *et al.* 1996b). This in turn leads to the stabilization of the GDP + P_i state relative to the GTP state. In contrast, in the case of F_1-ATPase (Štrajbl *et al.* 2003), the D→T conformational change has been demonstrated to stabilize

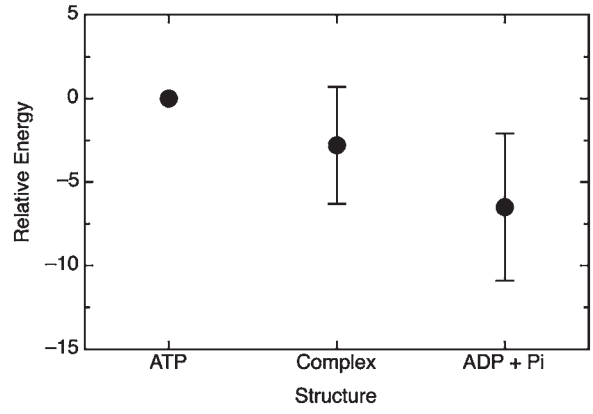

Fig. 36. Overall ΔG_{calc} for each of the steps of the reaction shown in Eq. (26), averaged over two different metal cations (Mg^{2+} and Ca^{2+}), as well as all three solvation models (QM/MM-FEP, COSMO and PCM), in the presence of explicit water molecules. All energies are given in kcal mol^{-1}, and the error bars show the standard deviation over three solvation models. This figure was originally presented in (Kamerlin & Warshel, 2009).

the *ATP* state, relative to its energy in D, and probably as a result of a shift of the electrostatic potential from P_β to halfway between P_β and P_γ. Now, the large electrostatic repulsion between the ADP + P_i fragments becomes relevant in the B→C process, where there exists an electrostatic cost toward assembling the fragments of the ADP + P_i system. This cost is paid for by the binding process, at which stage it is then rather simple for the protein to simply tune the energetics of the ATP state up or down, which can be achieved by increasing the electrostatic stabilization of the ATP state during the D→T transition, as mentioned above (note that this key was missed by (Yang *et al.* 2003), who explored free-energy changes during the B→C transition, and then purported to have found the missing link between structure and thermodynamics in F1-ATPase, but did not simulate the actual chemical reaction).

Now in order to explore the issue of the energetics of the dissociation of ATP into ADP + P_i further, we recently conducted a systematic study of the energetics of ATP hydrolysis in aqueous solution (Kamerlin & Warshel, 2009). Specifically, our aim was to model the energy conversion of the following reaction:

$$ATP \bullet M^{2+} + H_2O \rightarrow ADP \bullet P_i \bullet M^{2+} \rightarrow ADP \bullet M^{2+} + P_i \tag{26}$$

by means of several different solvation models (where M^{2+} indicates a divalent cation, that, in our case, was either Mg^{2+} or Ca^{2+}). Figure 36 shows the resulting overall ΔG_{calc} for each step of reaction (26), averaged over all solvation models, and over both metal cations. The error bars show the corresponding standard deviations in the calculations. It should be noted that despite being an apparently 'simple' process in aqueous solution, these are in fact *extremely* challenging calculations, as is illustrated by the large error bar observed in Fig. 36 (however we found that this problem could, to some extent, be mitigated by the inclusion of explicit water molecules in the calculations). This makes it therefore very difficult to obtain an overall free energy for this process in solution, a problem that is further compounded by the lack of available experimental data for this process. Nevertheless (Kamerlin & Warshel, 2009) still provides what is, at present, the most comprehensive and exhaustive exploration of the energetics of this process in aqueous solution, and, while we cannot reproduce exact free energies for this process, our calculations

clearly show that this is a downhill process with a dissociation energy that is sufficiently large that it should not be overlooked. Moreover, we examined the nature of the solvation effects in this process by a range of solvation models, and obtained a consensus result that not only is the bond breaking process making a significant contribution to the overall energetics, but also, this plays a significant part in the friction and dielectric compensation in this reaction, though more detailed careful entropic studies may be needed to explore this issue further.

Turning our attention back to the reaction in ATPases, we would like to point out to the reader that a consensus summary of the energetics of the different steps in the process in F1-ATPase is given in Table S1 of Mukherjee & Warshel, 2011. Obviously, this is still a work in progress; however, it also illustrates the need for careful experimental and theoretical analyses of this problem. In this respect, it may be useful to expand here on current QM/MM studies of ATP hydrolysis in proteins and in solution (Akola & Jones, 2006; Beke-Somfai *et al.* 2011; Dittrich *et al.* 2003; Grigorenko *et al.* 2007b; Li & Cui, 2004), some of which have not attempted to obtain quantitative results (e.g., Dittrich *et al.* 2003), and others of which have major problems due to not having included a sufficiently large part of the environment, and in having no sampling. In fact, the difficulties of reproducing the observed barrier led to the problematic message that has emerged from the work of (Grigorenko *et al.* 2007b). These works purported to obtain the same unreasonably low barrier of approximately 5 kcal mol^{-1} they obtained in the RasGAP system (see Section 4.1), but now without any higher barrier, and with a lower barrier step corresponding to the presumed PT between the two water molecules. This faces the problem that now the lowest barrier is much smaller than the value of approximately 14 kcal mol^{-1} that corresponds to the observed k_{cat}. To justify this, it was argued that the estimate from TS theory is a major overestimate, as 'the experimental rate constant incorporates contributions from conformational changes in the protein' from moving between the reactant and product. The confusion in this idea requires clear analysis, since it is related to the problems with using energy minimization. If the above idea means that the chemical barrier is lower than the conformational barrier, then it overlooks the fact that the chemical barrier associated with k_{cat} (or, more precisely, with the chemical step) has been unambiguously identified in unisite reactions (e.g., Weber & Senior, 1997) to be around 16–19 kcal mol^{-1}, and the fastest rate along the conformational path in F1 ATPase has been uniquely assigned to as being approximately 2 ms (with a barrier of about 14 kcal mol^{-1}, see also Section 4.2.4 below). However, if, as is more likely, the authors meant that the calculations miss the contribution of the protein structural changes to the chemical barrier, then we have here a realization that energy minimization approaches cannot capture the protein reorganization effect, which is an absolutely integral part of the barrier for the chemical step. Of course, one would also get an incorrect barrier if one freezes the protein (which is what is partially done by energy minimization studies) into a structure that corresponds to the TS conformation. However, this would simply be a poor modeling study, which provides no physically meaningful information.

It may also be useful to comment on attempts to consider the binding of ATP (Minehardt *et al.* 2002), or even ATP hydrolysis (Minehardt *et al.* 2003) by Car-Parrinello Molecular Dynamics (CPMD). Performing CPMD calculations in the protein led to ultrafast PT from an ionized lysine residue, due to insufficient stabilization of this group by the protein (see the discussion in Glennon *et al.* 2000 or a related problem in Futatsugi *et al.* 1999). However, even a more careful CPMD study (Akola & Jones, 2003), which explored the less challenging case of the reaction in water, has been problematic. These authors studied the hydrolysis of the highly charged reacting systems with only 54 water molecules over very few picoseconds, and concluded that '*The latter*

(*dissociative reaction*) *has an activation energy of 35 kcal mol*$^{-1}$*, where 25 kcal mol*$^{-1}$ *can be assigned to the P–O bond breaking and 10 kcal mol*$^{-1}$ *to the artificial stability of* PO_3^- *resulting from the small size and the short time scale of the simulation. The path and energy barrier* (*39 kcal mol*$^{-1}$) *of the less-favorable associative reaction suggest that it is possible only under conditions where the lytic water is already deprotonated to* OH^-'. Of course, presuming what the effect of the missing solvent would be is not a promising scientific approach for examining alternative mechanisms. Furthermore, the authors presume that the '*reactions discussed are very fast* (*9–13 ps*)' overlooking the fact that the reaction has an activation energy that leads to reaction rates of months, and in fact phosphate hydrolysis is one of the most thermodynamically inert reactions known (Wolfenden, 2011).

Subsequent studies by the same authors (Akola & Jones, 2003) tried to explore ATP hydrolysis in actin using MD simulations with an MM force field to generate starting points for cluster modeling. In this case, a perfectly valid S_N2 TS is called 'dissociative', and a reaction where another water serves as a base is defined as an 'associative' mechanism (the same problem mentioned in Section 3.5). The authors suggest that their work discriminates between different mechanisms, while presuming that their model includes a significant part of the protein electrostatics. Unfortunately, this work overlooks the fact that sampling with an MM force field that does not respond to the correct QM reaction charges cannot be effectively used in decent QM free-energy calculations of charge separation and transfer processes.

We must add that one cannot exclude at this point a PT at the associative (concerted) TS to another water molecule, but exploring this issue must involve a careful search in the direction of a direct PT to the phosphate oxygen (see Section 3.5).

4.2.3 The coupling between the subunits

One of the key unresolved questions about the action of F_1-ATPase is the nature of the coupling between the subunits. Attempts to determine the sequence of conformation/binding/chemical events have included studies that combined the available experimental information and kinetics (Gao *et al.* 2005) as well as simulation studies (Czub & Grubmueller, 2011; Koga & Takada, 2006; Pu & Karplus, 2008). However, the simulated directionality must result from the free-energy landscape itself, and not be imposed by targetted molecular dynamics (TMD) (Pu & Karplus, 2008) or by using switching functions to move between different structures (Koga & Takada, 2006). For example, the assertion in that the CG model reproduced the 80° rotation seems to reflect the fact that this rotation is observed in the X-ray of the target structure. In our point of view, studies that do not produce a free-energy surface for the whole process, considering the complete F_1-ATPase system, cannot be used to relate the protein structure to the directionality of the ATP hydrolysis/synthesis. Unfortunately, fully microscopic models are (presently) unable to capture the overall free-energy balance, with current computational resources. The problem is the difficulty of obtaining converging and meaningful free energies for the substantial conformational changes of large and highly charged systems.

In order to address this issue, we recently (Mukherjee & Warshel, 2011) used our CG model (Messer *et al.* 2010) to simulate the F_1-ATPase parts, consisting of three α/β-dimers and the γ-subunit. The CG model was first used to evaluate the overall electrostatic free-energy surface, without incorporating the effect of the ligands and/or chemistry, and the corresponding results are described in Fig. 37(*a*). As seen from the figure, the motion of the system is restricted to a least-energy path that produces the functional direction of the F_1-ATPase machine. The figure shows that the system undergoes an approximately 80° γ-rotation from the starting structure

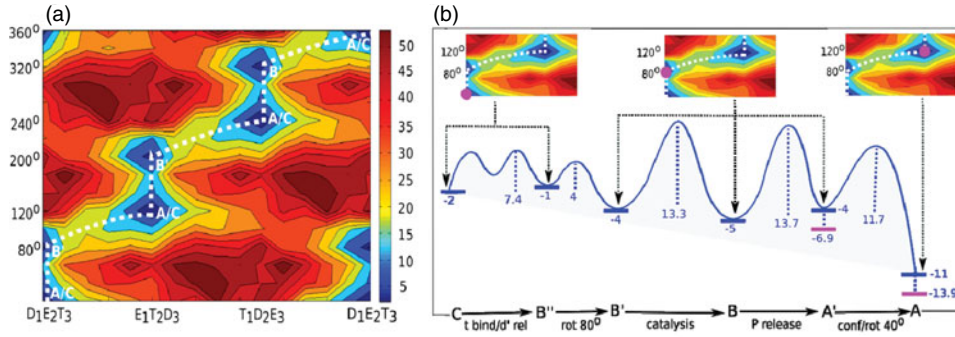

Fig. 37. Energetics of F1-ATPase in kcal mol^{-1}. Shown here are: (*a*) the calculated electrostatic free-energy surface in the space defined by the rotation of the central γ subunit (vertical coordinate) and the subunit alterations (horizontal coordinate). (*b*) Semiqualitative energetics (in the standard state without concentration effect and with the p concentration effect in magenta) along the least energy path based on the considerations in the Supporting Information of (Mukherjee & Warshel, 2011). This figure was originally presented in (Mukherjee & Warshel, 2011).

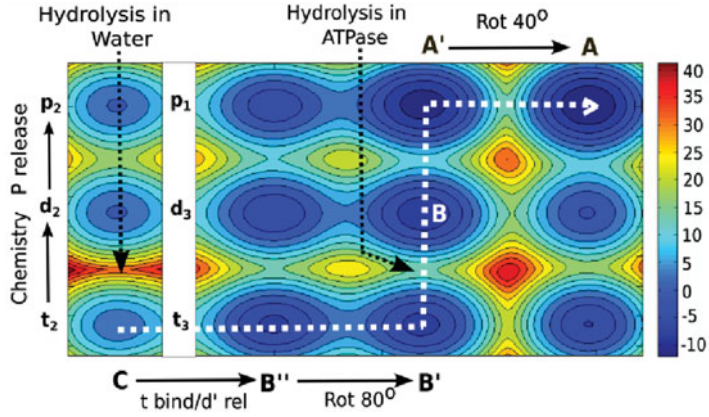

Fig. 38. The conformational/catalytic landscape of F1-ATPase. The figure describes the energetics of the system along the chemical coordinate of one nucleotide (bound to subunit 2 or 3) as well as the conformational coordinate of the combined system. The functional path is shown using a dashed white line. This figure was originally presented in (Mukherjee & Warshel, 2011).

(point C) followed by concerted γ-rotation and dimer conformational changes in the range of 40° (point A). Such an intermediate conformational barrier has not been reproduced by other studies in an unbiased way. Adding the energetic of the ligands allowed us to obtain the overall conformational/chemical landscape (Fig. 38). The figure incorporates key features of the chemical/conformational landscape, plus some elements of the binding/release coordinate, and yields a new fundamental insight. The key region for chemistry occurs at around region B', before the region where we found a high barrier that combines a conformational motion with a rotation of an additional approximately 40° (Fig. 37*a*). As this barrier is higher than the barrier for the hydrolysis, the system moves from t_3 to d_3 and probably releases the phosphate and then (or in concert) continues the motion from A′ to A at the upper part of the map (at this stage the conformational barrier decreases. It would seem to us that even though the treatment of Fig. 38 is somewhat tentative and simplified, nevertheless, even this advances our understanding of the

problem far beyond anything that has been obtained so far from other studies of the system, and can provide consistent insight into the events occurring in the F_1-ATPase system.

Now several works have proposed a steric mechanism (Koga & Takada, 2006; Pu & Karplus, 2008), or a somewhat related elastic spring model (Czub & Grubmueller, 2011), and asserted that such a mechanism emerged from their calculations. However, these works *forced* the system to move in the desired direction (by applying very large forces) rather than allowing the system to be driven by the chemical potential or by the proper external torque. It seems to us that the requirement for realistic modeling is having a surface that can reproduce, or at least approximate, the actual energy landscape. Apparently, the previous studies could not capture the energetics of the system, as is indicated by the fact that no energy values in a feasible range were provided. Moreover, these models cannot predict the observed unidirectionality but just impose it. The use of a microscopic model (Czub & Grubmueller, 2011), which only applied torque on the γ-subunits, has provided instructive information on the response of the other subunits to the γ-motion, but again this was done by use of a very large force.

One of the strongest arguments *against* the steric model is provided by the recent work of (Furuike *et al.* 2008), where it was established that the rotational motion occurs in the right direction even without most of the γ-subunit embedded inside the α/β-subunits. While this effect is reproduced by our electrostatic model (see the Supportive Information of Mukherjee & Warshel, 2011) it is unlikely to be reproduced by a consistent steric model, as most of the γ-subunit (which is packed inside the (α/β) subunit cavity in the native enzyme) is now absent.

An interesting hypothesis about 'phase shift catalysis' was put forward by Nordin and co-workers (Beke-Somfai *et al.* 2012) proposing that a free-energy synchronization at a macroscopic level may occur in F1-ATPase, leading to catalysis by destructive interference (e.g. at phonon level). It was suggested that such effects might be significant relative to normal (random) thermal fluctuations, thus helping in combining several sites in an energy-recycling mode. This idea is interesting but the relationship of the present model to concepts such as free energy and dissipation time (and the corresponding friction) remains to be established. Furthermore, the implicit proposal seems to suggest that the activation barrier in one site is overcame by a compensating effect in other sites, or the effect of the γ-subunit. It is not clear if such a condition can be established in a flexible protein system which seems to behave in the high friction limit (Pisliakov *et al.* 2009).

The use of the energy of ATP hydrolysis to drive and control key life processes is quite general (Berg *et al.* 2010). In this section, we have only focused on the F_1-ATPase, but the availability of even partial structural information now makes it possible to start to study the action of other ATP-driven vectorial processes, including the action of the helicase/translocase family. A recent example of such a system is provided by the Simian virus (SV40) large tumor antigen LTag, which acts as an efficient molecular machine that unwinds dsDNA (Borowiec *et al.* 1990; Fanning, 1992; Mastrangelo *et al.* 1989; Simmons, 2000). Our study of this system (Liu *et al.* 2009) used a CG model in which the DNA was represented by considering the electrostatic effect of its phosphate main chain, and similarly the helicase hexamer was represented by the corresponding electrostatic effect. The CG modeling produced a free-energy landscape with a clear vectorial direction for the coupling between the ATP hydrolysis in the six subunits and the translocation of the single strand DNA. However, since we do not at present have an X-ray structure for the system with its bound DNA, we are clearly not in the same advanced stages as with F_1-ATPase.

Another class of interesting ATPase activity is the machinery that is involve in the insertion of proteins through the membrane just after (or during) their synthesis by the ribosome (Dukanovic & Rapaport, 2011; White & von Heijne, 2008). This extremely interesting process is, in some cases helped by ATP driven processes, and might well involve a complex quasi-irreversible vectorial process (see the Supporting Information of Rychkova *et al.* 2010).

4.2.4 Concluding comments

Our recent work (Mukherjee & Warshel, 2011) has presented an actual structure-based landscape, and has provided powerful insight into the electrostatic origin of the conformational coupling, clarifying the principles used by nature to harness the ATP→ADP dissociation free energy. As shown in Fig. 38, it is very hard to harness the hydrolysis energy in water on biological time scales, since the corresponding barrier is approximately 30 kcal mol^{-1}, and since the chemical energy will not do any useful work if released in water. Now, in the ATPase, the barrier is reduced to approximately 14 kcal mol^{-1} and, as shown in Fig. 38, the movement from ATP to ADP and the subsequent release of chemical energy is coupled to the rotational motion and is thus capable of doing reversible work, when coupled to a proper load. This coupling of the chemistry with the conformational barrier releases the same energy that was stored in the water reaction, but now has a remarkable way to harness all the chemical free energy toward a vectorial process.

When asking why nature used this system, we note that having a triphosphate system transform into a diphosphate is basically a way to release electrostatic free energy upon cleavage of the bond holding the negative charges together. Now, in the case of a phosphate ester bond, the negatively charged phosphates are held together for a very long time in water, and the repulsive energy cannot be released. However, in the protein, the bond breaks fast, and the electrostatic energy released is used to do mechanical work, by electrostatic interactions between the protein units.

4.3 DNA polymerases and replication fidelity

4.3.1 Introduction

The reproduction and evolution of life depends on the accurate replication of the genome, which is facilitated by DNA polymerases (Kornberg & Baker, 1992). The synthesis of new DNA molecules would be impossible without these enzymes as they increase the rate of phosphodiester bond formation by many orders of magnitude compared with the corresponding uncatalyzed reaction in water (Schroeder *et al.* 2006; Wolfenden & Snider, 2001). Unlike the traditional image of enzyme catalysis proceeding with high selectivity for a single substrate, DNA polymerases are capable of selectively catalyzing the transformation of four structurally distinct substrates. A correct dNTP substrate (R) is selected for insertion by the polymerase depending on the identity of the templating base so that Watson–Crick base pairs are formed with high efficiency. The fidelity of DNA replication is in turn controlled by the polymerase active site where phosphodiester bond cleavage occurs and by the binding site of the incoming nucleotide, paired opposite a template base (see discussion in Beard & Wilson, 2003). The rate of incorporation of an incoming wrong nucleotide (W) is drastically slower than the corresponding rate of the right nucleotide (R), thereby ensuring high replication fidelity, as reviewed in

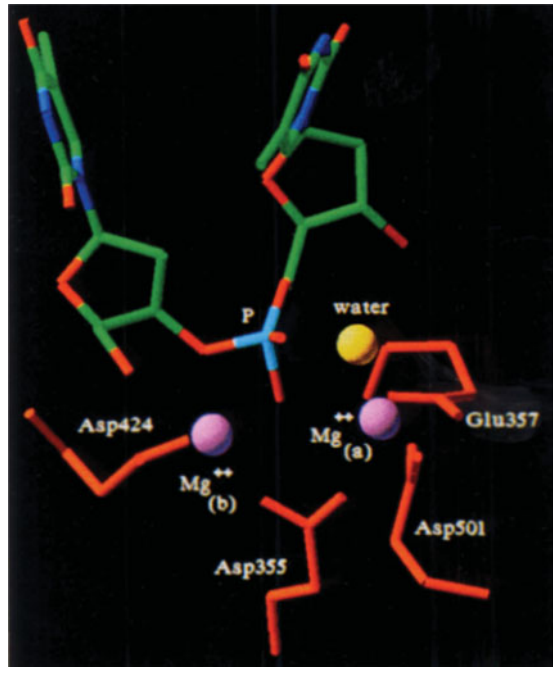

Fig. 39. An example of a typical polymerase active site. Shown here are key residues in the exo site of the Klenow fragment, outlining the position of the substrate, nucleophile, catalytic metal ions and key active site residues. This figure was originally presented in (Fothergill *et al.* 1995).

(Echols & Goodman, 1991; Johnson, 1993; Joyce & Benkovic, 2004). In exploring the function of DNA polymerases, it is essential to address several basic problems, which range from the mechanism of the phosphoryl transfer step, to the control of the fidelity through the modulation of the rate-limiting step. These issues will be addressed below.

The first detailed structure-based information about the mechanism of DNA polymerases emerged from structural studies (Steitz *et al.* 1994), which discovered the so-called 'two metal mechanism'. A typical polymerase active site is described in Fig. 39, where the precise reaction pathway and the contributions to the catalytic effect are not at all obvious from the crystal structure. The original structural studies were unable to resolve the question of whether the proton on the primer 3′ hydroxyl group goes to the carboxylate of either Asp^{256} or Glu^{357} (the authors proposed that there is no need for the participation of an amino acid sidechain for nucleophile activation in the chemical step), or what the precise role of the Mg^{2+} cations in the active site is (see also the controversy in Steitz *et al.* 1994). In fact, it was suggested by (Steitz *et al.* 1994) that Mg^{2+}(b) helps the reaction by steric effects, following Westheimer's idea (see discussion in Fothergill *et al.* 1995), but, the first theoretical study of this system (Fothergill *et al.* 1995) established that the effect of this magnesium cation is entirely electrostatic. Similarly, the theoretical study further clarified the problems with the idea that the carboxylate groups of the negatively charged amino acids are stabilizing the attacking hydroxyl group, pointing out that, in fact, the role of these negatively charged residues is most likely simply to position and shield the positive charge on the cationic Mg^{2+}, which will in turn make it much harder for the Mg^{2+} to lower the pK_a of the hydroxide, and therefore, that the most likely base is simply just another OH^- anion from the bulk. Subsequent theoretical studies (e.g. Florián *et al.* 2002, 2003a, 2003b;

Prasad & Warshel, 2011; Xiang *et al.* 2006) have shed more light on the mechanism of DNA polymerases, and this issue will be discussed in greater detail below.

The second key open question has been the control of the rate of nucleotide incorporation by the specific type of base pairing (which, in essence, is the control of fidelity). Here, despite significant experimental milestones in studies of the fidelity of DNA polymerases (e.g. Beard & Wilson, 2003; Echols & Goodman, 1991; Goodman, 2002; Goodman *et al.* 1993; Johnson, 1993; Joyce & Benkovic, 2004; Kunkel & Bebenek, 2000; Showalter *et al.* 2006), we still do not have a clear quantitative picture of the energetics of this progress. The progress made in using structural information to explore the question of fidelity will be reviewed in Section 4.3.5.

4.3.2 The chemical nature of the nucleotidyl transfer

Before embarking on a discussion of the fidelity (and other related issues), it is necessary to analyze the mechanism of the chemical reaction in the protein. In doing so, we assume that the reaction takes place in three generic steps (see Fig. 40). That is, a PT from the 3′ hydroxyl group of the primer is followed by inline nucleophilic attack of the now anionic oxygen of the primer on the dNTP α-phosphate (P_α), with the reaction being completed by the departure of the pyrophosphate leaving group (although, in principle, these steps could also be concerted in nature). For the PT step, there are also several possibilities: that is, the proton of the 3′-OH group of the primer could either transfer to the catalytic aspartic acid (e.g. Asp^{256} in Pol β and Asp^{654} in T7 polymerase) (Florián *et al.* 2003a), a non-bridging oxygen of the dNTP P_α), or a nearby water molecule. Based on previous simulation studies, it would seem that the dominant mechanism involves PT to the catalytic aspartic acid Florián *et al.* 2003a; Xiang *et al.* 2006), or to bulk water, followed by nucleophilic attack on P_α by the deprotonated 3′OH and the formation of a pentavalent intermediate in an associative mechanism (see Fig. 41).

When trying to understand and quantify the catalytic effect of an enzyme, it is absolutely essential to also perform a detailed study of the corresponding uncatalyzed reference reaction in solution, in order to understand the fundamental basic chemistry of the reaction (note that, per definition, in order to be contributing to catalysis, any effect observed in the enzyme has to be significantly different from that in the corresponding uncatalyzed reaction). Here, it is worth mentioning that the enzyme catalyzed reaction does not necessarily follow the same pathway as the preferred mechanism in aqueous solution, however, the *correct* reference state for an enzyme catalyzed reaction involves exactly the same mechanism and the same reacting atoms (for a general discussion of this issue see e.g. Warshel *et al.* 2006b), and for a discussion more specific to the reaction catalyzed by DNA polymerases, see Kamerlin *et al.* 2009c.

Section 3 introduced the substrate-as-base mechanism as a viable alternative for phosphate monoester hydrolysis in aqueous solution (see e.g. Florián & Warshel, 1997; Klähn *et al.* 2006), and such a mechanism is believed to operate in a number of phosphatases and even as a promiscuous side reaction in a sulfatase (e.g. Adamczyk & Warshel, 2011; Liljas *et al.* 2011; Luo *et al.* 2012a; Schweins *et al.* 1994). This type of mechanism can also operate in DNA polymerases (Florián *et al.* 2003a), and a version of this mechanism was used in a QM/MM energy minimization study of Pol β (Alberts *et al.* 2007; Radhakrishnan & Schlick, 2006), to explore PT from the 3′ oxygen of the primer to a non-bridging oxygen P_γ, through two intervening water molecules (following a 'Grotthuss-like' mechanism). It was further argued by the authors that this might even be the preferred pathway for other polymerases (Radhakrishnan & Schlick, 2006;

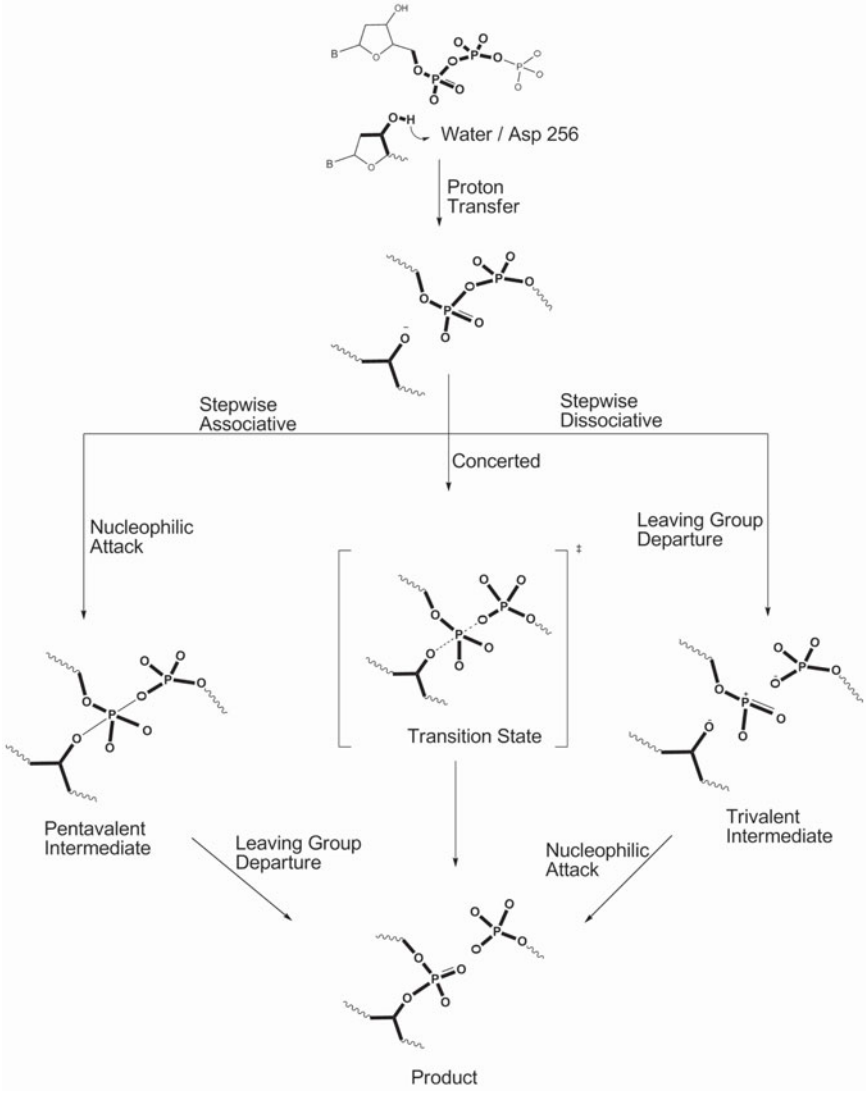

Fig. 40. A schematic description of the chemical steps of the nucleotide incorporation reaction catalyzed by Pol β. The initial deprotonation of the 3′-hydroxyl group of the primer is followed by nucleophilic attack on the P_α of the incoming dNTP. The formation of the O_3–P_α bond then results in the release of the pyrophosphate leaving group. As shown in this figure, after the initial PT step, the reaction may proceed *via* either a stepwise pathway involving either a trivalent (metaphosphate) or pentavalent (phosphorane) intermediate, or a concerted pathway which can in turn also be associative or dissociative in nature depending on the extent of bond formation to the incoming nucleophile and bond cleavage to the departing leaving group. However, it is worth noting that previous studies have suggested that the hydrolysis of a phosphate diester is unlikely to proceed via a stepwise dissociative mechanism. (Kamerlin *et al.* 2008b).

Wang & Schlick, 2008). Unfortunately, despite the elegance of these studies (e.g. Radhakrishnan & Schlick, 2006; Wang & Schlick, 2008), the corresponding conclusions are problematic. Even though there will be a pool of water molecules in the vicinity of the enzyme active site, the actual PT pathway depends on the electrostatic environment and the pK_a values of the donor and acceptor (which can also be a water molecule). This point has been argued repeatedly, and shown

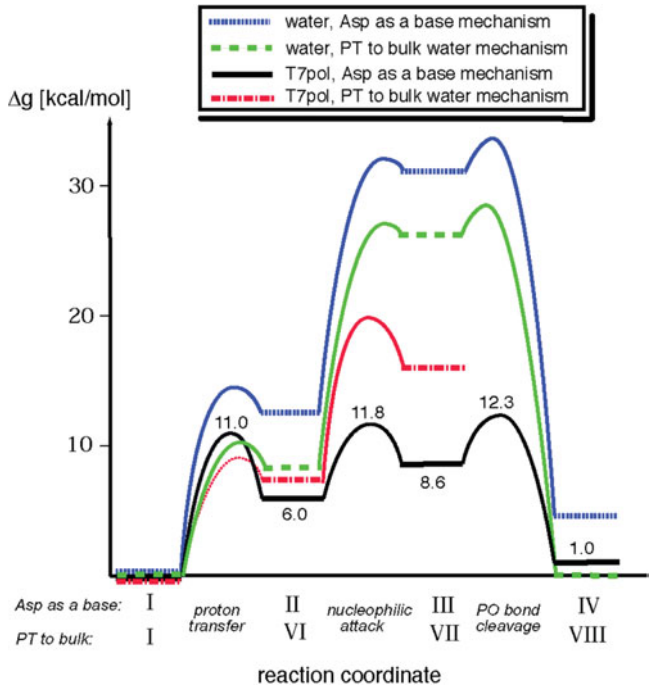

Fig. 41. A comparison of different mechanistic possibilities for the nucleotide transfer catalyzed by T7 polymerase. Shown here are the energetics of the 'asp-as-base' mechanism as well as a mechanism in which nucleophile activation occurs by PT from the 3′-OH group of the primer to bulk water. For a discussion of the mechanistic details, see the main text and (Florián *et al.* 2003a). This figure was originally presented in (Florián *et al.* 2003a).

several times in the literature (Braun-Sand *et al.* 2004, 2008). Hence, PT to a water molecule is possible *only* if its pK_a and the surrounding electrostatic environment favor this transfer. Unfortunately, as mentioned above, the afore-mentioned studies lack proper calibration and perform problematic energy minimization studies. Ideally, one needs to critically assess *all* the possible pathways before adopting one as the correct pathway. Here, it is also important to realize that the reported attempt (Wang & Schlick, 2008) to explore this path probably obtained a high barrier because it involved a large distance between the donor and acceptor, which is most probably entirely due to the use of energy minimization without proper relaxation of the protein (see the related prechemistry discussion below). In fact, energy minimization studies of PT have been show by us to be problematic (see e.g. Braun-Sand *et al.* 2004, 2008), even in cases with relatively careful treatments (e.g. Bondar *et al.* 2004a, 2004b). Finally, as discussed in Section 3.5 energy minimization searches tend to miss a direct PT from the attacking nucleophile to the phosphate oxygen. Thus careful studies of the type described in Figs. 19 and 20 will be needed here.

Overall, we have demonstrated that electrostatics, rather than the orientational effects, would drive the PT in biological systems, and that proper sampling is essential to obtain the relevant activation free energies (note that others are also coming to similar conclusions, see e.g. Riccardi *et al.* 2008). Similarly, we also have shown that when the distance requirements are not extreme (namely when the energy of bringing the donor and acceptor to a close enough distance is not very high), the PT process is unlikely to occur *via* several water molecules.

We also note that a related proposal of a concerted PT between several acidic groups (Radhakrishnan & Schlick, 2006) is problematic, as any proposal of PT between acidic groups is, at best, as reliable (or less so) than the corresponding accuracy with which it is possible to calculate the pK_a of the bridging groups. Correctly obtaining pK_a values by QM/MM calculations has only been accomplished by a few groups (e.g. Riccardi *et al.* 2005) and, in particular, in our recent study (Kamerlin *et al.* 2009b). Finally, it is crucial to realize that any proposal that suggests PT through several water molecules must be evaluated by also considering the corresponding path without any bridging water molecule(s), which is not usually the case.

Now one of the *main* requirements for elucidating the catalytic mechanism has been the ability to obtain free-energy surfaces that are not only sufficiently reliable to capture the correct reaction path, but also to be able to reproduce the proper free-energy surfaces for any possible 'pre-chemistry' events that occur prior to the phosphoryl transfer itself. This is an important issue, since, with the exception of (Xiang *et al.* 2008), previous fidelity studies (e.g. Alberts *et al.* 2007; Bojin & Schlick, 2007; Lin *et al.* 2006, 2008; Wang & Schlick, 2008) have used a relatively rigid protein model, coupled with energy minimization. Thus, it is not clear that such studies have been capable of capturing the relaxation free energy associated with the response of the protein to the motion from the incorrect (W) to the correct (R) base-paired system.

As stated above, the true TS for this reaction probably involves a concerted mechanism, however, in general, the catalytic effect of the protein can also be assessed by considering the stepwise mechanism. To establish this point, we have calculated the free-energy surface for the concerted nucleophilic attack of the 3′ oxyanion of the deoxyribose, on the P_α of the incoming nucleotide, and it was found that whether the nucleotidyl transfer reaction is concerted or stepwise has very little impact on our conclusions about fidelity and prechemistry, as long as both systems are modeled with the same EVB approach (see also Prasad & Warshel, 2011).

As an aside, it is important to address here previous QM/MM studies (e.g. Lin *et al.* 2006, 2008) which obtained a concerted reaction pathway in Pol β. This calculated path is probably reasonable. However, as a general note for workers in the field, we would like to point out that the surfaces being presented in these works are energy minimized potential surfaces rather than *free-energy* surfaces (which clearly leads to problems in determining the barrier to the pre-chemistry), and, as discussed in Klähn *et al.* 2006), there are other potential problems in using energy minimization approaches, specifically, the fact that this 'adiabatic mapping' approach is far too computationally expensive to at present be able to perform the extensive sampling required to obtain convergent results (rather than being a problem with *ab initio* QM calculations in and of themselves). However, our main concern with these studies (which is clearly not limited to Lin *et al.* 2006, 2008) is the fact that it the results have not been validated by reproducing the observed barriers for the solution reaction, and we find the tendency to assume that just performing *ab initio* QM/MM calculations in proteins should give the exact barrier quite risky.

4.3.3 Exploring the fidelity landscape

Some insight into effect that is relevant to the fidelity issue has been provided by a study (Yang *et al.* 2004) which used computer simulations to model the effect of mutations on the action of Pol β, and examined the corresponding electrostatic effects by a qualitative analysis. Additionally, theoretical studies of related problems have been recently reported (Feig *et al.* 2001; Oostenbrink & van Gunsteren, 2005; Orozco *et al.* 2003; Seibert *et al.* 2003). Molecular dynamics

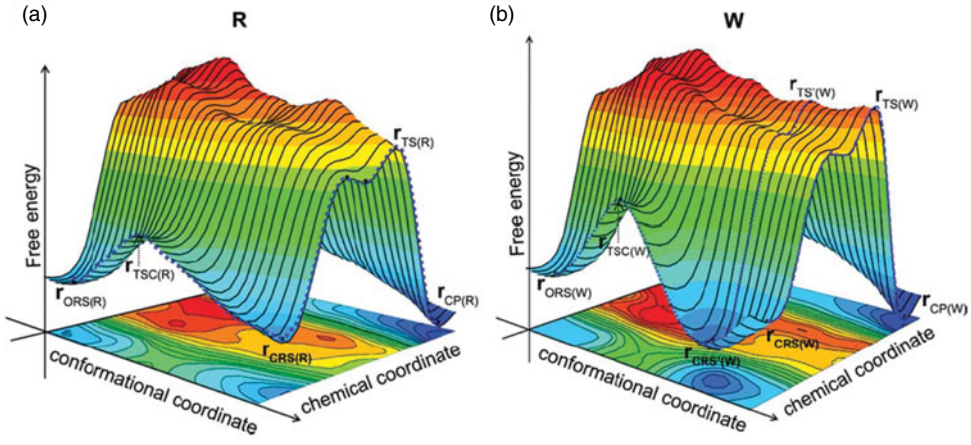

Fig. 42. Schematic free-energy surfaces, illustrating potential coupling between the conformational and chemical coordinates, for the insertion of (*a*) correct and (*b*) incorrect dNTP by a DNA polymerase. $\mathbf{r}_O$, and $\mathbf{r}_C$, designate the open and closed configurations respectively, and TSC designates the TS associated with the conformational motion. Finally, RS, TS, and PS designate the reactant (ES), transition and product states, respectively, and (R) and (W) designate correct and incorrect nucleotide insertion. This figure was originally presented in (Prasad & Warshel, 2011).

simulations have helped to delineate a possible sequence of conformational changes after dNTP binding in the catalytic cycle of Pol β (Arora *et al.* 2005; Radhakrishnan & Schlick, 2004, 2005). Simulation studies have also enabled us to test various catalytic proposals for nucleotide transfer reactions (Alberts *et al.* 2007; Florián *et al.* 2003a). Finally, an attempt to progress in a more quantitative direction in the comparison of calculated and observed mutational effects has been reported (Xiang *et al.* 2006). This study also introduced a useful approach for constructing interaction matrices and using them to probe the transfer of information between the binding and catalytic sites.

In order to move to a more direct analysis of the fidelity issue, we start with the landscape presented in Fig. 42. The early conceptual description of this figure has emerged from our previous studies of DNA polymerases (Florián *et al.* 2005; Xiang *et al.* 2008), in which we concluded (based on calculations of the barrier for the rate-determining chemical step) that the path along the chemical coordinate passes at a somewhat different value of the 'orthogonal' coordinate corresponding to the protein structural changes (which is typically considered to correspond to the motion between the open and closed forms). We would like to emphasize that this view is fundamentally different from related assumptions (e.g. Radhakrishnan & Schlick, 2004) that the barriers for the conformational changes between the open and closed forms contribute to the fidelity or to the reaction rate. The barriers for the motion from the open to the closed form in the reactant state of the R base pair (from $r_{ORS(R)}$ to $r_{CRS(R)}$ in Fig. 42) are unlikely to influence fidelity, or, for that matter, the reaction rate for both W and R nucleotide insertions, as long as they are lower than the chemical barrier (Kamerlin & Warshel, 2010). Following from this, the interesting possibility that the prechemistry barriers are different for W and R presented in (Radhakrishnan & Schlick, 2004) is not likely to influence fidelity, as long as those barriers are lower than the chemical barrier. In other words, what we are focusing on here is the path from the bound incoming dNTP to the product (i.e. $r_{CRS(R)}$ to $r_{CP(R)}$ and $r_{CRS'(W)}$ to $r_{CP(W)}$), and on its dependence on the protein conformation, assuming (based on experimental findings), that the

chemical step is rate determining. The relevant precise free-energy surface for this can be found in (Prasad & Warshel, 2011).

4.3.4 Examining the proposals of major prechemistry barriers

Recent advances in structural studies have led to several interesting theoretical works (e.g. Lin *et al.* 2006, 2008; Rucker *et al.* 2009). Some of these (Alberts *et al.* 2007; Lin *et al.* 2006, 2008; Radhakrishnan & Schlick, 2006; Radhakrishnan *et al.* 2006) involved an idea that has been a subject of some discussion of the barrier to potential 'prechemistry' events. For example, a QM/MM study on the barrier to prechemistry (before the nucleotide transfer reaction) in human DNA pol β with a G:A mismatch in the active site (Lin *et al.* 2008) was estimated to be about 14 kcal mol^{-1}. Based on this, the authors suggested that the free energy required for the formation of the prechemistry state is the major contributing factor to the decrease in the rate of incorrect nucleotide incorporation compared with correct nucleotide insertion, and therefore to the fidelity enhancement. Also, in a more recent QM/MM study on DNA polymerase IV (Dpo4) (Wang & Schlick, 2008), it was implied that the prechemical reorganization of the catalytic site (which brings the enzyme from its position in the X-ray structure to an active conformation, prior to the chemical step) would cost approximately 4·0–9·0 kcal mol^{-1} (although this was not based on actual barrier calculations). However, while elegant, both these studies employed a treatment that essentially amounts to energy minimization without proper relaxation and sampling, the problems with which have been discussed elsewhere (e.g. Kamerlin *et al.* 2009b; Klähn *et al.* 2006), and also, unlike e.g. (Prasad & Warshel, 2011; Xiang *et al.* 2008), this study was not properly validated by taking into account the corresponding reaction in water. In contrast, by calculating the PMF for this process using proper relaxation and extensive sampling (Prasad & Warshel, 2011), we obtained a much smaller barrier of 3·58 kcal mol^{-1} to the prechemistry step. More importantly, by using constraints with increasing force constraints from 0·5 to 3·0 kcal mol^{-1} $Å^{-2}$ on all residues within 10 Å of the reactive region (i.e., simulating the effect of non relaxed environment), we were also able to reproduce a trend of a correspondingly increasing barrier to this step. This highlighted the problems with not allowing the protein to relax properly, and demonstrating that the high prechemistry barriers obtained by (Lin *et al.* 2006; Wang & Schlick, 2008) are most likely overestimates due to the problems mentioned above.

Other workers (e.g. Joyce & Benkovic, 2004; Radhakrishnan *et al.* 2006) have invoked a somewhat related idea, suggesting that intermediates along the path toward the TS of the incorporation reaction can serves as kinetic 'checkpoints'. It was basically argued that although the presence of such checkpoints will not change the overall fidelity of the reaction, they *would* define the pathway by which that fidelity would be realized. In the case of the polymerase reaction, the intermediate steps (or checkpoints) would allow for the rejection of inappropriately paired dNTPs before the polymerase attempts phosphodiester bond formation of such substrates (see Joyce & Benkovic, 2004). Unfortunately, these arguments are not based on clear molecular concepts, and lack support from consistent simulations. In fact, even the common suggestion that the fidelity is due to an induced fit effect is not based on clear structure-energy analysis (see Xiang *et al.* 2008). However, if one does want to perform such an analysis, one has to first take into account the actual rates of the conformational transition *versus* the nucleotide insertion step (which, stopped flow fluorescence analysis has suggested is approximately 5 times slower than the conformational transition Bakhtina *et al.* 2005). If the activation barrier for the

conformational transition is smaller than that for the chemical step, then it is unlikely to help control the fidelity, and, in fact, a detailed study using our renormalization approach (Prasad & Warshel, 2011) demonstrated there is no effect of the motion along the conformational coordinate on the overall rate constant. Here, it is once again important to emphasize, however, that this refers *specifically* to a situation where the chemical step is the rate-limiting step (which is the case with Pol β (Ahn *et al.* 1998; Lin *et al.* 2008) and *obviously* not to a hypothetical scenario where the conformational transition is rate limiting.

Finally, to gain further insight into the catalytic landscape of the R and W systems (Prasad & Warshel, 2011), generated 2D surfaces for the free energy in the space defined by the chemical coordinate (which describes the nucleotide transfer reaction), and the coordinate describing the movement of the template 3′ primer. The calculated activation barriers (see Fig. 9 in Prasad & Warshel, 2011) clearly indicate that the attack of the $(O3')^-$ of the template 3′ primer is more favorable from r(ES(R)), where the primer (nucleophile) strongly interacts with the catalytic Mg^{2+} ion. Furthermore, the chemical barrier from the r(ES(R)) structure is lower for R than for W, reflecting the fact that the R system has better preorganization at r(ES(R)).

4.3.5 Exploring the molecular basis for the observed fidelity

In order to elucidate the origin of the calculated (and observed) fidelity, we evaluated the binding free energy of the incoming nucleotide at r(ES(R)) and r(ES(W)) in both the R and W systems. The corresponding binding energies are obtained by following the same strategy used in our recent study, where we reproduced the observed change in TS binding free energy in different mutants and upon changing the base (Rucker *et al.* 2009). In this approach, we evaluated the binding energy of the chemical (triphosphate) and the base sites separately, by running independent Protein Dipole/Langevin Dipole simulations (PDLD/S) using the LRA version (Lee *et al.* 1993; Sham *et al.* 2000), and using dielectric constants of 40 and 2 for the chemical (triphosphate) and base sites, respectively (the justification for this approach is presented in Rucker *et al.* 2009). The corresponding calculated binding free energies reproduced the balance between the binding of the base and the chemical sites, which is the essence of the allosteric effect that controls the fidelity of DNA polymerases. We demonstrated that the fidelity of DNA polymerases reflects allosteric competition between the binding of the base of the incoming dNTP to the base-binding site, and the binding of the TS to the catalytic site. In the case of incorrect dNTP insertion, poor preorganization in the base-binding site and poor binding of the base lead to structural rearrangements that are propagated to the catalytic site and destroy its preorganization (Xiang *et al.* 2006). This allosteric control of fidelity is described in Fig. 43, and this compensating allosteric effect reflects a rather simple interplay of free-energy contributions, and should not be described by such terms as 'population shifts', since the population shift is simply a result of the change in free energy associated with being in different configurations, and *not* a basis for any well-defined physical effect.

4.3.6 Exploring the idea of coupled motions and kinetic checkpoints

Recent ideas that the fidelity of DNA polymerases may involve 'gate keeping' dynamical effects (Radhakrishnan *et al.* 2006) (and related earlier ideas, Patel *et al.* 1991) are of interest, and should be examined in a well-defined way. Time-resolved fluorescence (Kim *et al.* 2003) and NMR (Bose-Basu *et al.* 2004; Kirby *et al.* 2005) studies have demonstrated that the motions (which have

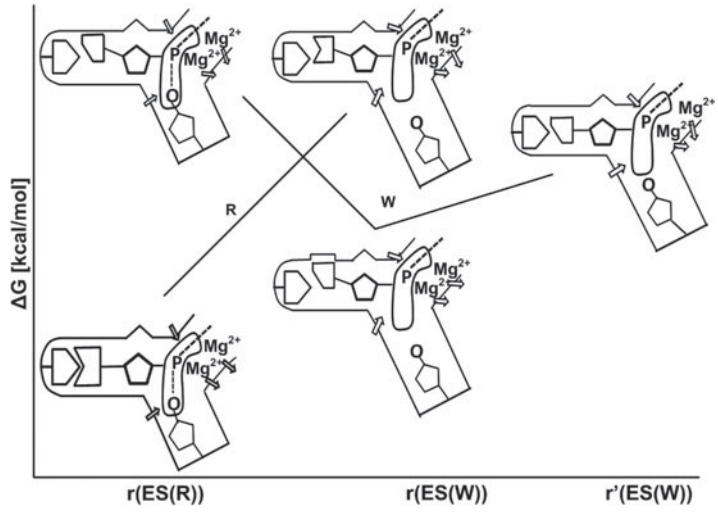

Fig. 43. Illustrating the origin of the allosteric effect. This figure uses the calculations from Fig. 10(C) of (Prasad & Warshel, 2011) and illustrates the relevant changes in the R and W systems at different protein configurations. Specifically, shown here are the protein rearrangement upon binding of R and W at r(ES(R)) and r(ES(W)). In the R case, the protein provides optimal sites for both the chemical and base-pairing parts of the process. On the other hand, in the W case, the protein has to relax in the base-pairing site in order to provide good binding, and this relaxation destroys the preorganization in the chemical part. The diagrams that represent the different configurations illustrate this allosteric effect, with the arrows indicating the protein dipoles. In the R state, these dipoles are preorganized in an optimal way, and thus provide maximum TS stabilization. In contrast, in the W state, the dipoles are forced to assume less effective preorganization, and thus yield correspondingly less TS stabilization. This figure was originally presented in (Prasad & Warshel, 2011).

often been referred to as the 'dynamics') of α-helix N are reduced when Watson–Crick hydrogen bonding occurs (i.e. correct nucleotide binding), but *not* when this hydrogen bonding is absent. In contrast, protein side-chain motions are increased in the vicinity of the active site aspartates when the closed complex is formed (Bose-Basu *et al.* 2004; Kirby *et al.* 2005). However, these motional effects are not likely to represent true dynamical effects, but rather merely equilibrium thermal fluctuations that follow the Boltzmann populations in different landscapes. In this case, we simply have well-defined entropic effects. An analysis and interpretation of these types of dynamical effects is presented in (Olsson *et al.* 2006b; Warshel *et al.* 2006a), and we have not found convincing theoretical or experimental evidence that dynamical effects play a significant role in catalysis (Kamerlin & Warshel, 2010; Olsson *et al.* 2006b). At this stage, we would like to point out that we have already provided clear illustrations of the problems with the dynamical proposals elsewhere (Kamerlin & Warshel, 2010; Pisliakov *et al.* 2009), as well as the fact that none of the arguments about dynamical effect have ever included a fully consistent study of the time dependence of such effects (Kamerlin & Warshel, 2010). In the present review, our aim is not to repeat our previous examination of the dynamical proposal but rather to relate it to the fidelity problem, an issue which has been discussed further in our recent work on this topic (Prasad & Warshel, 2011).

We have recently explored the coupling between the conformational and chemical steps, by constructing a free-energy landscape in the space defined by the conformational coordinate corresponding to the open→closed transition, and the chemical coordinate

(Prasad & Warshel, 2011). Here, we mainly want to emphasize that we can capture the main features of the catalytic landscape, and that as long as the conformational barrier is lower than the chemical barrier (as is the case in Pol β Ahn *et al.* 1998; Lin *et al.* 2008), we will not find any effect from the motion on the conformational coordinate on the overall rate constant (see below). This point can be further proven by running LD on this surface without any constraints, but the general concept has already been clearly established in (Pisliakov *et al.* 2009).

At this stage, it is possible to return to the discussion of the idea that dynamical effects help control the fidelity. Firstly, as pointed out above, we cannot obtain any significant dynamical coupling between the conformational motions (i.e. the transition from the open to closed conformation), and the chemical step. Furthermore, even if there *were* some form of dynamical coupling, it would be difficult to see how it would be advantageous to use this as a mechanism to control the replication fidelity. The fidelity is determined by the ratio between $k_{\mathrm{pol}}/K_{\mathrm{d}}$ of R and W. If the rate-limiting step is the conformational transition between the open and closed conformations, and the barrier for these states is much larger for W than for R, then there could, in principal, be conformational control of fidelity. However, this mechanism is in conflict with the chemical step being rate limiting (see Bakhtina *et al.* 2005; Florián *et al.* 2003b, 2005). If the chemical step is rate limiting, then it seems that the only way for conformational changes to control fidelity is that the TS for both W and R base pairing must occur in a different conformation than the RS, the barrier along the conformational coordinate must be higher than the chemical barrier, and this barrier must be higher for W than for R pairing.

It may be useful to expand on the discussion above and to comment on the implications of (Radhakrishnan & Schlick, 2004) that the prechemistry steps can have a major effect on the fidelity. Apparently, as long as the barriers for the prechemistry steps (e.g. the open to closed conformational transition) are significantly lower than the chemical barrier, they cannot change the kinetics and the corresponding fidelity (except in some particular saturation conditions). The reaction rate is determined by the difference between the energies of the TS and the E + S state for the chemical step. Having multiple barriers between the open and closed configurations is not going to change the kinetics as long as these barriers are smaller than the TS barrier. An additional insight is obtained for example from inspection of Fig. 1 of (Radhakrishnan & Schlick, 2005). This figure describes the calculated barriers for the prechemistry steps. Since the barriers for W are smaller than that for R pairing, it is very hard to see how these barriers could account for the observed fidelity. The obvious answer is that the real difference is in the chemical step.

It might also be useful here to mention a very recent work (Kirmizialtin *et al.* 2012) that performs elegant Milestone simulations of the conformational change in HIV reverse transcriptase, implying that DNA polymerases select the correct substrate by means of conformational dynamics. Now despite the technical accomplishments involved in the Miletone approach (noting, however, that our renormalization approach has also provided a very good estimate of the conformational barrier, Prasad & Warshel, 2011), we find the dynamical and conformational view to be problematic. More specifically, (Kirmizialtin *et al.* 2012) examined a case that might involve a rate-determining conformational barrier for the formation of the incorrect base pair (see the discussion of an identical case in T7 polymerase in (Prasad & Warshel, 2011)), where the relevant situation has been depicted here in Fig. 44. Now, the dynamical implications suffer from several problems. First (as was clarified in Prasad & Warshel, 2011), there is nothing (apart from evolutionary constraints) forbidding a protein from having a rate-determining conformational barrier, but, as we will explain below, this has very little to do with dynamics, and of course cannot be used as (even implicit) support of the checkpoint and prechemistry proposals, which

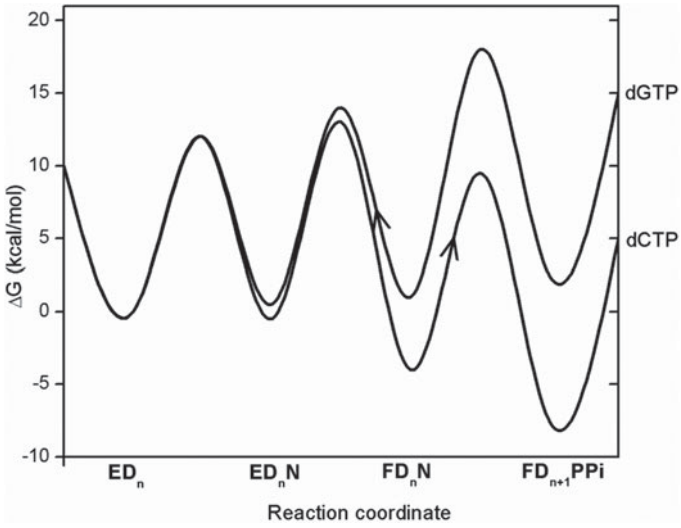

Fig. 44. A schematic illustration of free energies for correct (dCTP) and incorrect (dGTP) nucleotide incorporation reactions for the T7 DNA polymerase, Here, ED_nN represents the enzyme in complex with a DNA primer strand of n residues in length, N represents the incoming dNTP, and FD_nN represents the closed state of the enzyme with the DNA and nucleotide bound to it. This figure was originally presented in (Prasad & Warshel, 2011), and is adapted from (Johnson, 2008).

have been clearly stated to be related to the specific case where the chemical barriers are rate-limiting. Second, although the simulations reproduce the experimental conformational barriers, they have not considered the chemical barrier, and thus have not established the origin of the fidelity. We would also like to clarify (as was done in Prasad & Warshel, 2011) that even if the rate-determining barrier for R is the conformational step, this has nothing to do with the dynamical view that '*rapid motions in the specificity domain allow the enzyme to explore the substrate at the active site, probing for key interactions indicating a correctly bound substrate*', presented in (Kirmizialtin *et al.* 2012). Rather, all we have here is being entirely determined by the (conformation and chemical) activation barriers, and the depth of the binding free energy, which fully determines the kinetics of the system. The 'speed' of the motion is neither remembered, nor used to climb the chemical barrier (i.e. the system has enough time for fully stochastic motions at each of the minima; see also Pisliakov *et al.* 2009). The statement that '*only the correct substrate induces a rapid and energetically favorable alignment of catalytic residues to bring reactants together at the active site*' (Kirmizialtin *et al.* 2012) is as problematic as the proposals that the conformational motion is bringing the system to its preorganized structure (Benkovic & Hammes-Schiffer, 2003) (see discussion in Fig. 7 of Kamerlin & Warshel, 2010). This fact has been established in works that *actually* explore the dynamical proposal (e.g. Pisliakov *et al.* 2009), and has never been explored by other simulation studies. The correct physical picture is, in fact, very simple. In W, the conformational barrier is lower than the chemical barrier, and therefore the rate is fully determined by the height of the chemical TS (see Fig. 44) and the system will move over the chemical barrier according to the Boltzmann probability of this barrier. Furthermore, the system will also move *backwards* to the reactant state (ED_nN), over the conformational barrier, according to the Boltzmann probability of moving from ES to ED_nN. The chance of going backwards across the barrier is a trivial kinetic effect of having a lower barrier than the chemical step, and has nothing to do with dynamics. In the R case (if it is true that the conformational barrier is rate-limiting), the rate is

determined by the Boltzmann probability of crossing the conformational barrier once the system arrives at the ES state (FD_nN).

At this point, we note that the implication that the conformational barrier controls the fidelity is very problematic, since it is by far simpler to use the allosteric features depicted in Fig. 43 than to somehow have different conformational barriers for R and W. After all, evolution has had to work much harder to reduce the chemical barrier (from the very high barrier in the uncatalyzed reaction), than to change the almost trivial conformational barrier. In fact, the entire situation depicted in Fig. 44 reflects the following physics: in the R case, the system tries to get the fastest reaction, by pushing the top of the chemical TS (relative to the unbound state) up to the level of the conformational barrier (note that the reduction of the TS(R) barrier depicted in Fig. 44 is most likely an exaggeration). Now in the case of W, the fidelity is accomplished by pushing up the TS barrier, which can be done either by just destabilizing the TS(W), or by destabilizing both the TS(W) and RS(W), as is the situation illustrated in the figure. In fact, the conformational barriers are very similar for W and R, and the small increase in the conformational barrier is just a result of the destabilization of the ground state FD_nN, rather than the reason for any observed fidelity.

Finally, it is worth commenting on a new work (Hammes *et al.* 2011) that seems to present flexibility and cooperativity as key catalytic factors, despite the fact that modeling (or even just considering) such effects cannot account for the tremendous catalytic power of enzymes (see e.g. Adamczyk *et al.* 2011). This work could be taken as support of the checkpoint idea, as it states that '*the presence of multiple intermediates in enzyme mechanisms suggests that the enzyme functions by creating multiple steps with low standard free energies of activation, as contrasted to one-step, or few-step, reactions with relatively high standard free energies of activation. An enzyme is able to do this because of its structural flexibility or conformational adaptability. The enzyme can optimize its structure for each step in the reaction sequence.*' (Hammes *et al.* 2011) However, while it is true that the conformational motions may have several barriers, they cannot determine the overall rate, or be useful toward optimizing catalysis, *unless* they become as high as the rate-determining barrier. This issue should not be presented as a controversial issue, unless its proponents can produce a consistent model where this effect is actually reproduced, and can thus contradict our careful studies (e.g. Pisliakov *et al.* 2009).

4.3.7 What is the relationship between induced fit and fidelity?

In exploring the role of conformational changes in Pol β, it is important to discuss the idea that the fidelity of DNA polymerase is related to the induced-fit proposal (Arora & Schlick, 2004; Benitez *et al.* 2006; Koshland, 1995; Post & Ray, 1995; Sawaya *et al.* 1997). The fact is that substrate binding induces conformational changes that bring Pol β to its catalytic configuration can be described as an induced fit effect. However, this does not provide any information about the nature of the catalytic effect in the preorganized active site, nor does it tell us how the structure of the ES (or in our case the E-DNA–dNTP) complex determines the activation barrier of the chemical step. In other words, chemical catalysis does *not* depend on the structure of the enzyme before it binds the substrate (it is depends on the barrier going from the RS to the TS in the ES when the substrate is already bound), and, therefore, the induced fit concept does *not* describe the origin of the effect of the ES structure. In fact, the catalytic power of enzymes is largely due to the preorganized electrostatic environment of their active sites at the ES complex, and the enzymes inherent folding energy is invested in reorganizing its active site (Warshel *et al.* 2006b).

The use of the term 'induced fit' to describe the origin of the polymerase fidelity is also quite problematic. In fact, the use of kinetic diagrams of the type considered by (Post & Ray, 1995) does *not* provide a unique thermodynamic analysis, because they are not based on well-defined free-energy surfaces (e.g. surfaces of the type presented in Fig. 42). Armed with actual surfaces, we can define the relationship between structural changes upon binding, as well as the corresponding energetics, including the energy change of the enzyme upon moving from r(E) to r(ES) (this energy is defined as the reorganization energy upon binding, Muegge *et al.* 1997), the change in the enzyme substrate interaction upon moving from r(E) to r(ES) and the change in the activation barrier for the chemical step as the result of the binding induced structural change (this is defined in terms of the change in the reorganization energy of the chemical step and the work term, Liu & Warshel, 2007). Without such estimates, it is likely that the discussion will become circular.

As discussed and established by the logical analysis of (Xiang *et al.* 2008), it is almost impossible to conduct a logical analysis of the induced fit effect without describing the relevant free-energy surface. Apparently, even with a well-defined surface, we must define what the reference state for the induced fit effect is. When this is done (Xiang *et al.* 2008; Prasad and Warshel 2011), we get different results for different definitions. However, none of these results tell us how the induced fit determined the fidelity.

Although our discussion of the induced fit effect seems complex, we view it as an important step in illustrating the fact that there is no way to predict the trend in fidelity by using the typical induced fit concept, since it is ambiguous whether or not the fidelity might decrease or increase. In contrast, the calculations described in (Xiang *et al.* 2008) probe the actual effect of the structural change, and predict the consequence of the induced fit. Thus, accepting that the induced-fit concept does not explain fidelity, we can still investigate the consequence of this effect on fidelity in specific cases using computer simulation or experimental approaches.

4.3.8 Some points on the interpretation of LFER in polymerases

In Sections 2 and 3, we introduced the concept of LFER, and their traditional role as a mechanistic marker in studies of phosphoryl transfer reactions. We also highlighted some of the potential pitfalls of this approach, and, particularly the challenges in obtaining unique mechanistic interpretations from LFER (see also discussion in Åqvist *et al.* 1999; Klähn *et al.* 2006; Rosta *et al.* 2008, among other works). However, the main focal points of this discussion were on two central problems (1) the fact that both associative and dissociative pathways can give rise to similar LFER and therefore the LFER cannot distinguish between the two mechanistic possibilities (see Åqvist *et al.* 1999), and also, (2) the fact that the mechanism may gradually change across an LFER (from gradual changes in the nature of the TS, Alkherraz *et al.* 2010; Rosta *et al.* 2008) to full switches from associative to dissociative pathways (Klähn *et al.* 2006) with the increasing acidity of the leaving group, but *without* any break in the LFER. Therefore, not only do associative and dissociative TSs give rise to identical LFER, but also, provided the changes in mechanism are sufficiently gradual, LFER can mask mechanistic variation across a series of homologous substituted compounds.

Here, we would like to illustrate an example of the opposite problem – when LFER would appear to be telling us more about changes in mechanism than they actually are. Specifically, recent experimental studies (Sucato *et al.* 2007, 2008) presented a new probe for leaving group effects during the nucleotide transfer reaction catalyzed by Pol β, in which the β,γ-bridging

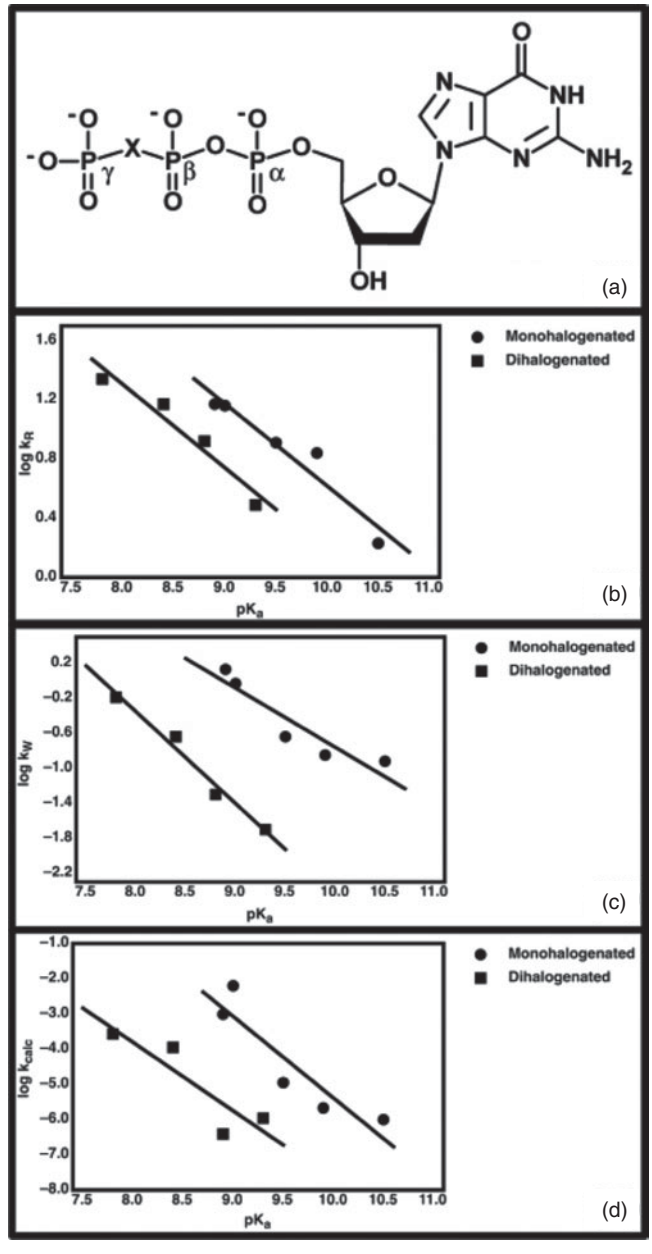

Fig. 45. An overview of (*a*) the dNTP analogs presented in (Kamerlin *et al.* 2009c; Sucato *et al.* 2007, 2008)], where X=O, CH_2, CHF, CHCl, CHBr ('monohalogenated' compounds, where for simplicity X=O and CH_2 are also under this label) or CF_2, CCl_2, CBr_2, CFCl ('dihalogenated' compounds), and the corresponding LFER for (*b*) correct, and (*c*) nucleotide insertion in DNA polymerase β, as well as (*d*) the corresponding reference reaction in aqueous solution.

oxygens of the dGTP were replaced by a series of substituted methylene groups (X=CYZ, where Y and Z=H, halogen or another substituent, Fig. 45). A LFER was then constructed by correlating the logarithm of the observed catalytic rate constant (k_{pol}) to the highest leaving group pK_a for each analog. Interestingly, however, while the data presented strong linear

correlations between the leaving group pK_a and k_{pol}, rather than corresponding to a single relationship, the data exhibited two independent correlations, one corresponding to the parent, methylated and monohalogenated compounds, and the other corresponding to the dihalogenated compounds. In addition to this, the difference between the two correlations appeared larger for incorrect (W) as opposed to correct (R) nucleotide insertion (Fig. 45). Several explanations were put forward to attempt to rationalize this effect, including the suggestion that leaving group elimination may potentially be hindered by intramolecular repulsion between the negative charge that develops on the α,β-oxygen and the Pro-S halogen (Sucato *et al.* 2008) or, alternatively, that the halogens are engaging in debilitating steric/electrostatic interactions with nearby amino acids in the enzyme active site (Sucato *et al.* 2008). An important issue however is whether the 'split' in the LFER reflects a trend that already exists in the background reaction in solution, and, in fact, a detailed theoretical study of the background reactivity suggested that this might indeed be the case (Kamerlin *et al.* 2009c). In the absence of experimental data on the corresponding background reaction, we evaluated complete free-energy surfaces for the hydrolysis of these compounds in aqueous solution, to obtain a theoretical prediction for the corresponding LFER. We verified our data using several alternative solvation models, and, in all cases, we found a split of the calculated LFER for the mono- and dihalogenated compounds into two near-parallel lines, in accordance with the experimental observation for the enzyme catalyzed reaction. Here, it is critical to point out that despite the split of the LFER into two separate lines, little qualitative change was observed in the overall free-energy surfaces (and therefore also the corresponding TSs). Rather, we believe that the split arises as a result of a generalized solvation phenomenon, which involves the overall electrostatic interactions between the substrates and their surroundings, and which would also be observed in polar solvents in the absence of the enzyme (assuming an identical mechanism in the absence of the enzyme). Additionally, it is worth emphasizing that in our study of the reference reaction, we capped the triphosphate with a methyl group and therefore the split is independent of the identity of the leaving group. However, the difference between the LFER lines for correct and incorrect nucleotide incorporation in the Pol β active site is still most likely due to differences in the electrostatic interaction between the TS charges in the polymerase active site. This study therefore demonstrates not only why it is extremely important to consider the uncatalyzed background reactivity as well as the enzymatic reaction, but also that there can exist situations where there are dramatic changes in the LFER despite negligible changes in mechanism, once again emphasizing the need for caution when attempting to make qualitative interpretations of LFER.

4.3.9 Concluding points

Enzymes are 'Nature's catalysts', accelerating the chemical reactions that control life from millions of years to the order of seconds (Warshel *et al.* 2006b; Wolfenden & Snider, 2001). There have been a number of recent high-profile papers that have claimed that Nature could have used arsenic as a substitute for biological phosphate (Wolfe-Simon *et al.* 2009), and, even more drastically, that there exist bacteria capable of replacing the phosphate in their genetic material with arsenic (Wolfe-Simon *et al.* 2010). There is, however, a major flaw with this argument. While there is not sufficient experimental data to be able to know the exact rate for the hydrolysis of arsenate esters, extrapolating from the existing data would suggest these compounds to be *extremely* unstable, with a half-life that could be as low as 0·06 s (Fekry *et al.* 2011). Even if this estimate were somewhat extreme, clearly, this would strongly suggest that it is impossible to

maintain the integrity of the genetic material using arsenic-oxygen linkages, as they would spontaneously hydrolyze. Phosphodiester linkages, in contrast, are among the most stable bonds known to man, with the half-life for the hydrolysis of the P–O bonds holding together DNA expected to be on the scale of 30 million years (Schroeder *et al.* 2006). Thus, it is no coincidence that the transformation used most commonly in biology is so extremely resistant to spontaneous hydrolysis. Without tight regulation, life as we know it would be impossible, and at the same time that enzymes are tremendous catalysts, they also ensure that transformations occur only when called for. Here, phosphate is a particularly desirable candidate for controlling the replication fidelity of the genetic material, as the high charge on the triphosphate of the nucleotides ensures that the system is inherently dependent on electrostatic stabilization during the actual nucleotide transfer step, and, following from this, extremely sensitive to subtle changes in charge. In this section, we have explored a number of different hypotheses that have been put forward to try to rationalize the fidelity of DNA replication, from rate-determining prechemistry steps to kinetic checkpoints and dynamical regulation. We demonstrate that, in all cases, while elegant, such rationales actually generate more problems than answers, and, in fact, despite the complexity of the replication process itself, all detailed theoretical studies and logical analyses point to the fact that the source of its regulation can be brought down to 'simple' electrostatic effects.

4.4 How do proteins communicate in signal transduction processes?

4.4.1 Introduction

The formation and cleavage of phosphate ester bonds plays a crucial role in biological signal transduction (Berg *et al.* 2010). Representative systems involved in signaling pathways include G-proteins, kinases, cyclic adenosine monophosphate (cAMP) and several phosphatases. Due to the breadth of the field, the present section will consider only a few typical examples, and discuss the factors that control the signal transfer.

4.4.2 Ras as an example of the role of phosphorylation in signaling

When the GTP substrate of Ras hydrolyzes to GDP, this results in a drop of the affinity of Ras toward its effector Raf, which, in turn, leads to a dissociation of the active growth complex, thereby terminating the cell growth signal (McCormick & Wittinghofer, 1996). The nature of this control is illustrated in Figs. 46 & 47.

Structural studies have provided crucial insight into the nature of these systems and the detailed protein/protein contacts that lead to signal transduction. For example, relevant developments in this area include the elucidation of the structure of RapA1-Raf (Nassar *et al.* 1995), showing a complex of Raf with Rap, which, for all practical considerations of the binding domain, is identical to Ras (see Fig. 48). Furthermore, the structure of the Ras-binding domain of RalGEF, another effector of Ras, was solved by NMR spectroscopy (Geyer *et al.* 1997), in addition to a variant of Ral-FEL called Rlf (Geyer *et al.* 1997). Recent years have seen a flurry of work in this area, so we would just like to mention a few more structural studies, including combined experimental/computational attempts to understand Ras-GDP interactions with mutants of Raf (Flichtinski *et al.* 2010), as well as structural studies of the Rap-RapGAP complex (Scrima *et al.* 2008), which is interesting due to its ability to perform GTP hydrolysis without the catalytic glutamine and arginine residues of RasGAP.

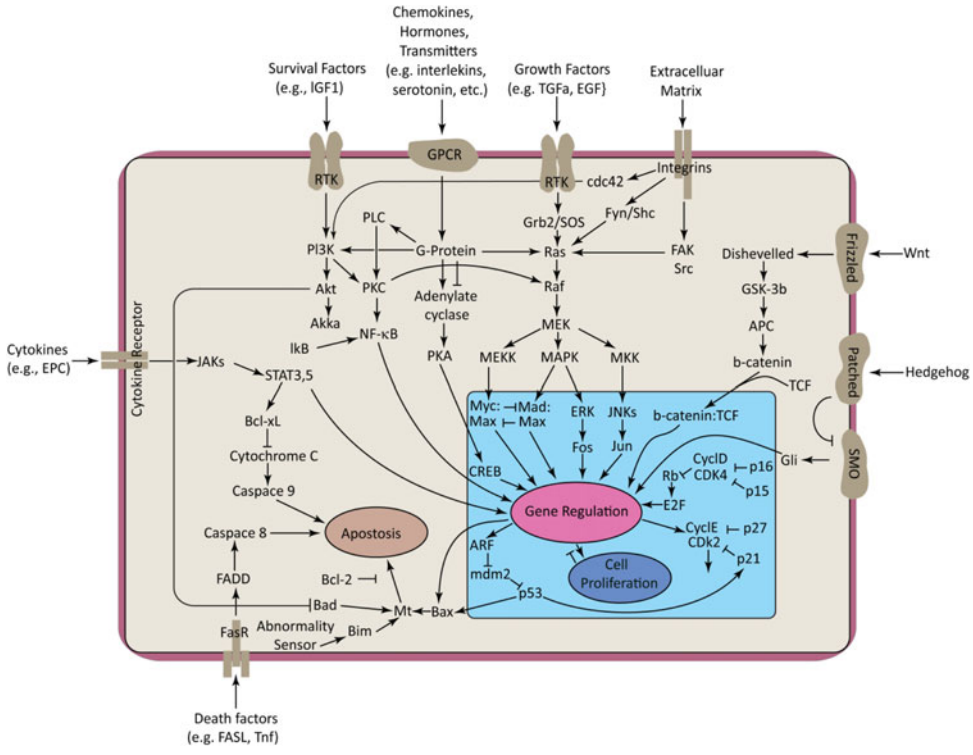

Fig. 46. An overview of cellular signal transduction pathways, highlighting the central role of Ras. Adapted from the Wikimedia Commons file 'Image: Signal_transduction_pathways.png' http://en.wikipedia.org/wiki/File:Signal_transduction_pathways.png

In any case, the availability of structural information for the RapA1/Raf complex allowed us to explore the key interactions between the two proteins (Muegge *et al.* 1998). Our careful study, which involved different levels of approximations, indicated that electrostatic interactions play a key role in protein/protein interactions. In particular, the calculations reproduced the observed effect of key mutations using a simple electrostatic model with a large effective dielectric constant (see Table 2).

Our ability to explore this system offers the option to develop drugs capable of mimicking one of the two partners, and thus breaking the signal transduction in monogenic Ras/RAF pairs. This idea is illustrated in Fig. 49.

Another interesting and relevant example is provided by the GEFs such as CDC25 (Broek *et al.* 1987) in yeast or SOS in mammalian cells (Bowtell *et al.* 1992; Wiesmuller & Wittinghofer, 1994). Exchange factors convert p21–GDP to p21–GTP by drastically reducing the affinity of both GDP and GTP in the ternary GEF–p21ras–GXP complex. Thus, a nucleotide free binary GEP–p21ras complex is formed that can bind GTP (the concentration of which in the cell is larger than that of GDP). This process 'activates' p21ras by generating the p21–GTP complex. Our study of the mode of GEF activation (Muegge *et al.* 1996) was performed before the availability of structural information on the complex (which became subsequently available, for review see Bos *et al.* 2007). In this work (Muegge *et al.* 1996), we evaluated the contribution of different residues of Ras to the binding of GTP and GDP (e.g. Fig. 50). We then established that a significant number of the residues with large electrostatic group contributions are those whose

Table 2. *Binding energies of Raf-RBD/Rap mutants**

	$\Delta\Delta G_{bind}$ (kcal/mol)						
	$\varepsilon_{in}=2$	$\varepsilon_{in}=3$	$\varepsilon_{in}=4$	$\varepsilon_{in}=5$	$\varepsilon_{in}=6$	$\varepsilon_{in}=25$	$\Delta\Delta G_{bind}^{obs}$† (kcal/mol)
Raf(R89L)	4·9	3·6	2·9	2·5	2·2	1·2	>5·5
Rap(D38A)	83·4	55·2	41·1	32·7	27·0	5·6	2·6
Raf(R59A)	58·5	39·3	29·7	24·0	20·1	5·6	2·0
Raf(K84A)	120·9	80·2	60·0	47·7	39·5	8·6	1·5
Raf(R67A)	−64·3	−41·6	−30·3	−23·5	−19·0	−1·8	1·3

*The calculations (Muegge *et al.* 1998) used a simplified effective dielectric treatment. The calculated binding free energies are given relative to the corresponding free energy of the wild type protein, evaluated using $\varepsilon_{in}=4$. This binding free energy is -17 kcal mol^{-1}.

†Taken from (Nassar *et al.* 1996) (Raf-mutants) and (Herrmann *et al.* 1995) (Ras mutant D38A).

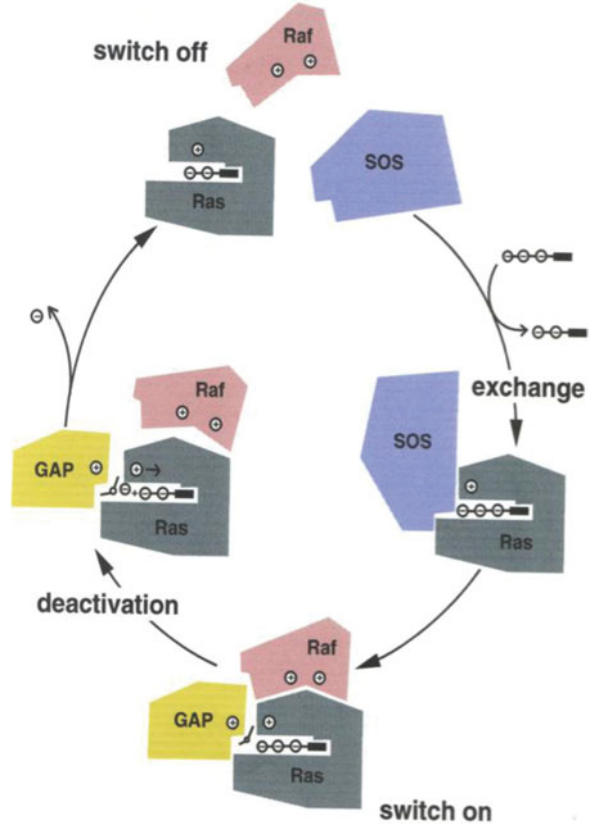

Fig. 47. A detailed description of the Ras activation and deactivation.

mutations with the action of GEFs. Interestingly, in considering the effect of GEF we had to account for the observation that the GEF reduces the binding affinity of GDP and GTP to p21ras by a similar amount. In doing so we concluded that the action of GEF might involve an interaction with Ras that pushes Gly^{60} toward Gly^{12} and that the resulting shift of Gly^{12} and its neighboring residues leads to a reduction in the interaction between the P-loop and the

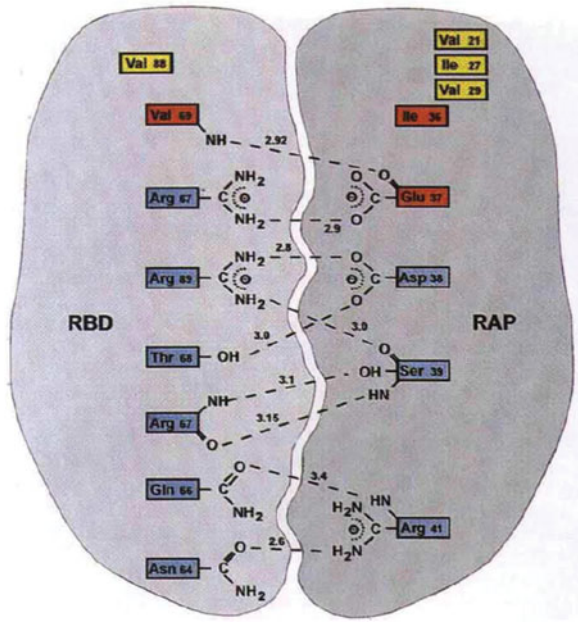

Fig. 48. Showing the ionized residues involved in the Rap/Raf interaction (Taken from (Nassar *et al.* 1995).

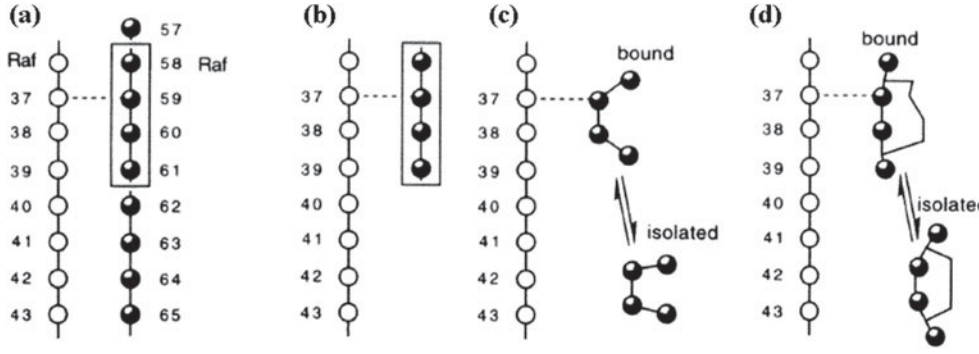

Fig. 49. Outlining a way for exploiting structural information from the Ras /Raf interface, for designing drugs that deactivate defected Ras. In this strategy we (a) identify the regions in Raf with strong contribution to binding, (b, c) examine if the corresponding short peptides also bind strongly, (d) and then examine if cyclic analogues of those peptides also retain strong binding. The resulting constructs would be good drug candidates.

nucleotide. More extensive studies with the new structural information would be very instructive and potentially useful for drug design (see e.g. Bos *et al.* 2007).

4.4.3 Other signal transducing systems

One of the most important prototypes for signal transduction is the action of kinases. The control of cellular signaling frequently involves posttranslational modification (PTM) of protein residues, such as phosphorylation (catalyzed by protein kinases), and dephosphorylation (catalyzed by protein phosphatases) (Hunter, 2000). Mutations of these enzymes have been implicated in many disease states, including cancer. As a result, there is strong interest in finding

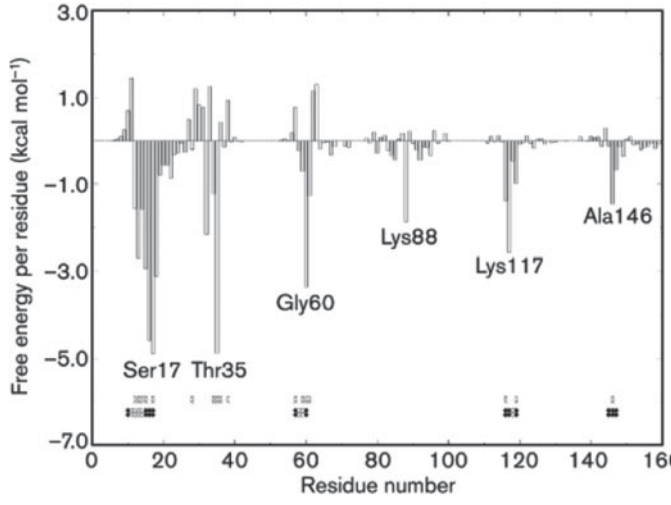

Fig. 50. Electrostatic contributions of individual residues to the binding of GTP to p^{21}ras. x indicates residues whose mutation leads to significant change in nucleotide-binding activity of p^{21}ras. Highly conserved residues are designated by filled diamonds, whereas those that belong to structural motifs involved in nucleotide binding, and are not that highly conserved, are designated by unfilled diamonds (below the x symbols) (Muegge *et al.* 1996).

effective inhibitors for these key enzymes (Blaskovich, 2009; Cohen, 2002). Significant effort has also been invested into structural and mechanistic studies (Adams, 2001; Taylor & Kornev, 2011; Zhang, 2003). There are more than 500 protein kinases encoded in the human genome, representing one of the largest protein families (Manning *et al.* 2002), and, despite sequence and structural diversity, many share nearly conserved catalytic domains (Taylor & Radzio-Andzelm, 1994). The chemical steps in some kinases are discussed in Section 4.5.3, and here, we would only like to comment that the details of the transfer of allosteric information from the effector binding site on one side of the membrane to the phosphorylation site is still a remarkable puzzle. We would like to point out, however, that in light of our experience with other systems, we are almost sure that this information transfer involves electrostatic control, which exploits both charge and changes in charge during reactions of the phosphate (see below).

Finally, due to space constraints, we would like to mention only a few other examples of the role of phosphorylation in signaling. The first of these is this is the epidermal growth factor receptor (EGFR), which is a tyrosine-specific protein kinase that catalyzes the phosphorylation of a number of endogenous membrane proteins (Cohen *et al.* 1980). EGFR is a cell-surface receptor, which normally exists in monomeric form. However, binding of EGF has been demonstrated to trigger dimerization, which causes the receptor to assume intracellular tyrosine kinase activity (Yarden & Schlessinger, 1987). This leads to the autophosphorylation of several tyrosine residues in the C-terminal domain of the receptor, which in turn causes several signal transduction cascades *via* the downstream activation of several other proteins that contain phosphotyrosine-binding domains (Downward *et al.* 1984), resulting in, among other things, DNA synthesis and cell proliferation (Oda *et al.* 2005). EFGR has also been identified as an oncogene, and mutations in this receptor have been shown to lead to a number of cancers (Drake *et al.* 2012; Larsen *et al.* 2011; Mimeault & Batra, 2011). In many cases, phosphorylation can also be linked to ubiquitination during PTM, in a form of 'crosstalk' (Hunter, 2007), which can be both positive (with one PTM serving as a signal for the addition or removal of a second

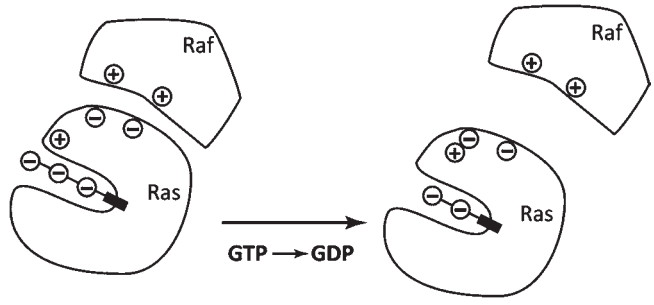

Fig. 51. Explaining the nature of the transfer of the signal from the transformation of GTP to GDP to the Ras/Raf interaction. The change from GTP to GDP involves a loss of a negative charge. This reduction of negative charge is transmitted to the Ras surface as a reduction of the Ras/Raf electrostatic interaction, thus leading to the dissociation of Raf.

PTM or for the binding of a protein that carries out the second modification) or negative (with direct competition for the modification of a single residue, or with one modification masking the recognition site for a second PTM). As outlined in detail in (Hunter, 2007), ubiquitination can be responsible for both the degradation and activation of protein kinases, as well as potentially for regulating protein phosphorylation, whereas phosphorylation can for instance regulate substrate localization and E3 ligase activity during ubiquitnation (see Fig. 1 of Hunter, 2007). Finally, caspases are a family of intracellular proteases that are responsible for regulating the process of programmed cell death (apoptosis) (Cardone *et al.* 1998). Figure 46 includes the caspase signaling pathway in detail, but we would like to conclude our discussion of other systems by mentioning that caspases can be directly regulated by protein phosphorylation, by blocking the protease activity (Cardone *et al.* 1998), and other kinases such as death-assisted protein kinase (DAPK) also play a direct role in apoptosis (Inbal *et al.* 2000).

4.4.4 Concluding remarks

Signal transduction plays a major role in life processes, and, in many cases, these processes are controlled by the formation and dissociation of phosphate ester bonds. Thus, trying to understand the general principle of signal transduction, we may ask why nature chose the phosphate ester bond for this purpose. In this respect, we believe that the effect of the γ-phosphate of GTP on the formation or dissociation of the RapA1-Raf complex in the Ras cycle depicted as a subset of Figs. 46 & 47 can be taken as a general example. Here, we find that even though the γ-phosphate is more than 12 Å away from the interface between the two proteins, this group is partly responsible for the increased affinity of the GTP-state of Ras relative to the corresponding GTP state. In other words, the move from three negatively-charged phosphate centers to two (upon the departure of the γ-phosphate) reduces the interaction of the overall Ras subunit with the positively charged Raf unit (see Fig. 51), providing a perfect example of the electrostatic control of signaling.

4.5 Other systems

As there exist so many systems that are controlled by the cleavage of phosphate ester bonds, space constraints prevent us from visiting all of them. However, below, we will touch on some of the most significant.

4.5.1 The electrostatic basis for catalytic promiscuity in the alkaline phosphatase (AP) superfamily

In recent years, it has became increasingly apparent that most (if not all) enzymes are capable of some degree of catalytic promiscuity (Khersonsky *et al.* 2006; Khersonsky & Tawfik, 2010; O'Brien & Herschlag, 1999) (where one enzyme is capable of catalyzing the reactions of more than one substrate). It has been suggested that such promiscuity is important for the evolution of function in enzyme superfamilies (Jensen, 1976; O'Brien & Herschlag, 1999), as well as potentially playing an important role in artificial enzyme design (Khersonsky & Tawfik, 2010; Toscano *et al.* 2007), guiding the insertion of novel functionality based on existing protein scaffolds. An excellent example of catalytic promiscuity can be seen in the AP superfamily, several members of which are capable of cross-catalyzing each other's native reactions (Jonas & Hollfelder, 2009; O'Brien & Herschlag, 2001). For instance, the name giving member of the family, AP, is a phosphomonoesterase that is also capable of catalyzing phosphodiester, phosphonate monoester and sulfate monoester hydrolysis (O'Brien & Herschlag, 1998; O'Brien & Herschlag, 2001; Stec *et al.* 2000b). In contrast, nucleotide pyrophosphatase/phosphodiesterase (NPP) is a phosphodiesterase that can also catalyze phosphate and sulfate monoester hydrolysis (Lassila & Herschlag, 2008; Zalatan *et al.* 2006), and, more recently, it has been demonstrated that an arylsulfatase from *Pseudomonas aeruginosa* as well as phosphonate monoester hydrolases from *Rhizobium leguminosarum* and *Burkholderia caryophilli* are capable of hydrolyzing a wide range of phosphate, sulfate, and phosphonate esters (Babtie *et al.* 2009; Jonas *et al.* 2008; Olguin *et al.* 2008; van Loo *et al.* 2010) (for detailed discussion and classification, see e.g. (Jonas & Hollfelder, 2009). The key issue, however, is that despite the fact that these enzymes all catalyze a similar set of reactions, their specificity patterns vary (at times quite dramatically), as do their metal requirements and active site architectures. Therefore, despite being part of the same superfamily, these enzymes are actually quite different, with different active site preorganization. The question then becomes one of understanding both the molecular basis for promiscuity in individual superfamily members, as well as the reason for cross-promiscuity between different members of the superfamily, and insight into this can be significant both for understanding the evolution of function in enzyme superfamilies, as well as for driving future enzyme design effort.

There has been extensive experimental and theoretical work invested into understanding the mechanisms of specificity and promiscuity in the AP superfamily (see e.g. Catrina *et al.* 2007; Galperin *et al.* 1998; Galperin & Jedrzejas, 2001; Hollfelder & Herschlag, 1995a; Hou & Cui, 2012; Jonas & Hollfelder, 2009; Jonas *et al.* 2008; Kim & Wyckoff, 1991; Lassila & Herschlag, 2008; Lopéz-Canut *et al.* 2011; Luo *et al.* 2012a; Luo *et al.* 2012b; O'Brien & Herschlag, 1998, 2001; Simopoulos & Jencks, 1994; Stec *et al.* 2000a; van Loo *et al.* 2010; Zalatan *et al.* 2006; Zalatan & Herschlag, 2006), to name just a few examples). One hypothesis that has been put forward to explain the promiscuity in AP and related members of this superfamily is the idea that, despite the presence of metal ions in the active site (which one would expect to facilitate a more compact TS, Humphry *et al.* 2004; Kamerlin & Wilkie, 2007; Klähn *et al.* 2006; Williams, 2004b), AP and NPP do not alter the nature of the TS relative to the corresponding reaction in aqueous solution, but rather, that they can identify and stabilize different TS for different reactions. This hypothesis has been put forward based on the interpretation of a variety of experimental techniques, including data from examining LFER, and from KIEs (see e.g. Catrina *et al.* 2007; Hollfelder & Herschlag, 1995a; Lassila *et al.* 2011; O'Brien & Herschlag, 1999;

Zalatan *et al.* 2006; Zalatan & Herschlag, 2006), and discussion in (Hou & Cui, 2012; Lassila *et al.* 2011). More specifically, the reasoning for the above hypothesis starts with the assertion that AP does not alter the nature of the enzymatic reaction compared with the reaction in solution. This is based on the argument that the gradients of LFER for a range of substrates catalyzed by AP would suggest that the TS is comparable with that in aqueous solution (Nikolic-Hughes *et al.* 2004; O'Brien & Herschlag, 1998, 2002) (an observation which was claimed to be supported by recent theoretical studies, Hou & Cui, 2012). Additionally, as it is (incorrectly) believed (see Klähn *et al.* 2006 and Section 3) that the size of the TS can be correlated to the charge distribution, it was further argued that the similar LFER for phosphorothioate and sulfate hydrolysis in the AP active site indicates 'negligible effects of active site interactions on charge distribution in the TS' (Nikolic-Hughes *et al.* 2004).

The challenges with qualitative interpretation of LFER have already been highlighted in great detail elsewhere in this review (and, in particular, in Sections 2 and 3), but we would nevertheless like to retiterate here a number of key problems with the assumption that similar LFER can be used to rationalize the promiscuity and play down the role of electrostatic interactions. The first of these is that, as highlighted in Section 3, and illustrated by e.g. (Åqvist *et al.* 1999; Florián *et al.* 1998; Kamerlin, 2011; Kamerlin *et al.* 2008a, 2009c; Klähn *et al.* 2006; Rosta *et al.* 2008), traditional mechanistic markers such as LFER, activation entropies and KIEs do not have unique mechanistic interpretations, with both different mechanisms giving rise to similar experimental observables, and also gradual changes in mechanism being masked by LFER. Therefore, in and of itself, none of the experimental evidence presented for the idea that the TS does not change upon moving from solution to the enzyme active site is conclusive without further support from careful computational studies that can reproduce all relevant available experimental observables. Additionally, just because the overall gradient of the LFER is very similar (Nikolic-Hughes *et al.* 2004), this does not automatically mean that the TSs are also similar, if all compounds on the LFER have been perturbed by a similar amount (maintaining the same overall trend in rate constants – as a reminder, the gradient of the LFER is measuring a *relative* and not an absolute effect). The second and much more problematic issue is the implicit idea that the catalysis is associated with the TS structure and that the role of electrostatics is minimal (Nikolic-Hughes *et al.* 2004), despite the fact that countless simulation studies have shown that electrostatic preorganization is the *most* important determinant of the catalytic power of enzymes (Kamerlin *et al.* 2010; Warshel, 1978; Warshel *et al.* 2006b), and also when rationalizing mutational effects (e.g. Adamczyk *et al.* 2011; Liu & Warshel, 2007) and is therefore the most likely explanation for the observed promiscuity (as was shown in the cases of human aldose reductase, Várnai & Warshel, 2000) and *P. aeruginosa* arylsulfatase, Luo *et al.* 2012b).

There have been a number of recent computational studies that have attempted to explore the issue of the comparative size of the TS for the solution and AP/NPP catalyzed hydrolysis of phosphate diesters (Hou & Cui, 2012; López-Canut *et al.* 2010, 2011), which have obtained sharply contrasting results. López-Canut and co-workers (López-Canut *et al.* 2010, 2011) have proposed that in both NPP and AP, phosphodiester hydrolysis proceeds through far more dissociative TS in the enzyme active site than in solution. However, these works employed quite a low level of theory, and, as was also highlighted by (Hou & Cui, 2012), the Zn^{2+}–Zn^{2+} distance in these simulations increased to an unphysical value of as high as 7·0 Å (Lopéz-Canut *et al.* 2011), in sharp contrast to the crystal structure and also to data from Extended X-ray Absorption Fine Structure (EXAFS) and X-ray crystallography, which suggests little change in the binuclear Zn^{2+} catalytic site during the course of the reactions catalyzed by AP and NPP

(Bobyr *et al.* 2012). In contrast to this, a more careful (but, as argued below, not yet fully reliable) study by (Hou & Cui, 2012) obtains a qualitatively similar phosphodiester TS in aqueous solution to that obtained by our earlier study (Rosta *et al.* 2008), coupled with the more reasonable observation of a significant tightening of the TS in the enzyme active site (see also Luo *et al.* 2012b).

Before moving the discussion on the above analysis, we may note that, although (Hou & Cui, 2012) recognizes the importance of validating their calculations by studies of the reference reaction in aqueous solution, the QM/PB calculation used is still not able to reproduce the experimental solution results. Furthermore, the correction by *ab initio* calculations both in the enzyme and solution are not determined at the *ab initio* least energy path. The fact that the SCC-DFTBPR QM/MM approach is not fully calibrated on consistent solution surface, as well as the fact that the solution study is not done by the same approach as the enzyme calculation may in part explain why (Hou & Cui, 2012) could not reproduce the correct trend in the observed mutational effects. Additionally, it should be pointed out that, despite the fact that (Hou & Cui, 2012) claims that their calculations suggest that AP and NPP do not alter the nature of the TS relative to aqueous solution, their own data (Fig. 3 and Table 4 of Hou & Cui, 2012) demonstrate a *significant* shift toward a more associative TS (as was also observed in the related case of PAS, Luo *et al.* 2012a; Luo *et al.* 2012b). This could partially be due to the fact that the authors appear to overlook the fact that concerted TSs for phosphoryl transfer reactions lie on a spectrum from more expansive/dissociative to more compact/associative TSs (Kamerlin & Wilkie, 2007; Kamerlin *et al.* 2008a; Rosta *et al.* 2008), and the discussion in (Hou & Cui, 2012) focuses on the idea that a change in mechanism necessarily would constitute a shift toward a more dissociative TS, whereas the authors are clearly presenting a shift toward a more *associative* TS, which is still a shift in mechanism compared with aqueous solution.

In any case, all the focus on whether the enzyme alters the TS compared with aqueous solution overlooks a key point about enzyme catalysis. The issue is not the *size* of the TS, but rather its energetics, and, in fact, in the case of e.g. GTP hydrolysis, Ras p21 shows a similar catalytic effect for both associative and dissociative pathways (Glennon *et al.* 2000) (and therefore the preferred pathway is determined by the corresponding energetics in solution, which in the case of GTP hydrolysis is far lower for an associative than for a dissociative mechanism, see also discussion in Glennon *et al.* 2000). Similarly, in the case of a related member of the superfamily (see below), the substrates toward which this enzyme showed the greatest catalytic proficiency out of a range of chemically distinct promiscuous substrates were those with both the most compact and the most expansive TS, suggesting that electrostatics overrides the size of the TS (provided that the active site is sufficiently plastic to allow the substrate to bind) (Luo *et al.* 2012b). The importance of electrostatics in determining the specificity was further demonstrated by examining the electrostatic contribution of different active site residues ('groups') for the hydrolysis of a range of substrates by PAS, and showing that, in all cases, the group contributions are quantitatively different, reflecting the differing requirements for efficiently catalyzing the substrates examined, but, qualitatively, they overlay almost perfectly (Fig. 52), illustrating that its electrostatic interactions with the same residues that are driving the selectivity and promiscuity. Finally, tying in with this, the argument of Fig. 11 of Hou & Cui, 2012, which essentially tries to correlate TS binding free energies with the size of the TS has no quantitative basis, as the TS binding free energy is not determined by the size of the TS, but rather by electrostatic interactions with given residues in the active site, which are determined by the active site preorganization, which the enzyme has evolutionarily optimized.

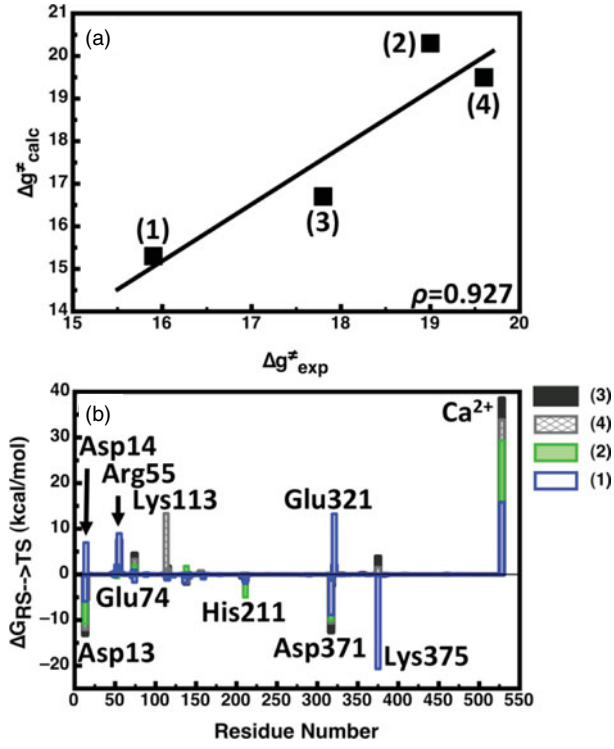

Fig. 52. (*a*) The correlation between theoretical and calculated activation barriers and (*b*) overlay of the electrostatic group contributions for the hydrolysis of (1) *p*-nitrophenyl sulfate, (*b*) ethyl-*p*-nitrophenyl phosphate, (*c*) bis-*p*-nitrophenyl phosphate and (*d*) *p*-nitrophenyl phosphate by the arylsulfatase from *Pseudomonas aeruginosa*. This figure was originally presented in (Luo *et al.* 2012b).

4.5.2 Staphylococcal nuclease (SNase) and the effect of metal substitution

EVB simulations of SNase (Åqvist & Warshel, 1989) provide arguably the first simulation of enzyme catalyzed phosphate hydrolysis. The catalytic effect of SNase is described in Fig. 53, which considers the catalytic reaction of this enzyme. Essentially all the tremendous catalytic effect in this case can be attributed to the electrostatic interaction between the Ca^{2+} ion and the TS. SNase has significantly been used to explore the effect of metal substitution on enzyme catalysis that is, (Åqvist & Warshel, 1990) quantitatively reproduced the effect of Ca^{2+} substitution on phosphate hydrolysis, in terms of the change of the electrostatic effect of different metal ions (which were in turn represented by the proper change in ionic radius).

In order to appreciate the power of this computational analysis, it might be useful to consider related experimental studies of SNase. That is, insightful analysis of the different contributions to catalysis for different residues were reported by Mildvan and co-workers (Mildvan, 1979), which, although the contributions are not additive, still provide a very useful benchmark for theoretical studies. However, in this respect, it is important to clarify a major misconception that emerged from some such studies (e.g. Hale *et al.* 1993; Wilde *et al.* 1998), that mutated the general base, Glu^{43}, to an aspartate. Although these works found a reduction in the catalytic effect of the enzyme by a factor of 2000, which (as outlined in chapter 9 of Warshel, 1991) is fully consistent with the increase in distance to the base, it has been argued that the observed effect does not prove that Glu^{43} is the base, since the mutation involves significant structural changes.

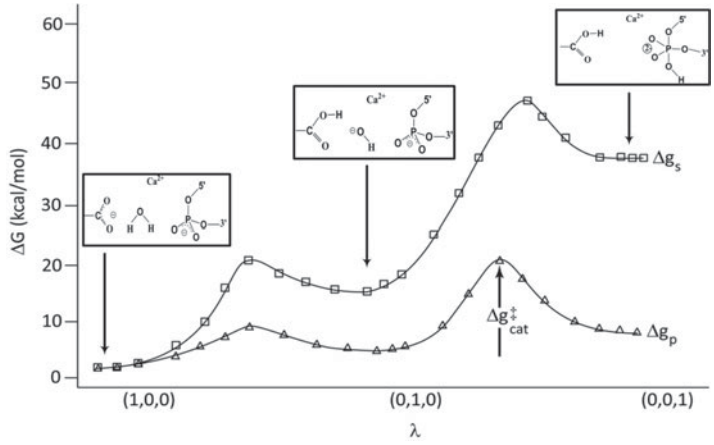

Fig. 53. Calculated free-energy profiles for the reaction catalyzed by SNase in aqueous solution (squares) and in the enzyme active site (triangles). Shown here are also the corresponding valence bond structures for the different reacting states. This figure was originally presented in (Åqvist & Warshel, 1989).

The problem here is two-fold. First, the best (and perhaps most valid) proof that Glu^{43} is the base should come from calculations (which should, of course, be consistent with mutation experiments). Secondly, all mutations involve structural change (this is an integral part of the mutation), but, the change in the effect of each residue should be evaluated while carefully considering the change in the active site preorganization. Unless the mutation completely destroys the local folding, its effect should be assessed by free-energy calculations, and *not* by assuming that all the protein is fixed, and simply presuming what would happen. In fact, the reorganization of the enzyme following the mutation is simply part of the effective dielectric of the system, which, again, cannot be assessed by inspection without electrostatic calculations.

The problem along the above lines continued in (Hale *et al.* 1993), where it was argued that Glu^{43} could not be a general base, since the pH profile of the reaction provided an effective pK_a of approximately 7·5, whereas the neighboring Ca^{2+} should presumably reduce the pK_a of Glu^{43}. This is a typical argument that looks superficially reasonable, reflecting the assumption (see also the case of DNA polymerase described in Section 4.3) that pK_a changes in proteins can be determined by looking at subsystems e.g., the Glu-Ca^{2+} system. In fact, while the Ca^{2+} probably reduces the pK_a of Glu^{43} (although its effect may be balanced by the phosphate charges), it is likely that the observed pH dependence reflects the pK_a of the $Glu^- \cdots HOH \rightarrow GluH \cdots OH^-$ system (where the p*Ka* of the Ca^{2+} bound water is strongly reduced). The key issue is the free-energy cost of the PT in this system (which was evaluated by Åqvist & Warshel, 1989). Of course, one may also examine the origin of the pH dependence, but one should not use completely qualitative presumptions in dismissing quantitative studies. Here, the best way to judge this system is probably to ask whether an inspection-based analysis that tries to tell us what the key factors in catalysis could make any prediction of the actual catalytic effect (to which the answer is no). In doing so, one must realize that reliable EVB calculations (Åqvist & Warshel, 1989) have reproduced both the overall catalytic effects and the effect of mutants, as well as the effect of metal-substitution. In fact, the main catalytic effect of SNase (see Åqvist & Warshel, 1989) comes from the Ca^{2+} ion, which contributes approximately 15 kcal mol^{-1} catalysis, whereas base-catalysis by Glu^{43} is a chemical effect that is already included in the reference reaction in water, and thus is not a true catalytic effect for the given mechanism.

4.5.3 Kinases

The control of cellular signaling frequently involves PTM of protein residues, such as phosphorylation (catalyzed by protein kinases), and dephosphorylation (phosphatases) (Hunter, 2000, see also Section 4.4, and references cited therein). Theoretical studies of the mechanisms of kinases have not been extensive, and therefore we will only briefly consider some examples and directions in this section. One instructive case is the recent study of Smith and co-workers (Smith *et al.* 2011), which attempted to explore the mechanism of the phosphoryl transfer reaction catalyzed by a cyclin-dependent kinase, CDK2, by means of QM/MM. The authors concluded that the reaction involves an aspartate-as-base mechanism, which is in contrast to the conclusion of previous studies (Cavalli *et al.* 2003; De Vivo *et al.* 2007), which explored a catalytic mechanism in which the nucleophilic attack of a serine on the phosphate proceeds *via* a substrate-as-base mechanism (see also Section 4.1 for analogous discussion). Unfortunately, in the absence of truly rigorous theoretical studies (for what we define as a rigorous study, see the Conclusions section of this review), it is still unclear what the mechanism of CDK2 actually is. However, it is instructive to point out some of the problems with Smith *et al.* 2011. The first (and most alarming) is the fact that Fig. 4 of this work shows a clear concerted S_N2 TS, which the authors then describe as a metaphosphate dissociative TS. This problematic analysis is recurrent in the field, from qualitative interpretations of LFER (Lassila *et al.* 2011), to the repeated description of surfaces that are clearly dominated by bond formation to the nucleophile as proceeding through dissociative pathways (Akola & Jones, 2006; Lopéz-Canut *et al.* 2011). Furthermore, even when ignoring this issue, one could ask what the computational evidence for the alleged aspartic acid-as-base mechanism actually is. We find it truly problematic to see that the mechanism with PT from serine to the phosphate oxygen was rejected based on a one-dimensional profile in which the energy went up by 40 kcal mol^{-1} (similarly to the analogous Ras case, Grigorenko *et al.* 2005). Of course, the problem here is one of an improper electrostatic treatment. More specifically, one can always check the limiting values of the energetics with simple pK_a considerations. For instance, stepwise PT from serine to the γ-phosphate cannot be too different from the energetics obtained from the pK_a difference in solution, and the corresponding ΔpK_a is most probably less than 20 kcal mol^{-1}, and definitely not 40 kcal mol^{-1}. The obvious reason for obtaining the high 40 kcal mol^{-1} barrier is the fact that the calculations have not allowed the protein to relax properly in response to the new charge distribution. In fact, the QM/MM sampling is unlikely to be able to reproduce any reasonable pK_a in a protein interior.

The energetics can then be further analyzed by examining the stepwise pathways for PT (which was never examined by Smith *et al.* 2011), where one can compare a mechanism (Fig. 54) with PT from Ser to Asp, followed by the attack of the Ser on the negatively charged phosphate, or the corresponding mechanism with PT from Ser to the phosphate oxygen and nucleophilic attack on a protonated phosphate (which has a lower barrier than attack on the non-protonated phosphate). Since the energy of PT from Ser to Asp and from Ser to the phosphate are expected to be similar even in the protein (to within 5 kcal mol^{-1}), it is hard to see why the barrier for the phosphate-as-base mechanism was found to be so much higher than that for the Asp-as-base mechanism. The most obvious conclusion from this is that this QM/MM study (as many similar works) provides reasonable energies for pathways that do not involve large changes in the charge of the reacting fragments, but cannot capture the energetics of pathways that involve large changes in the charge distribution, and require much more sampling. In fact,

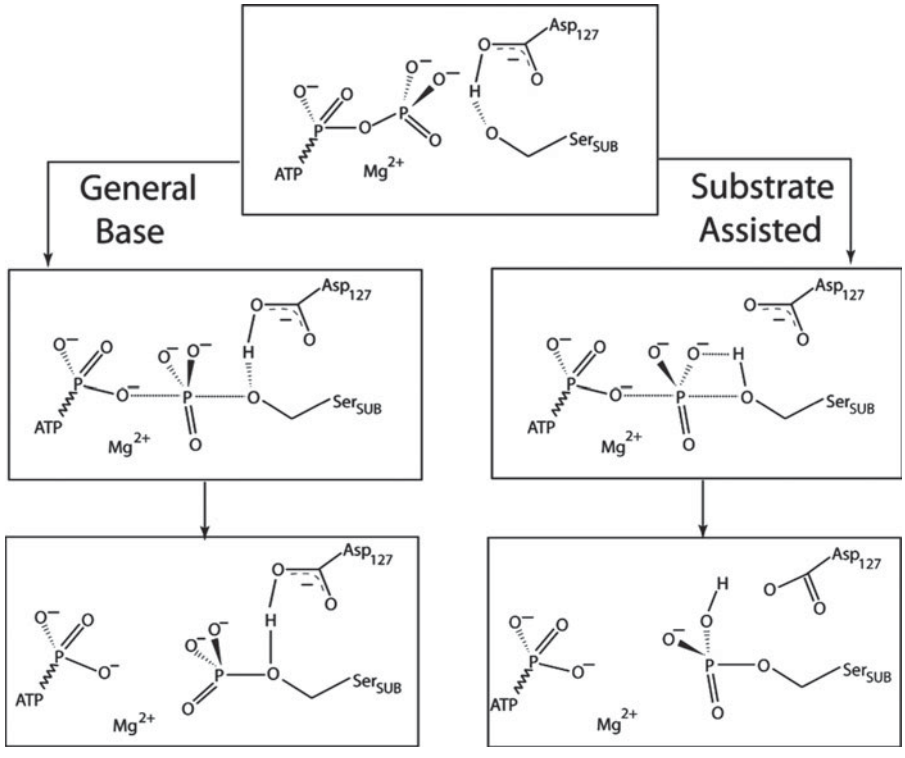

Fig. 54. Different mechanistic possibilities for the reaction catalyzed by cyclin-dependent kinase 2. This figure is adapted from (Smith *et al.* 2011).

even though (Smith *et al.* 2011) talks about having a PMF, it is basically a long MM run with a charge distribution that does not correspond to the QM/MM charges, and then a short QM/MM run near the MM structure. Of course, this would never capture the energetics of paths that involve significant relaxation of the protein electrostatic environment, and calculations that do not involve proper electrostatic relaxation will always favor the least charge change (missing the key features of protein electrostatics). Finally, the experimental finding that mutations of Asp to Ala only reduce the rate by three orders of magnitude (approximately 5 kcal mol^{-1}) indicates that the difference between the alternative mechanisms cannot be 20 kcal mol^{-1}.

Theoretical studies of kinases have been limited, and mainly involved partial analysis of the reaction, missing the relevant process, or using oversimplified model systems (see for instance Cavalli *et al.* 2003; De Vivo *et al.* 2007; Henzler-Wildman *et al.* 2007a, 2007b), and therefore the only study of the actual chemical catalysis that we are aware of is our very preliminary study in (Pisliakov *et al.* 2009) (see Fig. 55). An area where kinases have attracted major interest, however, involves the exploration of the coupling between enzyme dynamics and catalysis. For example, several very high profile studies (Henzler-Wildman *et al.* 2007a, 2007b; Wolf-Watz *et al.* 2004) presented NMR and single-molecule studies which presumably prove that conformational dynamics plays a crucial role in chemical catalysis. Note in this respect that, despite the potential for claims that this was not the intention of the authors, this was exactly what was proposed (see discussion and examples in Kamerlin & Warshel, 2010). At any rate, in addition to demonstrating the logical flaws in the dynamical coupling idea (Warshel & Parson, 2001), we

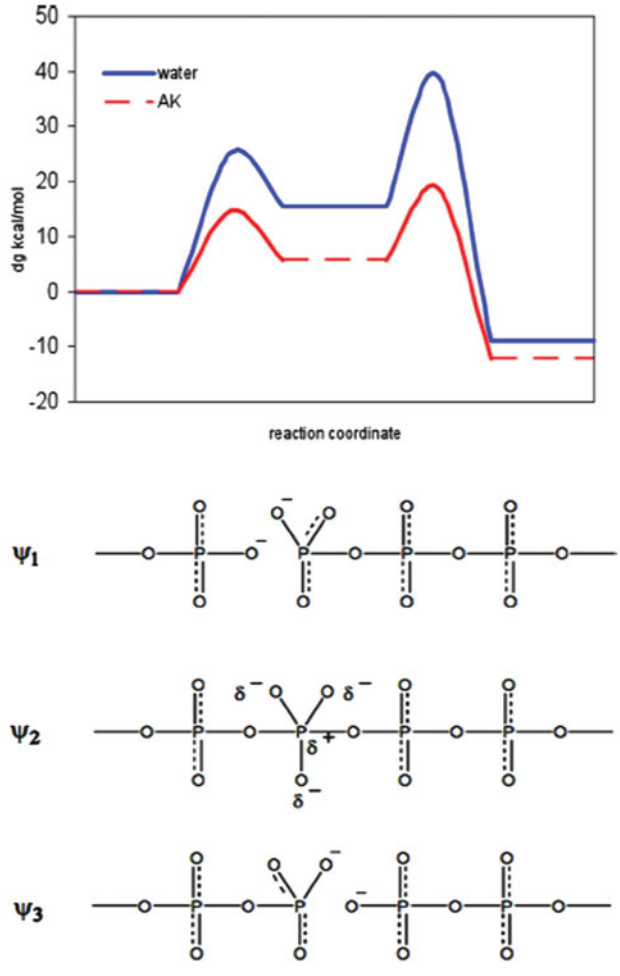

Fig. 55. Calculated free-energy profiles for the reaction catalyzed by adenylate kinase in aqueous solution (blue) and in the enzyme active site (red). Shown here are also the corresponding valence bond structures for the different reacting states. This figure was originally presented in (Pisliakov *et al.* 2009).

were able to develop a theoretical approach (the renormalization model) (Pisliakov *et al.* 2009) that could reach the long timescale relevant to the above experiments, as well as exploring the free-energy landscape in the conformational and configurational space (see Fig. 56). This allowed us to show that all the conformational dynamics are dissipated long before the system is able to surmount the chemical barrier (by Boltzmann controlled fluctuations). This provides clear evidence that the free-energy landscape, and thus the height of the chemical barrier, control the rate of the chemical step in ADP. In light of our findings, we believe that one has to explore the activation of kinases by asking what the allosteric factors that change the activation barrier for the phosphoryl transfer step are, rather than invoking exotic dynamical effects.

4.5.4 Protein tyrosine phosphatases (PTPs)

Following on from our discussion in Section 4.5.3 on kinases, we now turn to a brief introduction of PTPs. Traditionally, protein tyrosine kinases have received the lion's share of

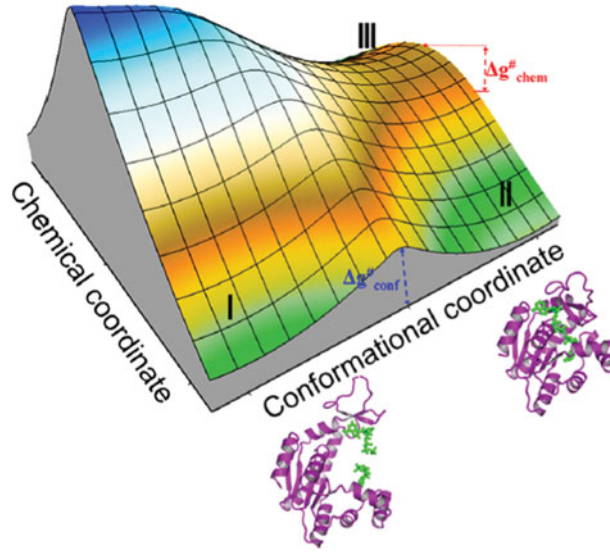

Fig. 56. Effective 2D catalytic landscape for adenylate kinase, evaluated in the space defined the conformational and chemical coordinates. Here, I, II, and III correspond to the open reactant state, the closed reactant state and the closed product state respectively. This figure was originally presented in (Pisliakov *et al.* 2009).

attention, and there are numerous reviews on these enzymes (e.g. Hunter, 1995; Hunter & Cooper, 1985). However, PTPs have also recently gained dramatically in popularity, in large part due to increasing awareness of their role as potential drug targets (Hoofs van Huijsduijnen *et al.* 2002), particularly for treating type II diabetes and obesity (Johnson *et al.* 2002; Pei *et al.* 2004). Additionally, they have been shown to play an important role in any process governed by protein tyrosine phosphorylation (Hunter, 1995), giving them a central role in many signaling processes. Therefore, the availability of potent and selective inhibitors of this enzyme is useful both for therapeutic purposes, as well as for further elucidation of cellular signaling pathways.

PTPs are traditionally classified into four subfamilies (Kolmodin & Åqvist, 2001): (i) tyrosine-specific phosphatase (which form the main group of this family), (ii) dual-specificity phosphatase (DSP), (iii) low molecular weight PTPs and (iv) dual-specific cdc25 phosphatases. These enzymes differ in sequence and topology, and also whether they are cytosolic or membrane bound, but, with the exception of the cdc25 phosphatases, which have one residue substitution, all contain the characteristic (H/V)CX_5R(S/T) sequence (Kolmodin & Åqvist, 2001). These enzymes utilize a two-step 'ping-pong' mechanism, in which a nucleophilic cysteine attacks the bound phosphate, allowing for leaving group departure in a concerted fashion, with the assistance of an active site aspartate that acts as a general acid, resulting in a covalently bound phosphoenzyme intermediate (Guan & Dixon, 1991; Pannifer *et al.* 1998). In the second step, the same aspartate now changes role to a general base, activating a water molecule for nucleophilic attack on the bound phosphate, to give P_i and the restored enzyme. Residues 177–185 of this enzyme comprise a flexible loop, termed the 'WPD' loop, which exists in an open conformation in the native, unliganded form, and subsequently closes over the bound substrate, forming a tight-binding pocket for the substrate. The closure of this loop is essential for the catalytic mechanism of PTP1B (Kolmodin & Åqvist, 2001), as it positions the catalytic Asp in the relevant position to act

as a general acid/base, and it has been demonstrated that the enzyme can be selectively inhibited by allosteric inhibitors (Wiesmann *et al.* 2004) that block the loop in a semi-closed position (Kamerlin *et al.* 2007), therefore preventing the catalytic Asp from reaching the active site. It should be noted as an aside that the catalytic Asp is missing from the cdc25 phosphatases, however, it has also been suggested that in this case the role of general acid is fulfilled by the substrate protein, cdk2–pTpY/CycA (Chen *et al.* 2000). Also, the conformational change necessary for catalysis in this enzyme has been correctly predicted from MD simulations (Kolmodin & Åqvist, 1999).

In determining the catalytic mechanism for PTPs, there are a number of questions of importance. The first is the pK_as of the nucleophilic cysteine, and therefore whether a general base would be required for nucleophile activation. Experiment has suggested that the pK_a of this residue could be as low as below 4 (Evans *et al.* 1996), and therefore that the residue is likely to be anionic in the resting state of the enzyme. There have been several attempts to calculate this pK_a, e.g. by using Poisson–Boltzmann approaches (Peters *et al.* 1998; You & Bashford, 1995), however, these have been discouraging due to errors of up to 10 pK_a units compared with experiment (for a discussion of challenges involved in calculating pK_as of ionizable groups in proteins, see e.g. Kamerlin *et al.* 2009b). However, the calculations also suggest depressed pK_a for this residue, and more detailed calculations EVB calculations that dissected both the energetics and the structural features of PT between the cysteine thiol and the phosphate suggested that the protein environment significantly stabilizes the thiolate ion, preparing it for nucleophilic attack (Hansson *et al.* 1997). A parallel problem, therefore, is then the ionization state of the substrate itself. In principle, the phosphate group could bind in either its mono- or dianionic forms, although the presence of a negatively charged nucleophile makes the dianionic form less likely QM/MM simulations on the low-molecular weight PTP (LMPTP) using both mono- and dianionic forms of the substrate predicted similar activation barriers (Alhambra *et al.* 1998), but ignored the substrate binding. On the other hand, EVB calculations also yielded similar barriers for both steps (Kolmodin *et al.* 1999), but took into account the free energy of the substrate binding, and found significant destabilization of the enzyme substrate complex for the fully ionized state. Use of a monoanionic substrate was also able to reproduce the energetics of the entire reaction pathway for both wild-type LMPTP as well as the D218A mutant (Kolmodin & Åqvist, 1999), providing further evidence in favor of a monoanionic substrate. Finally, it should be noted that there do exist several elegant experimental investigations of the relevant pH rate profiles for these systems (e.g. Evans *et al.* 1996; Zhang *et al.* 1994, 1999). However, as discussed also in (Kolmodin & Åqvist, 2001), the qualitative interpretation of such profiles is extremely challenging, as multiple ionizable groups are involved in the process, making it hard to make unambiguous assignments. Nevertheless, the work of e.g. Kolmodin *et al.* 1999 once again shows the enormous predictive power of careful computational (and, in particular EVB) calculations in resolving complex mechanistic questions.

4.5.5 β-Phosphoglucomutase (βPGM): metal fluoride TS analogs as probes for phosphoryl transfer

In discussing phosphoryl transfer, it is worth mentioning βPGM that has been the topic of great interest, following the publication of a structure of an apparent high-energy pentacoordinated phosphorane intermediate (Lahiri *et al.* 2003), with elongated P–O distances

of 2·0 and 2·1 Å to the oxygens of the substrate's glucose ring and the side-chain carboxylate of the aspartate to which the phosphoryl group is being transferred. However, as was pointed out by Blackburn & Williams (2003), the structure presented by Lahiri and coworkers corresponds to a species with a half-life of $<10^{-11}\,s^{-1}$ for its decomposition even at a temperature as low as 93 K, and also, more importantly, it had already been demonstrated that MgF_3^- is not only a good mimic of the TS for phosphoryl transfer reactions (Graham *et al.* 2002), but also that it can form under the crystallization conditions employed by Lahiri and co-workers. This was then later examined by detailed NMR studies that also provided insight in the molecular mechanism of fluoride inhibition by βPGM (Baxter *et al.* 2006), as well as further (steady and pre-steady state) kinetic analysis of both the reactivity of βPGM and its inhibition by magnesium fluoride, which essentially ruled out the existence of a stable pentavalent phosphorane intermediate in the active site of βPGM (Golicnick *et al.* 2009). Recent theoretical studies would also seem to support this finding (Marcos *et al.* 2010). Nevertheless, even if we do not have a TS intermediate, the alternative existence of a metal fluoride TSA in the active site of the enzyme is of great interest in and of itself, as it allows for exploration of the enzyme–TSA complex by ^{19}F NMR, the chemical shifts from which are is highly susceptible to the local electrostatic environment, providing both a sensitive probe for subtle electrostatic changes in the enzyme at a level that cannot be provided by any other experimental technique at present, as well as a means to resolve structural assignments that are indistinguishable by X-ray crystallography (see discussion in Baxter *et al.* 2006).

4.5.6 Ribozymes and ribonuclease

The discovery that RNA molecules can function as enzymes (Guerrier-Takeda *et al.* 1983; Kruger *et al.* 1982) revolutionized the concept of early life by showing that biological catalysts do not require proteins. The progress in structural studies of ribozymes (e.g. Pley *et al.* 1994; Pyle, 1993; Scott, 1999), as well as the corresponding biochemical studies (Narlikar *et al.* 1995; Piccirilli *et al.* 1993; Scott, 1999) set the stage for theoretical studies of these systems. Due to space limitations, we will not expand on these systems, and rather simply point the readers in the direction of several theoretical studies of ribozymes and relevant model systems (Dejaegere & Karplus, 1993; Leclerc & Karplus, 2006; Lopez *et al.* 2006; Mayaan *et al.* 2004; Mlynsky *et al.* 2011; Schiemann *et al.* 2003; Torres *et al.* 2003; Uchimaru *et al.* 1993; Zhou *et al.* 1996). We would also like to clarify that despite the interest in exotic intermediates (see e.g. Scott, 1999) that were presumably indicated by *ab initio* calculations (e.g. Dejaegere & Karplus, 1993; Zhou & Taira, 1998), they actually reflect irrelevant extremely high-energy studies, that examine possible bumps in the TS region rather than the actual TS energy. We only want to point out that the catalytic power of the ribosome is likely to be due to the same electrostatic reorganization factors that control all other phosphate hydrolysis processes outlined in this review. They are, of course, unlikely to be due to dynamical effects or deviations from TS theory (Scott, 1999), for the same reasons that such effects do not contribute to enzyme catalysis (Kamerlin & Warshel, 2010). As for other effects and mechanistic implications, we suggest that the reader keeps the same critical perspective recommended in other sections, and bear in mind that such evidence as isotope effects or LFER are far from providing unique mechanistic markers.

The study of ribozymes has recently progressed with several QM/MM studies of the hammerhead ribozyme and related reactions (Wong *et al.* 2011, 2012). One of these works (Wong *et al.* 2011) provided a reasonable free-energy surface, using an AM1-type model. This is

encouraging, though the lack of calibration to (or validation on) solution experiments has been a clear problem, and, in particular, something that has been missing is the evaluation or reproduction of the observed catalytic effect (which is the whole issue with regard to the action of ribozymes). The other of these (Wong *et al.* 2012) is a solution study of a transphosphorylation in the methanolysis of ethylene phosphate in solution. One set of calculations was based on PCM calculations, which seem reasonable, although they do not involve an evaluation of a 2D energy surface. Overall, (Wong *et al.* 2012) report reasonable agreement between having a late TS and the calculated KIE for the methanolysis of ethylene phosphate, and inconsistent results for the S3′ system. Another set of calculations involved *ab initio* QM/MM DFT free-energy calculations. Unfortunately, this study, which would be a major advance over the previous AM1-type studies, seems to be a very preliminary study, since no details are given on what it took to overcome such a major sampling challenge, and, in light of the computational cost involved in each individual DFT call, it is hard to accept the quoted 1 kcal mol^{-1} error range as a reasonable estimate of the convergence problems in such a highly charged system, with what must have been limited sampling. It is also interesting to clarify that the path integral approached used for calculating the isotope effect cannot be applied to QM/MM studies, where one has to use our QCP approach (Hwang & Warshel, 1993, 1996; Olsson *et al.* 2006a), which allows one to consider the system's fluctuations.

The ribonucleases (RNAses) are a class of enzymes that very effectively catalyze the cleavage of phosphodiester bonds in RNA (Blackburn & Moore, 1982; Perreault & Anslyn, 1997; Raines, 1998; Richards & Wyckoff, 1971) (for instance, RNAse A, in particular, provides an estimated 15 orders of magnitude rate acceleration). RNAse A has been one of the most studied of all enzymes, reflecting in part the fact that this enzyme was among the first enzymes whose structure was solved (see e.g. Richards & Wyckoff, 1971), as well as from interest in the mechanism of the transesterification reaction (which is same class of reaction catalyzed by ribozymes) as well as from the early ground state five-membered-ring strain proposal of Westheimer (Westheimer, 1968). However, the ring strain effect turned out to be small, where it was found that the TS is better solvated than the ground state in *water* (Dejaegere & Karplus, 1993). Unfortunately, the preoccupation with the ring issue turned out to be related to the wrong question, and caused significant confusion in the field. That is, although mutational studies (see e.g. Raines, 1998) were instructive in identifying catalytically important residues, they could not resolve the origin of the overall catalysis, which was, in some cases (Thompson *et al.* 1995), attributed to ground state desolvation effects and claiming that enzymes work by reducing the ground state solvation. In fact, the popular idea of ground state destabilization has been shown to basically never work in enzymes (see discussion in Kamerlin *et al.* 2010; Warshel *et al.* 1989, 2000, 2006b). The confusion has stemmed from the belief that (Dejaegere & Karplus, 1993) were able to explore the catalytic effect of the enzyme, whereas they only studied the reaction in water, and basically addressed Westheimer's strain proposal (Westheimer, 1968), which had nothing to do with the effect of the enzyme, but rather with the energetics of the reference reaction in water. Apparently, it was an EVB study (Glennon & Warshel, 1998) that provided the first clear assessment of the origin of the catalytic effect in two feasible mechanisms, showing that it is due to electrostatic preorganization (and thus improved TS solvation) rather than desolvation. This is fascinating, since no metal is involved in the large electrostatic effect.

In considering the overwhelming electrostatic effect in the catalytic reaction of ribonuclease, it might also be useful to point out works that assumed that the electrostatic stabilization effect of Lys^{41} is actually due to a low-barrier hydrogen bond (LBHB) (see e.g. Gerlt & Gassman, 1993).

Here, we must again point out that such catalytic issues must be judged by calculations, rather than by simply presuming what the effect is, and, in any case, it has been shown that the LBHB proposal is anticatalytic (e.g. Schutz & Warshel, 2004; Warshel & Papazyan, 1996).

5. Summary and perspectives

This review has covered key issues about the mechanism and chemistry of reactions that involve the cleavage and formation of phosphoester bonds in aqueous solution and in relevant biological systems. This is achieved by exploiting the insight obtained from different computational approaches, which should, from our point of view, provide a way to combine the available experimental information to provide a consistent mechanistic picture. Considering the length of the review, it is useful to conclude by summarizing its key points.

5.1 What is the actual mechanism of phosphate hydrolysis and related reactions?

The importance of phosphoryl transfer reactions was realized quite early on, which led to several decades of extensive and careful mechanistic studies by very prominent research groups (e.g. Butcher & Westheimer, 1955; Di Sabato & Jencks, 1961; Hengge, 1999, 2002; Herschlag & Jencks, 1989a, 1989b; Kirby & Jencks, 1965a; Kirby & Varvoglis, 1967a, 1968b; Kirby & Younas, 1970b; Lad *et al.* 2003b; Stockbridge & Wolfenden, 2009; Westheimer, 1968, 1981; Williams, 2004a; Williams & Wyman, 2001; Wolfenden *et al.* 1998). The experimental studies provided valuable insight, but, in some cases led to dogmatic views on issues that could not be resolved, such as the idea that phosphate monoester hydrolysis cannot involve initial PT to the substrate from an attacking water molecule, followed by hydroxide attack on a monoester (Sections 3 and 4), or the idea that LFER can prove the mechanism to be dissociative (Section 3).

The emergence of theoretical approaches has allowed for the re-examination of mechanistic issues, despite at times strong objections from some quarters. Interestingly, the objections have been rather minimal (as can be judged by the ease of publication) toward studies that have attempted to address the issue using rather irrelevant gas-phase calculations (which focused on whether there is a small bump at the TS rather than the energy height of the actual TS, see e.g. Dejaegere *et al.* 1994; Zhou & Taira, 1998). The situation has become more controversial, however, with the emergence of a series of careful quantum mechanical studies of the reaction in solution (Alkherraz *et al.* 2010; Florián & Warshel, 1997, 1998; Kamerlin, 2011; Kamerlin *et al.* 2008a, 2008b; Klähn *et al.* 2006; Rosta *et al.* 2008), which reproduced the observed experimental results, but sometimes with a different mechanistic picture than those accepted by many workers. While this of course does not mean that the current computational picture is necessarily perfect, at the end of the day, the correct mechanistic picture will emerge from theoretical studies (but only from those that are capable of reproducing all the relevant available experimental information, while also telling us about the actual reaction surface).

Similar (or even more serious) problems have concerned studies of phosphate hydrolysis in enzymes, where there appears to be a tendency to adopt unproven (and in fact incorrect) assumptions from solution reactions, such as that we have a dissociative reaction in aqueous solution (Cleland & Hengge, 2006; Di Sabato & Jencks, 1961; Hengge, 2002; Herschlag & Jencks, 1989a, 1989b; Lassila *et al.* 2011), or the idea that electrostatic interactions with metal ions does not make a major contribution to catalysis (Admiraal & Herschlag, 1995; Herschlag & Jencks, 1987; Lassila *et al.* 2011), and to then project these assumptions on to the corresponding

enzymatic reactions. The problem has then been compounded when the arguments involve unjustified assertions, as is the case in the discussion about the phosphate-as-base mechanism (see Section 4.1). Here, works (e.g. Lassila *et al.* 2011) that do not mention the fact that those who proposed this mechanism (Schweins *et al.* 1994) have also proven that the alternative glutamine-as-base mechanism is invalid (Langen *et al.* 1992), are omitting a very relevant point. The same is true when the analysis ignores the fact that the water proton ultimately ends up on the phosphate (see more problems with this issue in the discussion in Section 4.1). This problem is further compounded by a tendency to take isotope (Hengge, 2002) and entropic (Kirby & Jencks, 1965a) effects, which are qualitatively difficult to interpret, and to use them as unique mechanistic markers supporting a dissociative mechanism, as was done in e.g. (Hoff & Hengge, 1998b; Lassila *et al.* 2011). In such cases, it may be useful to note that careful theoretical studies (Åqvist *et al.* 1999; Kamerlin *et al.* 2008a; Klähn *et al.* 2006; Rosta *et al.* 2008) have identified the risks and inconsistencies associated with the qualitative interpretation of traditional experimental markers. This is perhaps not surprising, recalling that experimental markers provide hard physical data while on the other hand, interpreting such data can at times become a game of Jeopardy, where one is faced with facts but has to find the question. In fact, our recent analysis of the complications associated with the 1W and 2W paths clearly established that we have here a complex situation that could not be envisioned or analyzed by the early experimental markers. In view of the above problems, it is important to combine experimental work with detailed theoretical studies (e.g. QM or QM/MM studies) that can reproduce the relevant observed experimental markers, allowing for discrimination between different mechanistic possibilities. Perhaps the best way to realize what the future direction in this field should be is to compare catalytic ideas based on examining the mechanism in solution and qualitative considerations (see e.g. Lassila *et al.* 2011 and references cited therein) to systematic computer simulations. It seem reasonable to accept that the way to progress is to use calculations that can not only reproduce the observed catalytic effect (see e.g. Warshel *et al.* 2006b and references cited therein), but also the effects that mutations are observed to have on the catalysis (as has been done in countless computational studies on a wide range of systems, see e.g. Adamczyk *et al.* 2011; Adamczyk & Warshel, 2011; Frushicheva *et al.* 2011; Liu & Warshel, 2007; Roca *et al.* 2009; Shurki & Warshel, 2004). This is the only way one can not only characterize the mechanism of the enzyme-catalyzed reactions, but also understand the molecular basis for the observed effect.

5.2 The requirements for meaningful theoretical studies of phosphate hydrolysis

The previous section mentioned the importance we attribute to theoretical studies as the ultimate way to resolve the mechanism of phosphate hydrolysis (and, by extension, related group transfer reactions). Here, we are faced with a very serious issue if we would like to move forward from mechanistic ideas based on powerful but qualitative physical organic chemistry, to increased reliance on theoretical studies. While it is tempting to assume that very high level *ab initio* quantum mechanical calculations would provide the ultimate answer, one nevertheless has to be very careful (although, eventually, such calculations will lead to *definite* answers). Here, the most promising strategy appears to be to calibrate to both available experimental data as well as calculations on the background reactivity in aqueous solution. This serves a two-fold purpose. The first is purely technical, as a way to carefully validate the approach subsequently being used to study the enzymatic reaction. But, just as importantly, this is essential in order to understand the fundamental chemistry of the reaction being considered in the absence of a catalyst, without

which it is impossible to understand the role of the catalyst. Following from this, the extrapolation to proteins can at present be best done by the EVB approach (Hwang *et al.* 1988; Warshel & Weiss, 1980), which is a highly robust semi-empirical QM/MM approach that is both sufficiently fast to allow for the extensive sampling required for meaningful convergent free-energy calculations, as well as carrying a tremendous wealth of chemical information, allowing for a meaningful description of chemical processes involving bond cleavage and bond formation. It should be noted as an aside that many popular semi-empirical approaches are problematic when attempting to study phosphoryl transfer due to the fact that they have not generally been parameterized for this (examples of the problems with correctly reproducing mutational effects using a semi-empirical approach, see (Hou & Cui, 2012) and also, the low-lying d-orbitals of the phosphorus atom pose a significant computational challenge. In addition to this, their poor treatment of transition metals can lead to unphysical results (e.g. Lopéz-Canut *et al.* 2011). Clearly, more parameterization here would be helpful, and there have been attempts to resolve this issue, particularly in order to take into account the effect of the d-orbitals using a reference quantum dataset (Lopez & York, 2003). Other more successful recent attempts to resolving the accuracy issues with semi-empirical QM/MM calculations include performing initial extensive sampling using a semi-empirical approach, and then adding on a high-level QM correction (e.g. Lonsdale *et al.* 2012). In contrast, the EVB approach has been repeatedly demonstrated to be a highly powerful tool for studying chemical reactivity in solution and in enzymes, particularly if one wants to obtain a quantitative understanding of the molecular basis for observed catalytic effects (several examples of this are presented in Warshel *et al.* 2006b). Therefore, this approach allows one to not only explore the myriad of mechanistic possibilities available to phosphoryl transfer reactions (which are already sufficiently complicated in aqueous solution as outlined in Section 3, before bringing in the added complexity of the enzyme) while at the same time performing a meaningful energetic analysis.

Quantifying the role of the metal ion(s) found in many phosphatase active sites is also critical, and gradual improvements in the description of the metals is needed. Just placing the metals in the QM region is not so useful, since charge transfer to the ligands is an important factor (and also, as mentioned above, both semi-empirical and to some extent DFT approaches provides very poor descriptions of transition metals, Siegbahn, 2006). Additionally, including the metal in the quantum region is very expensive. Here, we have promising strategies that have also been used in the case of, for instance, DNA polymerase β (Oelschlaeger *et al.* 2007). Furthermore, the emerging paradynamics (PD) approach (Plotnikov *et al.* 2011; Rosta *et al.* 2006), that uses the EVB as a reference potential for *ab initio* QM/MM calculations *should* provide a powerful way to further quantify the EVB results. However, we must clarify that the sampling issue is an *absolutely critical* issue, and that several current studies that overlooked this issue led to results that are, in some cases, problematic (e.g. (Alberts *et al.* 2007; Grigorenko *et al.* 2005, 2006), discussed in Section 4.3). Another frequently overlooked issue is the requirement of having a complete free-energy surface instead of just a TS search (Alberts *et al.* 2007; Grigorenko *et al.* 2005, 2006). Not considering a full surface, or examining alternative mechanisms (as is done in the more complete analysis of Figs 19 and 20, and in Prasad *et al.* 2012), tends to lead to ideas such as water wires (e.g. see Section 4.3), or even more problematic results such as that of (Cavalli & Carloni, 2002), who concluded that Gln^{61} is the base in the catalytic reaction of the Ras-GAP complex (for related problems, see Section 4.1), based on very short MD simulations in which the distance between the nucleophilic oxygen and the γ-phosphate of the GTP were constrained to 1·8 Å, resulting unsurprisingly in a collapse of the water proton to Gln^{61} (as, pushing the attacking

nucleophile to within bonding distance of the central phosphorus atom would result in the proton being ejected to the closest base, which, in this case happened to be Gln^{61} due to the artificial constraint placed to the γ-phosphate). Additionally, the authors' simulation system included only Gln^{61}, a few other residues, and the attacking water molecule (the orientation of which was already problematic). Clearly, a meaningful study would require taking into account the full system, and performing both proper equilibration and long free-energy perturbation simulations. Similar problems surfaced in Car-Parinello studies of ATP hydrolysis (discussed in Section 4.2.2). Overall, many of the problems mentioned above are reminiscent of the entrance of *ab initio* workers to studies of enzymes, using gas phase calculations (see the discussion in Warshel, 2003), which are today considered inadequate for studies of enzymes. Here we obviously have to move to more adequate QM/MM approaches. However, this involves the risk of giving the impression to both the public and the workers themselves that the use of *ab initio* QM/MM approaches guarantees the correct results. In fact, this is a problematic conclusion, since QM/MM without very extensive sampling cannot provide reasonable energetics for reactions that involve large charge separation or transfer. Using MD sampling with an MM potential that does not reflect the QM change in charge as preparation for the QM/MM minimization would lead to artificial results, including cases with low barriers due to starting from QM configurations that are not true minima on the QM surface. Conversely, non-trivial reaction paths such as the path for PT from the attacking water to the phosphate in the GTP hydrolysis (Klähn *et al.* 2006) or the existence of multiple, equi-energetic concerted TSs (Kamerlin *et al.* 2008a) are likely to be overlooked by simple energy minimization procedures, and the energy for PT to the phosphate oxygen is likely to be very hard to evaluate without proper electrostatic calculations. Since this point may not be obvious, we must point out that the error in calculation of the energetics of reaction mechanisms using QM/MM calculations is as large as the error of QM/MM calculations of pK_as in proteins (which is typically very large, as discussed in Kamerlin *et al.* 2009b).

We must emphasize that our aim with the discussion above is not to criticize various workers, who have clearly invested tremendous effort into their work, but rather, to serve as a clear warning to both experimentalists and theoreticians in this challenging field that one cannot judge all theoretical studies equally. It is far too easy to obtain erroneous results due to the sheer complexity of the problem and the demands placed on reasonable theoretical approaches, and, therefore, in our view, only works that reproduce observed facts and are validated by reproducing relevant findings in solution are likely to provide meaningful mechanistic insights. In fact, resolving the central controversies is not helped by the problematic tendency (e.g. Akola & Jones, 2003; Grigorenko *et al.* 2007a; Smith *et al.* 2011) to call perfect associative or S_N2 pathways 'dissociative', or to describe a system where the distance between central phosphorus atom and the oxygen atom of the attacking nucleophile is 1·8 Å as involving a 'metaphosphate' species (Grigorenko *et al.* 2006). Nevertheless, we are encouraged to see that, in addition to systematic works using implicit solvent models, the QM/MM community is also starting to get reasonable results for phosphate hydrolysis in solution (Wenjin *et al.* 2012), which incidentally reproduces our results using an implicit solvent (Kamerlin *et al.* 2008a; Klähn *et al.* 2006). This is relevant to the resolution of the controversy surrounding the substrate-as-base mechanism, as, even when exploring the landscape in three dimensions, and in explicit solvent, the preferred pathway is clearly an A_ND_N pathway with the phosphate itself acting as a 'bridge' for the transfer of a proton from the nucleophile to the leaving group (see also the discussion in Kamerlin *et al.* 2008b). Furthermore, another recent work (Branduardi *et al.* 2011) obtained the free-energy surface for the same system by means of MD sampling, using a model that involves explicit

solute (reacting) atoms and an implicit (PCM) solvent model. The resulting surface is also similar to that obtained in our work, although it seems to have feature near the TS that indicate very significant hysteresis. Clearly, we are pleased that these high-level calculations support our earlier study. However, we would like to touch on an important point as an aside – that is, there is currently a dangerous trend in the theoretical community to believe that 'bigger is better', and, a tremendous amount of computational resources were invested into the work of (Wenjin *et al.* 2012). However, while solving complex problems is extremely demanding in terms of computational resources, the issue is not how computationally demanding the simulations are, but rather, whether one is asking the correct questions, and whether one can then formulate it in the simplest possible (physically meaningful) way (Kamerlin *et al.* 2011; Kamerlin & Warshel, 2011b). Here, we would like to point out that it is important to us that our previous results with implicit solvent model were reproduced by use of a far more detailed computational approach, however, ultimately, the results obtained were not changed by simply using more computational resources. Here again we point out to the need of very careful search of different reaction paths as is done in Figs. 19 and 20 and eventually further validated by ab initio QM/MM with extensive sampling (see below).

5.3 Mechanistic insights into phosphate hydrolysis in proteins

Studies of phosphate hydrolysis in different proteins, and, in particular, G-proteins, have provided excellent test cases for resolving central controversies with regard to the mechanism of phosphoryl transfer in a range of systems. As outlined in Section 4.1, this has led to lively dispute not only about the viability of the phosphate itself as a base, but also, given the clearest arguments whether the preferred pathway for GTP hydrolysis is associative or dissociative in nature. Despite extensive evidence to the contrary (e.g. Adamczyk & Warshel, 2011; Florián & Warshel, 1997; Klähn *et al.* 2006; Langen *et al.* 1992), this is still a controversial topic (Admiraal & Herschlag, 2000; Cheng *et al.* 2001; Du *et al.* 2004; Iché-Tarrat *et al.* 2007; Lassila *et al.* 2011). Particularly, the most extreme recent argument has appeared in (Lassila *et al.* 2011), where it has been claimed that structural studies were unable to determine the identity of the base (despite the fact that it was almost uniformly accepted that Gln^{61} is the general base, see Section 4.1.2), and that our substrate-as-base mechanism (Langen *et al.* 1992) was the paradigm shift that established that glutamine cannot be the base. The argument was presented as if our only purpose for proposing such a mechanism was to offer a way to generate hydroxide ion as a nucleophile, and that the fact that there is no big catalytic advantage to doing so (which overlooks the fact that the k_0 for this reaction corresponds not to water attack on a dianion, but to hydroxide attack on a monoanion, see discussion in Section 3). The overwhelming problems with these arguments have been discussed in detail in Section 4.1.2.

The other issue is the evolving story about whether the preferred mechanism is associative or dissociative. It has been repeatedly argued that phosphate monoester hydrolysis will proceed through 'loose' (i.e. dissociative) TSs (Lassila *et al.* 2011). In fact, the term loose is in itself a downgrade of earlier works (Admiraal & Herschlag, 2000; Zalatan & Herschlag, 2006), which pushed for a dissociative, 'metaphosphate-like' TS, despite the extensive theoretical evidence to the contrary. This was most notable in (Admiraal & Herschlag, 1995), which presented an energy diagram (see Section 4.1) which clearly indicated that any associative pathway is likely to be by far higher in energy than a dissociative pathway, an argument that was clearly demonstrated to be wrong by subsequent theoretical studies (Florián & Warshel, 1998; Kamerlin *et al.* 2008a;

Klähn *et al.* 2006). Similarly, the argument that the charge shift in the dissociative mechanism can be used as a rationale for the action of GAP is extremely problematic, especially when coupled with overlooking our earlier and clearer demonstration of the nature of this effect (see Section 4.1).

Finally, we may clarify our general perspective by noting for example the argument (Lassila *et al.* 2011) that it has been demonstrated how the understanding of the mechanism in solution help to understand the enzyme catalysis (or, specifically, that '*the Ras-GAP -GTPase example illustrates how understanding of nonenzymatic phosphoryl transfer reactions can inform and enrich investigations of important biological enzymes*'. In fact, while it is important to have the correct reference solution reaction to quantify catalysis, knowing about the reference reaction does not tell us much, about the enzyme's catalytic effect. What is needed are calculations of the *actual* effect of the enzyme environment on the given charge distribution of the solution reaction (with possible charge modification by the protein environment). Of course, the difference between our approach and the qualitative approach promoted by e.g. (Lassila *et al.* 2011) can also be realized by the fact that we have been actually able to reproduce the effect of different mutations (Adamczyk *et al.* 2011; Adamczyk & Warshel, 2011; Frushicheva *et al.* 2011; Liu & Warshel, 2007; Roca *et al.* 2009; Shurki & Warshel, 2004), whereas (Lassila *et al.* 2011) could only argue that they suggested the direction of the effect of the Arg finger, without being able to even hint at what the magnitude of those effects may be, and were also forced to invoke factors such as water positioning, thereby overlooking the key effect of the P-loop and the Mg^{2+} ion, as well as the fact that catalysis by water positioning has long been shown to have only a minor catalytic effect (both as NAC and as an entropic effect, see discussion in Warshel *et al.* 2006b and references cited therein).

Note that the proposal of the effect of water positioning has emerged in other systems, including the highly popular proposal for the catalytic effect of the exonuclease domain in the Klenow fragment (e.g. Steitz & Steitz, 1993). In this case, it was suggested that the catalysis comes from the orientation of the nucleophilic hydroxide ion by Glu^{357} and Tyr^{397}, which is problematic, as discussed in (Fothergill *et al.* 1995), since the two negatively charged species would repel each other. Similarly, it has been suggested that the role of the Mg^{2+} ion in the active site is to stabilize the correct hybridization (following Westheimer's influential proposal, Westheimer, 1968). However, careful analysis in (Fothergill *et al.* 1995) has demonstrated that the effect of the Mg^{2+} ion is to due to electrostatic stabilization, once again emphasizing the crucial need for an energy-based analysis. A similar problem can be seen in the 'Cheshire cat' conjecture (Yarus, 1993) for RNA catalysis, which provided an elegant argument for the importance of metal ions in RNA catalysts, paralleled with analogous metalloenzymes, but which completely missed the central role of electrostatic effect of the metal ions in the catalysis.

It may also be useful to comment here on the challenges involved in drawing conclusions about the nature of the TS in phosphate hydrolysis from X-ray structures of TSAs (e.g. Bao *et al.* 2011). On the one hand, such structures can provide a powerful starting point for subsequent experimental and theoretical work. However, sometimes the conclusion is that we have a dissociative TS based on the unjustified assumption that the distance between the nucleophile and leaving group can be used to classify a dissociative TS, or just on the belief that solution experiments point to such TS. The problem is that TSAs cannot tell us *directly* about the actual TS (Barbany *et al.* 2003). In other words, the correct strategy must involve reliable computations of both the real TS and the TSA, followed by use of the calculated difference as well as the observed TSA structure as a guide for the 'observed' TS structure.

In conclusion to this and concluding this section, we would like to remark that, since the first quantitative study of phosphate ester hydrolysis in a protein (Åqvist & Warshel, 1989), we have

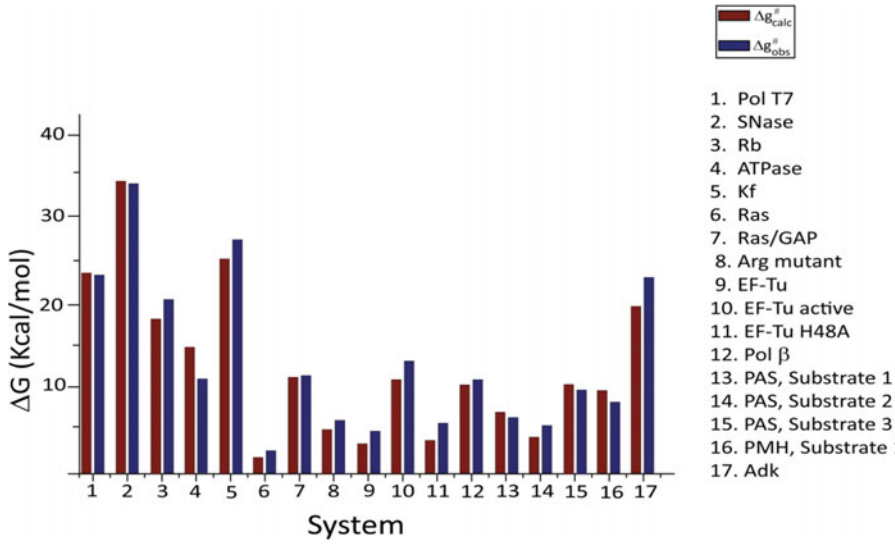

Fig. 57. The catalytic effect (in kcal mol^{-1}) on phosphate hydrolysis in different enzymes. The type of the system considered and the source of the analysis are: (1) T7 DNA polymerase (Florián *et al.* 2003a), (2) SNase (Åqvist & Warshel, 1989), (3) Ribonuclease (dianionic intermediate mechanism) (Glennon & Warshel, 1998), (4) ATPase (Štrajbl *et al.* 2003), (5) The exonuclease activity of the Klenow fragment or DNA polymerase I (Fothergill *et al.* 1995), (6–8) Ras GTPase alone, in complex with GAP and the arginine mutant respectively (Shurki & Warshel, 2004), (9–11) EF–Tu before activation, after activation and the H48A mutant respectively (Adamczyk & Warshel, 2011), (12) DNA polymerase β (Prasad & Warshel, 2011), (13–15). *P. aeruginosa* arylsulfatase (Luo *et al.* 2012a), Substrate 1 = *p*-nitrophenyl phosphate monoester, Substrate 2 = ethyl-*p*-nitrophenyl phosphate, Substrate 3 = bis-*p*-nitrophenyl phosphate (Luo *et al.* in press). Note that all are promiscuous activities, (16) Phosphonate monoester hydrolase from *R. leguminosarum* (Substrate 1, Unpublished results), (17) Adenylate kinase (Pisliakov *et al.* 2009).

witnessed a gradual progress in the computational studies. They range from irrelevant gas-phase calculations, to even recent puzzling QM/MM results providing a barrier of 70 kcal mol^{-1} for a completely impossible mechanism (Lu *et al.* 2011), to a decent *ab initio* QM/MM study of phosphate hydrolysis in solution (Wenjin *et al.* 2012) mentioned above. Presently, it would seem that the most effective way forward has been to study the solution reaction by an implicit solvation model (e.g. Langevin dipoles, PCM, or COSMO), combined with experimental information (see e.g. Alkherraz *et al.* 2010; Kamerlin, 2011; Kamerlin *et al.* 2008a, 2008b; Klähn *et al.* 2006; Rosta *et al.* 2008), and then evaluating the effect of moving from solution to the protein using the EVB approach. The effectiveness of this strategy is evident from its uniform success in studies of different phosphoryl transfer reactions in different proteins, as well as from the reproduction of the (at times enormous) observed catalytic effects. Some of these are summarized in Fig. 57.

Finally, in light of the strength of the EVB approach as a powerful tool for consistently moving from aqueous solution to protein active sites, we would like to point out an interesting relationship between the EVB concept, and a conceptual adaptation in the work of (Lassila *et al.* 2011; Zalatan & Herschlag, 2006) and (Hou & Cui, 2012): Fig. 11 of (Hou & Cui, 2012) presents an interesting diagram (inspired by the work of Herschlag and coworkers, see e.g. Lassila *et al.* 2011) where the energy associated with TS structural changes (moving from more compact to more 'dissociative' concerted TSs) in solution is added to the assumed effect of the protein, in an attempt to assess the energy of structural changes in the protein. A problem, however, is that the solution surface is not evaluated in the most reliable way (see Section 4.5), and the QM/MM

approach used to evaluate the energy in the protein is also not properly calibrated on the solution reaction. Thus, we are dealing here more with a qualitative demonstration of a principle rather than a powerful quantitative computational concept (such as the EVB). The problem may be even more serious when, instead of using solution calculations, one is using the *presumed* energy difference in solution (which can be very problematic, as demonstrated in Fig. 34). At any rate, the problems with this qualitative analysis, which lacks a sound quantitative basis, proves our point on the importance of using careful studies of the corresponding background reaction in solution as the basis for informed analysis of the enzymatic reaction.

Obviously, with increasingly available computer time it is important to move into the evaluation of QM/MM *ab initio* (QM (ai)/MM) free-energy surfaces in solutions and proteins. Here we believe that our PD approach using the EVB as a reference potential (Plotnikov *et al.* 2011; Plotnikov & Warshel, 2012) should provide one of the most effective strategies, due to its great efficiency. In this context it is important to note that the alternative MTD approach recently used to study the hydrolysis of phosphate monoester dianion in solution (Glaves *et al.* 2012). As in several other problematic studies it was assumed that a fully associative/concerted calculated path is dissociative and then concluded that the reaction goes through the PT through two water molecules shown in Fig. 20. Here again we face the problem that frequently appears in minimization studies, the automated MTD study involving most probably an assumed reaction coordinate that has not explored the direct PT (the 1W path in Figs. 19 and 20). It seem to us that PD studies that will allow one to look systematically at different TSs could be very useful for examining the mechanistic issue in a much more systematic way. The issue of obtaining careful QM(ai)/MM mapping will become more serious when applied the phosphate hydrolysis in proteins.

5.4 Why Nature *really* used phosphate

Westheimer, in his 1987 review, asking 'Why nature chose phosphates' (Westheimer, 1987) a number of requirements were suggested in order to rationalize this choice, including the need for a compound that is at least divalent (to hold together e.g. DNA linkages), carries a negative charge (so it can stay in membranes), is not excessively reactive, and sufficiently mechanistically versatile to be able to decompose through mechanisms other than regular nucleophlic attack where necessary. In fact, Westheimer's insightful paper has identified an advantage of using phosphate in biology, recognizing the fact that possibly alternatives (such as arsenc acid) are unstable in solution, and thus have too low a barrier to hydrolysis (resulting in kinetic instability). However, some other assumptions of this work, which were considered to be very reasonable at the time, have subsequently appeared to be less certain. This includes the assumption that the dissociative (metaphosphate) mechanism is advantageous and plays a major role in the hydrolysis, and the idea that the negatively charged phosphate makes attack by OH^- very slow (both assumptions have been shown to be problematic in Sections 3 and 4.1). Most significantly, however, is the key point that was not examined by him, which is the electrostatic origin of the catalytic effect of the proteins (see below).

There has been a wealth of research in this field since 1987, and one of the purposes of this review has been to critically examine and explore this issue, readdressing this key question. As outlined in the introduction, the cleavage of phosphorus–oxygen bonds is arguably the most important biological reaction, being involved in everything, from signaling and energy transduction to protein synthesis and maintaining the integrity of the genetic material. Therefore, understanding phosphoryl transfer reactions has tremendous therapeutic implications, as most

disease pathways involve the malfunctioning of proteins that regulate this 'simple' biological reaction. Which brings us back to the same question: why *did* nature choose phosphates? There are a number of reasons for this. The first of these is that fact that, as hinted at by Westheimer (although he cannot have at the time envision the tremendous rate accelerations required), catalyzing phosphoryl transfer is notoriously demanding, with half-lives for the cleavage of phosphorus–oxygen bonds extending into millions of years (Wolfenden & Snider, 2001). This immediately gives phosphate esters an advantage as a biologically relevant compound, as the existence of life as we know it depends on the tight regulation of interrelated chemical reactions occurring *in vivo*, where a simple difference in rate of a factor of 2 can be the difference between life and death. Such tight regulation is impossible if the reactions being regulated could occur spontaneously, which is why we are also highly skeptical of the idea that arsenate can replace phosphate in biological systems (see the discussion in Section 4.3). However, despite the tremendous rate accelerations enzymes need to achieve to cleave phosphorus–oxygen bonds, phosphate esters are highly versatile. As outlined in this review, they can proceed through a multitude of different mechanisms, and also, equally importantly, can act as their own proton sink for nucleophile activation, giving them an immediate evolutionary advantage over, for instance, sulfuryl transfer reactions, where the pK_a of the non-bridging oxygens on the sulfate is sufficiently low to require the recruitment of an enzymatic residue as a base (see e.g. Luo *et al.* 2012a). This is doubly important in light of the fact that sulfates fulfill many of the same biological functions as phosphates; however, it seems to appear that they are even more inert than phosphate esters in aqueous solution (it has been suggested that S–O cleaving enzymes as the most proficient known to date, Edwards *et al.* 2012), and, the requirement of an 'external' base makes it more challenging to evolve for sulfatase activity than for phosphatase activity, which perhaps explains the prevalence of phosphate over sulfate esters in biology. Despite this, however, it is the same multitude of potential pathways that makes phosphate esters so versatile to catalyze also so difficult to study, and there are a large number of open questions remaining in the field, which need urgent addressing in light of the biological importance of phosphates. We believe that, the only way the key controversies in the field will be resolved is by means of detailed and careful computational studies (that, of course, can reproduce all the available experimental data). However, the answer to why nature *really* chose phosphate seems to be almost completely associated with electrostatic effects. As illustrated in Section 4, proteins exploit the charge shift upon motion from the reactant to TS in phosphate hydrolysis, to convert a very high-activation barrier in solution to a relatively low activation barrier, and thus gain complete kinetic control on many biological processes (see e.g. the discussion of ATP hydrolysis in Section 4). Next, the electrostatic repulsion between negatively charged phosphate fragments is used as a source of biological energy (again appoint that was not clearly analyzed before). Finally, having negatively charged phosphates and changing the magnitude of the negative charge during the reaction (e.g. GTP being hydrolyzed to GDP) provides an ideal way for allosteric control of protein function and signal transduction (see Section 4.5).

5.5 Future directions

Phosphoryl transfer has been an intensely studied research area for several decades, and this review obviously does not provide the last word in studies of phosphate ester hydrolysis. While we on the one hand have provided compelling arguments that support the overwhelming trend of our calculated surfaces, as well as clear analysis of the problems with less consistent studies, we

expect theoretical studies in the field to continue. Particularly, we expect increased focus on *ab initio* QM/MM calculations, but with a gradual realization of the crucial role of proper sampling and free-energy calculations, and believe that using reference potentials (like the EVB coupled with the PD approach, Plotnikov *et al.* 2011; Rosta *et al.* 2006) in such calculations will lead the way in terms of the reliability of the sampling process. At some stage, there will be a more uniform consensus, but not necessarily in the immediate future, as it is increasingly easy to run standardized simulation packages in a relatively uncritical way, and, sadly, the belief that the results must automatically be correct is more prevalent than the realization that it is always crucial to validate and calibrate the calculations on well-defined experiments in aqueous solution.

The points above are, of course, very relevant to the analysis of the catalytic effect in phosphoryl transfer reactions. As was clarified and illustrated above, only carefully calibrated studies that take the protein environment into account and perform sufficient sampling can be used to draw conclusions about the catalytic effect in protein active sites. Here, we expect a growing realization of the importance of free-energy calculations, and also a gradual consensus to be reached in the field, upon the emergence of more *ab initio* QM/MM free-energy perturbation calculations.

Additionally, as discussed in Section 3, phosphoryl transfer is not only an extremely versatile reaction, but also, the enzymes that catalyze it are often highly promiscuous, catalyzing more than one related group transfer reactions, all of which are very demanding to catalyze, but with highly different requirements for efficient catalysis of a substrates with a wide range of sizes and charges. The existing multifunctionality of these enzymes, combined with the tremendous demands of catalyzing even the native reactions, provides a template for efficient enzyme design, and a training set for inserting novel functionality into pre-existing systems.

Finally, we should comment about the connection between phosphate hydrolysis and the development of drugs for various diseases. As established above, phosphate hydrolysis is involved in all cellular functions (and thus also in their malfunctions), and this applies to signaling, replication, transcription and other processes. The problem of course is to translate molecular understanding to selective drugs. Blocking the active site of defective G-proteins and DNA polymerases may not be so specific, and should require exploitation of the unique features of the specific active site. Thus, it may be useful to interfere with phosphate hydrolysis by exploiting allosteric factors that change the active site preorganization (e.g. by interfering with protein–protein interactions), but this would also require and understanding of the TS stabilization by the given active site.

6. Acknowledgements

This work was supported by grants from the NIH (GM24492) NSF (MCB-0342276) and NCI (1U19CA105010) to AW, and the Swedish Research Council (VR, 2010-5026) and Carl Tryggers Foundation for Scientific Research (CTS 11:226) to SCLK. The authors would like to also thank Shayanti Mukherjee for helpful comments on the manuscript.

7. References

Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. (1994). Structure at 2·8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature* **370**, 621–628.

Adachi, K., Oiwa, K., Nishizaka, T., Furuike, S., Noji, H., Itoh, H., Yoshida, M. & Kinosita, K. (2007). Coupling of rotation and catalysis in F_1-ATPase

revealed by single-molecule imaging and manipulation. *Cell* **130**, 309–321.

Adamczyk, A. J., Cao, J., Kamerlin, S. C. L. & Warshel, A. (2011). Catalysis by dihydrofolate reductase and other enzymes arises from electrostatic preorganization, not conformational motions. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 14115–14120.

Adamczyk, A. J. & Warshel, A. (2011). Converting structural information into an allosteric-energy-based picture for elongation factor Tu activation by the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9827–9832.

Adams, J. A. (2001). Kinetic and catalytic mechanisms of protein kinases. *Chemical Reviews* **101**, 2271–2290.

Admiraal, S. J. & Herschlag, D. (1995). Mapping of the transition state for ATP hydrolysis: implications for enzymatic catalysis. *Chemistry & Biology* **2**, 729.

Admiraal, S. J. & Herschlag, D. (2000). The substrate-assisted general base catalysis model for phosphate monoester hydrolysis: evaluation using reactivity comparisons. *Journal of the American Chemical Society* **122**, 2145–2148.

Agirrezabala, X. & Frank, J. (2009). Elongation in translation as a dynamic interaction among the ribosome, tRNA, and elongation factors EF-G and EF-Tu. *Quarterly Reviews of Biophysics* **42**, 159–200.

Aguilar-Pérez, F., Gómez-Tagle, P., Collado-Fregoso, E. & Yatsimirsky, A. K. (2006). Phosphate ester hydrolysis by hydroxo complexes of trivalent lanthanides stabilized by 4-imidazolecarboxylate. *Inorganic Chemistry* **45**, 9502–9517.

Ahmadian, M. R., Stege, P., Scheffzek, K. & Wittinghofer, A. (1997). Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nature Structural Biology* **4**, 686–689.

Ahn, J. W., Kraynov, V. S., Zhong, X. J., Werneburg, B. G. & Tsai, M. D. (1998). DNA polymerase β: effects of gapped DNA substrates on dNTP specificity, fidelity, processivity and conformational changes. *The Biochemistry Journal* **331**, 79–87.

Akola, J. & Jones, R. O. (2003). ATP hydrolysis in water – A density functional study. *Journal of Physical Chemistry B* **107**, 11774–11783.

Akola, J. & Jones, R. O. (2006). Density functional calculations of ATP systems. 2. ATP hydrolysis at the active site of actin. *Journal of Physical Chemistry B* **110**, 8121–2129.

Alberts, I. L., Wang, Y. & Schlick, T. (2007). DNA polymerase β catalysis: are different mechanisms possible? *Journal of the American Chemical Society* **129**, 11100–11110.

Alcolombri, U., Elias, M. & Tawfik, D. S. (2011). Directed evolution of sulfotransferases and paraoxonases by ancestral libraries. *Journal of Molecular Biology* **411**, 837–853.

Alhambra, C., Wu, L., Zhang, Z.-Y. & Gao, J. (1998). Walden-inversion-enforced transition-state stabilization in a protein tyrosine phosphatase. *Journal of the American Chemical Society* **120**, 3858–3866.

Alkherraz, A., Kamerlin, S. C. L., Feng, G., Sheikh, Q. I., Warshel, A. & Williams, N. H. (2010). Phosphate ester analogues as probes for understanding enzyme catalysed phosphoryl transfer. *Faraday Discussions* **145**, 281–299.

Allin, C., Ahmadian, M. R., Wittinghofer, A. & Gerwert, K. (2001). Monitoring the GAP catalyzed H-Ras GTPase reaction at atomic resolution in real time. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7754–7759.

Allin, C. & Gerwert, K. (2001). Ras catalyzes GTP hydrolysis by shifting negative charges from γ- to β-phosphate as revealed by time-resolved FTIR difference spectroscopy. *Biochemistry* **40**, 3037–3046.

Anderson, V. E., Cassano, A. G. & Harris, M. E. (2006). *Isotope Effects in Chemistry and Biology*. Boca Raton, FL: Taylor and Francis.

Åqvist, J. & Feierberg, I. (2002). The catalytic power of ketosteroid isomerase investigated by computer simulation. *Biochemistry* **41**, 15728–15735.

Åqvist, J., Kolmodin, K., Florián, J. & Warshel, A. (1999). Mechanistic alternatives in phophsate monoester hydrolysis: what conclusions can be drawn from available experimental data? *Chemistry & Biology* **6**, R71–R80.

Åqvist, J. & Warshel, A. (1989). Calculations of free energy profiles for the staphylococcal nuclease catalyzed reaction. *Biochemistry* **28**, 4680–4689.

Åqvist, J. & Warshel, A. (1990). Free energy relationships in metalloenzyme-catalyzed reactions. Calculations of the effects of metal ion substitutions in staphylococcal nuclease. *Journal of the American Chemical Society.* **112**, 2860–2868.

Åqvist, J. & Warshel, A. (1993). Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chemical Reviews* **93**, 2523–2544.

Arantes, G. M. & Chaimovich, H. (2005). Thiolysis and alcoholysis of phosphate tri- and monoesters with alkyl and aryl leaving groups. An *ab initio* study in the gas phase. *Journal of Physical Chemistry A* **109**, 5625–5635.

Arndt, J. W., Gong, W. M., Zhong, X. J., Showalter, A. K., Liu, J., Dunlap, C. A., Lin, Z., Paxson, C., Tsai, M. D. & Chan, M. K. (2001). Insight into the catalytic mechanism of DNA polymerase β: structures of intermediate complexes. *Biochemistry* **40**, 5368–5375.

Arora, K., Beard, W. A., Wilson, S. H. & Schlick, T. (2005). Mismatch-induced conformational distortions in polymerase support an induced-fit mechanism for fidelity. *Biochemistry* **44**, 13328–13341.

Arora, K. & Schlick, T. (2004). In silico evidence for DNA polymerase-β's substrate-induced conformational change. *Biophysical Journal* **87**, 3088–3099.

Atkins, P. W. (1998). *Physical Chemistry*, 6 edn, Oxford: Oxford University Press.

Averbuch-Pouchot, M. T. & Durif, A. (1996). *Topics in Phosphate Chemistry*, Singapore: World Scientific Publishing Co. Pte. Ltd.

Ba-Saif, S. A., Davis, A. M. & Williams, A. (1989a). Effective charge distribution for attack of phenoxide ion on aryl methyl phosphate monoanion: studies related to the action of ribonuclease. *The Journal of Organic Chemistry* **54**, 5483–5486.

Ba-Saif, S. A., Davis, A. M. & Williams, A. J. (1989b). Effective charge distrtibution for attack of phenoxide ion on aryl methyl phosphate monoanion: studies related to the action of ribonuclease. *The Journal of Organic Chemistry* **54**, 5483–5486.

Ba-Saif, S. A., Waring, M. A. & Williams, A. (1990). Single transition states in the transfer of a neutral phosphoryl group between phenoxide ion nucleophiles in aqueous solution. *Journal of the American Chemical Society*. **112**, 8115–8120.

Ba-Saif, S. A., Waring, M. A. & Williams, A. (1991). Dependence of transition-state structure on nucleophile in the reaction of aryl oxide anions with aryl diphenylphosphate esters. *Journal of the American Chemical Society Perkin Transactions* **2**, 1653–1659.

Babtie, A. C., Bandyopadhyay, S., Olguin, L. F. & Hollfelder, F. (2009). Efficient catalytic promiscuity for chemically distinct reactions. *Angewandte Chemie* (*International ed. in English*) **48**, 3692–3694.

Bakhtina, M., Lee, S., Wang, Y., Dunlap, C., Lamarche, B. & Tsai, M.-D. (2005). Use of viscogens, dNTPαS, and Rhodium(III) as probes in stopped-flow experiments to obtain new evidence for the mechanism of catalysis by DNA polymerase β. *Biochemistry* **44**, 5177–5187.

Bao, Z. Q., Jacobsen, D. M. & Young, M. A. (2011). Briefly bound to active: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. *Structure* **19**, 675–690.

Barbacid, M. (1987). Ras Genes. *Annual Review of Biochemistry* **56**, 779–827.

Barbany, M., Gutierrez-De-teran, H., Sanz, F., Villà-Freixa, J. & Warshel, A. (2003). On the generation of catalytic antibodies by transition state analogues. *ChemBioChem: A European Journal of Chemical Biology* **4**, 277–285.

Barnard, P. W. C., Bunton, C. A., Llewellyn, D. R., Vernon, C. A. & Welch, V. A. (1961). The reactions of organic phosphates. Part V. The hydrolysis of triphenyl and trimethyl phosphates. *Journal of the Chemical Society* 2670–2676.

Barnes, J. A., Wilkie, J. & Williams, I. H. (1994). Transition-state structural variation and mechanistic change. *Journal of the Chemical Society Faraday Transactions* **90**, 1709–1714.

Baxter, N. J., Olguin, L. F., Golicnick, M., Feng, G., Hounslow, A. M., Bermel, W., Blackburm, G. M., Hollfelder, F., Waltho, J. P. & Williams, N. H. (2006). A Trojan horse transition state analogue generated by MgF3- formation in an enzyme active site. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 14732–14737.

Beard, W. A. & Wilson, S. H. (2003). Structural insights into the origins of DNA polymerase fidelity. *Structure* **11**, 489–496.

Beatty, R. (2001). *The Elements: Phosphorus*. New York: Marshall Cavendish Corporation.

Beke-Somfai, T., Feng, B. & Nordén, B. (2012). Energy phase shift as mechanism for catalysis. *Chemical Physics Letters* **535**, 169–172.

Beke-Somfai, T., Lincoln, P. & Norden, B. (2011). Double-lock ratchet mechanism revealing the role of αSER-344 in F_0F_1 ATP synthase. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4828–4833.

Benitez, B. A. S., Arora, K. & Schlick, T. (2006). *In silico* studies of the African swine fever virus DNA polymerase X support an induced-fit mechanism. *Biophysical Journal* **90**, 42–56.

Benkovic, S. J. & Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. *Science* **301**, 1196–1202.

Bennett, C. H. (1977). Molecular dynamics and transition state theory: the simulation of infrequent events. In *Alogrithms for Chemical Computations, Americal Chemical Society*, vol. 46 (Ed. R. Christofferson), pp. 63–97. Washington, D.C.

Benson, S. W. (1965). Bond energies. *Journal of Chemical Education* **42**, 502–518.

Berg, J. M., Tymoczko, J. L. & Stryer, L. (2010). *Biochemistry*, 7th edn. New York: W. H. Freeman and Co.

Bilgin, N. & Ehrenberg, M. (1994). Mutations in 23-S ribosomal-RNA perturb transfer-RNA selection and can lead to streptomycin dependence. *Journal of Molecular Biology* **235**, 813–824.

Blackburn, G. M. & Williams, N. H. (2003). Comment on 'The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction'. *Science* **301**, 1184.

Blackburn, P. & Moore, S. (1982). Pancreatic ribonuclease. In *The Enzymes*, vol. 15 (Ed. P. D. Boyer), pp. 317–433. New York: Academic Press.

Blaskovich, M. A. (2009). Drug discovery and protein tyrosine phosphatases. *Current Medicinal Chemistry* **16**, 2095–2176.

Bobyr, E., Lassila, J. K., Wiersma-Koch, H. I., Fenn, T. D., Lee, J. J., Nikolic-Hughes, I., Hodgson, K. O., Rees, D. C., Hedman, B. & Herschlag, D. (2012). High-resolution analysis of Zn(2+) coordination in the

alkaline phosphatase superfamily by EXAFS and X-ray crystallography. *Journal of Molecular Biology* **415**, 102–117.

Bohm, A., Gaudet, R. & Sigler, P. B. (1997). Structural aspects of heterotrimeric G-protein signaling. *Current Opinion in Biotechnology* **8**, 480–487.

Bojin, M. D. & Schlick, T. (2007). A Quantum mechanical investigation of possible mechanisms for the nucleotidyl transfer reaction catalyzed by DNA polymerase β. *Journal of Physical Chemistry B* **111**, 11244–11252.

Bondar, A.-N., Fischer, S., Smith, J. C., Elstner, M. & Suhai, S. (2004a). Key role of electrostatic interactions in bacteriorhodopsin proton transfer. *Journal of the American Chemical Society* **126**, 14668–14677.

Bondar, A. N., Elstner, M., Suhai, S., Smith, J. C. & Fischer, S. (2004b). Mechanism of primary proton transfer in bacteriorhodopsin. *Structure* **12**, 1281–1288.

Borowiec, J. A., Dean, F. B., Bullock, P. A. & Hurwitz, J. (1990). Binding and unwinding—how T antigen engages the SV40 origin of DNA replication. *Cell* **60**, 181–184.

Bos, J. (1989). Ras oncogenes in human cancer: a review. *Cancer Research* **49**, 4682–4689.

Bos, J. L., Rehmann, H. & Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865–877.

Bose-Basu, B., Derose, E. F., Kirby, T. W., Mueller, G. A., Beard, W. A., Wilson, S. H. & London, R. E. (2004). Dynamic characterization of a DNA repair enzyme: NMR studies of [methyl-C-13]methionine-labeled DNA polymerase β. *Biochemistry* **43**, 8911–8922.

Bourne, H. R., Sanders, D. A. & Mccormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117–127.

Bourne, N. & Williams, A. (1984). Effective charge on oxygen in phosphoryl (-PO_3^{2-}) group transfer from an oxygen donor. *The Journal of Organic Chemistry* **49**, 1200–1204.

Bowtell, D., Fu, P., Simon, M. & Senior, P. (1992). Identification of murine homologues of the drosophila son of sevenless gene. Potential activators of ras. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 6511–6515.

Boyer, P. D. (1993). The binding change mechanism for ATP synthase – some probabilities and possibilities. *Biochimica et Biophysica Acta* **1140**, 215–250.

Boyer, P. D. (1997). The ATP synthase – a splendid molecular machine. *Annual Review of Biochemistry* **66**, 717–749.

Branduardi, D., De Vivo, M., Rega, N., Barone, V. & Cavalli, A. (2011). Methyl phosphate dianion hydrolysis in solution characterized by path collective variables coupled with DFT-based enhanced sampling simulations. *Journal of Chemical Theory and Computation* **7**, 539–543.

Braun-Sand, S., Sharma, P. K., Tchu, Z., Pisliakov, A. V. & Warshel, A. (2008). The energetics of the primary proton transfer in bacteriorhodopsin revisited: it is a sequential light-induced charge separation after all. *Biochimica et Biophysica Acta* **1777**, 441–452.

Braun-Sand, S., Strajbl, M. & Warshel, A. (2004). Studies of proton translocations in biological systems: simulating proton transport in carbonic anhydrase by EVB based models. *Biophysical Journal* **87**, 2221–2239.

Broek, D., Toda, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Powers, S. & Wigler, M. (1987). The *S. cerevisiae* CDC25 gene product regulates the RAS/adenylate cyclase pathway. *Cell* **48**, 789–799.

Bruice, T. C., Tsubouchi, A., Dempcy, R. O. & Olson, L. P. (1996). One- and two-metal ion catalysis of the hydrolysis of adenosine 3′-alkyl phosphate esters. Models for one- and two-metal ion catalysis of RNA hydrolysis. *Journal of the American Chemical Society* **118**, 9867–9875.

Buchwald, S. L., Friedman, J. M. & Knowles, J. R. (1984). Stereochemistry of nucleophilic displacement on two phosphoric monoesters and a phosphoguanidine: the role of metaphosphate. *Journal of the American Chemical Society* **106**, 4911–4916.

Bunton, C. A., Llewwellyn, D. R., Oldham, K. G. & Vernon, C. A. (1958). The reaction of organic phosphates. Part I. The hydrolysis of methyl dihydrogen phosphate. *Journal of the Chemical Society* 3574–3587.

Bunton, C. A., Mhala, M. M., Oldham, K. G. & Vernon, C. A. (1960). The reactions of organic phosphates. Part II. The hydrolysis of dimethyl phosphate. *Journal of the Chemical Society* 3293–3301.

Butcher, W. W. & Westheimer, F. (1955). The lanthanum hydroxide gel promoted hydrolysis of phosphate esters. *Journal of the American Chemical Society* **77**, 2420–2424.

Cardone, M. H., Roy, N., Stennicke, H. R., Salvesen, G. S., Franke, T. F., Stanbridge, E., Frisch, S. & Reed, J. C. (1998). Regulation of cell death protease caspase-9 by phosphorylation. *Science* **282**, 1318–1321.

Carpenter, B. K. (2005). Nonstatistical dynamics in thermal reactions of polyatomic molecules. *Annual Review of Physical Chemistry* **56**, 57–89.

Cassano, A. G., Anderson, V. E. & Harris, M. E. (2002). Evidence for direct attack by hydroxide in phosphodiester hydrolysis. *Journal of the American Chemical Society* **124**, 10964–10965.

Catrina, I., O'Brien, P. J., Purcell, J., Nikolic-Hughes, I., Zalatan, J. G., Hengge, A. C. & Herschlag, D. (2007). Probing the origin of the compromised catalysis of *E. coli* alkaline phosphatase in its promiscuous sulfatase reaction. *Journal of the American Chemical Society* **129**, 5760–5765.

Cavalli, A. & Carloni, P. (2002). Enzymatic GTP Hydrolysis: insights from an *ab initio* molecular dynamics study. *Journal of the American Chemical Society* **124**, 3763–3768.

Cavalli, A., De Vivo, M. & Recanatini, M. (2003). Density functional study of the enzymatic reaction catalyzed by a cyclin-dependent kinase. *Chemical Communications* (*Cambridge, England*) **7**, 1308–1309.

Chakrabarti, P. P., Daumke, O., Suveyzdis, Y., Kotting, C., Gerwert, K. & Wittinghofer, A. (2007). Insight into catalysis of a unique GTPase reaction by a combined biochemical and FTIR approach. *Journal of Molecular Biology* **367**, 983.

Chen, W., Wilborn, M. & Rudolph, J. (2000). Dual-specific Cdc25B phosphatase: in search of the catalytic acid. *Biochemistry* **39**, 10781–10789.

Cheng, H., Sukal, S., Callender, R. & Leyh, T. S. (2001). γ-phosphate protonation and pH-dependent infolding of the Ras·GTP·Mg^{2+} complex A vibrational spectroscopy study. *Journal of Molecular Biology* **276**, 9931–9935.

Cherepanov, D. A., Mulkidjanian, A. Y. & Junge, W. (1999). Transient accumulation of elastic energy in proton translocating ATP synthase. *FEBS Letters* **449**, 1–6.

Chin, J., Banaszcyk, M., Jubian, V. & Zou, X. (1989a). Cobalt(III) complex-promoted hydrolysis of phosphate diesters: comparison in reactivity of rigid *cis*-diaquo(tetraaza)cobalt(III) complexes. *Journal of the American Chemical Society* **111**, 186–190.

Chin, J., Banaszczyk, M., Jubian, V. & Zou, X. (1989b). Cobalt(III) complex-promoted hydrolysis of phosphate diesters: comparison in reactivity of rigid cis-diaquo(tetraaza)cobalt(III) complexes. *Journal of the American Chemical Society* **111**, 186–190.

Chung, H.-H., Benson, D. R., Cornish, V. W. & Schultz, P. G. (1993a). Probing the role of loop 2 in Ras function with unnatural amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 10145–10149.

Chung, H.-H., Benson, D. R. & Schultz, P. G. (1993b). Probing the structure and mechanism of Ras protein with an expanded genetic code. *Science* **259**, 806–809.

Cleland, W. W. & Hengge, A. C. (2006). Enzymatic mechanisms of phosphate and sulfate transfer. *Chemical Reviews* **106**, 3252–3278.

Cohen, P. (2002). Protein kinases – the major drug targets of the twenty-first century? *Nature Reviews. Drug Discovery* **1**, 309–315.

Cohen, S., Carpenter, G. & King, L. J. (1980). Epidermal growth factor-receptor-protein kinase interactions. Co-purification of receptor and epidermal growth factor-enhanced phopshorylation activity. *The Journal of Biological Chemistry* **255**, 4834–4842.

Coleman, D. E. & Sprang, S. R. (1999). Reaction dynamics of G-protein catalyzed hydrolysis of GTP as viewed by X-ray crystallographic snapshots of $G_{i\alpha1}$. *Methods in Enzymology* **308**, 70–92.

Cool, R. H. & Parmeggiani, A. (1991). Substitution of histidine-84 and the GTPase mechanism of elongation factor-Tu. *Biochemistry* **30**, 362–366.

Cottrel, T. L. (1958). *The Strengths of Chemical Bonds*, 2nd edn. London: Butterworths.

Czub, J. & Grubmueller, H. (2011). Torsional elasticity and energetics of F_1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7408–7413.

Dall'acqua, W. & Carter, P. (2000). Substrate-assisted catalysis: molecular basis and biological significance. *Protein Science: A Publication of the Protein Society* **9**, 1–9.

Davis, A. M., Hall, A. D. & Williams, A. (1988). Charge description of base-catalyzed alcoholysis of Aryl phosphodiesters – a ribonuclease model. *Journal of the American Chemical Society* **110**, 5105–5108.

Daviter, T., Wieden, H. J. & Rodnina, M. V. (2003). Essential role of histidine 84 in elongation factor Tu for the chemical step of GTP hydrolysis on the ribosome. *Journal of Molecular Biology* **332**, 689–699.

De Grauuw, M., Hensbergen, P. & Van Der Water, B. (2006). Phospho-proteomic analysis of cellular signaling. *Electrophoresis* **27**, 2676–2686.

De Vivo, M., Cavalli, A., Carloni, P. & Recanatini, M. (2007). Computationl study of the phosphoryl transfer catalyzed by a cyclin-dependent kinase. *Chemistry* **13**, 8437–8444.

Dejaegere, A. & Karplus, M. (1993). Hydrolysis rate difference between cyclic and acyclic phosphate esters: solvation *versus* strain. *Journal of the American Chemical Society* **115**, 5316–5317.

Dejaegere, A., Liang, X. & Karplus, M. (1994). Phosphate ester hydrolysis: calculation of gas-phase reaction paths and solvation effects. *Journal of the Chemical Society Faraday Transactions* **90**, 1763–1770.

Di Sabato, G. & Jencks, W. P. (1961). Mechanism and catalysis of reactions of acyl phosphates. II. Hydrolysis. *Journal of the American Chemical Society* **83**, 4400–4405.

Dittrich, M., Hayashi, S. & Schulten, K. (2003). On the mechanism of ATP hydrolysis in F_1-ATPase. *Biophysical Journal* **85**, 2253–2266.

Downward, J., Parker, P. & Waterfield, M. D. (1984). Autophosphorylation sites on the epidermal growth factor receptor. *Nature* **311**, 483–485.

Drake, J. M., Graham, N. A., Stoyanova, T., Sedghi, A., Goldstein, A. S., Cai, H., Smith, D. A., Zhang, H., Komisopoulou, E., Huang, J., Graeber, T. G. & Witte, O. N. (2012). Oncogene-specific activation of tyrosine kinase networks during prostate cancer progression. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 1643–1648.

Du, X. L., Black, G. E., Lecchi, P., Abramson, F. P. & Sprang, S. R. (2004). Kinetic isotope effects in Ras-catalyzed GTP hydrolysis: evidence for a loose transition state. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8858–8863.

Du, X. L., Frei, H. & Kim, S. H. (2000). The mechanism of GTP hydrolysis by Ras probed by Fourier transform infrared spectroscopy. *The Journal of Biological Chemistry* **275**, 8492–8500.

Dukanovic, J. & Rapaport, D. (2011). Multiple pathways in the integration of proteins into the mitochondrial outer membrane. *Biochimica et Biophysica Acta – Biomembranes* **1808**, 971–980.

Echols, H. & Goodman, M. F. (1991). Fidelity mechanisms in DNA replication. *Annual Review of Biochemistry* **60**, 477–511.

Edwards, D. R., Lohman, D. C. & Wolfenden, R. (2012). Catalytic proficiency: the extreme case of S–O cleaving sulfatases. *Journal of the American Chemical Society* **134**, 525–531.

Ehrenberg, M. (2009). Nobel prize in chemistry 2009 – scientific background, http://nobelprize.Org/nobel_prizes/chemistry/laureates/2009/sci.html

Emsley, J. (2000). *The shocking history of phosphorus: a biography of the devil's element*, pp. 133–158. London, England: Macmillan publishers Ltd.

Evans, B., Tishmack, P. A., Pokalsky, C., Zhang, M. & Van Etten, R. L. (1996). Site-directed mutagenesis, kinetic and spectroscopic studies of the P-loop residues in a low molecular weight protein tyrosine phosphatase. *Biochemistry* **35**, 13609–13617.

Fanning, E. (1992). Simian virus 40 large T antigen: the puzzle, the pieces, and the emerging picture. *Journal of Virology* **66**, 1289–1293.

Fardilha, M., Esteves, S. L., Korrodi-Gregório, L., Da Cruz E Silva, O. A. & Da Cruz E Silva, F. F. (2010). The physiological relevance of protein phosphatase 1 and its interacting proteins to health and disease. *Current Medicinal Chemistry* **17**, 3996–4017.

Feig, M., Zacharias, M. & Pettitt, B. M. (2001). Conformations of an adenine bulge in a DNA octamer and its influence on DNA structure from molecular dynamics simulations. *Biophysical Journal* **81**, 352–370.

Fekry, M. I., Tipton, P. A. & Gates, K. S. (2011). Kinetic consequences of replacing the internucleotide phosphorus atoms in DNA with arsenic. *ACS Chemistry & Biology* **6**, 127–130.

Feng, G., Tanifum, E. A., Adams, H., Hengge, A. C. & Williams, N. H. (2009). Mechanism and transition state structure of aryl methylphosphonate esters doubly coordinated to a dinuclear cobalt(III) center. *Journal of the American Chemical Society* **131**, 12771–12779.

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

Flaks, J. G., Erwin, M. J. & Buchanan, J. M. (1957). Biosynthesis of the purines: XVI. The synthesis of adenosine 5′-phosphate and 5-amino-4 imidazolecarboxamide ribotide by a nucleotide pyrophosphorylase. *The Journal of Biological Chemistry* **228**, 201–213.

Flichtinski, D., Sharabi, O., Ruppel, A., Vetter, I. R., Herrmann, C. & Shifman, J. M. (2010). What makes Ras an efficient molecular switch: a computational, biophysical and structural study of Ras-GRP interactions with mutants of Raf. *Journal of Molecular Biology* **399**, 422–435.

Florián, J., Åqvist, J. & Warshel, A. (1998). On the reactivity of phosphate monoester dianions in aqueous solution: Brønsted linear free-energy relationships do not have an unique mechanistic interpretation. *Journal of the American Chemical Society* **120**, 11524–11525.

Florián, J., Goodman, M. F. & Warshel, A. (2002). Theoretical investigation of the binding free energies and key substrate-recognition components of the replication fidelity of human DNA polymerase β. *Journal of Physical Chemistry B* **106**, 5739–5753.

Florián, J., Goodman, M. F. & Warshel, A. (2003a). Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *Journal of the American Chemical Society* **125**, 8163–8177.

Florián, J., Goodman, M. F. & Warshel, A. (2003b). Computer simulation studies of the fidelity of DNA polymerases. *Biopolymers* **68**, 286–299.

Florián, J., Goodman, M. F. & Warshel, A. (2005). Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6819–6824.

Florián, J. & Warshel, A. (1997). A fundamental assumption about OH- attack in phosphate hydrolysis is not fully justified. *Journal of the American Chemical Society* **119**, 5473–5474.

Florián, J. & Warshel, A. (1998). Phosphate ester hydrolysis in aqueous solution: associative *versus* dissociative mechanisms. *Journal of Physical Chemistry B* **102**, 719–734.

Florián, J. & Warshel, A. (1999). Calculation of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the Langevin dipole solvation model. *Journal of Physical Chemistry B* **103**, 10282–10288.

Fothergill, M., Goodman, M. F., Petruska, J. & Warshel, A. (1995). Structure-energy analysis of the role of metal ions in phosphodiester bond hydrolysis by DNA polymerase I. *Journal of the American Chemical Society* **117**, 11619–11627.

Friedman, J. M., Freeman, S. & Knowles, J. R. (1988). The quest for free metaphosphate in solution: racemization at phosphorus in the transfer of the phospho group from aryl phosphate monoesters to tert-butyl alcohol in acetonitrile or in tert-butyl alcohol. *Journal of the American Chemical Society* **110**, 1268–1275.

Frushicheva, M. P., Cao, J. & Warshel, A. (2011). Challenges and advances in validating enzyme design proposals: the case of Kemp eliminase catalysis. *Biochemistry* **50**, 3849–3858.

Furuike, S., Hossain, M. D., Maki, Y., Adachi, K., Suzuki, T., Kohori, A., Itoh, H., Yoshida, M. & Kinosita J. R. K. (2008). Axle-less F_1-ATPase rotates in the correct direction. *Science* **319**, 955–958.

Futatsugi, N., Hata, M., Hoshino, T. & Tsuda, M. (1999). *Ab initio* study of the role of lysine 16 for the molecular switching mechanism of Ras protein p21. *Biophysical Journal* **77**, 3287–3292.

Galperin, M. Y., Bairoch, A. & Koonin, E. V. (1998). A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases. *Protein Science: A Publication of the Protein Society* **7**, 1829–1835.

Galperin, M. Y. & Jedrzejas, M. J. (2001). Conserved core structure and active site residues in alkaline phosphatase superfamily enzymes. *Proteins: Structure Function and Bioinformatics* **45**, 318–324.

Gani, D. & Wilkie, J. (1995). Stereochemical, mechanistic, and structural features of enzyme-catalysed phosphate monoester hydrolyses. *Chemical Society Reviews* **24**, 55–63.

Gao, Y. Q., Yang, W. & Karplus, M. (2005). A structure-based model for the synthesis and hydrolysis of ATP by F_1-ATPase. *Cell* **123**, 195–205.

Gerlt, J. A. & Gassman, P. G. (1993). Understanding the rates of certain enzyme-catalyzed reactions: proton abstraction from carbon acids, Acyl-transfer reactions, and displacement reactions of phosphodiesters. *Biochemistry* **32**, 11943.

Geyer, M., Herrmann, C., Wohlgemuth, S., Wittinghofer, A. & Kalbitzer, H. R. (1997). Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. *Nature Structural Biology* **4**, 694–699.

Gideon, P., John, J., Frech, M., Lautwein, A., Clark, R., Scheffler, J. E. & Wittinghofer, A. (1992). Mutational and kinetic analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal domain of GAP is not sufficient for full activity. *Molecular and Cellular Biology* **12**, 2050–2056.

Glaves, R., Mathias, G. & Marx, D. (2012). Mechanistic insights into the hydrolysis of a nucleoside triphosphate model in neutral and acidic solution. *Journal of the American Chemical Society* **134**, 6995–7000.

Glennon, T. M., Villà, J. & Warshel, A. (2000). How does GAP catalyze the GTPase reaction of Ras? A computer simulation study. *Biochemistry* **39**, 9641–9651.

Glennon, T. M. & Warshel, A. (1998). Energetics of the catalytic reaction of ribonuclease A: a computational study of alternative mechanisms. *Journal of the American Chemical Society* **120**, 10234–10247.

Glusker, J. R., Katz, A. K. & Bock, C. W. (1999). Metal ions in biological systems. *Rigaku Journal* **16**, 8–16.

Golicnick, M., Olguin, L. F., Feng, G., Baxter, N. J., Waltho, J. P., Williams, N. H. & Hollfelder, F. (2009). Kinetic analysis of β-phosphoglucomutase and its inhibition by magnesium fluoride. *Journal of the American Chemical Society* **131**, 1575–1588.

Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annual Review of Biochemistry* **71**, 17–50.

Goodman, M. F., Creighton, S., Bloom, L. B. & Petruska, J. (1993). Biochemical basis of DNA replication fidelity. *Critical Reviews in Biochemistry and Molecular Biology* **28**, 83–126.

Graham, D. L., Lowe, P. N., Grime, G. W., Marsh, M., Rittinger, K., Smerdon, S. J., Gamblin, S. J. & Eccleston, J. F. (2002). MgF_3^- as a transition state analog of phopshoryl transfer. *Chemistry & Biology* **9**, 375–381.

Graziano, M. P. & Gilman, A. G. (1989). Synthesis in *Escherichia coli* of GTPase-deficiect mutants of $G_{s\alpha}$. *The Journal of Biological Chemistry* **264**, 15475–15482.

Grigorenko, B. L., Nemukhin, A. V., Shadrina, M. S., Topol, I. A. & Burt, S. K. (2007a). Mechanisms of guanosine triphosphate hydrolysis by Ras and Ras-GAP proteins as rationalized by *ab initio* QM/MM simulations. *Proteins: Structure Function and Bioinformatics* **66**, 456–466.

Grigorenko, B. L., Nemukhin, A. V., Topol, I. A., Cachau, R. E. & Burt, S. K. (2005). QM/MM modeling the Ras-GAP catalyzed hydrolysis of guanosine triphosphate. *Proteins: Structure Function and Bioinformatics* **60**, 495–503.

Grigorenko, B. L., Rogov, A. V. & Nemukhin, A. V. (2006). Mechanism of triphosphate hydrolysis in aqueous solution: QM/MM simulations in water clusters. *Journal of Physical Chemistry B* **110**, 4407–4412.

Grigorenko, B. L., Rogov, A. V., Topol, I. A., Burt, S. K., Martinez, H. M. & Nemukhin, A. V. (2007b). Mechanism of the myosin catalyzed hydrolysis of ATP as rationalized by molecular modeling. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7057–7061.

Grigorenko, B. L., Shadrina, M. S., Topol, I. A., Collins, J. R. & Nemukhin, A. V. (2008). Mechanism of the chemical step for the guanosine triphosphate (GTP) hydrolysis catalyzed by elongation factor Tu. *Biochimica et Biophysica Acta – Proteins and Proteomics* **1784**, 1908–1917.

Grunwald, E. (1985). Reaction mechanism from structure-energy relations. 1. Base-catalyzed addition of alcohols to formaldehyde. *Journal of the American Chemical Society* **107**, 4710–4715.

Guan, K. L. & Dixon, J. E. (1991). Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. *The Journal of Biological Chemistry* **266**, 17026–17030.

Guerrier-Takeda, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857.

Guthrie, J. P. (1977). Hydration and dehydration of phosphoric acid derivatives: free energies of formation of the pentacoordinate intermediates of phosphate ester hydrolysis and of monomeric metaphosphate. *Journal of the American Chemical Society* **99**, 3991–4001.

Guthrie, J. P. (1996). Multidimensional Marcus theory: an analysis of concerted reactions. *Journal of the American Chemical Society* **118**, 12878–12885.

Hale, S. P., Poole, L. B. & Gerlt, J. A. (1993). Mechanism of the reaction catalyzed by staphylococcal nuclease; identification of the rate-determining step. *Biochemistry* **32**, 1479.

Hall, C. R. & Inch, T. D. (1980). Phosphorus stereochemistry: mechanistic implications of the observed stereochemistry of bond forming and breaking processes at phosphorus in some 5- and 6-membered ring phosphorus esters. *Tetrahedron* **36**, 2059–2095.

Hammes, G. G., Benkovic, S. J. & Hammes-Schiffer, S. (2011). Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. *Biochemistry* **50**, 10422–10430.

Hansson, T., Nordlund, P. & Åqvist, J. (1997). Energetics of nucleophile activation in a protein tyrosine phosphatase. *Journal of Molecular Biology* **265**, 118–127.

Hausner, T. P., Atmadja, J. & Nierhaus, K. H. (1987). Evidence that the G2661 region of 23S ribosomal-RNA is located at the ribosomal-binding sites of both elongation-factors. *Biochimie* **69**, 911–923.

Heesen, H. T., Gerwert, K. & Schlitter, J. (2007). Role of the arginine finger in Ras·RasGAP revealed by QM/MM calculations. *FEBS Letters* **581**, 5677–5684.

Henchman, M., Viggiano, A. A., Paulson, J. F., Freedman, A. & Wormhoudt, J. (1985). Thermodynamic and kinetic properties of the metaphosphate anion, PO^{3-} in the gas phase. *Journal of the American Chemical Society* **107**, 1453–1455.

Hengge, A. C. (1999). *Insights from Heavy-Atom Isotope Effects on Phosphoryl and Thiophosphoryl Transfer Reactions*. Amsterdam: IOS Press.

Hengge, A. C. (2002). Isotope effects in the study of phosphoryl and sulfuryl transfer reactions. *Accounts of Chemical Research* **35**, 105–112.

Hengge, A. C. & Cleland, W. W. (1991). Phosphoryl-transfer reactions of phosphodiesters: characterization of transition states by heavy-atom isotope effects. *Journal of the American Chemical Society* **113**, 5835–5841.

Hengge, A. C., Edens, W. A. & Elsing, H. (1994). Transition-state structures for phosphoryl-transfer reactions of p-nitrophenyl phosphate. *Journal of the American Chemical Society* **116**, 5045–5049.

Hengge, A. C., Tobin, A. E. & Cleland, W. W. (1995). Studies of transition-state structures in phosphoryl transfer reactions of phosphodiesters of p-nitrophenol. *Journal of the American Chemical Society* **117**, 5919–5926.

Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M. & Kern, D. (2007a). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916.

Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hubner, C. G. & Kern, D. (2007b). Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844.

Herrmann, C., Martin, G. A. & Wittinghofer, A. (1995). Quantitative analysis of the complex between p21ras and the Ras-binding domain of the human Raf-1 protein kinase. *The Journal of Biological Chemistry* **270**, 2901.

Herschlag, D. & Jencks, W. P. (1987). The effect of divalent metal ions on the rate and transition-state structure of phosphoryl-transfer reactions. *Journal of the American Chemical Society* **109**, 4665–4674.

Herschlag, D. & Jencks, W. P. (1989a). Evidence that metaphosphate monoanion is not an intermediate in solvolysis reactions in aqueous solution. *Journal of the American Chemical Society* **111**, 7579–7586.

Herschlag, D. & Jencks, W. P. (1989b). Phosphoryl transfer to anionic oxygen nucleophiles. Nature of the transition state and electrostatic repulsion. *Journal of the American Chemical Society* **111**, 7587–7596.

Hoff, R. H. & Hengge, A. C. (1998a). Entropy and enthalpy contributions to solvent effects on phosphate monoester solvolosys. The importance of entropy effects in the dissociative transition state. *The Journal of Organic Chemistry* **63**, 6680–6688.

Hoff, R. H. & Hengge, A. C. (1998b). Entropy and enthalpy contributions to solvent effects on phosphate monoester solvolysis. The importance of entropy effects in the dissociative transition state. *The Journal of Organic Chemistry* **63**, 6680–6688.

Hoff, R. H., Larsen, P. & Hengge, A. C. (2001). Isotope effects and medium effects on sulfuryl transfer reactions. *Journal of the American Chemical Society* **123**, 9338–9344.

Hollfelder, F. & Herschlag, D. (1995a). The nature of the transition state for enzyme-catalyzed phopshoryl transfer. Hydrolysis of O-arylphosphorothioates by alkaline phosphatase. *Biochemistry* **34**, 12255–12264.

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

Holtz, K. M., Cartrine, I. E., Hengge, A. C. & Kantrowitz, E. (2000). General acid-base catalysis of complex reactions in water. *Biochemistry* **39**, 9451–9458.

Hoofs Van Huijsduijnen, R., Wälchi, S., Ibberson, M. & Harrenga, A. (2002). Protein tyrosine phosphatases as drug targets: PTP1B and beyond. *Expert Opinion on Therapeutic Targets* **6**, 637–647.

Hou, G. & Cui, Q. (2012). QM/MM analysis suggests that alkaline phosphatase (AP) and nucleotide pyrophosphatase/phosphodiesterase slightly tighten transition state for phosphate diester hydrolysis relative to solution: implication for catalytic promiscuity in the AP

superfamily. *Journal of the American Chemical Society* **134**, 229–246.

Hu, C.-H. & Brinck, T. (1999). Theoretical studies of the hydrolysis of the methyl phosphate anion. *Journal of Physical Chemistry A* **103**, 5379–5386.

Hübscher, U., Maga, G. & Spadari, S. (2002). Eukaryotic DNA polymerases. *Annual Review of Biochemistry* **71**, 133–163.

Huheey, J. E., Keitler, E. A. & Keitler, R. L. (1993). *Inorganic Chemistry*, 4th edn. New York: Harper Collins.

Humphry, T., Forconi, M., Williams, N. H. & Hengge, A. C. (2004). Altered mechanisms of reactions of phosphate esters bridging a dinuclear metal center. *Journal of the American Chemical Society* **126**, 11864–11869.

Hunter, T. (1995). Protein kinases and phosphatases: the Yin and Yang of protein phosphorylation and signalling. *Cell* **80**, 225–236.

Hunter, T. (2000). Signaling–2000 and beyond. *Cell* **100**, 113–127.

Hunter, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Molecular Cell* **28**, 730–738.

Hunter, T. & Cooper, J. A. (1985). Protein-tyrosine kinases. *Annual Review of Biochemistry* **54**, 897–930.

Hwang, J.-K., King, G., Creighton, S. & Warshel, A. (1988). Simulation of free energy relationships and dynamics of S_N2 reactions in aqueous solution. *Journal of the American Chemical Society* **110**, 5297–5311.

Hwang, J.-K. & Warshel, A. (1993). A quantized classical path approach for calculations of quantum mechanical rate constants. *Journal of Physical Chemistry* **97**, 10053–10058.

Hwang, J.-K. & Warshel, A. (1996). How important are quantum mechanical nuclear motions in enzyme catalysis? *Journal of the American Chemical Society* **118**, 11745–11751.

Iche-Tarrat, N., Barthelat, J. C., Rinaldi, D. & Vigroux, A. (2005). Theoretical studies of the hydroxide-catalyzed P–O cleavage reactions of neutral phosphate triesters and diesters in aqueous solution: examination of the changes induced by H/Me substitution. *Journal of Physical Chemistry B* **109**, 22570–22580.

Iché-Tarrat, N., Ruiz-Lopez, M., Barthelat, J.-C. & Vigroux, A. (2007). Theoretical evaluation of the substrate-assisted catalysis mechanism for the hydrolysis of phosphate monoester dianions. *Chemistry–A European Journal* **13**, 3617–3629.

Inbal, B., Shani, G., Cohen, O., Kissil, J. & Kimchi, A. (2000). Death-associated protein kinase-related protein 1, a novel serine/threonine kinase involved in apoptosis. *Molecular and Cellular Biology* **20**, 1044–1054.

Jacquet, E. & Parmeggiani, A. (1988). Structure-function relationships in the GTP binding domain of EF-Tu – mutation of Val20, the residue homologous to position 12 in p21. *The EMBO Journal* **7**, 2861–2867.

Jankowski, S., Quin, L. D., Paneth, P. & O'Leary, M. H. (1994). Kinetic isotope effects on ethyl metaphosphate transfer from a phosphoramidate to ethanol. *Journal of the American Chemical Society* **116**, 11675–11677.

Jencks, W. P. (1969). *Catalysis in Chemistry and Enzymology*. New York: McGraw-Hill.

Jencks, W. P. (1972). General acid-base catalysis of complex reactions in water. *Chemical Reviews* **72**, 705–718.

Jencks, W. P. (1980). When is an intermediate not an intermediate? Enforced mechanisms of general acid-base, catalyzed, carbocation, carbanion, and ligand exchange reaction. *Accounts of Chemical Research* **13**, 161–169.

Jencks, W. P. (1985). A primer for the bema hapothle. An empirical approach to the characterization of changing transition-state structures. *Chemical Reviews* **85**, 511–527.

Jencks, W. P. (1987). *Catalysis in Chemistry and Enzymology*. New York: Dover.

Jensen, R. A. (1976). Enzyme recruitment in evolution of new function. *Annual Reviews of Microbiology* **30**, 409–425.

Johnson, K. A. (1993). Conformational coupling in DNA polymerase fidelity. *Annual Review of Biochemistry* **62**, 685–713.

Johnson, K. A. (2008). Role of induced fit in enzyme specificity: a molecular forward/reverse switch. *The Journal of Biological Chemistry* **283**, 26297–26301.

Johnson, T. O., Ermolieff, J. & Jirousek, M. R. (2002). Protein tyrosine phosphatase 1B inhibitors for diabetes. *Nature Reviews. Drug Discovery* **1**, 696–708.

Jonas, S. & Hollfelder, F. (2009). Mapping catalytic promiscuity in the alkaline phosphatase superfamily. *Pure and Applied Chemistry* **81**, 731–742.

Jonas, S., Van Loo, B., Hyvönen, M. & Hollfelder, F. (2008). A new member of the alkaline phosphatase superfamily with a formylglycine nucleophile: structural and kinetic characterisation of a phosphate monoester hydrolyase/phosphodiesterase from Rhizobium leguminosarum. *Journal of Molecular Biology* **384**, 120–136.

Joyce, C. M. & Benkovic, S. J. (2004). DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry* **43**, 14317–14324.

Kamerlin, S. C. L. (2011a). Theoretical comparison of p-nitrophenyl phosphate and sulfate hydrolysis in aqueous solution: implications for enzyme catalyzed sulfuryl transfer. *The Journal of Organic Chemistry* **76**, 9228–9238.

Kamerlin, S. C. L., Florián, J. & Warshel, A. (2008a). Associative *versus* dissociative mechanisms of phosphate monoester hydrolysis: on the interpretation of activation entropies. *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry* **9**, 1767–1773.

Kamerlin, S. C. L., Haranczyk, M. & Warshel, A. (2009a). Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models. *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry* **10**, 1125–1134.

Kamerlin, S. C. L., Haranczyk, M. & Warshel, A. (2009b). Progress in *ab initio* QM/MM free-energy simulations of electrostatic energies in proteins:

accelerated QM/MM studies of pKa, redox reactions, and solvation free energies. *Journal of Physical Chemistry B* **113**, 1253–1272.

Kamerlin, S. C. L., Mckenna, C. E., Goodman, M. F. & Warshel, A. (2009c). A computational study of the hydrolysis of dGTP analogues with halomethylene-modified leaving groups in solution: implications for the mechanism of DNA polymerases. *Biochemistry* **48**, 5963–5971.

Kamerlin, S. C. L., Rucker, R. & Boresch, S. (2007). A molecular dynamics study of WPD-loop flexibility in PTP1B. *Biochemical and Biophysical Research Communications* **356**, 1011–1016.

Kamerlin, S. C. L., Sharma, P. K., Chu, Z. T. & Warshel, A. (2010). Ketosteroid isomerase provides further support for the idea that enzymes work by electrostatic preorganization. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4075–4080.

Kamerlin, S. C. L., Vicatos, S. & Warshel, A. (2011). Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Annual Review of Physical Chemistry* **62**, 41–64.

Kamerlin, S. C. L. & Warshel, A. (2009). On the energetics of ATP hydrolysis in solution. *Journal of Physical Chemistry B* **113**, 15692–15698.

Kamerlin, S. C. L. & Warshel, A. (2010). At the dawn of the 21st century: is dynamics the missing link for understanding enzyme catalysis? *Proteins: Structure Function and Bioinformatics* **78**, 1339–1375.

Kamerlin, S. C. L. & Warshel, A. (2011a). The empirical valence bond model: theory and applications. *Computational Molecular Science* **1**, 30–45.

Kamerlin, S. C. L. & Warshel, A. (2011b). Multiscale modeling of biological problems. *Physical Chemistry Chemical Physics* **13**, 10401–10411.

Kamerlin, S. C. L. & Wilkie, J. (2007). The role of metal ions in phosphate ester hydrolysis. *Organic & Biomolecular Chemistry* **5**, 2098–2108.

Kamerlin, S. C. L. & Wilkie, J. (2011). The effect of leaving group on mechanistic preference in phosphate monoester hydrolysis. *Organic & Biomolecular Chemistry* **9**, 5394–5406.

Kamerlin, S. C. L., Williams, N. H. & Warshel, A. (2008b). Dineopentyl phosphate hydrolysis: evidence for stepwise water attack. *The Journal of Organic Chemistry* **73**, 6960–6969.

Khan, S. A. & Kirby, A. J. (1970). The reactivity of phosphate esters. Multiple structure–reactivity correlations for the reactions of triesters with nucleophiles. *Journal of the Chemical Society* ***B*** 1172–1182.

Khersonsky, O., Roodveldt, C. & Tawfik, D. S. (2006). Enzyme promiscuity: evolutionary and mechanistic aspects. *Current Opinion in Chemistry & Biology* **10**, 498–508.

Khersonsky, O. & Tawfik, D. S. (2010). Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annual Review of Biochemistry* **79**, 471–505.

Kim, E. E. & Wyckoff, H. W. (1991). Reaction mechanism of alkaline phosphatase based on crystal structures: two metal ion catalysis. *Journal of Molecular Biology* **218**, 449–464.

Kim, S. J., Beard, W. A., Harvey, J., Shock, D. D., Knutson, J. R. & Wilson, S. H. (2003). Rapid segmental and subdomain motions of DNA polymerase β. *The Journal of Biological Chemistry* **278**, 5072–5081.

Kirby, A. J. & Jencks, W. P. (1965a). The reactivity of nucleophilic reagents toward the p-nirophenyl phosphate dianions. *Journal of the American Chemical Society* **87**, 3209–3216.

Kirby, A. J. & Varvoglis, A. G. (1967a). The reactivity of phosphate esters. Monoester hydrolysis. *Journal of the American Chemical Society* **89**, 415–423.

Kirby, A. J. & Varvoglis, A. G. (1968a). The reactivity of phosphate esters: reactions of monoesters with nucleophiles. Nucleophilicity independent of basicity in a bimolecular substitution reaction. *Journal of the American Chemical Society B* 135–141.

Kirby, A. J. & Varvoglis, A. G. (1968b). The reactivity of phosphate esters: reactions of monoesters with nucleophiles. Nucleophilicity independent of basicity in a bimolecular substitution reaction. *Journal of the American Chemical Society B* 135–141.

Kirby, A. J. & Younas, M. (1970a). The reactivity of phosphate esters. Reactions of diesters with nucleophiles. *Journal of the American Chemical Society B* 510–513.

Kirby, A. J. & Younas, M. (1970b). The reactivity of phosphate esters. Reactions of diesters with nucleophiles. *Journal of the American Chemical Society B* 1165–1172.

Kirby, J. A. & Jencks, W. P. (1965b). The reactivity of nucleophillic reagents towards the p-nitrophenyl phosphate dianion. *Journal of the American Chemical Society* **87**, 3209–3216.

Kirby, J. A. & Varvoglis, A. G. (1967b). The reactivity of phosphate monoester hydrolysis. *Journal of the American Chemical Society* **89**, 415–423.

Kirby, J. A. & Younas, M. (1970c). The reactivity of phosphate esters. Diester hydrolysis. *Journal of the American Chemical Society B* 510–513.

Kirby, T. W., Derose, E. F., Beard, W. A., Wilson, S. H. & London, R. E. (2005). A thymine isostere in the templating position disrupts assembly of the closed DNA polymerase β ternary complex. *Biochemistry* **44**, 15230–15237.

Kirmizialtin, S., Nguyen, V., Johnson, K. A. & Elber, R. (2012). How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations. *Structure* **20**, 618–627.

Klähn, M., Rosta, E. & Warshel, A. (2006). On the mechanism of hydrolysis of phosphate monoesters dianions in solutions and proteins. *Journal of the American Chemical Society* **128**, 15310–15323.

KNIGHT, W. B., WEISS, P. M. & CLELAND, W. W. (1986). Determination of equilibrium ^{18}O isotope effects on the deprotonation of phosphate and phosphate esters and the anomeric effect on deprotonation of glucose 6-phosphate. *Journal of the American Chemical Society* **108**, 2579–2761.

KNUDSEN, C., WIEDEN, H. J. & RODNINA, M. V. (2001). The importance of structural transitions of the switch ii region for the functions of elongation factor Tu on the ribosome. *The Journal of Biological Chemistry* **276**, 22183–22190.

KOGA, N. & TAKADA, S. (2006). Folding-based molecular simulations reveal mechanisms of the rotary motor F_1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5367–5372.

KOLMODIN, K. & ÅQVIST, J. (1999). Computational modeling of the rate limiting step in low molecular weight protein tyrosine phosphatase. *FEBS Letters* **456**, 301–305.

KOLMODIN, K. & ÅQVIST, J. (2001). The catalytic mechanism of protein tyrosine phosphatases revisited. *FEBS Letters* **498**, 208–213.

KOLMODIN, K., NORDLUND, P. & AQVIST, J. (1999). Mechanism of substrate dephosphorylation in low M-r protein tyrosine phosphatase. *Proteins* **36**, 370–379.

KORNBERG, A. & BAKER, T. A. (1992). *DNA Replication*. New York: W. H. Freeman.

KORNBERG, A., LIEBERMAN, I. & SIMMS, E. S. (1955). Enzymatic synthesis and properties of 5-phosphoribosylpyrophosphatase. *The Journal of Biological Chemistry* **215**, 389–402.

KOSHLAND, D. E. (1995). The key-lock theory and the induced fit theory. *Angewandte Chemie* (*International ed. in English*) **33**, 2375–2378.

KOSLOFF, M., ZOR, T. & SELINGER, Z. (2000). Substrate-assisted catalysis: implications for biotechnology and drug design. *Drug Development & Research* **50**, 250–257.

KRAB, I. M. & PARMEGGIANI, A. (1999). Mutagenesis of three residues, isoleucine-60, threonine-61, and aspartic acid-80, implicated in the GTPase activity of *Escherichia coli* elongation factor Tu. *Biochemistry* **38**, 13035–13041.

KRAUT, D. A., SIGALA, P. A., PYBUS, B., LIU, C. W., RINGE, D., PETSKO, G. A. & HERSCHLAG, D. (2006). Testing electrostatic complementarity in enzyme catalysis: hydrogen bonding in the ketosteroid isomerase oxyanion hole. *PLoS Biology* **4**, 0501–0519.

KRENGEL, U., SCHLICHTING, L., SCHERER, A., SCHUMANN, R., FRECH, M., JOHN, J., KABSCH, W., PAI, E. F. & WITTINGHOFER, A. (1990). Three-dimensional structures of H-Ras P21 mutants – molecular basis for their inability to function as signal transduction switch molecules. *Cell* **62**, 539–548.

KRUGER, K., GRABOWSKI, P. J., ZAUG, A. J., SANDS, J., GOTTSCHLING, D. E. & CECH, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31**, 147–157.

KUMAMOTO, J., COX, J. R., JR. & WESTHEIMER, F. H. (1956). Barium ethylene phosphate. *Journal of the American Chemical Society* **78**, 4858–4860.

KUNKEL, T. A. & BEBENEK, K. (2000). DNA replication fidelity. *Annual Review of Biochemistry* **69**, 497–529.

KURZ, J. L. (1978). The relationship of barrier shape to 'linear' free energy slopes and curvatures. *Chemical Physics Letters* **57**, 243–246.

LAD, C., WILLIAMS, N. H. & WOLFENDEN, R. (2003a). The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5607–5610.

LAD, C., WILLIAMS, N. H. & WOLFENDEN, R. (2003b). The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5607–5610.

LAHIRI, S. D., ZHANG, G. F., DUNAWAY-MARIANO, D. & ALLEN, K. N. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science* **299**, 2067–2071.

LAMBRIGHT, D. G., NOEL, J. P., HAMM, H. E. & SIGLER, P. B. (1994). Structural determinants for activation of the α-subunit of a heterotrimeric G protein. *Nature* **369**, 621–628.

LANCASTER, L., LAMBERT, N. J., MAKLAN, E. J., HORAN, L. H. & NOLLER, H. F. (2008). The sarcin–ricin loop of 23S rRNA is essential for assembly of the functional core of the 50S ribosomal subunit. *RNA – A Publication of the RNA Society* **14**, 1999–2012.

LANG, T. M., MAITRA, M., STARCEVIC, D., LI, S. X. & SWEASY, J. (2004). A DNA polymerase β mutant from colon cancer cells induces mutations. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6074–6079.

LANGEN, R., SCHWEINS, T. & WARSHEL, A. (1992). On the mechanism of guanosine triphosphate hydrolysis in ras p^{21} proteins. *Biochemistry* **31**, 8691–8696.

LARSEN, A. K., OUARET, D., EL OUADRANI, K. & PETITPREZ, A. (2011). Targeting EGFR and VEGF(R) pathway cross-talk in tumor survival and angiogenesis. *Pharmacology & Therapeutics* **131**, 80–90.

LASSILA, J. K. & HERSCHLAG, D. (2008). Promiscuous sulfatase activity and thio-effects in a phosphodiesterase of the alkaline phosphatase superfamily. *Biochemsitry* **47**, 12853–12859.

LASSILA, J. K., ZALATAN, J. G. & HERSCHLAG, D. (2011). Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annual Review of Biochemistry* **80**, 669–702.

LECLERC, F. & KARPLUS, M. (2006). Two-metal-ion mechanism for hammerhead-ribozyme catalysis. *Journal of Physical Chemistry B* **110**, 3395–3409.

Lee, F. S., Chu, Z. T., Bolger, M. B. & Warshel, A. (1992). Calculations of antibody antigen interactions – microscopic and semimicroscopic evaluation of the free-energies of binding of phosphorylcholine analogs to McPC603. *Protein Engineering* **5**, 215–228.

Lee, F. S., Chu, Z. T. & Warshel, A. (1993). Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. *Journal of Computational Chemistry* **14**, 161–185.

Leffler, J. E., & Grunwald, E. (1963). *Rates and Equilibria of Organic Reactions*. New York: Wiley and Sons.

Li, G. & Zhang, X. C. (2004). GTP Hydrolysis Mechanism of Ras-like GTPases. *Journal of Molecular Biology* **340**, 921.

Li, G. H. & Cui, Q. (2004). Mechanochemical coupling in myosin: a theoretical analysis with molecular dynamics and combined QM/MM reaction path calculations. *Journal of Physical Chemistry B* **108**, 3342–3357.

Liao, X., Anjaneyulu, P. S. R., Curley, J. F., Hsu, M., Boehringer, M., Caruthers, M. H. & Piccirilli, J. A. (2001a). The tetrahymena ribozyme cleaves a 5′-methylene phosphonate monoester ~102-fold faster than a normal phosphate diester: implications for enzyme catalysis of phosphoryl transfer reactions. *Biochemistry* **40**, 10911–10926.

Liao, X., Anjaneyulu, P. S. R., Curley, J. F., Hsu, M., Boehringer, M., Caruthers, M. H. & Piccirilli, J. A. (2001b). The tetrahymena ribozyme cleaves a 5′-ethylene phosphonate monoester 102-fold faster than a normal phosphate diester: implications for enzyme catalysis of phosphoryl transfer reactions. *Biochemistry* **40**, 10911–10926.

Liljas, A. (2004). *Structural Aspects of Protein Synthesis*. Singapore: World Scientific Publishing Co.

Liljas, A., Ehrenberg, M. & Åqvist, J. (2011). Comment on 'the mechanism for activation of GTP hydrolysis on the ribosome'. *Science* **333**, 37.

Lima, F. S., Chaimovich, H. & Cuccovia, I. M. (2012). Kinetics and product distribution of p-nitrophenyl phosphate dianion solvolysis in ternary DMSO/alcohol/water mixtures are compatible with metaphosphate formation. *Journal of Physical Organic Chemistry* **25**, 9–13.

Lin, P., Batra, V. K., Pedersen, L. C., Beard, W. A., Wilson, S. H. & Pedersen, L. G. (2008). Incorrect nucleotide insertion at the active site of a G: a mismatch catalyzed by DNA polymerase β. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5670–5674.

Lin, P., Pedersen, L. C., Batra, V. K., Beard, W. A., Wilson, S. H. & Pedersen, L. G. (2006). Energy analysis of chemistry for correct insertion by DNA polymerase β. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13294–13299.

Liu, H., Shi, Y., Chen, X. S. & Warshel, A. (2009). Simulating the electrostatic guidance of the vectorial translocations in hexameric helicases and translocases. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7449–7454.

Liu, H. & Warshel, A. (2007). The catalytic effect of dihydrofolate reductase and its mutants is determined by reorganization energies. *Biochemistry* **46**, 6011–6025.

Liu, Y., Gregersen, B. A., Lopez, X. & York, D. M. (2006). Density functional study of the in-line mechanism of methanolysis of cyclic phosphate and thiophosphate esters in solution: insights into thio effects in RNA transesterification. *Journal of Physical Chemistry B* **109**, 19987–20003.

Lonsdale, R., Hoyle, S., Grey, T., Ridder, L. & Mulholland, A. J. (2012). Determinants of reactivity and selectivity in soluble epoxide hydrolase from quantum mechanics/molecular mechanics modeling. *Biochemistry* **51**, 1774–1786.

López-Canut, V., Roca, M., Bertrán, J., Moliner, V. & Tuno, I. (2010). Theoretical study of phosphodiester hydrolysis in nucleotide pyrophosphatase/phosphodiesterase. Environmental effects on the reaction mechanism. *Journal of the American Chemical Society* **132**, 6955–6963.

Lopéz-Canut, V., Roca, M., Bertrán, J., Moliner, V. & Tuñon, I. (2011). Promiscuity in alkaline phosphatase superfamily. Unraveling evolution through molecular simulations. *Journal of the American Chemical Society* **133**, 12050–12062.

Lopez, X., Dejaegere, A., Leclerc, F., York, D. M. & Karplus, M. (2006). Nucleophilic attack on phosphate diesters: a density functional study of in-line reactivity in dianionic, monoanionic, and neutral systems. *Journal of Physical Chemistry B* **110**, 11525–11539.

Lopez, X. & York, D. M. (2003). Parameterization of semiempirical methods to treat nucleophilic attack to biological phosphates: AM1/d parameters for phosphorus. *Theoretical Chemistry Accounts* **109**, 149–159.

Lowy, D. R. & Willumsen, B. M. (1993). Function and regulation of Ras. *Annual Review of Biochemistry* **62**, 851–891.

Lu, Q., Nassar, N. & Wang, J. (2011). A mechanism of catalyzed GTP hydrolysis by Ras protein through magnesium ion. *Chemical Physics Letters* **516**, 233–238.

Luo, J., Loo, B. V. & Kamerlin, S. C. L. (2012a). Examining the promiscuous phosphatase activity of *Pseudomonas aeruginosa* arylsulfatase: a comparison to analogous phosphatases. *Proteins: Structure Function and Bioinformatics* **80**, 1211–1226.

Luo, J., Van Loo, B. & Kamerlin, S. C. L. (2012b). Catalytic promiscuity in Pseudomonas aeruginosa arylsulfatase as an example of chemistry-driven protein evolution. *FEBS Letters* **586**, 1622–1630.

Maegley, K. A., Admiral, S. J. & Herschlag, D. (1996). Ras-catalyzed hydrolysis of GTP: a new perspective from model studies. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 8160–8166.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. (2002a). The protein kinase complement of the human genome. *Science* **298**, 1912–1934.

Marcos, E., Field, M. J. & Creheut, R. (2010). Pentacoordinated phosphorus revisited by high-level QM/MM calculations. *Proteins: Structure Function and Bioinformatics* **78**, 2405–2411.

Maskill, H. (1999). *Structure and Reactivity in Organic Compounds*. Oxford, UK: Oxford University Press.

Mastrangelo, I. A., Hough, P. V. C., Wall, J. S., Dodson, M., Dean, F. B. & Hurwrtz, J. (1989). ATP-dependent assembly of double hexamers of SV40T antigen at the viral origin of DNA replication. *Nature* **338**, 658–662.

Mayaan, E., Range, K. & York, D. M. (2004). Structure and binding of Mg(II) ions and di-metal bridge complexes with biological phosphates and phosphoranes. *Journal of Biological Inorganic Chemistry: JBIC: A Publication of the Society of Biological Inorganic Chemistry* **9**, 807–817.

Mccormick, F. & Wittinghofer, A. (1996). Interactions between Ras proteins and their effectors. *Current Opinion in Biotechnology* **7**, 449–456.

Menz, R. I., Walker, J. E. & Leslie, A. G. W. (2001). Structure of bovine mitochondrial F_1-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. *Cell* **106**, 331–341.

Mercero, J. M., Barrett, P., Lam, C. W., Fowler, J. E., Ugalde, J. M. & Pedersen, L. G. (2000). Quantum mechanical calculations on phosphate hydrolysis reactions. *Journal of Computational Chemistry* **21**, 43–51.

Messer, B. M., Roca, M., Chu, Z. T., Vicatos, S., Kilshtain, A. V. & Warshel, A. (2010). Multiscale simulations of protein landscapes: using coarse-grained models as reference potentials to full explicit models. *Proteins: Structure Function and Bioinformatics* **78**, 1212–1227.

Milburn, M., Tong, L., Devos, A. M., Bruenger, A., Yamaizumi, Z., Nishimura, S. & Kim, S. H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic Ras proteins. *Science* **247**, 939–945.

Mildvan, A. S. (1979). The role of metals in enzyme-catalyzed substitutions at each of the phosphorus atoms of ATP. *Advances in Enzymology and Related Areas of Molecular Biology* **49**, 103–126.

Mildvan, A. S. (1997). Mechanisms of signaling and related enzymes. *Proteins* **29**, 401–416.

Mimeault, M. & Batra, S. K. (2011). Complex oncogenic signaling networks regulate brain tumor-initiating cells and their progenies: pivotal roles of wild-type EGFR, EGFRvIII mutant and hedgehog cascades and novel multitargeted therapies. *Brain Pathology* **21**, 479–500.

Minehardt, T., Cooke, R., Marzari, N., Car, R. & Pate, E. (2003). Car-Parrinello simulations of ATP hydrolysis in myosin. In *225th National Meeting of the American-Chemical-Society*, New Orleans, LA.

Minehardt, T. J., Marzari, N., Cooke, R., Pate, E., Kollman, P. A. & Car, R. (2002). A classical and *ab initio* study of the interaction of the myosin triphosphate binding domain with ATP. *Biophysical Journal* **82**, 660–675.

Mlynsky, V., Banás, P., Walter, N. G., Sponer, J. & Otyepka, M. (2011). QM/MM studies of hairpin ribozyme self-cleavage suggest the feasibility of multiple competing reaction mechanisms. *Journal of Physical Chemistry B* **115**, 13911–13924.

Moazed, D., Robertson, J. M. & Noller, H. F. (1988). Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature* **334**, 362–364.

Mohr, D., Wintermeyer, W. & Rodnina, M. V. (2002). GTPase activation of elongation factors Tu and G on the ribosome. *Biochemistry* **41**, 12520–12528.

Moore, P. B. (2012). How should we think about the ribosome? *Annual Review of Biophysics* **41**, 18.11–18.19.

More O'Ferrall, R. A. (1970). Relationships between E2 and E1cB mechanisms of β-elimination. *Journal of the American Chemical Society B* 274–277.

Muegge, I., Qi, P. X., Wand, A. J., Chu, Z. T. & Warshel, A. (1997). The reorganization energy of cytochrome *c* revisited. *Journal of Physical Chemistry B* **101**, 825–836.

Muegge, I., Schweins, T., Langen, R. & Warshel, A. (1996). Electrostatic control of GTP and GDP binding in the oncoprotein p^{21} ras. *Structure* **4**, 475–489.

Muegge, I., Schweins, T. & Warshel, A. (1998). Electrostatic contributions to protein-protein binding affinities: application to Rap/Raf interaction. *Proteins: Structure Function and Bioinformatics* **30**, 407–423.

Mukherjee, S. & Warshel, A. (2011). Electrostatic origin of the mechanochemical rotary mechanism and the catalytic dwell of F_1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20550–20555.

Narlikar, G. J., Gopalakrishnan, V., Mcconnell, T. S., Usman, N. & Herschlag, D. (1995). Use of binding energy by an RNA enzyme for catalysis by positioning and substrate destabilization. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3668–3672.

Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, A. & Wittinghofer, A. (1996). Ras/Rap Effector Specificity Determined by Charge Reversal. *Nature Structural Biology* **3**, 723–729.

Nassar, N., Horn, G., Herrmann, C., Scherer, A., Mccormick, F. & Wittinghofer, A. (1995). The 2·2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature* **375**, 554–560.

Nikolic-Hughes, I., Rees, D. & Herschlag, D. (2004). Do electrostatic interactions with positively charged active site groups tighten the transition state for enzymatic phosphoryl transfer? *Journal of the American Chemical Society* **126**, 11814–11819.

Northrop, D. B. (1975). Steady-state analysis of kinetic isotope effects in enzymic reactions. *Biochemistry* **14**, 2644–2651.

O'brien, P. J. & Herschlag, D. (1998). Sulfatase activity of E-coli alkaline phosphatase demonstrates a functional link to arylsulfatases, an evolutionarily related enzyme family. *Journal of the American Chemical Society* **120**, 12369–12370.

O'brien, P. J. & Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. *Chemistry & Biology* **6**, R91–R105.

O'brien, P. J. & Herschlag, D. (2001). Functional interrelationships in the alkaline phosphatase superfamily: phosphodiesterase activity of *E-coli* alkaline phosphatase. *Biochemistry* **40**, 5691–5699.

O'Brien, P. J. & Herschlag, D. (2002). Alkaline phosphatase revisited: hydrolysis of alkyl phosphates. *Biochemistry* **41**, 3207–3225.

Oda, K., Matsuoka, Y., Funahashi, A. & Kitano, A. H. (2005). A comprehensive pathway map of epidermal growth factor receptor signaling. *Molecular Systems Biology* **1**, 1–17.

Oelschlaeger, P., Klahn, M., Beard, W. A., Wilson, S. H. & Warshel, A. (2007). Magnesium-cationic dummy atom molecules enhance representation of DNA polymerase β – in molecular dynamics simulations: improved accuracy in studies of structural features and mutational effects. *Journal of Molecular Biology* **366**, 687–701.

Ogle, J. M., Brodersen, D. E., Clemons, W. M., Jr., Tarry, M. J., Carter, A. P. & Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902.

Okuno, D., Fujisawa, R., Iino, R., Hirono-Hara, Y., Imamura, H. & Nojia, H. (2008). Correlation between the conformational states of F_1-ATPase as determined from its crystal structure and single-molecule rotation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 20722–20727.

Olguin, L. F., Askew, S. E., O'Donoghue, A. C. & Hollfelder, F. (2008). Efficient catalytic promiscuity of an enzyme superfamily: an arylsulfatase shows a rate acceleration of 1013 for phosphate monoester hydrolysis. *Journal of the American Chemical Society* **130**, 16547–16555.

Olsson, M. H. M., Mavri, J. & Warshel, A. (2006a). Transition state theory can be used in studies of enzyme catalysis: lessons from simulations of tunnelling and dynamical effects in lipoxygenase and other systems. *Philosophical Transactions of the Royal Society B-Biological Sciences* **361**, 1417–1432.

Olsson, M. H. M., Parson, W. W. & Warshel, A. (2006b). Dynamical contributions to enzyme catalysis: critical tests of a popular hypothesis. *Chemical Reviews* **106**, 1737–1756.

Oostenbrink, C. & Van Gunsteren, W. F. (2005). Efficient calculation of many stacking and pairing free energies in DNA from a few molecular dynamics simulations. *Chemistry – A European Journal* **11**, 4340–4348.

Orozco, M., Pérez, A., Noy, A. & Luque, F. J. (2003). Theoretical methods for the simulation of nucleic acids. *Chemical Society Reviews* **32**, 350–364.

Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. & Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-Ras P^{21} at 1·35 Å resolution: implications for the mechanism of GTP hydrolysis. *The EMBO Journal* **9**, 2351–2359.

Pannifer, A. D. B., Flint, A. J., Tonks, N. K. & Barford, D. (1998). Visualization of the cysteinyl-phosphate intermediate of a protein-tyrosine phosphatase by X-ray crystallography. *The Journal of Biological Chemistry* **273**, 10454–10462.

Pasqualato, S. & Cherfils, J. (2005). Crystallographic evidence for substrate-assisted GTP hydrolysis by a small GTP binding protein. *Structure* **13**, 533–540.

Patel, J. S., Branduardi, D., Masetti, M., Rocchia, W. & Cavalli, A. (2011). Insights into ligand–protein binding from local mechanical response. *Journal of Chemical Theory and Computation* **7**, 3368–3378.

Patel, S. S., Wong, I. & Johnson, K. A. (1991). Pre-steady-state kinetic-analysis of processive dna-replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* **30**, 511–525.

Pauling, L. (1960). *The Nature of the Chemical Bond*. Ithaca, N. Y.: Cornell University Press.

Pei, Z., Liu, G., Lubben, T. H. & Szczepankiewicz, B. G. (2004). Inhibition of protein tyrosine phosphatase 1B as a potential treatment of diabetes and obesity. *Current Pharmaceutical Design* **10**, 3481–3504.

Perreault, D. M. & Anslyn, E. V. (1997). Unifying the current data on the mechanism of cleavage – transesterification of RNA. *Angewandte Chemie* (*International ed. in English*) **36**, 432–450.

Peters, G. H., Frimurer, T. M. & Olsen, O. H. (1998). Electrostatic evaluation of the signature motif (H/V)CX_5R(S/T) in protein-tyrosine phosphatases. *Biochemistry* **37**, 5383–5393.

Piccirilli, J. A., Vyle, J. S., Caruthers, M. G. & Cech, T. R. (1993). Metal ion catalysis in the tetrahymena ribozyme reaction. *Nature* **361**, 85.

Pisliakov, A. V., Cao, J., Kamerlin, S. C. L. & Warshel, A. (2009). Enzyme millisecond conformational dynamics do not catalyze the chemical step. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 17359–17364.

Pley, H. W., Flaherty, K. M. & Mckay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature* **372**, 68–74.

Plotnikov, N. V., Kamerlin, S. C. L. & Warshel, A. (2011). Paradynamics: an effective and reliable model for *ab initio* QM/MM free-energy calculations and related tasks. *Journal of Physical Chemistry B* **115**, 7950–7962.

Plotnikov, N. V. & Warshel, A. (2012). Exploring, refining and validating the paradynamics QM/MM sampling. *Journal of Physical Chemistry B* **116**, 10342–10356.

Post J. R. C. B. & Ray, W. J. (1995). Reexamination of induced fit as a determinant of substrate specificity in enzymatic reactions. *Biochemistry* **34**, 15881–15885.

Prasad, B. R., Plotnikov, N. V. & Warshel, A. (2012). Resolving uncertainties about phosphate hydrolysis pathways by careful free energy mapping. In Press (DOI: 10.1021/jp309778n), J. Phys. Chem. B, 2012.

Prasad, B. R. & Warshel, A. (2011). Prechemistry *versus* preorganization in DNA replication fidelity. *Proteins: Structure Function and Bioinformatics* **79**, 2900–2919.

Preiss, J. & Handler, P. (1958). Biosynthesis of diphosphopyridine nucleotide: II. Enzymatic aspects. *The Journal of Biological Chemistry* **233**, 493–500.

Pu, J. & Karplus, M. (2008). How subunit coupling produces the γ-subunit rotary motion in F_1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1192–1197.

Pyle, A. M. (1993). Ribozymes: a distinct class of metalloenzymes. *Science* **261**, 709.

Radhakrishnan, R., Arora, K., Wang, Y. L., Beard, W. A., Wilson, S. H. & Schlick, T. (2006). Regulation of DNA repair fidelity by molecular checkpoints: 'Gates' in DNA polymerase β's substrate selected. *Biochemistry* **45**, 15142–15156.

Radhakrishnan, R. & Schlick, T. (2004). Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β's closing. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 5970–5975.

Radhakrishnan, R. & Schlick, T. (2005). Fidelity discrimination in DNA polymerase β: differing closing profiles for a mismatched (G: A) *versus* matched (G: C) base pair. *Journal of the American Chemical Society* **127**, 13245–13252.

Radhakrishnan, R. & Schlick, T. (2006). Correct and incorrect nucleotide incorporation pathways in DNA polymerase β. *Biochemical and Biophysical Research Communications* **350**, 521–529.

Radzicka, A. & Wolfenden, R. (1995). A proficient enzyme. *Science* **267**, 90–93.

Raines, R. T. (1998). Ribonuclease A*Chemical Reviews* **98**, 1045–1066.

Ramakrishnan, V. (2008). What we have learned from ribosome structures. *Biochemical Society Transactions* **36**, 567–574.

Ramakrishnan, V. (2010). Unraveling the structure of the ribosome (Nobel Lecture). *Angewandte Chemie* (*International ed. in English*) **49**, 4355–4380.

Riccardi, D., Konig, P., Guo, H. & Cui, Q. (2008). Proton transfer in carbonic anhydrase is controlled by electrostatics rather than the orientation of the acceptor. *Biochemistry* **47**, 2369–2378.

Riccardi, D., Schaefer, P. & Cui, Q. (2005). pK_a calculations in solution and proteins with QM/MM free energy perturbation simulations: a quantitative test of QM/MM protocols. *Journal of Physical Chemistry B* **109**, 17715–17733.

Richards, F. M. & Wyckoff, H. W. (1971). In *The Enzymes*, vol. 4 (Ed. P. D. Boyer), pp. 647–806. New York: Academic Press.

Roca, M., Vardi-Kilshtai, A. & Warshel, A. (2009). Toward accurate screening in computer aided enzyme design. *Biochemistry* **48**, 3046–3056.

Rodnina, M. V. & Wintermeyer, W. (2009). Recent mechanistic insights into eukaryotic ribosomes. *Current Opinion in Cell Biology* **21**, 435–443.

Rosta, E., Kamerlin, S. C. L. & Warshel, A. (2008). On the interpretation of the observed linear free energy relationship in phosphate hydrolysis: a thorough computational study of phosphate diester hydrolysis in solution. *Biochemistry* **47**, 3725–3735.

Rosta, E., Klähn, M. & Warshel, A. (2006). Towards accurate *ab initio* QM/MM calculations of free-energy profiles of enzymatic reactions. *Journal of Physical Chemistry B* **110**, 2934–2941.

Rowell, R. & Gorenstein, D. G. (1981). Multiple structure-reactivity correlations in the hydrolysis of epimeric 2-aryloxy-2-oxy-dioxaphosphorinanes. Stereoelectronic effects. *Journal of the American Chemical Society* **103**, 5894–5902.

Rucker, R., Oelschlaeger, P. & Warshel, A. (2009). A binding free energy decomposition approach for accurate calculations of the fidelity of DNA polymerases. *Proteins: Structure Function and Bioinformatics* **78**, 671–680.

Rychkova, A., Vicatos, S. & Warshel, A. (2010). On the energetics of translocon-assisted insertion of charged transmembrane helices into membranes. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17598–17603.

Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. & Pelletier, H. (1997). Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205–11215.

Scarano, G., Krab, I. M., Bocchini, V. & Parmeggiani, A. (1995). Relevance of histidine-84 in the elongation-factor Tu GTPase activity and in poly(phe) synthesis – its substitution by glutamine and alanine. *FEBS Letters,* **365**, 214–218.

Scheffzek, K. & Ahmadian, M. (2005). GTPase activating proteins: structural and functional insights 18 years after discovery. *Cellular and Molecular Life Science* **62**, 3014–3038.

Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F. & Wittinghofer, A. (1997). The Ras-RasGAP complex – structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338.

Schiemann, O., Fritscher, J., Kisseleva, N., Sigurdsson, S. T. & Prisner, T. F. (2003). Structural investigation of a high affinity MnII binding site in the hammerhead ribozyme by EPR spectroscopy and DFT calculations. Effects of neomycine-B on metal-ion bindings. *Chemical & Biological Chemistry* **4**, 1057.

Schmeing, T. M. & Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234–1242.

Schmeing, T. M., Voorhees, R. M., Kelley, A. C., Gao, Y.-G., Iv, F. V. M., Weir, J. R. & Ramakrishnan, V. (2009). The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688–694.

Schroeder, G. K., Lad, C., Wyman, P., Williams, N. H. & Wolfenden, R. (2006). The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4052–4055.

Schuette, J.-C., Iv, F. V. M., Kelley, A. C., Weir, J. R., Giesebrecht, J., Connell, S. R., Loerke, J., Mielke, T., Zhang, W., Penczek, P. A., Ramakrishnan, V. & Spahn, C. M. T. (2009). GTPase activation of elongation factor EF-Tu by the ribosome during decoding. *The EMBO Journal* **28**, 755–765.

Schutz, C. N. & Warshel, A. (2004). The low barrier hydrogen bond (LBHB) proposal revisited: the case of the Asp … His pair in serine proteases. *Proteins* **55**, 711–723.

Schwans, J. P., Kraut, D. A. & Herschlag, D. (2009). Determining the catalytic role of remote substrate binding interactions in ketosteroid isomerase. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14271–14275.

Schweins, T. (1991). Untersuchung des reaktionsmechanismusses von ras-p21 mittels computer-modeling. *Diplomarbeit Thesis*, University of Southern California.

Schweins, T., Geyer, M., Kalbitzer, H. R., Wittinghofer, A. & Warshel, A. (1996a). Linear free energy relationships in the intrinsic and GTPase activating protein-stimulated guanosine 5′-triphosphate hydrolysis of p21 ras. *Biochemistry* **35**, 14225–14231.

Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H. R. & Wittinghofer, A. (1995). Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21 ras and other GTP-binding proteins. *Nature Structural Biology* **2**, 36–44.

Schweins, T., Langen, R. & Warshel, A. (1994). Why have mutagenesis studies not located the general base in ras p21. *Nature Struct. Biol* **1**, 476–484.

Schweins, T., Wittinghofer, A. & Warshel, A. (1996b). Mechanistic analysis of the observed linear free energy relationships in $p21^{ras}$ and related systems. *Biochemistry* **35**, 14232–14243.

Scott, W. G. (1999). Biophysical and biochemical investigations of RNA catalysis in the hammerhead ribozyme. *Quarterly Reviews of Biophysics* **32**, 241–284.

Scrima, A., Thomas, C., Deaconescu, D. & Wittinghofer, A. (2008). The Rap-RapGAP complex: GTP hydrolysis without catalytic glutamine and arginine residues. *The EMBO Journal* **27**, 1145–1153.

Scudder, P. H. (1990). Use of reaction cubes for generation and display of multiple mechanistic pathways. *The Journal of Organic Chemistry* **55**, 4238–4240.

Seewald, M. J., Körner, C., Wittinghofer, A. & Vetter, I. R. (2002). RanGAP mediates GTP hydrolysis without an arginine finger. *Nature* **415**, 662–666.

Seibert, E., Ross, J. B. A. & Osman, R. (2003). Contribution of opening and bending dynamics to specific recognition of DNA damage. *Journal of Molecular Biology* **330**, 687–703.

Sham, Y. Y., Chu, Z. T., Tao, H. & Warshel, A. (2000). Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease. *Proteins: Structure Function and Bioinformatics* **39**, 393–407.

Sharma, P. K., Xiang, Y., Kato, M. & Warshel, A. (2005). What are the roles of substrate assisted catalysis and proximity effects in peptide bond formation by the ribosome? *Biochemistry* **44**, 11307–11314.

Shimo-Kon, R., Muneyuki, E., Sakai, H., Adachi, K., Yoshida, M. & Kinosita, K. (2010). Chemo-mechanical coupling in F_1-ATPase revealed by catalytic site occupancy during catalysis. *Biophysical Journal* **98**, 1227–1236.

Showalter, A. K., Lamarche, B. J., Bakhtina, M., Su, M. I., Tang, K. H. & Tsai, M. D. (2006). Mechanistic comparison of high-fidelity and error-prone DNA polymerases and ligases involved in DNA repair. *Chemical Reviews* **106**, 340–360.

Shriver, D. & Atkins, P. (2006). *Inorganic Chemistry*. New York: Oxford University Press.

Shurki, A. & Warshel, A. (2004). Why does the Ras switch 'break' by oncogenic mutations? *Proteins: Structure Function and Bioinformatics* **55**, 1–10.

Siegbahn, P. (2006). The performance of hybrid DFT for mechanisms involving transition metal complexes in enzymes. *Journal of Biological Inorganic Chemistry: JBIC: A Publication of the Society of Biological Inorganic Chemistry* **11**, 695–701.

Sigala, P., Kraut, D., Caaveiro, J., Pybus, B., Ruben, E., Ringe, D., Petsko, G. & Herschlag, D. (2008). Testing geometrical discrimination within an enzyme active site: constrained hydrogen bonding in the ketosteroid isomerase oxyanion hole. *Journal of the American Chemical Society* **130**, 13696–13708.

Simmons, D. T. (2000). SV40 large T antigen functions in DNA replication and transformation. *Advances In Virus Research* **55**, 75–134.

Simopoulos, T. T. & Jencks, W. P. (1994). Alkaline phosphatase is an almost perfect enzyme. *Biochemistry* **33**, 10375–10380.

Smith, G. K., Ke, Z., Guo, H. & Hengge, A. C. (2011). Insights into the phosphoryl transfer mechanism of

cyclin-dependent protein kinases from *ab initio* QM/MM free-energy studies. *Journal of Physical Chemistry B* **115**, 13713–13722.

Smith, J. P., Brown, W. E. & Lehr, J. R. (1955). Structure of crystalline phosphoric acid. *Journal of the American Chemical Society* **77**, 2728–2730.

Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E. & Sigler, P. B. (1994). GTPase mechanism of G proteins from the 1·7 Å crystal structure of transducinα-GDP-AlF_4^-. *Nature* **372**, 276–279.

Spandidos, D. A. (1989). *Ras Oncogenes*. New York: Plenum Press.

Sprang, S. R. (1997a). G protein mechanisms: insights from structural analysis. *Annual Review of Biochemistry* **66**, 639–678.

Sprang, S. R. (1997b). G proteins, effectors and GAPs: structure and mechanism. *Current Opinion in Structural Biology* **7**, 849–856.

Stec, B., Holtz, K. M. & Kantrowitz, E. R. (2000a). A revised mechanism for the alkaline phosphatase reaction involving three metal ions. *Journal of Molecular Biology* **299**, 1303–1311.

Stec, B., Holtz, K. M. & Kantrowitz, E. R. (2000b). A revised mechanism for the alkaline phosphatase reaction involving three metal ions. *Journal of Molecular Biology* **299**, 1303–1311.

Steitz, T. A. (2010). From the structure and function of the ribosome to new antibiotics (Nobel Lecture). *Angewandte Chemie* (*International ed. in English*) **49**, 4381–4398.

Steitz, T. A., Smerdon, S. J., Jäger, J. & Joyce, C. M. (1994). A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266**, 2022–2025.

Steitz, T. A. & Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 6498–6502.

Stockbridge, R. B. & Wolfenden, R. (2009). Phosphate monoester hydrolysis in cyclohexane. *Journal of the American Chemical Society* **131**, 18248–18249.

Strajbl, M., Sham, Y. Y., Villa, J., Chu, Z. T. & Warshel, A. (2000). Calculation of activation entropies of chemical reactions in solution. *Journal of Physical Chemistry B* **104**, 4578–4584.

Štrajbl, M., Shurki, A. & Warshel, A. (2003). Converting conformational changes to electrostatic energy in molecular motors: the energetics of ATP synthase. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14834–14839.

Sucato, C. A., Upton, T. G., Kashemirov, B. A., Batra, V. K., Martinek, V., Xiang, Y., Beard, W. A., Pedersen, L. C., Wilson, S. H., Mckenna, C. E., Florian, J., Warshel, A. & Goodman, M. F. (2007). Modifying the β,γ leaving-group bridging oxygen alters nucleotide incorporation efficiency, fidelity, and the catalytic mechanism of DNA polymerase β. *Biochemistry* **46**, 461–471.

Sucato, C. A., Upton, T. G., Kashemirov, B. A., Osuna, J., Oertell, K., Beard, W. A., Wilson, S. H., Florián, J., Warshel, A., Mckenna, C. E. & Goodman, M. F. (2008). DNA polymerase β fidelity: halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry* **47**, 870–879.

Sunahara, R. K., Tesmer, J. J. G., Gilman, A. G. & Sprang, S. R. (1997). Crystal structure of the adenylyl cyclase activator $G_{s\alpha}$. *Science* **278**, 1943.

Taylor, S. S. & Kornev, A. P. (2011). Protein kinases: evolution of dynamic regulatory proteins. *Trends in Biochemical Sciences* **36**, 65–77.

Taylor, S. S. & Radzio-Andzelm, E. (1994). Three protein kinase structures define a common motif. *Structure* **2**, 345–355.

Thompson, J. E., Kutateladze, T. G., Schuster, M. C., Venegas, F. D., Messmore, J. M. & Raines, R. T. (1995). Limits to catalysis by Ribonuclease A. *Bioorganic Chemistry* **23**, 471–481.

Tinoco, I. & Wen, J. D. (2009). Simulation and analysis of single-ribosome translation. *Physical Biology* **6**, 10.

Todd, A. (1959). Some aspects of phosphate chemistry. *Proceedings of the National Academy of Sciences of the United States of America* **45**, 1389–1397.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics*. London, UK: Oxford University Press.

Topol, I. A., Cachau, R. E., Nemukhin, A. V., Grigorenko, B. L. & Burt, S. K. (2004). Quantum chemical modeling of the GTP hydrolysis by the RAS-GAP protein complex. *Biochimica et Biophysica Acta – Proteins and Proteomics* **1700**, 125–136.

Torres, R. A., Himo, F., Bruice, T. C., Noodleman, L. & Lovell, T. (2003). Theoretical examination of Mg2+-mediated hydrolysis of a phosphodiester linkage as proposed for the hammerhead ribozyme. *Journal of the American Chemical Society* **125**, 9861–9867.

Toscano, M. D., Woycechowsky, K. J. & Hilvert, D. (2007). Minimalist active-site redesign: teaching old enzymes new tricks. *Angewandte Chemie* (*International ed. in English*) **46**, 4468–4470.

Trushkov, I. V., Zhdankin, V. V., Koz'min, A. S. & Zefirov, N. S. (1990). Cubic reaction coordinate diagram in the nucleophilic substitution process. *Tetrahedron Letters* **31**, 3199–3200.

Uchimaru, T., Uebayasi, M., Tanabe, K. & Taira, K. (1993). Theoretical analyses on the role of Mg^{2+} ions in ribozyme reactions. *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* **7**, 137.

Van Loo, B., Jonas, S., Babtie, A. C., Benjdia, A., Berteau, O., Hyvönen, M. & Hollfelder, F. (2010). An efficient, multiply promiscuous hydrolase in the alkaline phosphatase superfamily. *Proceedings of the National*

Academy of Sciences of the United States of America **107**, 2740–2745.

Várnai, P. & Warshel, A. (2000). Computer simulation studies of the catalytic mechanism of human aldose reductase. *Journal of the American Chemical Society* **122**, 3849–3860.

Vetter, I. R. & Wittinghofer, A. (1999). Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Quarterly Reviews of Biophysics* **32**, 1–56.

Villa, E., Sengupta, J., Trabuco, L. G., Lebarron, J., Baxter, W. T., Shaikh, T. R., Grassucci, R. A., Nissen, P., Ehrenberg, M., Schulten, K. & Frankd, J. (2009). Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1063–1068.

Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. (2010). The mechanism for activation of GTP hydrolysis on the ribosome. *Science* **330**, 835–838.

Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. (2011). Response to comment on 'The mechanism for activation of GTP hydrolysis on the ribosome'. *Science* **333**, 37.

Wang, H. & Oster, G. (1998). Energy transduction in the F_1 motor of ATP synthase. *Nature* **396**, 279–282.

Wang, L. M., Patel, U., Ghosh, L. & Banerjee, S. (1992). DNA polymerase beta mutations in human colorectal cancer. *Cancer Research* **52**, 4284–4827.

Wang, Y. & Schlick, T. (2008). Quantum mechanics/molecular mechanics investigation of the chemical reaction in Dpo4 reveals water-dependent pathways and requirements for active site reorganization. *Journal of the American Chemical Society* **130**, 13240–13250.

Wang, Y. N., Topol, I. A., Collins, J. R. & Burt, S. K. (2003). Theoretical studies on the hydrolysis of monophosphate and tri-phosphate in gas phase and aqueous solution. *Journal of the American Chemical Society* **125**, 13265–13273.

Warshel, A. (1978). Energetics of enzyme catalysis. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 5250–5254.

Warshel, A. (1991). *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. New York: John Wiley and Sons.

Warshel, A. (2003). Computer simulations of enzyme catalysis: methods, progress, and insights. *Annual Review of Biophysics and Biomolecular Tructure* **32**, 425–443.

Warshel, A., Åqvist, J. & Creighton, S. (1989). Enzymes work by Solvation substitution rather than by Desolvation. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 5820–5824.

Warshel, A. & Florian, J. (2004). The empirical valence bond (EVB) method. In *The Encyclopedia of Computational Chemistry*, (Eds P. von Ragué Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III and P. R. Schreiner), Chichester, UK: John Wiley and Sons.

Warshel, A., Hwang, J. K. & Åqvist, J. (1992a). Computer-simulations of enzymatic-reactions – examination of linear free-energy relationships and quantum-mechanical corrections in the initial proton-transfer step of carbonic-anhydrase. *Faraday Discussions* **93**, 225–238.

Warshel, A., Hwang, J. K. & Åqvist, J. (1992b). Computer simulations of enzymatic reactions: examination of linear free energy relationships and quantum mechanical corrections in the initial proton transfer step of carbonic anyhdrase. *Faraday Discussions* **93**, 225–238.

Warshel, A. & Papazyan, A. (1996). Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13665–13670.

Warshel, A. & Parson, W. W. (2001). Dynamics of biochemical and biophysical reactions: insight from computer simulations. *Quarterly Reviews of Biophysics* **34**, 563–670.

Warshel, A., Schweins, T. & Fothergill, M. (1994). Linear free energy relationships in enzymes. Theoretical analysis of the reaction of tyrosyl-tRNA synthase. *Journal of the American Chemical Society* **116**, 8437–8442.

Warshel, A., Sharma, P. K., Chu, Z. T. & Aqvist, J. (2007). Electrostatic contributions to binding of transition state analogues can be very different from the corresponding contributions to catalysis: phenolates binding to the oxyanion hole of ketosteroid isomerase. *Biochemistry* **46**, 1466–1476.

Warshel, A., Sharma, P. K., Kato, M. & Parson, W. W. (2006a). Modeling electrostatic effects in proteins. *Biochimica et Biophysica Acta* **1764**, 1647–1676.

Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H. & Olsson, M. H. M. (2006b). Electrostatic basis for enzyme catalysis. *Chemical Reviews* **106**, 3210–3235.

Warshel, A., Villà, J., Štrajbl, M. & Florián, J. (2000). Remarkable rate enhancement of orotidine 5′-monophosphate decarboxylase is due to transition state stabilization rather than ground state destabilization. *Biochemistry* **39**, 14728–14738.

Warshel, A. & Weiss, R. M. (1980). An empirical valence bond approach for comparing reactions in solutions and in enzymes. *Journal of the American Chemical Society* **102**, 6218–6226.

Weber, J. & Senior, A. E. (1997). Catalytic mechanism of F_1-ATPase. *Biochimica et Biophysica Acta* **1319**, 19–58.

Weiss, P. M., Knight, W. B. & Cleland, W. W. (1986). Secondary ^{18}O isotope effects on the hydrolysis of glucose 6-phosphate. *Journal of the American Chemical Society* **108**, 2761–2762.

Wenjin, Li., Rudack, T., Gerwert, K., Grater, F. & Schlitter, J. (2012). Exploring the Multidimensional

Free Energy Surface of Phosphoester Hydrolysis with Constrained QM/MM Dynamics. *Journal of Chemical Theory Computation* 8, 3596–3604.

Westheimer, F. H. (1968). Pseudo-rotation in the hydrolysis of phosphate esters. *Accounts of Chemical Research* **1**, 70–78.

Westheimer, F. H. (1981). Monomeric metaphosphates. *Chemical Reviews* **81**, 313–326.

Westheimer, F. H. (1987). Why nature chose phosphates. *Science* **235**, 1173–1178.

White, S. H. & Von Heijne, G. (2008). How translocons select transmembrane helices. *Annual Review of Biophysics* **37**, 23–42.

Wiesmann, C., Barr, K. J., Kung, J., Zhu, J., Erlanson, D. A., Shen, W., Fahr, B. J., Zhong, M., Taylor, T., Randall, M., Mcdowell, R. S. & Hansen, S. K. (2004). Allosteric inhibition of protein tyrosine phosphatase 1B. *Nature Structural and Molecular Biology* **11**, 730–737.

Wiesmuller, L. & Wittinghofer, A. (1994). Signal transduction pathways involving Ras. *Cellular Signalling* **6**, 247–267.

Wilde, J. A., Bolton, P. H., Dell'Acqua, M., Hibler, D. W., Pourmotabbed, T. & Gerlt, J. A. (1998). Identification of residues involved in a conformational change accompanying substitutions for glutamate-43 in staphylococcal nuclease. *Biochemistry* **27**, 4127–4132.

Wilkie, J. & Gani, D. (1996). Comparison of inline and non-inline associative and dissociative reaction pathways for model reactions of phosphate monoester hydrolysis. *Journal of the American Chemical Society, Perkin Transactions* **2**, 783–787.

Williams, A. (1984). Effective charge and Leffler's Index as mechanistic tools for reactions in solution. *Accounts of Chemical Research* **17**, 425–430.

Williams, A. (1992). Effective charge and transition-state structure in solution. *Advances in Physical Organic Chemistry* **27**, 1–55.

Williams, N. H. (2004a). Models for biological phosphoryl transfer*Biochimica et Biophysica Acta* **1697**, 279–287.

Williams, N. H. (2004b). Models for biological phosphoryl transfer. *Biochimica et Biophysica Acta* **1697**, 279–287.

Williams, N. H., Cheung, J. & Chin, J. (1998a). Reactivity of phosphate diesters doubly coordinated to a dinuclear cobalt(III) complex: dependence of the reactivity on the basicity of the leaving group. *Journal of the American Chemical Society* **120**, 8079–8087.

Williams, N. H., Cheung, W. & Chin, J. (1998b). Reactivity of phosphate diesters doubly coordinated to a dinuclear cobalt(III) complex: dependence of the reactivity on the basicity of the leaving group. *Journal of the American Chemical Society* **120**, 8079–8087.

Williams, N. H. & Wyman, P. (2001). Base catalysed phosphate diester hydrolysis. *Chemical Communications* 1268–1269.

Wittinghofer, A. (2006). Phosphoryl transfer in Ras proteins, conclusive or elusive? *Trends in Biochemical Sciences* **31**, 20–23.

Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z. & Kern, D. (2004). Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural and Molecular Biology* **11**, 945–949.

Wolfe-Simon, F., Blum, J. S., Kulp, T. R., Gordon, G. W., Hoeft, S. E., Pett-Ridge, J., Stolz, J. F., Webb, S. M., Weber, P. K., Davies, P. C. W., Anbar, A. D. & Oremland, R. S. (2010). A bacterium that can grow by using arsenic instead of phosphorus. *Science* **332**, 1163–1166.

Wolfe-Simon, F., Davies, P. C. & Anbar, A. D. (2009). Did nature also use arsenic? *International Journal of Astrobiology* **8**, 69–74.

Wolfenden, R. (2006). Degrees of difficulty of water-consuming reactions in the absence of enzymes. *Chemical Reviews* 106, 3379–3396.

Wolfenden, R. (2011). Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annual Review of Biophysics* **80**, 645–667.

Wolfenden, R., Ridgeway, C. & Young, G. (1998). Spontaneous hydrolysis of ionised phosphate monoesters and diesters and the proficiencies of the phosphatases and phosphodiesterases as catalysts. *Journal of the American Chemical Society* **120**, 833–834.

Wolfenden, R. & Snider, M. J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Accounts of Chemical Research* **34**, 938–945.

Wong, K.-Y., Gu, H., Zhang, S., Piccirilli, J. A., Harris, M. E. & York, D. M. (2012). Characterization of the reaction path and transition States for RNA transphosphorylation models from theory and experiment. *Angewandte Chemie* (*International ed. in English*) **51**, 647–651.

Wong, K.-Y., Lee, T.-S. & York, D. M. (2011). Active participation of the Mg^{2+} ion in the reaction coordinate of RNA self-cleavage catalyzed by the hammerhead ribozyme. *Journal of Chemical Theory and Computation* **7**, 1–3.

Xiang, Y., Goodman, M. F., Beard, W. A., Wilson, S. H. & Warshel, A. (2008). Exploring the role of large conformational changes in the fidelity of DNA polymerase β. *Proteins: Structure Function and Bioinformatics* **70**, 231–247.

Xiang, Y., Oelschlaeger, P., Florian, J., Goodman, M. F. & Warshel, A. (2006). Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in human pol β. *Biochemistry* **45**, 7036–7048.

Yang, L. J., Beard, W. A., Wilson, S. H., Broyde, S. & Schlick, T. (2004). Highly organized but pliant active site of DNA polymerase β: compensatory mechanisms

in mutant enzymes revealed by dynamics simulations and energy analyses. *Biophysical Journal* **86**, 3392–3408.

YANG, W., GAO, Y. Q., CUI, Q., MA, J. & KARPLUS, M. (2003). The missing link between thermodynamics and structure in F_1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 874–879.

YARDEN, Y. & SCHLESSINGER, J. (1987). Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. *Biochemistry* **26**, 1443–1451.

YARUS, M. (1993). How many catalytic RNAs? Ions and the Cheshire cat conjecture. *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* **7**, 31–39.

YONATH, A. (2010). Polar bears, antibiotics, and the evolving ribosome (Nobel Lecture). *Angewandte Chemie* (*International ed. in English*) **49**, 4340–4354.

YOU, T. J. & BASHFORD, D. (1995). Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophysical Journal* **69**, 1721–1733.

ZALATAN, J. G., FENN, T. D., BRUNGER, T. A. & HERSCHLAG, D. (2006). Structural and functional comparisons of nucleotide pyrophosphate/phosphodiesterase and alkaline phosphatase: implications for mechanism and evolution. *Biochemistry* **45**, 9788–9803.

ZALATAN, J. G. & HERSCHLAG, D. (2006). Alkaline phosphatases mono- and diesterase reactions: comparative transition state analysis. *Journal of the American Chemical Society* **128**, 1293–1303.

ZEIDLER, W., EGLE, C., RIBEIRO, S., WAGNER, A., KATUNIN, V., KREUTZER, R., RODNINA, M., WINTERMEYER, W. & SPRINZL, M. (1995). Site-directed mutagenesis of *Thermus thermophilus* elongation factor Tu – replacement of His85, Asp81 and Arg300. *European Journal of Biochemistry* **229**, 596–604.

ZHANG, H., ZHA, X., TAN, Y., HORNBECK, P. V., MASTRANGELO, A. J., ALESSI, D. R., POLAKIEWICZ, R. D. & COMB, M. J. (2002). Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *The Journal of Biological Chemistry* **277**, 39379–39387.

ZHANG, Z.-Y. (2003). Chemical and mechanistic approaches to the study of protein tyrosine phosphatases. *Accounts of Chemical Research* **36**, 385–392.

ZHANG, Z. Y., HARMS, E. & VAN ETTEN, R. L. (1994). Asp129 of low molecular weight protein tyrosine phosphatase is involved in leaving group protonation. *The Journal of Biological Chemistry* **269**, 25947–25950.

ZHANG, Z. Y., WANG, Y. & DIXON, J. E. (1999). Dissecting the catalytic mechanism of protein-tyrosine phosphatases. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1624–1627.

ZHOU, D.-M., KUMAR, P. K. R., ZHANG, L.-H. & TAIRA, K. (1996). Ribozyme mechanism revisited: evidence against direct coordination of a Mg2+ ion with the Pro-R oxygen if the scissile phosphate in the transition state of a hammerhead ribizyme-catalyzed reaction. In *EMBO Workshop*, Xanten, Germany.

ZHOU, D.-M. & TAIRA, K. (1998). The hydrolysis of RNA: from theoretical calculations to the hammerhead ribozyme-mediated cleavage of RNA. *Chemical Reviews* **98**, 991–1026.