Review Article

Applications of Potential Energy Surfaces in the Study of Enzymatic Reactions

Eric A. C. Bushnell, WenJuan Huang, and James W. Gauld

Department of Chemistry and Biochemistry, University of Windsor, Windsor, ON, Canada N9B 3P4

Correspondence should be addressed to James W. Gauld, gauld@uwindsor.ca

From a generated PES, one can determine the relative energies of species involved, the sequence in which they occur, and the activation barrier(s) associated with individual steps or the overall mechanism. Furthermore, they can provide more insights than a simple indication of a path of sequential mechanistic structures and their energetic relationships. The investigation into the activation of O_2 by alpha-ketoglutarate-dependent dioxygenase (AlkB) clearly shows the opportunity for spin inversion, where one can see that the lowest energy product may be formed via several possible routes. In the investigation of uroporphyrinogen decarboxylase III (UROD), the use of QM/MM methods allowed for the inclusion of the anisotropic protein environment providing greater insight into the rate-limiting barrier. Lastly, the mechanism of 6-phospho-α-glucosidase (GlvA) was discussed using different active site models. In particular, a continuum model PES was compared to the gas-phase PES.

1. Introduction

For a chemical reaction, enzymatic or nonenzymatic, the reactants, intermediates, and products all exist on a multidimensional surface. With this surface, a reaction is perfectly described by the statistical average of all possible paths from reactants to products via all possible intermediates [1]. However, a system with N atoms would require computing a (3N-6)-dimensional surface. Hence, if X number of points are to be computed for each of the (3N-6) dimensions, then X^{3N-6} calculations must be done [1]. Thus, such an undertaking is usually only computationally feasible for chemical models that consist of a few atoms. In contrast, studies on enzymatic mechanisms often necessarily require large chemical models consisting of important active site functional groups and cofactors and the substrate. Indeed, the cluster-based approach for investigating biochemical reactions typically use chemical models containing 200 atoms or more [2]. Thus, for such systems, it is impossible to determine the complete PES. Instead, a "slice" of the surface is typically constructed that involves only two coordinates, energy and reaction coordinate, and is commonly referred to as the PES [3].

The use and applicability of such "reduced-dimensionality" PESs reflects the fact that chemists and biochemists are usually only interested in key, mechanistically relevant structures including, for example, the (i) isolated reactants, (ii) reactive complex, (iii) transition structures (TS), (iv) intermediate(s), (v) product complex, and (vi) the separated products (Figure 1) [1]. It is noted that in general, computational studies on enzymatic reactions are investigated with the enzyme-substrate complex already formed. This is due in part to the inherent difficulties associated with modeling substrate binding and product release such as, for instance, consideration of clathrate waters. Furthermore, as previously stated [4], in some cases substrate binding is "seldom the most interesting part of the potential surface".

A "reduced dimensionality" PES can provide a considerable amount of information and insight into a chemical system. For example, a common goal of computational chemistry is the accurate determination of the thermochemistry of a particular system. From a PES, one can determine the relative energies of species involved, the sequence in which they occur, and the activation barrier(s) associated with individual steps or the overall mechanism. Consequently, one can determine if a reaction step or pathway is
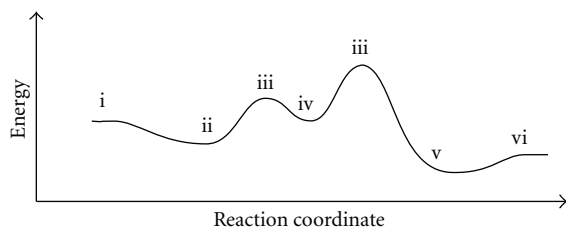
FIGURE 1: A generic PES showing (i) isolated reactants, (ii) reactive complex, (iii) TSs, (iv) intermediate complex, (v) product complex, and (vi) separated products.

feasible under the given conditions. With the appropriate corrections, these potential energies can be converted to enthalpies or free energies as has been previously discussed in detail [3]. For enzyme-catalysed reactions, free energy barriers between 55–75 kJ mol^{-1} are typical with the upper thermodynamic limit generally held to be 84–105 kJ mol^{-1} [4]. For barriers greater than 40 kJ mol^{-1}, transition state theory is sufficiently accurate to enable determination of the associated rate constant, albeit perhaps to within several orders of magnitude (1) [4]

$$\mathbf{k} = \frac{\mathbf{k_B T}}{\mathbf{h}} \mathbf{e}^{-\Delta G^{\ddagger}/RT}. \tag{1}$$

The reliability and accuracy of any PES is dependent on the computational model used; that is, the choice of computational method plus chemical model. A common approach, in particular for investigations of enzymatic processes, is the use of "cluster chemical models" in combination with a reliable quantum mechanical (QM) or density functional theory (DFT) methods. Typically, in such cases, only those regions of the active site-substrate complex involved in bond making and breaking processes and active site residues or functional groups that play a direct role are included [5]. For the study of biocatalytic processes, DFT methods, in particular, B3LYP, are presently the most widely used [6–8]. This is due to the fact that they include electron correlation effects yet are computationally less expensive than more conventional wave function-based electron correlation methods. Hence, they can be applied to comparatively large chemical models [4, 9]. Furthermore, they have been shown to often be able to provide accurate and reliable structures and thermochemistry. Indeed, B3LYP has been shown to often be highly accurate [5] with relative energy errors for 1st and 2nd row atom and transition metal-containing systems of just ~13.0 and ~21.0 kJ mol^{-1}, respectively [4]. The surrounding protein environment, more specifically its general polarity, is then included via use of a polarizable continuum model (e.g., IEF-PCM) [10, 11].

More recently, QM/MM methods have been increasingly applied to the elucidation of enzymatic reaction PESs. This is due in part to the fact that (i) they are able to model a greater portion of the enzyme through their combined use of QM and molecular mechanics (MM) methods, and thus (ii) are able to model the polar and steric nonhomogeneity of the protein environment surrounding the active site. Furthermore, the computed energetics of the system have

been shown to often converge faster with increasing chemical model size than for the above cluster/QM-based approach [12]. In addition, the role of the residues and mechanism is also less sensitive to increasing or varying model size [13]. However, QM/MM-based approaches suffer some of the same limitations as the alternative cluster/DFT approach such as the fact that the dynamic behaviour of the enzyme is not fully taken into account [13, 14]. The application of these methods to the elucidation of enzymatic pathways has been previously reviewed; see, for example, Senn and Thiel [15] and Llano and Gauld [16].

There are computational techniques available to determine a statistical average of possible alternative pathways of an enzymatic reaction [1, 3, 12, 17]. One common approach is to use a semiempirical (SE) method to describe the QM layer, while an MM method is used to describe the outer layer. The use of SE methods is due to the fact that extensive sampling of the protein, substrate, and solvent must be performed in order to calculate the free energies. This sampling requires large numbers of calculations that also include the inner layer, hence the need for highly efficient methods that describe this region. Alternatively, the CPMD method typically describes the inner layer using DFT methods. However, due to the added computational costs, simulation timescales are limited to 0.1 ns, significantly shorter than those noted above using SE methods [18]. In addition, QM/MM MD can also be used to determine free energies. Similar to CPMD, QM/MM MD uses DFT methods to describe the inner layer, where the complete QM/MM model is simulated for a period of time [19]. For electron transfer (ET) processes, a combined QM/MM and MD approach can be used to obtain free energies [20]. More specifically, structures of the ET reactant and product are obtained using QM/MM. An MD simulation is then run on these structures, where the equilibrium averages of the MD trajectories are used to calculate the free energies. It is noted that Zhang et al. [13] investigated the effects of structural fluctuations on the reaction energy barrier in the catalytic mechanism of acetylcholinesterase. Using eight different starting conformations obtained from a 1 ns MD simulation, they found that while the enzyme-substrate structural fluctuations led to differences in the calculated barrier of approximately ±8 kJ mol^{-1}, the mechanistic details remained very consistent.

Potential energy surfaces, however, can provide more insights than a simple indication of a path of sequential mechanistic structures and their energetic relationships. For example, species found along an enzymatic mechanism may have more than one electronic state that lie close in energy and within the boundaries imposed by the "enzyme thermodynamic limit" (see above), particularly in the case of metalloenzymes. Often, such species contain one or more unpaired electrons. Importantly, however, it is possible that as a reaction progresses, the energy differences between states may vary, and in fact, may even change in their energetic ordering. Thus, a mechanism may proceed via "switching" between states, that is, undergoing a spin inversion (SI). Such reactions are termed multistate reactions (MSR). Hence, in the case of MSR pathways, a fundamental property is that
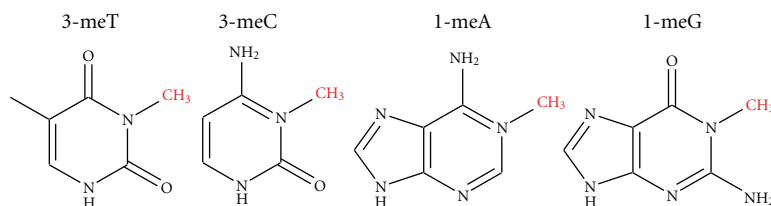
SCHEME 1: Examples of alkylated nucleobases known to be repaired under physiological conditions by the AlkB family of enzymes [21].

the PES for each state may cross or may remain in close proximity due to the near degeneracy of the d-orbitals of metals within the complex [22]. The ability in changing of states may have a significant impact on the reaction. For example, for a given transition metal complex, a particular reaction may have a significant barrier, and thus may not feasibly occur. However, for another possible state of the complex, this reaction may have a markedly lower barrier such that it may now be feasible. While SI is typically forbidden, it has been suggested that it may be common in transition metal chemistry [23] and by extension, metalloenzymes. The points at which SI may occur on a full-dimensionality PES are challenging to describe accurately due to the possibility of strong spin-orbit coupling between the two states. However, these crossing points are not generally at stationary points [23]. Fortunately, modern computational tools are good at locating stationary points [23]. Thus, one can map a "reduced-dimensionality" PES for each state, and where they cross indicates a possible region in which SI may occur [23].

In this present paper, some of the ways in which PESs may be exploited to provide insights into enzymatic mechanisms and some factors that may influence their accuracy and reliability are discussed. These are illustrated through a review of several enzymatic systems that we have recently examined. More specifically, this article highlights three major topics in the application of PES in the elucidation of enzymatic mechanisms through examples: (i) the use of multistate reactivity PESs in elucidating possible pathways of O_2 activation in the nonheme iron DNA repair metalloenzyme alpha-ketoglutarate-dependent dioxygenase (AlkB) and the movement of electrons within the active site during the reactions progression [21], (ii) how differences in model and method (QM versus QM/MM) choice influence the mechanistic PES of Uroporphyrinogen Decarboxylase III (UROD), a key "protein-only" enzyme in porphyrin biosynthesis, [24, 25], and (iii) the influence of solvation and the use of "broken" PESs in investigations on the mechanism of the NAD^+ and Mn^{2+} dependent metalloenzyme 6-phospho-α-glucosidase (GlvA) [26].

2. Multistate Reactivity Potential Energy Surfaces: Activation of O_2 in AlkB

An organisms "blueprint" is encoded in its DNA, in particular the sequence of its nucleobases. Thus, the fidelity and integrity of these DNA components is essential to its proper functioning and genetic transmission. However, both internal (e.g., metabolic byproducts) and external (e.g.,

radiation) factors can damage nucleobases via depurination, oxidation or deamination processes [27–34]. Alternatively, they may be alkylated at their oxygen or nitrogen centres [35]. Consequently, cells have developed several approaches to mediate or repair such alkylation damage [36–40]. In particular, the AlkB family of enzymes have evolved a novel mechanism by which they oxidatively dealkylate several such damaged nucleobases, specifically 1-meA, 3-meC, 3-meT, and 1-meG (Scheme 1).

The AlkB proteins belong to the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily but are the only members to catalyze oxidative dealkylation [41, 42]. More specifically, they are nonheme iron containing proteins in which the Fe(II) ion is ligated by a facial triad of residues consisting of two histidines (His131, and His187) and an aspartate (Asp133) [43, 44]. In addition, they also require α-ketoglutarate (α-KG) as a cosubstrate [41, 42] that is converted to succinate during the mechanistic activation of O_2.

The overall mechanism of AlkB is thought to occur in three stages: [21] (i) O_2 activation to give an ferryl-oxo ($Fe^{IV}= O$) moiety, (ii) reorientation of the $Fe^{IV} = O$ oxygen from the axial to an equatorial position, and (iii) oxidative dealkylation of the alkylated nucleobase [45]. It is noted that a reorientation of the $Fe^{IV} = O$ has been suggested to occur in other enzymes of this superfamily such as the clavaminate synthase [46], while the dealkylation step has been previously computationally studied [21].

The overall proposed mechanism by which the O_2 moiety is activated with concomitant oxidation of α-KG to succinate and CO_2 is shown in Scheme 2. In particular, following binding of α-KG and O_2, a self-redox process results in formation of a ferric-superoxide Fe^{III}-$O_2^{\bullet-}$ species [47, 48]. This then nucleophilically attacks the α-keto group of α-KG, resulting in its decarboxylation (i.e., loss of CO_2) and the formation of succinate with concomitant generation of an "axial-positioned" $Fe^{IV} = O$ intermediate. However, the exact details by which the electrons released from the α-KG cosubstrate are used in the O_2 activation process, as well as the influence of the systems electronic state upon the choice of preferred pathway, remain unclear.

We examined [21] this key stage using a cluster/DFT-based computational model. In particular, optimized geometries were obtained at the PB-SCRF-B3LYP/LACVP(d) level of theory with a dielectric constant of 4.0. Frequency calculations were performed on these structures in order to characterize them as minima or transition states. Relative total energies were calculated via single-point energy
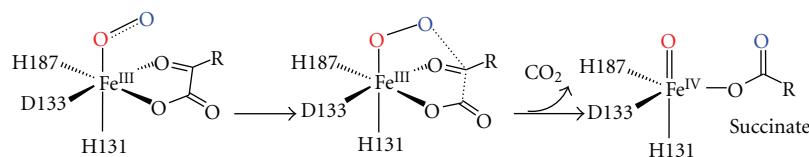
Scheme 2: Proposed O_2 activation mechanism as catalyzed by AlkB [21].
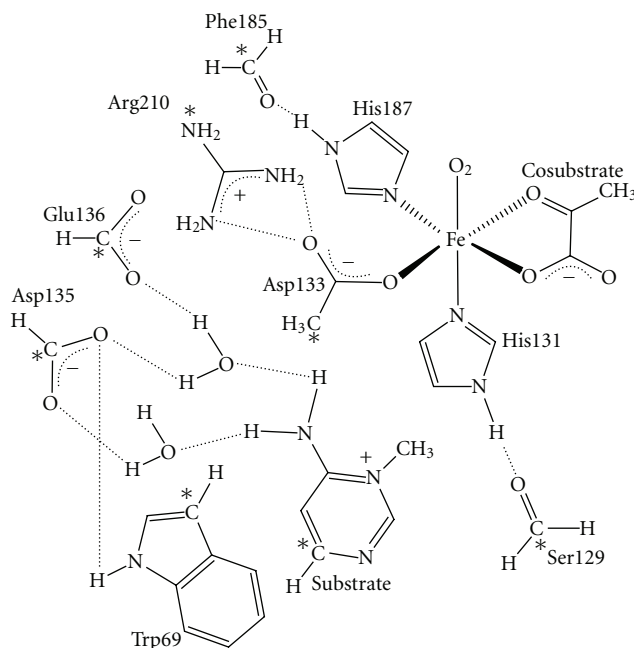
Scheme 3: Chemical model used in the study on the mechanism of AlkB. The substrate and cosubstrate (α-KG) are modeled as 3-methyl-4-amino pyrimidinyl cation and pyruvate, respectively. Atoms marked by an * were held fixed at their crystal structure (PDB: 2FD8) positions [21].

calculations at the PB-SCRF/B3LYP/LACV3P+(d,p) level of theory based on the above geometries. All calculations were performed using the *Gaussian 03* [49] and *Jaguar 5.5* [50] software packages. For the chemical model, the coordinates of the key active site residues and groups were obtained from a crystal structure of AlkB (PDB: 2FD8) and is shown in Scheme 3 [45]. In order to maintain integrity of the model, a minimum number of atoms remote from the reactive regions were held fixed at their crystal structure positions. Complete computational details are provided in the article by Liu et al. [21].

The potential energy surfaces obtained for activation of O_2 in AlkB to give the highly reactive oxo- and oxyl-type intermediates (**IC4**) are shown in Figure 2. It should be noted that the oxidation state of the Fe centre in each complex was in part determined by calculating its spin density [21]. In addition, in order to provide further insights into electron transfer processes, a stylized electronic representation of the bond making/breaking region is also shown for the minimum energy path. Assuming the commonality of spin changes in transition metal chemistry, this minimum energy path involves spin inversion from the septet to quintet surface [23].

The preferred mode for O_2 binding to the Fe(II) centre is end-on (η^1). Four overall spin combinations are possible for the resulting dioxygen-bound α-KG$-$Fe(II) active-site complex (**RC**): singlet, triplet, quintet, and septet. The overall spin-triplet complex, ^{3}**RC**, was chosen to be the reference energy level throughout this work.

In both the singlet and triplet states, O_2 binding has occurred without electron transfer from the Fe(II) to O_2 moiety. The spin-singlet state, ^{1}**RC**, lies significantly higher in energy than ^{3}**RC** by 57.1 kJ mol^{-1} (not shown). This is likely due to less favourable antiferromagnetic alignment of the unpaired electrons of $^1O_2(^1\Sigma_g^+)$ with the low-spin Fe(II) centre in ^{1}**RC**, compared to $^3O_2(^3\Sigma_g^-)$ with the high-spin Fe(II) centre in ^{3}**RC** [21]. In contrast, in the quintet and septet spin states, O_2 binding occurs with charge transfer. That is, a high-spin sextet Fe(III) ion ($S = 5/2$) is formed with concomitant reduction of $^3O_2$ (valence electron configuration $\sigma_x^2\pi_y^2\pi_z^2\pi_y^{*1}\pi_z^{*1}\sigma_x^{*0}$ in a fragment orbital picture) to $O_2^{\bullet-}$ ($\sigma_x^2\pi_y^2\pi_z^2\pi_y^{*2}\pi_z^{*1}\sigma_x^{*0}$). Importantly, the spin-septet ^{7}[Fe^{III}-$O_2^{\bullet-}$] complex (^{7}**RC**) lies 12.5 kJ mol^{-1} *lower* in energy than ^{3}**RC**, while the spin-quintet complex ^{5}**RC** is just 5.7 kJ mol^{-1} higher in energy than ^{3}**RC**.
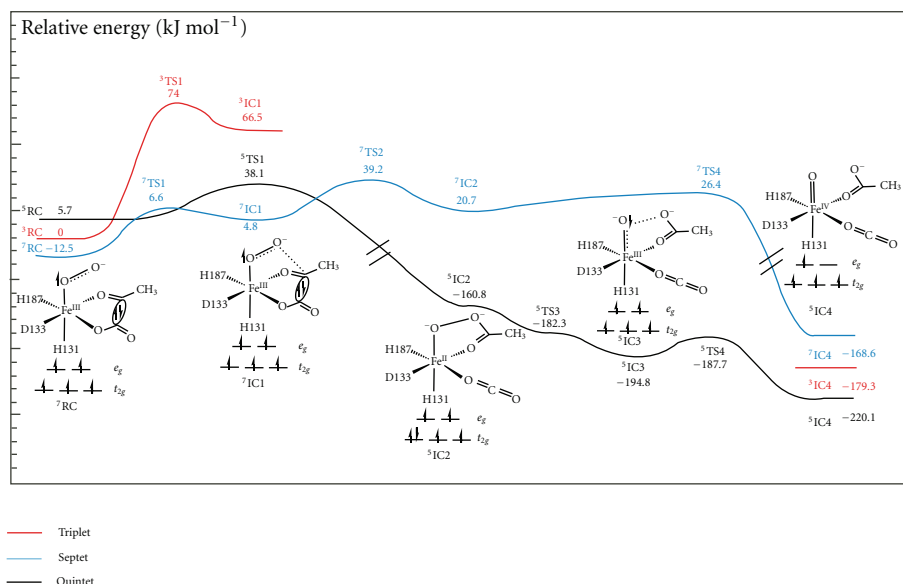
Figure 2: Potential energy surfaces for activation of oxygen through formation of the ferryl-oxo $Fe^{IV}=O$ and ferric-oxyl $Fe^{III}-O^{\bullet}$ compounds.

While ^{3}**RC** was chosen as the "reference relative energy" for all complexes, the mechanistic details of the spin-triplet PES are not discussed herein due to its considerably higher activation barriers compared to the spin-quintet and spin-septet surfaces (**Figure 2**).

For ^{7}**RC**, the lowest energy reactant complex, the first step is nucleophilic attack of the ferric-superoxide ($^7[Fe^{III}-O_2^{\bullet-}]$) group at the pyruvate's C2-carbon. It is noted that the ability of $Fe^{III}-O_2^{\bullet-}$ to act as a strong nucleophile has been previously investigated by both experimental [51–54] and theoretical [55–58] studies. This step proceeds via ^{7}**TSI** with a barrier of only 19.1 kJ mol^{-1} to give the intermediate complex ^{7}**IC1** lying just 1.8 kJ mol^{-1} lower in energy than ^{7}**TSI** (**Figure 2**). In ^{7}**IC1**, the O_2 moiety forms a peroxy bridge between Fe(III) and the pyruvate cosubstrates C2-centre [21]. In the subsequent step, the cosubstrate's C2–COO$^-$ bond is cleaved with the electrons formally moving into the Fe-peroxo moiety. As a result, the iron is reduced to Fe(II), while the peroxy moiety now has a charge of −2, and consequently, its O–O bond lengthens considerably. This step proceeds via ^{7}**TS2** with a barrier of 34.4 kJ mol^{-1} to give the ferric-peroxide intermediate ^{7}**IC2** lying 15.9 kJ mol^{-1} higher than ^{7}**IC1**. The peroxide O–O bond is then cleaved with concomitant two-electron oxidation of the Fe(II) centre to Fe(IV) via ^{7}**TS4** at a cost of just 5.7 kJ mol^{-1} with respect to ^{7}**IC2**. The resulting Fe(IV)-oxo-type product complex ^{7}**IC4** is considerably lower in energy than ^{3}**RC** by 168.6 kJ mol^{-1}. Thus, overall, this pathway is enzymatically feasible.

On the quintet-state PES, the first step is also found to be nucleophilic attack of the $^5[Fe^{III}-O_2^{\bullet-}]$ group at the cosubstrate's C2 centre. However, this now occurs with decarboxylation of the pyruvate and concomitant reduction of both Fe(III) and $O_2^{\bullet-}$ to Fe(II) and O_2^{2-}, respectively. Importantly, this reaction proceeds via ^{5}**TSI** with a barrier of only 32.4 kJ mol^{-1} to give the energetically very low lying intermediate ^{5}**IC2** (−160.8 kJ mol^{-1}). Based on their structural similarity, it is suggested that ^{7}**IC1** could be regarded as the reaction-coordinate equivalent of ^{5}**TSI** [21]. Subsequently, ^{5}**IC2** can undergo a stepwise two-electron transfer from Fe(II) to the O–O unit [21]. The first electron transfer occurs via ^{5}**TS3** without a barrier. It is noted that optimized structures are obtained by minimization of the electronic energy at 0 K. As a result, on flat PESs the inclusion of energy corrections via, for example, single-point calculations, enthalpy, or Gibb's free energy corrections can result in TSs having lower relative energies than the reactant or product which it interconnects. This is typically taken to indicate that the particular reaction essentially occurs without a barrier. In the resulting intermediate ^{5}**IC3**, the electron has transferred from the Fe(II) into the σ_{o-o}^* orbital. As a result, the peroxo moiety, in a fragment MO picture, has the electronic configuration $\sigma_x^2 \pi_y^2 \pi_z^2 \pi_y^{*2} \pi_z^{*2} \sigma_x^{*1}$. The second electron transfer proceeds via ^{5}**TS4** at a cost of 7.1 kJ mol^{-1} to give the ferryl-oxo product complex. ^{5}**IC4** is the ferryl-oxo compound and is found to lie 25.3 kJ mol^{-1} lower in energy than ^{5}**IC3**. This one-electron oxidation of Fe(III) again promotes an electron into to the σ_{o-o}^* orbital, thus giving the peroxo moiety the highly unstable electron configuration $\sigma_x^2 \pi_y^2 \pi_z^2 \pi_y^{*2} \pi_z^{*2} \sigma_x^{*2}$ and resulting in scission of the O–O linkage. The complex ^{5}**IC4** lies significantly lower in energy than ^{3}**RC** by 220.1 kJ mol^{-1}. Importantly, however, it is also markedly lower in energy than both ^{7}**IC4** and ^{3}**IC4** by at least 40.8 kJ mol^{-1} (**Figure 2**).

Thus, by elucidating the pathways for the various possible spin states one can see that the lowest energy product ^{5}**IC4** may in fact be formed via several possible low energy routes. In particular, it may occur either at the beginning of the mechanism via an initial spin inversion from ^{7}**RC** to ^{5}**RC**, and thus providing a pathway with an overall barrier of 50.6 kJ mol^{-1}. Alternatively, spin inversion may occur along the pathway, for example, as ^{7}**IC1** reacts and proceeds towards ^{7}**TS2** and thus to directly give ^{5}**IC2**, allowing for
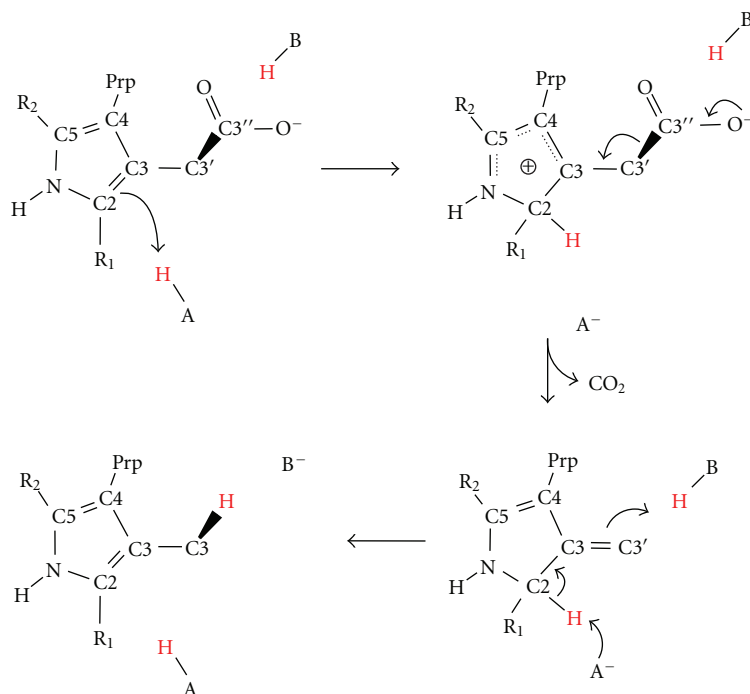
Scheme 4: Proposed general acid-base mechanism for the UROD-catalyzed decarboxylation of the acetates of URO-III [59–61]. HA and HB represent general acids.

cleavage of the peroxo O–O bond with a possible pathway barrier of less than 50.6 kJ mol^{-1}. Importantly, spin inversion would enable the mechanism to access and proceed via considerably more exoergic intermediates, for example, ^{5}IC2 lies 181.3 kJ mol^{-1} lower in energy than ^{7}IC2 [21]. The probability of SI occurring at a crossing-point is determined by the strength of spin-orbit coupling between the two states. Fortunately, as noted in the introduction, such crossing-points are not generally stationary points [23]. In addition, it is rarely found that the spin-orbit coupling is so strong that it affects the adiabatic energies of the PESs of the two states [23].

3. Effects of Explicitly Modeling the Protein Environment on the Catalytic Mechanism of Uroporphyrinogen Decarboxylase III

Porphyrin is an important biomolecule for all organisms. It plays a range of diverse key roles in, for example, proteins and enzymes involved in ligand transport, electron transfer, light harvesting, and redox mechanisms [62–65]. Within cells, its biosynthesis occurs via a multistage multienzymatic process in which Uroporphyrinogen Decarboxylase III (UROD) catalyses the first branching point, the fifth step. Specifically, it catalyses the sequential nonsymmetric decarboxylation of the four acetates of uroporphyrinogen III (URO-III) to give coproporphyrinogen III (CP-III) [62–64, 66].

UROD exists as a homodimer, and it has been shown experimentally that each active site is independent from the other [24]. Three electrostatic regions have been identified

within each active site: a negative, a polar-positive, and a hydrophobic region [62]. The negatively charged region contains an invariant aspartyl (hUROD: Asp86) that is thought to help orientate the substrate for catalysis as well as stabilize various mechanistic intermediates [62, 67]. Helping in substrate binding and recognition, the polar-positive region contains several residues that interact with the carboxylates of URO-III [67, 68]. Importantly, of these polar-positive residues, it has been found that one or more active-site arginyl residues (hUROD: Arg37, Arg41 and Arg50) are catalytically essential [62, 68–71]. While all four acetate decarboxylations are believed to involve the same catalytic residues, unfortunately, the exact mechanistic role of various active site residues remains unclear.

A general mechanism has been proposed for decarboxylation of the pyrrole-acetate and is shown in Scheme 4 [59–61]. Specifically, two acids are thought to be involved. In the first step, an acid (HA) protonates the C2 carbon of the pyrrole ring. This destabilizes the carboxylate group, resulting in cleavage of the C3′–C3″OO$^-$ bond, that is, decarboxylation. The second acid (HB) then protonates the newly formed methylene C3′ carbon, while the first acid, now in its conjugate base form (A$^-$), abstracts a proton from the–C2H$_2$$^+$–group, thus regenerating HA.

Several experimental X-ray crystal structure and kinetic studies, based on this "blueprint", have proposed possible identities for the two mechanistic acids. Specifically, from enzyme X-ray structures they obtained coupled with manual docking of the substrate, Martins et al. [64] suggested that HA is Asp86, while HB is Tyr164. In contrast, based on a X-ray crystal structure of a UROD$\cdots$product complex,
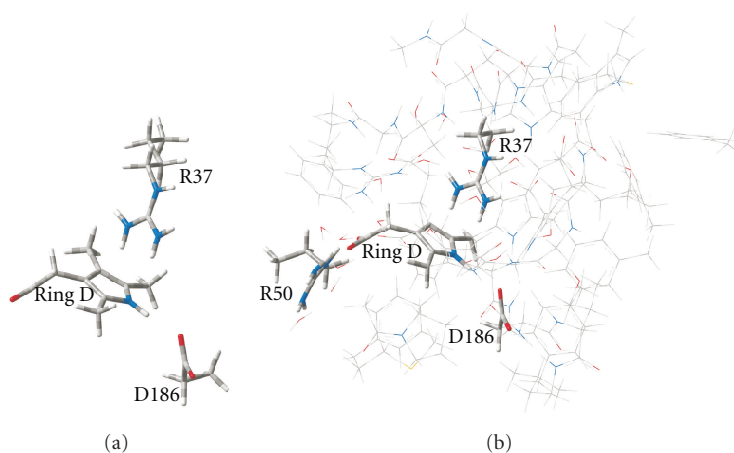
(a)   (b)

FIGURE 3: Representative images for the (a) DFT/Cluster + PCM model used by and Silva and Ramos [25] and (b) QM/MM model used by Bushnell et al. [24].

Phillips et al. [67] suggested that HA is an H_2O but were unable to conclusively determine HB. More recently, Lewis and Wolfenden, [69] using kinetics and pK_a measurements, suggested that HA and HB are Asp86 and Arg37, respectively.

Silva and Ramos [25] performed a computational investigation on the first acetate decarboxylation of URO-III as catalysed by UROD. More specifically, they used a cluster/DFT-type approach with a chemical model consisting of the aspartyl (Asp86) and arginyl (Arg37) R-groups and 1,3,4-methyl-2-acetyl pyrrole for the URO-II substrate (Figure 3). This model was derived from a crystal structure of a UROD···product complex [67]. A PCM-solvation approach was used to model the general affects of the surrounding protein environment. Using this computational model, they concluded that decarboxylation could proceed in accordance with the "blueprint" mechanism and that it was thermodynamically and enzymatically feasible with an overall barrier of $89.5\,kJ\,mol^{-1}$ [25]. Notably, this energy corresponded to the initial step, that is, proton transfer from Arg37 onto C2 of the substrate pyrrole ring. Thus, the R-group guanidinium of the active-site residue Arg37 was concluded to be both suitably positioned and capable of acting as the general acid HA that protonates the substrate's C2 centre. They further concluded that the second proton, from HB, is donated by the solvent. However, based on the experimental observation that stereochemistry is retained at the C3′ centre from which the CO_2 is lost, it has been concluded that HB must be an active site residue [59, 60, 72].

We reexamined the catalytic mechanism of UROD using an alternate computational approach. In particular, it was investigated [24] through combined QM and MM methods in the ONIOM formalism with mechanical embedding [73–81] as implemented in the *Gaussian 03* [44] program suite. Such an approach enabled us to consider the role of a second arginyl (Arg50) residue as well to take into account the anisotropic environment surrounding the active site. The resulting QM/MM chemical model included the substrate URO-III and all active site residues immediately surrounding it, that is, first-shell residues (Figure 3). In

addition, for those portions of the substrate exposed to solvent, the first solvation shell was retained. A subset of the complete model centered on the reactive region of the active site was then selected for the high-level QM treatment consisting of two arginyls (Arg37 and 50), an aspartyl (Asp86) residue, and the first substrate pyrrole that is decarboxylated [24]. Optimized geometries were obtained at the ONIOM(B3LYP/6-31G(d):AMBER94) level of theory [82]. However, in order to obtain more reliable calculated relative Gibbs free energies, single-point energy calculations on the above optimized structures were performed at the ONIOM(B3LYP/6-311+G(2df,p):AMBER94) level of theory. The free energies were obtained by adding the necessary energy corrections calculated at standard ambient temperature and pressure (SATP) as implemented in Gaussian 03.

Using this larger and more complete model, the initial step is again proton transfer from the guanidinium of Arg37 onto the C2 centre of the substrate pyrrole. However, this step now occurs via **TS1** at a cost of $43.1\,kJ\,mol^{-1}$ relative to the initial substrate-bound active site complex **RC** (Figure 4). This is approximately half the size of the barrier for this initial step obtained by Silva and Ramos [25] using a cluster/DFT approach. Experimentally, human UROD has been measured to have a k_{cat} value of $0.16\,s^{-1}$ [5, 69, 83], corresponding to an overall barrier of $77.4\,kJ\,mol^{-1}$. Thus, the smaller cluster/DFT approach would appear to give better agreement than the above larger and more extensive QM/MM-based approach. However, as summarized by Juárez et al. [84], decarboxylation of URO-III generating the 7-carboxylate intermediate, that is, the first decarboxylation, is most likely not the rate-limiting step. In fact, for several variants of UROD, the rate-limiting step appears to be the decarboxylation of the 7-carboxylate intermediate, that is, the second acetate decarboxylation. It is noted that for UROD from various species, the experimentally determined barriers for the first decarboxylation lie in the range of 8.4–$51.5\,kJ\,mol^{-1}$ [68, 84]. Thus, the QM/MM calculated barrier is in fact in better agreement with related experimentally reported values for the decarboxylation of ring D.
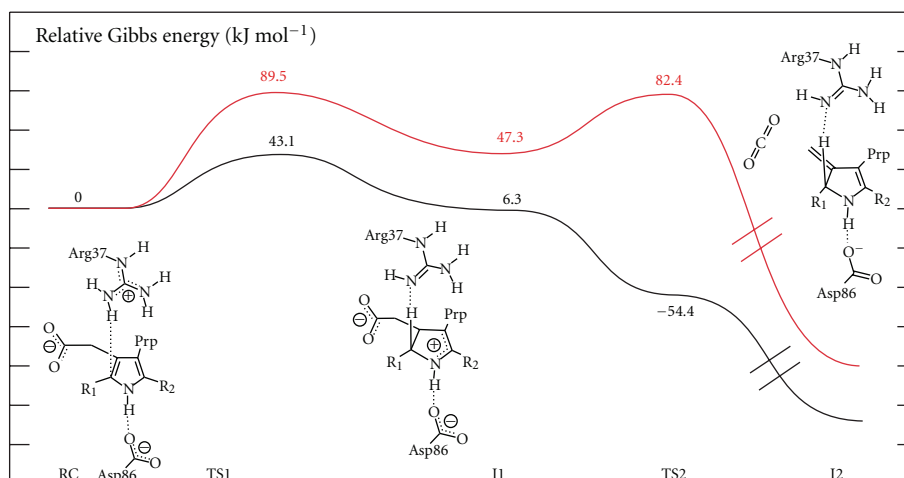
FIGURE 4: Free energy PES for the initial proton transfer and decarboxylation of ring **D** obtained using a QM/MM model (black surface with relative energies) and cluster/DFT model (red surface with relative energies) [24, 25].

The resulting C2-protonated intermediate (**I1**) lies higher in energy than **RC** by just 6.3 kJ mol^{-1}. Importantly, with the explicit addition of the protein environment (as opposed to the use of a PCM approach), a stabilizing effect occurs. This stabilizing effect is seen to reduce the relative energy of **I1** by 41.0 kJ mol^{-1} in comparison to that obtained using the cluster/DFT approach in which **I1** was calculated to lie higher in energy than **RC** by 47.3 kJ mol^{-1} [25].

Following protonation of C2 by Arg37, the next step is the decarboxylation of the acetate moiety (Scheme 4). Using a cluster/DFT-based approach, Silva and Ramos [25] obtained a barrier for this step of 82.4 kJ mol^{-1}. Furthermore, based on the optimized structures obtained, they concluded that the release of CO_2 was hindered by the hydrogen bond interaction between the carboxylate of Asp86 and the pyrrole ring's HN moiety. In contrast, using a QM/MM-based approach, we found that this occurs via **TS2** with a relative energy barrier *lower* than that of **I1** (again due to inclusion of single-point energy and Gibb's free energy corrections; see section on Multistate Reactivity PESs) [24]. That is, the loss of CO_2 from the acetate essentially occurs without a barrier. Furthermore, a more complete and explicit inclusion of the protein environment enables a lengthening of the Asp86$\cdots$C2-protonated pyrrole interaction. In turn, this leads to a destabilization of the C2-protonated pyrrole and thus enhancement of the rate of decarboxylation [25].

The subsequent and final step is protonation of the newly formed methylene carbon from which the CO_2 was lost. Unfortunately, direct comparison of the results obtained using the cluster/DFT-based [25] and QM/MM-based [24] approaches for this step is not possible. Using the former approach, the proton transferred to C′3 was proposed to originate from the solvent. In contrast, for the latter QM/MM-based study, it was found that a second arginyl active site residue (Arg50) could act as the second required mechanistic acid HB (Scheme 4) that protonates the methylene carbon. Furthermore, the barrier for this step was just 3.1 kJ mol^{-1} [24]. It is noted that the identification of an active site residue as HB is supported by experimental conclusions based on the observed retention of stereochemistry at the pyrrole-CH3 group formed [59–61].

4. Elucidating the Catalytic Redox Mechanism and "Driving Force" in 6-Phospho-α-Glucosidase (GlvA)

In aqueous solution under standard conditions, glycosidic bonds are remarkably resistant to hydrolysis [85–89]. Glycoside hydrolases (GH) are the family of enzymes that catalyse their hydrolytic cleavage within organisms and, in fact, are amongst some of the most efficient enzymes known [90–100]. In general, these enzymes are highly stereospecific; however, the GH4 subfamily is able to catalyse the hydrolysis of either or both α- and β-glycosidic bonds [101, 102]. The enzyme 6-phospho-α-glucosidase (GlvA) is a member of this subfamily that uses NAD^+ and a divalent Mn^{2+} as cofactors in its redox mechanism of glycosidic bond cleavage [97, 103, 104]. In particular, the mechanism has been proposed to proceed via two half-reactions shown in Scheme 5 [99, 105–107]. The first is thought to be initiated by proton transfer from the C3-OH group to an Mn^{2+} bound hydroxyl and hydride transfer from C3–H to the NAD^+ moiety, ultimately resulting in cleavage of the glycosidic bond at C1. The cleaved product moiety ROH is then replaced within the active site by a water molecule. The second half-reaction is believed to essentially be the reverse of the first with the H_2O moiety in place of the ROH.

In order to investigate the proposed [99, 105–107] catalytic mechanism of GlvA, we employed [26] a cluster/DFT-based approach in which a large active site model was used with the surrounding protein simply modeled using a PCM-based solvation method. In particular, optimized geometries were obtained at the B3LYP/6-31G(d) 5D level of theory. Relative energies were obtained at the B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) 5D level of theory (i.e., gas-phase). Corrections for the surrounding environment were
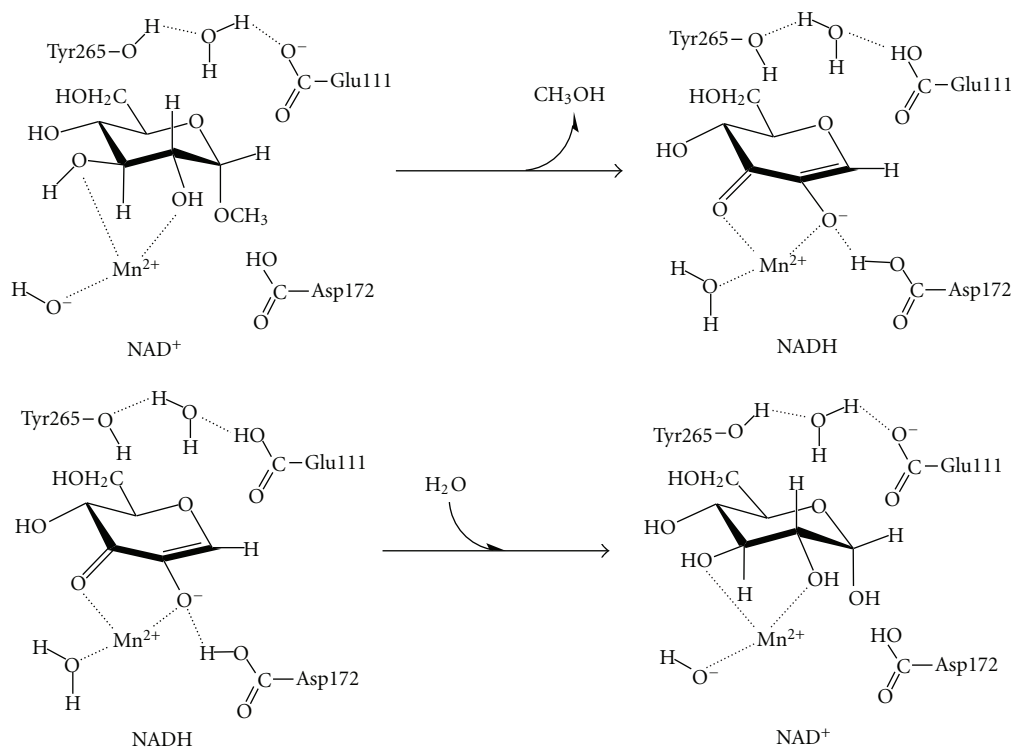
Scheme 5: Schematic illustration of the two half-reactions in the proposed catalytic mechanism of GlvA [99, 105–107].

obtained at the same level of theory using the integral equation formalism (IEF-) PCM method with a dielectric constant of $\varepsilon = 4.0$; a value typically used to model the polarity of an internal protein environment [26]. The chemical model used (Scheme 6) was derived from an X-ray crystal structure of GlvA cocrystallized with the substrate analog 6-phospho-β-glucose bound within its active site (PDB ID: 1U8X) [100]. The coordinates of proposed key mechanistic residues were then extracted, and a minimum number of atoms remote from the reaction region were held fixed at these positions in order to maintain integrity of the model during computations. Full computational details are provided in [22].

The PESs obtained for the two half-reactions are illustrated in Figure 5. The relative energies given in parenthesis were obtained in the gas phase, that is, without use of the PCM-based approach to model the general effects of the surrounding polar protein environment, and are included for comparison (see below). The catalytic mechanism of GlvA is initiated by an oxidation of the glucopyranose ring at the 3-position. Specifically, the 3′–OH group first transfers its proton to the metal bound hydroxide ligand $(Mn^{2+} \cdots OH^-)$ to give a metal-bound water and a now negatively charged 3′–O$^-$ moiety on the sugar ring (IC1). This step occurs with a barrier of $22.7\,kJ\,mol^{-1}$ when within the protein environment ("solution phase"), while the resulting intermediate IC1 lies $46.3\,kJ\,mol^{-1}$ lower in energy than the RC. This is then followed by a hydride transfer from the 3′C–H group to the pro-R face of the C4 centre of NAD$^+$ [26] cofactor via TS2 at a cost of $80.8\,kJ\,mol^{-1}$ with

respect to IC1. This results in the formation of the 3-keto intermediate IC2 which lies only slightly higher in energy than IC1 by $7.2\,kJ\,mol^{-1}$.

With the formation of IC2, a double bond has now been introduced into the substrate via the formation of a keto group. In the subsequent steps, the double bond is then effectively shifted around the ring via a series of keto-enol tautomerizations coupled with proton transfers. First, a glutamyl-tyrosyl catalytic diad deprotonates the 2′C–H group of IC2 via TS3 to form the enolate $(2'C = 3'C–O^-)$ containing intermediate IC3 (Figure 5). The barrier for this process is $54.0\,kJ\,mol^{-1}$, while IC3 lies $14.6\,kJ\,mol^{-1}$ lower in energy than RC. This is subsequently followed by proton abstraction from the C2–OH group by an adjacent aspartyl (Asp172) residue. Simultaneously, however, the Asp172 donates its proton to the leaving CH$_3$O$^-$ moiety. Importantly, it is in this step that the actual cleavage of the glycosidic bond occurs via TS4 at cost of only $58.9\,kJ\,mol^{-1}$ with respect to IC3. NBO and second-order perturbation analyses were used to analyse the "driving force" that leads to this greatly reduced barrier for heterolytic glycosidic bond cleavage and the regioselectivity of GlvA [22]. In particular, it was found that in the first half-reaction, electronic destabilizing effects within the sugar ring are enhanced due to several electron delocalizations that strongly favour the heterolytic cleavage of glycosidic bonds. Furthermore, this effect was greater in the case of axial (α) compared to equatorial (β) glycosidic bonds. It is noted that the first half-reaction is exothermic as the resulting product complex IC4$\cdots$MeOH lies $29.9\,kJ\,mol^{-1}$, respectively, lower in energy than RC.
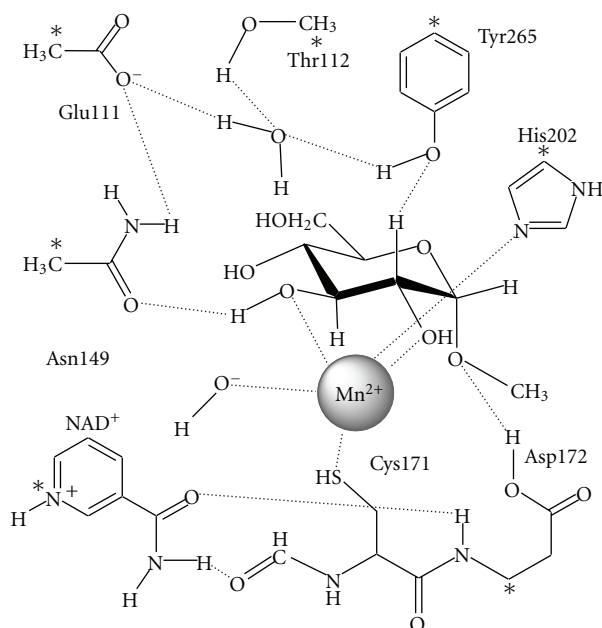
SCHEME 6: The substrate-bound active site model derived from the X-ray crystal structure PDB ID: 18UX and used in our computational investigation on GlvA. Atoms marked by an * were held fixed at their crystal structure positions.

For the second half-reaction, the cleaved MeOH moiety must first be replaced by a water to give $IC4\cdots H_2O$. By comparing the energies of $[IC4\cdots MeOH + H_2O]$ versus $[IC4\cdots H_2O + MeOH]$, it is found that this replacement is slightly exothermic by $7.3\,kJ\,mol^{-1}$. Thus, while this exchange requires a "splitting" of the calculated overall PES, one is still able to maintain an energetic relationship between the two surfaces. With the formation of $IC4\cdots H_2O$ the second half-reaction is effectively the reverse of the first half-reaction [26]. For instance, it is initiated by nucleophilic attack of the water oxygen at the C1 centre of IC4. This occurs via TS5 at a cost of $77.7\,kJ\,mol^{-1}$ in which the attacking H_2O simultaneously transfers a proton to the Asp172 residue which itself transfers its proton onto the $C2$–O^- oxyanion centre. Similar to the analogous complex IC3, the resulting intermediate IC5 lies only slightly lower in energy than RC by $18.4\,kJ\,mol^{-1}$. In the subsequent step, the C2 centre is reprotonated by the catalytic tyrosyl-glutamyl diad via TS6 at a cost of $79.5\,kJ\,mol^{-1}$ to give the keto intermediate IC6 lying $39.6\,kJ\,mol^{-1}$ lower in energy than RC. The last two steps in the overall mechanism begin with a hydride transfer from the NADH moiety onto the C3 centre via TS7, with a barrier of $72.2\,kJ\,mol^{-1}$, to give the oxyanionic intermediate IC7. Notably, both IC7 and the preceding intermediate IC6 are essentially thermoneutral with their analogous first half-reaction complex IC2 and IC1, respectively (see Figure 5). The next and final step is a proton transfer from the metal-bound H_2O onto the newly formed $C3$–O^- group via TS8 at a cost of $98.6\,kJ\,mol^{-1}$. The overall mechanism is slightly exothermic, with the product-bound active site complex PC lying just $4.0\,kJ\,mol^{-1}$ lower in energy than RC.

Comparison of the above "solution-phase" relative energies with the gas-phase energies enables one to gain insight the role and effect of the general protein environment. It can be seen from Figure 5 that the effect of the polar environment is not systematic. For example, inclusion of the polar environment stabilizes IC1 and TS7 by 17.3 and $42.3\,kJ\,mol^{-1}$, respectively, but destabilizes IC2 and TS4 by 0.5 and $10.0\,kJ\,mol^{-1}$. In general, however, larger effects are observed for transition structures; the observed absolute changes in relative energies for TSs range from 3.3–$42.3\,kJ\,mol^{-1}$ while for the intermediates and product complexes, they range from 0.5–$27.3\,kJ\,mol^{-1}$.

More importantly, however, the inclusion of the general protein environment can have marked effect on the catalytic mechanism. For example, in the gas phase the overall mechanism is endothermic with the PC lying $11.4\,kJ\,mol^{-1}$ higher in energy than RC. However, in the "solution-phase", it is now slightly exothermic by $4.0\,kJ\,mol^{-1}$ (Figure 5). Moreover, with the inclusion of the protein environment the rate-limiting step has changed. Specifically, in the gas phase, the proton transfer from the Mn^{2+}-bound water to the $3'C$–O^- group via TS8 represents the rate-limiting step with a reaction barrier relative to RC of $78.3\,kJ\,mol^{-1}$. However, with inclusion of the protein environment (i.e., use of a PCM-based approach), the energy of TS8 is reduced significantly by $\sim26.2\,kJ\,mol^{-1}$ to $52.1\,kJ\,mol^{-1}$ respect to RC. As a consequence, TS6 ($61.1\,kJ\,mol^{-1}$) now has the highest energy relative to RC. Thus, reprotonation of C2 by the catalytic tyrosyl-glutamyl diad has become the rate-limiting step. In agreement with experiment, glycosidic bond cleavage is not the rate-controlling step.
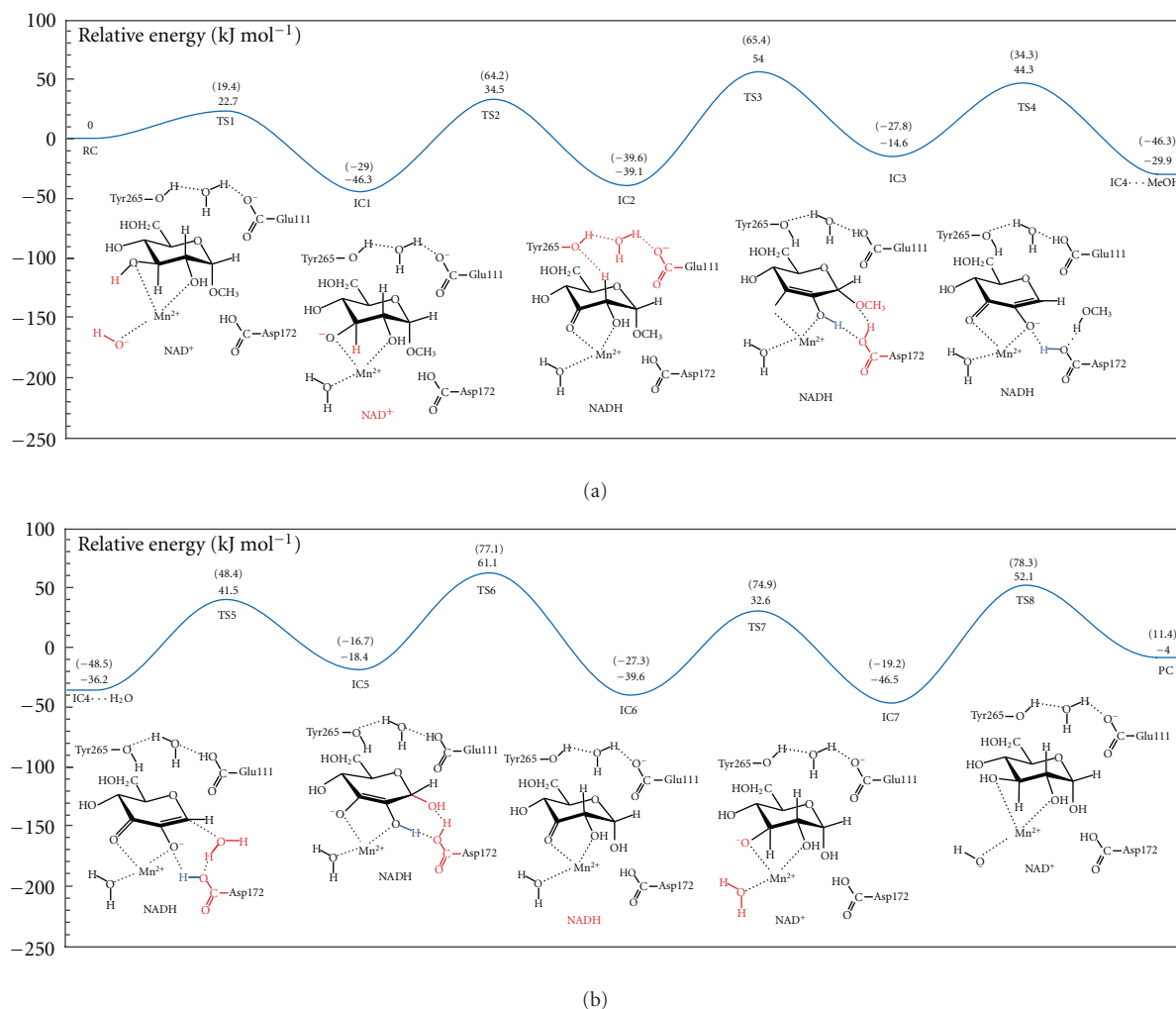
(a)

(b)

FIGURE 5: Potential energy surfaces for the catalytic mechanism of redox hydrolysis of αGlc1Me by GlvA. Single-point energies calculated in solvation and gas phase (parenthesis).

5. Summary

Potential energy surfaces (PESs) are powerful tools for elucidating the catalytic mechanisms of enzymes. Indeed, they can be more than a simple indication of a path of sequential mechanistic structures and their energetic relationships. In this paper the application of the tools and techniques of computational chemistry in combination with PES investigations has been described. Through a review of several of our recently published studies on enzymatic systems, how such an approach may be applied in order to provide deeper insights into the movement of electrons during the course of a reaction, the mechanistic role of active site residues and the surrounding environment, and the utility of multistate PESs in metalloenzymes has been highlighted. In addition, some of the challenges that these studies may encounter, such as appropriate choice of computational model on the reliability and accuracy of a PES, have also been discussed.

For example, the catalytic activation of O_2 by AlkB to give the highly reactive oxo- and oxyl-type intermediates was examined for several possible spin states. It was found that

at least in the initial steps of the process, multiple states were close in energy with each other and, furthermore, within the enzymatic thermochemical limit. As a result, and considering the likelihood of spin changes in transition metal chemistry, several minimum energy pathways to give the lowest energy product complex $^5\mathbf{IC4}$ were identified, each involving spin inversion (SI) from a septet to quintet PES [23]. Importantly, SI enables the catalytic mechanism of AlkB to access and proceed via highly exoergic intermediates.

For the enzyme UROD, a comparison was made of its catalytic mechanism as elucidated using a combined QM and MM method in the ONIOM formalism with mechanical embedding and a cluster/DFT+PCM approach. [25] Use of a QM/MM approach enables one to more directly and extensively model the anisotropic environment surrounding an enzyme active site. In the case of UROD, with the QM/MM-based approach, it was found that the stabilizing effect of the surrounding environment provided a significant reduction in both the calculated rate-limiting proton transfer as well as the barrier for acetate decarboxylation. In fact, the more accurate inclusion of the environment enabled us

[19] to obtain an energy for the rate-limiting barrier in good agreement with related experimentally determined values.

The applicability of cluster/DFT-based approaches to the study of enzymatic systems was examined further using the case of the glycosidic hydrolase GlvA. In particular, such an approach was shown to be able to provide reliable and accurate insights when used in combination with a well-chosen chemical model. Inclusion of the general effects (polarity) of the environment surrounding the active site was shown to be important in choosing a suitable model. Indeed, its inclusion had significant effects on the mechanism obtained such as shifting it from an overall endothermic to exothermic process. More importantly, however, the rate-limiting step changed. Specifically, in the gas phase, a proton transfer from a metal-bound water to the $3'$C–O$^-$ group represented the rate-limiting step. However, with the inclusion of the polar protein environment using a PCM-based approach, the rate-limiting step was instead protonation of the sugar ring's C2 centre by the catalytic tyrosyl-glutamyl diad.

Acknowledgments

References

[1] D. Young, *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*, Wiley-Interscience, New York, NY, USA, 2001.

[2] P. E. M. Siegbahn and F. Himo, "Recent developments of the quantum chemical cluster approach for modeling enzyme reactions," *Journal of Biological Inorganic Chemistry*, vol. 14, no. 5, pp. 643–651, 2009.

[3] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons, New York, NY, USA, 2002.

[4] P. E. M. Siegbahn and T. Borowski, "Modeling enzymatic reactions involving transition metals," *Accounts of Chemical Research*, vol. 39, no. 10, pp. 729–738, 2006.

[5] F. Himo, "Quantum chemical modeling of enzyme active sites and reaction mechanisms," *Theoretical Chemistry Accounts*, vol. 116, no. 1–3, pp. 232–240, 2006.

[6] A. D. Becke, "A new mixing of Hartree-Fock and local density-functional theories," *The Journal of Chemical Physics*, vol. 98, no. 2, pp. 1372–1377, 1993.

[7] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.

[8] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, vol. 37, no. 2, pp. 785–789, 1988.

[9] P. E. M. Siegbahn, "The performance of hybrid DFT for mechanisms involving transition metal complexes in enzymes," *Journal of Biological Inorganic Chemistry*, vol. 11, no. 6, pp. 695–701, 2006.

[10] B. Mennucci and J. Tomasi, "Continuum solvation models: a new approach to the problem of solute's charge distribution and cavity boundaries," *Journal of Chemical Physics*, vol. 106, no. 12, pp. 5151–5158, 1997.

[11] J. Tomasi, B. Mennucci, and E. Cancès, "The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level," *Journal of Molecular Structure: THEOCHEM*, vol. 464, no. 1–3, pp. 211–226, 1999.

[12] L. Hu, P. Söderhjelm, and U. Ryde, "On the convergence of QM/MM energies," *Journal of Chemical Theory and Computation*, vol. 7, no. 3, pp. 761–777, 2011.

[13] Y. Zhang, J. Kua, and J. A. McCammon, "Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies," *Journal of Physical Chemistry B*, vol. 107, no. 18, pp. 4459–4463, 2003.

[14] M. Klähn, S. Braun-Sand, E. Rosta, and A. Warshel, "On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions," *Journal of Physical Chemistry B*, vol. 109, no. 32, pp. 15645–15650, 2005.

[15] H. M. Senn and W. Thiel, "QM/MM methods for biomolecular systems," *Angewandte Chemie International Edition*, vol. 48, no. 7, pp. 1198–1229, 2009.

[16] J. Llano and J. W. Gauld, *Mechanistics of Enzyme Catalysis: From Small to Large Active-Site Models*, vol. 2, Wiley-VCH, Weinheim, Germany, 2010.

[17] O. Acevedo and W. L. Jorgensen, "Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions," *Accounts of Chemical Research*, vol. 43, no. 1, pp. 142–151, 2010.

[18] M. Dal Peraro, P. Ruggerone, S. Raugei, F. L. Gervasio, and P. Carloni, "Investigating biological systems using first principles Car-Parrinello molecular dynamics simulations," *Current Opinion in Structural Biology*, vol. 17, no. 2, pp. 149–156, 2007.

[19] Y. Hagiwara, M. J. Field, O. Nureki, and M. Tateno, "Editing Mechanism of Aminoacyl-tRNA synthetases operates by a hybrid ribozyme/protein catalyst," *Journal of the American Chemical Society*, vol. 132, no. 8, pp. 2751–2758, 2010.

[20] J. Blumberger, "Free energies for biological electron transfer from QM/MM calculation: method, application and critical assessment," *Physical Chemistry Chemical Physics*, vol. 10, no. 37, pp. 5651–5667, 2008.

[21] H. Liu, J. Llano, and J. W. Gauld, "A DFT study of nucleobase dealkylation by the DNA repair enzyme AlkB," *Journal of Physical Chemistry B*, vol. 113, no. 14, pp. 4887–4898, 2009.

[22] S. Shaik, H. Hirao, and D. Kumar, "Reactivity of high-valent iron-oxo species in enzymes and synthetic reagents: a tale of many states," *Accounts of Chemical Research*, vol. 40, no. 7, pp. 532–542, 2007.

[23] R. Poli and J. N. Harvey, "Spin forbidden chemical reactions of transition metal compounds. New ideas and new computational challenges," *Chemical Society Reviews*, vol. 32, no. 1, pp. 1–8, 2003.

[24] E. A. C. Bushnell, E. Erdtman, J. Llano, L. A. Eriksson, and J. W. Gauld, "The first branching point in porphyrin biosynthesis: a systematic docking, molecular dynamics and quantum mechanical/molecular mechanical study of substrate binding and mechanism of uroporphyrinogen-III decarboxylase," *Journal of Computational Chemistry*, vol. 32, no. 5, pp. 822–834, 2011.

[25] P. J. Silva and M. J. Ramos, "Density-functional study of mechanisms for the cofactor-free decarboxylation performed

by uroporphyrinogen III decarboxylase," *Journal of Physical Chemistry B*, vol. 109, no. 38, pp. 18195–18200, 2005.

[26] W. Huang, J. Llano, and J. W. Gauld, "Redox mechanism of glycosidic bond hydrolysis catalyzed by 6-phospho-α-glucosidase: a DFT study," *Journal of Physical Chemistry B*, vol. 114, no. 34, pp. 11196–11206, 2010.

[27] R. De Bont and N. van Larebeke, "Endogenous DNA damage in humans: a review of quantitative data," *Mutagenesis*, vol. 19, no. 3, pp. 169–185, 2004.

[28] Y. Wang, "Bulky DNA lesions induced by reactive oxygen species," *Chemical Research in Toxicology*, vol. 21, no. 2, pp. 276–281, 2008.

[29] J. S. Taylor, "Unraveling the molecular pathway from sunlight to skin cancer," *Accounts of Chemical Research*, vol. 27, no. 3, pp. 76–82, 1994.

[30] H. Kamiya, S. Iwai, and H. Kasai, "The (6-4) photoproduct of thymine-thymine induces targeted substitution mutations in mammalian cells," *Nucleic Acids Research*, vol. 26, no. 11, pp. 2611–2617, 1998.

[31] M. Lysetska, A. Knoll, D. Boehringer, T. Hey, G. Krauss, and G. Krausch, "UV light-damaged DNA and its interaction with human replication protein A: an atomic force microscopy study," *Nucleic Acids Research*, vol. 30, no. 12, pp. 2686–2691, 2002.

[32] W. L. Neeley and J. M. Essigmann, "Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products," *Chemical Research in Toxicology*, vol. 19, no. 4, pp. 491–505, 2006.

[33] E. C. Friedberg, L. D. McDaniel, and R. A. Schultz, "The role of endogenous and exogenous DNA damage and mutagenesis," *Current Opinion in Genetics and Development*, vol. 14, no. 1, pp. 5–10, 2004.

[34] T. Lindahl, "Instability and decay of the primary structure of DNA," *Nature*, vol. 362, no. 6422, pp. 709–715, 1993.

[35] F. Drabløs, E. Feyzi, P. A. Aas et al., "Alkylation damage in DNA and RNA—repair mechanisms and medical significance," *DNA Repair*, vol. 3, no. 11, pp. 1389–1407, 2004.

[36] Y. Mishina, E. M. Duguid, and C. He, "Direct reversal of DNA alkylation damage," *Chemical Reviews*, vol. 106, no. 2, pp. 215–232, 2006.

[37] B. Sedgwick, P. A. Bates, J. Paik, S. C. Jacobs, and T. Lindahl, "Repair of alkylated DNA: recent advances," *DNA Repair*, vol. 6, no. 4, pp. 429–442, 2007.

[38] T. Lindahl, B. Sedgwick, M. Sekiguchi, and Y. Nakabeppu, "Regulation and expression of the adaptive response to alkylating agents," *Annual Review of Biochemistry*, vol. 57, pp. 133–157, 1988.

[39] A. K. McCullough, M. L. Dodson, and R. S. Lloyd, "Initiation of base excision repair: glycosylase mechanisms and structures," *Annual Review of Biochemistry*, vol. 68, pp. 255–285, 1999.

[40] P. J. O'Brien and T. Ellenberger, "Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines," *Biochemistry*, vol. 42, no. 42, pp. 12418–12429, 2003.

[41] S. C. Trewick, T. F. Henshaw, R. P. Hausinger, T. Lindahl, and B. Sedgwick, "Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage," *Nature*, vol. 419, no. 6903, pp. 174–178, 2002.

[42] P. Ø. Falnes, R. F. Johansen, and E. Seeberg, "AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli," *Nature*, vol. 419, no. 6903, pp. 178–182, 2002.

[43] T. Duncan, S. C. Trewick, P. Koivisto, P. A. Bates, T. Lindahl, and B. Sedgwick, "Reversal of DNA alkylation damage by two human dioxygenases," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 26, pp. 16660–16665, 2002.

[44] L. Aravind and E. V. Koonin, "The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases," *Genome biology*, vol. 2, no. 3, Article ID RESEARCH0007, 2001.

[45] B. Yu, W. C. Edstrom, J. Benach et al., "Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB," *Nature*, vol. 439, no. 7078, pp. 879–884, 2006.

[46] Z. Zhang, J.-S. Ren, K. Harlos, C. H. McKinnon, I. J. Clifton, and C. J. Schofield, "Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centred rearrangements," *FEBS Letters*, vol. 517, no. 1–3, pp. 7–12, 2002.

[47] R. Nakajima and I. Yamazaki, "The mechanism of oxyperoxidase formation from ferryl peroxidase and hydrogen peroxide," *Journal of Biological Chemistry*, vol. 262, no. 6, pp. 2576–2581, 1987.

[48] G. I. Berglund, G. H. Carlsson, A. T. Smith, H. Szöke, A. Henriksen, and J. Hajdu, "The catalytic pathway of horseradish peroxidase at high resolution," *Nature*, vol. 417, no. 6887, pp. 463–468, 2002.

[49] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., *Gaussian 03, Revision D.02*, Gaussian, Inc., Wallingford, Conn, USA, 2004.

[50] *Jaguar*, Schrodinger, L. L. C., Portland, Ore, USA, 5.5 edition, 1991–2003.

[51] J. San Filippo Jr., L. J. Romano, C. I. Chern, and J. S. Valentine, "Cleavage of esters by superoxide," *Journal of Organic Chemistry*, vol. 41, no. 3, pp. 586–588, 1976.

[52] J. San Filippo Jr., C. I. Chern, and J. S. Valentine, "Oxidative cleavage of α-keto, α-hydroxy, and α-halo ketones, esters, and carboxylic acids by superoxide," *Journal of Organic Chemistry*, vol. 41, no. 6, pp. 1077–1078, 1976.

[53] Y. M. Chiou and L. Que, "Models for α-keto acid-dependent non-heme iron enzymes: structures and reactivity of $[Fe^{II}CL(O_2CCOPh)](ClO_4)$ complexes," *Journal of the American Chemical Society*, vol. 117, no. 14, pp. 3999–4013, 1995.

[54] M. P. Mehn, K. Fujisawa, E. L. Hegg, and L. Que, "Oxygen activation by nonheme iron(II) complexes: α-keto carboxylate versus carboxylate," *Journal of the American Chemical Society*, vol. 125, no. 26, pp. 7828–7842, 2003.

[55] T. Borowski, A. Bassan, and P. E. M. Siegbahn, "Mechanism of dioxygen activation in 2-oxoglutarate-dependent enzymes: a hybrid DFT study," *Chemistry: A European Journal*, vol. 10, no. 4, pp. 1031–1041, 2004.

[56] T. Borowski, A. Bassan, and P. E. M. Siegbahn, "A hybrid density functional study of O-O bond cleavage and phenyl ring hydroxylation for a biomimetic non-heme iron complex," *Inorganic Chemistry*, vol. 43, no. 10, pp. 3277–3291, 2004.

[57] A. Bassan, T. Borowski, and P. E. M. Siegbahn, "Quantum chemical studies of dioxygen activation by mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad," *Dalton Transactions*, no. 20, pp. 3153–3162, 2004.

[58] T. Borowski, A. Bassan, and P. E. M. Siegbahn, "4-hydroxyphenylpyruvate dioxygenase: a hybrid density functional study of the catalytic reaction mechanism," *Biochemistry*, vol. 43, no. 38, pp. 12331–12342, 2004.

[59] G. F. Barnard and M. Akhtar, "Stereochemistry of porphyrinogen carboxy-lyase reaction in haem biosynthesis," *Journal of the Chemical Society, Chemical Communications*, no. 13, pp. 494–496, 1975.

[60] G. F. Barnard and M. Akhtar, "Stereochemical and mechanistic studies on the decarboxylation of uroporphyrinogen III in haem biosynthesis," *Journal of the Chemical Society, Perkin Transactions 1*, pp. 2354–2360, 1979.

[61] M. Akhtar, "The modification of acetate and propionate side chains during the biosynthesis of haem and chlorophylls: mechanistic and stereochemical studies," *Ciba Foundation symposium*, vol. 180, pp. 131–152, 1994.

[62] J. Fan, Q. Liu, Q. Hao, M. Teng, and L. Niu, "Crystal structure of uroporphyrinogen decarboxylase from Bacillus subtilis," *Journal of Bacteriology*, vol. 189, no. 9, pp. 3573–3580, 2007.

[63] I. U. Heinemann, M. Jahn, and D. Jahn, "The biochemistry of heme biosynthesis," *Archives of Biochemistry and Biophysics*, vol. 474, no. 2, pp. 238–251, 2008.

[64] B. M. Martins, B. Grimm, H. P. Mock, R. Huber, and A. Messerschmidt, "Crystal structure and substrate binding modeling of the uroporphyrinogen-III decarboxylase from Nicotiana tabacum: implications for the catalytic mechanism," *Journal of Biological Chemistry*, vol. 276, no. 47, pp. 44108–44116, 2001.

[65] H. P. Mock, L. Trainotti, E. Kruse, and B. Grimm, "Isolation, sequencing and expression of cDNA sequences encoding uroporphyrinogen decarboxylase from tobacco and barley," *Plant Molecular Biology*, vol. 28, no. 2, pp. 245–256, 1995.

[66] A. H. Jackson, H. A. Sancovich, A. M. Ferramola et al., "Macrocyclic intermediates in the biosynthesis of porphyrins," *Philosophical Transactions of the Royal Society of London. Series B*, vol. 273, no. 924, pp. 191–206, 1976.

[67] J. D. Phillips, F. G. Whitby, J. P. Kushner, and C. P. Hill, "Structural basis for tetrapyrrole coordination by uroporphyrinogen decarboxylase," *The EMBO Journal*, vol. 22, no. 23, pp. 6225–6233, 2003.

[68] G. Chaufan, M. C. Ríos de Molina, and L. C. San Martín de Viale, "How does hexachlorobenzene treatment affect liver uroporphyrinogen decarboxylase?" *International Journal of Biochemistry and Cell Biology*, vol. 33, no. 6, pp. 621–630, 2001.

[69] C. A. Lewis and R. Wolfenden, "Uroporphyrinogen decarboxylation as a benchmark for the catalytic proficiency of enzymes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 45, pp. 17328–17333, 2008.

[70] S. Billi de Catabbi, M. C. Ríos de Molina, and L. C. San Martín de Viale, "Studies on the active centre of rat liver porphyrinogen carboxylase in vivo effect of hexachlorobenzene," *International Journal of Biochemistry*, vol. 23, no. 7-8, pp. 675–679, 1991.

[71] K. M. Jones and P. M. Jordan, "Purification and properties of the uroporphyrinogen decarboxylase from Rhodobacter sphaeroides," *Biochemical Journal*, vol. 293, no. 3, pp. 703–712, 1993.

[72] M. Akhtar, *New Comprehensive Biochemistry: Biosynthesis of Tetrapyrroles*, vol. 19, Elsevier, London, UK, 1991.

[73] M. J. Bearpark, F. Ogliaro, T. Vreven et al., "CASSCF calculations for excited states of large molecules: choosing when to use the RASSCF, ONIOM and MMVB approximations," in *Computation in Modern Science and Engineering, Parts A and B*, T. E. Simos and G. Maroulis, Eds., vol. 2, pp. 583–585, Amer Inst Physics, Melville, NY, USA, 2007.

[74] S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch, "A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives," *Journal*

of Molecular Structure: THEOCHEM*, vol. 461-462, pp. 1–21, 1999.

[75] S. Humbel, S. Sieber, and K. Morokuma, "The IMOMO method: integration of different levels of molecular orbital approximations for geometry optimization of large systems: test for *n*-butane conformation and S_N2 reaction: $RCl + Cl^-$," *Journal of Chemical Physics*, vol. 105, no. 5, pp. 1959–1967, 1996.

[76] F. Maseras and K. Morokuma, "IMOMM—a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states," *Journal of Computational Chemistry*, vol. 16, no. 9, pp. 1170–1179, 1995.

[77] K. Morokuma, D. G. Musaev, T. Vreven, H. Basch, M. Torrent, and D. V. Khoroshun, "Model studies of the structures, reactivities, and reaction mechanisms of metalloenzymes," *IBM Journal of Research and Development*, vol. 45, no. 3-4, pp. 367–395, 2001.

[78] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, "ONIOM: a multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and $Pt(P(t-Bu)_3)_2 + H_2$ oxidative addition," *Journal of Physical Chemistry*, vol. 100, no. 50, pp. 19357–19363, 1996.

[79] T. Vreven, K. S. Byun, I. Komáromi et al., "Combining quantum mechanics methods with molecular mechanics methods in ONIOM," *Journal of Chemical Theory and Computation*, vol. 2, no. 3, pp. 815–826, 2006.

[80] T. Vreven and K. Morokuma, "On the application of the IMOMO (integrated molecular orbital + molecular orbital) method," *Journal of Computational Chemistry*, vol. 21, no. 16, pp. 1419–1432, 2000.

[81] T. Vreven, K. Morokuma, O. Farkas, H. B. Schlegel, and M. J. Frisch, "Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints," *Journal of Computational Chemistry*, vol. 24, no. 6, pp. 760–769, 2003.

[82] D. A. Case, T. E. Cheatham, T. Darden et al., "The Amber biomolecular simulation programs," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1668–1688, 2005.

[83] H. de Verneuil, S. Sassa, and A. Kappas, "Purification and properties of uroporphyrinogen decarboxylase from human erythrocytes. A single enzyme catalyzing the four sequential decarboxylations of uroporphyrinogens I and III," *Journal of Biological Chemistry*, vol. 258, no. 4, pp. 2454–2460, 1983.

[84] A. B. Juárez, C. Aldonatti, M. S. Vigna, and M. C. Ríos de Molina, "Studies on uroporphyrinogen decarboxylase from Chlorella kessleri (Trebouxiophyceae, Chlorophyta)," *Canadian Journal of Microbiology*, vol. 53, no. 2, pp. 303–312, 2007.

[85] R. Wolfenden, X. Lu, and G. Young, "Spontaneous hydrolysis of glycosides," *Journal of the American Chemical Society*, vol. 120, no. 27, pp. 6814–6815, 1998.

[86] D. L. Zechel and S. G. Withers, "Glycosidase mechanisms: anatomy of a finely tuned catalyst," *Accounts of Chemical Research*, vol. 33, no. 1, pp. 11–18, 2000.

[87] J. M. Stubbs and D. Marx, "Glycosidic bond formation in aqueous solution: on the oxocarbenium intermediate," *Journal of the American Chemical Society*, vol. 125, no. 36, pp. 10960–10962, 2003.

[88] J. M. Stubbs and D. Marx, "Aspects of glycosidic bond formation in aqueous solution: chemical bonding and the

role of water," *Chemistry: A European Journal*, vol. 11, no. 9, pp. 2651–2659, 2005.

[89] K. Das, D. Ray, N. R. Bandyopadhyay, T. Ghosh, A. K. Mohanty, and M. Misra, "A study of the mechanical, thermal and morphological properties of microcrystalline cellulose particles prepared from cotton slivers using different acid concentrations," *Cellulose*, vol. 16, no. 5, pp. 783–793, 2009.

[90] Q. P. Liu, G. Sulzenbacher, H. Yuan et al., "Bacterial glycosidases for the production of universal red blood cells," *Nature Biotechnology*, vol. 25, no. 4, pp. 454–464, 2007.

[91] B. Capon, "Mechanism in carbohydrate chemistry," *Chemical Reviews*, vol. 69, no. 4, pp. 407–498, 1969.

[92] B. Henrissat, "A classification of glycosyl hydrolases based on amino acid sequence similarities," *Biochemical Journal*, vol. 280, no. 2, pp. 309–316, 1991.

[93] J. D. McCarter and S. G. Withers, "Mechanisms of enzymatic glycoside hydrolysis," *Current Opinion in Structural Biology*, vol. 4, no. 6, pp. 885–892, 1994.

[94] B. Henrissat and G. Davies, "Structural and sequence-based classification of glycoside hydrolases," *Current Opinion in Structural Biology*, vol. 7, no. 5, pp. 637–644, 1997.

[95] M. L. Sinnott, "Catalytic mechanisms of enzymic glycosyl transfer," *Chemical Reviews*, vol. 90, no. 7, pp. 1171–1202, 1990.

[96] G. J. Davies, T. M. Gloster, and B. Henrissat, "Recent structural insights into the expanding world of carbohydrate-active enzymes," *Current Opinion in Structural Biology*, vol. 15, no. 6, pp. 637–645, 2005.

[97] B. P. Rempel and S. G. Withers, "Covalent inhibitors of glycosidases and their applications in biochemistry and biology," *Glycobiology*, vol. 18, no. 8, pp. 570–586, 2008.

[98] J. Thompson, C. R. Gentry-Weeks, N. Y. Nguyen, J. E. Folk, and S. A. Robrish, "Purification from Fusobacterium mortiferum ATCC 25557 of a 6-phosphoryl-O-α-D-glucopyranosyl:6-phosphoglucohydrolase that hydrolyzes maltose 6-phosphate and related phospho-α-D-glucosides," *Journal of Bacteriology*, vol. 177, no. 9, pp. 2505–2512, 1995.

[99] A. Varrot, V. L. Y. Yip, Y. Li et al., "NAD$^+$ and metal-ion dependent hydrolysis by family 4 glycosidases: structural insight into specificity for phospho-β-D-glucosides," *Journal of Molecular Biology*, vol. 346, no. 2, pp. 423–435, 2005.

[100] J. Thompson, S. B. Ruvinov, D. I. Freedberg, and B. G. Hall, "Cellobiose-6-phosphate hydrolase (CelF) of Escherichia coli: characterization and assignment to the unusual family 4 of glycosylhydrolases," *Journal of Bacteriology*, vol. 181, no. 23, pp. 7339–7345, 1999.

[101] F. Kunst, N. Ogasawara, I. Moszer et al., "The complete genome sequence of the gram-positive bacterium Bacillus subtilis," *Nature*, vol. 390, no. 6657, pp. 249–256, 1997.

[102] J. Thompson, A. Pikis, S. B. Ruvinov, B. Henrissat, H. Yamamoto, and J. Sekiguchi, "The gene glvA of Bacillus subtilis 168 encodes a metal-requiring, NAD(H)-dependent 6-phospho-α-glucosidase. Assignment to family 4 of the glycosylhydrolase superfamily," *Journal of Biological Chemistry*, vol. 273, no. 42, pp. 27347–27356, 1998.

[103] B. G. Hall, A. Pikis, and J. Thompson, "Evolution and biochemistry of family 4 glycosidases: implications for assigning enzyme function in sequence annotations," *Molecular Biology and Evolution*, vol. 26, no. 11, pp. 2487–2497, 2009.

[104] B. I. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat, "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics," *Nucleic Acids Research*, vol. 37, no. 1, pp. D233–D238, 2009.

[105] S. S. Rajan, X. Yang, F. Collart et al., "Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD$^+$/Mn^{2+}-dependent phospho-α-glucosidase from Bacillus subtilis," *Structure*, vol. 12, no. 9, pp. 1619–1629, 2004.

[106] V. L. Y. Yip, J. Thompson, and S. G. Withers, "Mechanism of GlvA from Bacillus subtilis: a detailed kinetic analysis of a 6-phospho-α-glucosidase from glycoside hydrolase family 4," *Biochemistry*, vol. 46, no. 34, pp. 9840–9852, 2007.

[107] V. L.Y. Yip and S. G. Withers, "Family 4 glycosidases carry out efficient hydrolysis of thioglycosides by an α,β-elimination mechanism," *Angewandte Chemie International Edition*, vol. 45, no. 37, pp. 6179–6182, 2006.