Lawan, N., Ranaghan, K. E., Manby, F. R., & Mulholland, A. J. (2014). Comparison of DFT and ab initio QM/MM methods for modelling reaction in chorismate synthase. Chemical Physics Letters, 608, 380-385. DOI: 10.1016/j.cplett.2014.06.010

Comparison of DFT and ab initio QM/MM methods for modelling reaction in chorismate synthase

Narin Lawan^,‡, Kara E. Ranaghan‡, Frederick R. Manby and Adrian J. Mulholland*

Centre for Computational Chemistry,

School of Chemistry,

University of Bristol,

Bristol BS8 1TS, UK.

Fax: +44(0)1179250612;

Tel: +44(0)1179289097;

E-mail: Adrian.Mulholland@bristol.ac.uk

‡These authors contributed equally to this work.

^Current address:

Computational Simulation and Modelling Laboratory (CSML),

Department of Chemistry,

Faculty of Science,

Chiang Mai University,

Chiang Mai, Thailand 50200

Abstract

Quantum mechanics / molecular mechanics (QM/MM) methods are a popular tool in the investigation of enzyme reactions. Here, we compare B3LYP functional ab initio QM/MM methods for modelling the conversion of 5-enolpyruvylshikimate-3-phosphate to chorismate in chorismate synthase. Good agreement with experimental data is only obtained at the SCS-MP2/CHARMM27 level for a reaction mechanism in which phosphate elimination precedes proton transfer. B3LYP predicts reaction energetics that are qualitatively wrong, stressing the need for ab initio QM/MM methods, and caution in interpretation of DFT results for this enzyme.

Keywords

Quantum mechanics / molecular mechanics, Chorismate synthase, Density functional theory, Møller-Plesset perturbation theory

The award of the 2013 Nobel Prize for Chemistry to Martin Karplus, Michael Levitt and Arieh Warshel, recognises the development of combined quantum mechanical / molecular mechanical (QM/MM) methods and the important role these methods can play in our understanding of biological systems [1]. Many QM codes now allow the inclusion of point charges and other MM energy evaluations, making QM/MM calculations easier than ever to carry out. The term QM/MM method covers a wide range of implementations, where the extent to which the two regions 'feel' each other is an important factor in their utility [2-4]. Several different schemes exist for the coupling between QM and MM regions, ranging from mechanical embedding to full polarization of both QM and MM regions. Due to computational expense most QM/MM methods adopt a scheme where the QM atoms are polarized by the MM atoms, but the MM atoms have fixed parameters throughout a calculation. The accuracy of the QM method employed is also another important factor in the assessment of QM/MM methods. It is now possible to approach chemical accuracy with QM/MM calculations using coupled cluster methods for the QM region [5, 6], but these methods remain expensive particularly for large QM regions. Density functional theory (DFT) provides a good compromise between accuracy and computational expense and is widely used for the QM region in QM/MM calculations. Unlike ab initio methods, DFT is not systematically improvable and the choice of functional for a particular application is not always obvious. B3LYP has become the functional of choice for many studies QM/MM studies, but mechanistic conclusions based on this level of theory alone may not always be reliable. Here we present a QM/MM study of the conversion of 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismate catalysed by chorismate synthase (CS). The results of B3LYP QM/MM calculations are compared with those from ab initio methods, showing that the B3LYP method fails to model the energetics of this reaction accurately and indeed may lead to qualitatively incorrect mechanistic conclusions.

CS is part of the (non-mammalian) shikimate pathway and is an attractive target for new antibiotics and herbicides [7]. The development of a new generation of CS inhibitors would benefit from knowledge of the enzyme structure, active site conformational flexibility, and reaction mechanism. The reaction catalysed by CS involves the cleavage of the C-O and C-H bonds of 5-EPSP in an anti-1,4 elimination reaction [8], using a flavin mononucleotide (FMN) cofactor. Transient kinetic studies are consistent with a stepwise mechanism, rather than a concerted mechanism [9]. The observation of a V/K isotope effect at both C2 and C4 for this irreversible reaction suggests that

both bond-breaking steps (the phosphate elimination step and the proton transfer step) are partially rate-limiting and that is the second step that is irreversible. It has been suggested that the reaction involves an unstable intermediate, formed by the initial loss of phosphate, which may be of a radical nature [10], but a cationic mechanism cannot yet be ruled out.

FMN, the cofactor for the reaction, can exist in two forms differing in the protonation of N35 (see Scheme 1): $FMNH_2$ (N35 protonated) or $FMNH^-$ (unprotonated). The cofactor must exist as $FMNH^-$ to accept a proton from EPSP during the reaction. The ternary complex of CS, EPSP and oxidized FMN is stable and has been characterised crystallographically [8]. Deprotonation of $FMNH_2$ at the N35 position could occur by transfer of the proton to Asp339 or Arg45, residues identified as being located close to N35 in the CS binding site in *Streptococcus pneumoniae* [8]. Mutagenesis studies of the equivalent Asp367 in CS from *Neurospora crassa* have been interpreted as suggesting that it plays a role as a base during the reaction [11]. FMN has been shown to bind to *apo*-CS in the reduced form [12] but it is thought to become protonated by His110 due to conformational changes after EPSP binding, thereby increasing its reduction potential [8]. There is also evidence from the study of flavin-dependent α-hydroxy acids that FMN could exist as $FMNH^-$ due to a shift in the pK_a of the N35 position caused by surrounding residues in the active site [13]. From this information, the production mechanism for chorismate from EPSP in CS can be written in several feasible ways. All the proposed mechanisms involve (i) proton transfer from EPSP to $FMNH^-$ and (ii) phosphate elimination. Several proposals are based the assumption that FMN will exist as $FMNH_2$. These mechanisms [8, 9] propose that the reaction proceeds with phosphate elimination as the first step, followed by deprotonation of $FMNH_2$ at the N35 position either through a water molecule to Asp339 [9] or directly by Asp339. The last step is then a proton transfer from EPSP to $FMNH^-$. Bornemann and co-workers suggest that the intermediate formed after the elimination of phosphate is of radical character [9], whereas Maclean and Ali suggest that the intermediate is cationic in nature [8]. One other possibility is that the proton transfer from $FMNH_2$ to Asp339 is the first step of the reaction or that it is concerted with the proton transfer from EPSP to FMN.

The crystal structure of CS co-crystallised with EPSP and FMN has been determined [8] and is used as the basis of our model system. Analysis of the structure using the WHATIF [14] and ProPKa [15] web interfaces indicated that Asp339 should be deprotonated and that Arg45 should be

protonated at neutral pH. His10 and His110, also in the active site were predicted to be doubly and singly protonated, respectively. Arg45 forms a hydrogen bond to N35 of $FMNH^-$, the position that would be protonated in $FMNH_2$. Arginine has been known to act as a base in other enzymes [16, 17] and could play a similar role here. Given this information, we have created a model of CS with $FMNH^-$ already in the reduced and deprotonated form without a separate step for its formation. We have then used reaction coordinates to model the proton transfer and phosphate elimination steps simultaneously to determine their order in the reaction mechanism.

Methods

The 2.0 Å structure of CS co-crystallised with EPSP and oxidized FMN (PDB code 1QXO) was taken from the Protein Databank [8]. The tetrameric structure of CS (1QXO) protein has 4 identical active sites and the fourth active site located in chain D was selected to be the centre of the model system created for this study. FMN was modified to $FMNH^-$ and all atom types in the topology files were assigned based on the CHARMM27 parameter set [18]. The MM atomic partial charges of $FMNH^-$ and EPSP used in the initial setup of the model were based on fitting point charges to the electrostatic potential (HF/6-31G(d) CHELPG) [19, 20].There are four different forms of the phosphate anion: PO_4^{3-}, HPO_4^{2-}, $H_2PO_4^-$, and H_3PO_4. Under neutral conditions, the phosphate anion could exist in either the $H_2PO_4^-$ or the HPO_4^{2-} forms, so the protonation state of the phosphate group of EPSP is an important consideration for the setup of the model system. Tests were carried out at the AM1/CHARMM27 level of theory, indicating that protonation of the phosphate group of EPSP is important for the stability of the phosphate anion [21]. We have therefore modelled the reaction with O28 of EPSP protonated (see Scheme 1 for atom numbering). However, it is also possible that the reaction could occur within a population containing more than one protonation state of EPSP.

The positions of hydrogen atoms were placed in the protein using the CHARMM (version c27b2) procedure HBUILD [22], considering the predictions made by the WHATIF web interface [14]. As a result of these predictions, His10 in the active site is modelled in the protonated form and His110 is neutral (HSE form, proton on the NE2 atom) in our model system. The model was then solvated with a pre-equilibrated box of CHARMM-type TIP3P [23, 24] water molecules and then truncated to a 25 Å sphere centered on the carbon atom C4 of EPSP (see Scheme 1). This procedure has been applied previously to model several other enzymes [19, 25, 26]. Combined QM/MM methods are a

valuable tool in the growing area of computational enzymology [2]. For the CS model, the QM region was defined as the EPSP molecule and part of $FMNH^-$ molecule. $FMNH^-$ was partitioned across the bond between atoms C1' and C2' using a link atom [27]. The QM region therefore consisted of 63 atoms (including 1 hydrogen 'link' atom), and had a net charge of −4*e* (see Scheme 1).

Hybrid density functional theory QM/MM calculations (B3LYP/6-31G(d)/CHARMM27) were used here to model the proposed mechanism of CS. QoMMMa [28], an in-house interface between QM packages such Jaguar [29] and Molpro [30] and the Tinker MM program [31] was used to perform the QM/MM calculations. Polarization effects on the QM region were accounted for by including the MM atomic partial charges in the QM Hamiltonian. All atoms further than 16 Å away from the central C4 atom were held fixed during the QM/MM calculations. A minimized structure of the substrate complex from lower level (AM1/CHARMM27) QM/MM calculations was used as the starting structure for the generation of the potential energy surface. Two reaction coordinates were defined: (i) the reaction coordinate for proton transfer $R_{H+\ transfer}$ = *d*(C2–H8) − *d*(H8–N35) Å and (ii) the reaction coordinate for phosphate elimination $R_{PO\ elimination}$ = *d*(C4 – O26) Å (see Scheme 1). The potential energy surface for the reaction was modelled by restraining the two reaction coordinates simultaneously. $R_{H+\ transfer}$ was sampled at 0.2 Å intervals for reaction coordinate values between −1.3 to 1.3 Å. $R_{PO\ elimination}$ was run at 0.1 Å intervals for the reaction coordinate between 1.4 to 3.1 Å. The force constant for the reaction coordinate restraint in all simulations was 1000 kcal $mol^{-1}Å^{-2}$ [32, 33].

The B3LYP method has several well-known limitations, including the lack of dispersion, and it often (but not always) underestimates reaction barriers [5]. To further investigate the energetic of the reaction further (and to test the performance of B3LYP for this system), single point energy calculations on the structures generated by B3LYP/6-31G(d)/CHARMM27 were carried out at the MP2/aVDZ and SCS-MP2/aVDZ levels (i.e. MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 and SCS-MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 levels calculations), using the Molpro program [30]. The acronym aVDZ indicates that the aug-cc-pVDZ basis set was used for all atoms [34, 35]. The MM partial charges of protein and water molecules were taken into account as the environment, i.e. these charges polarize the QM region. The acronym SCS in the SCS-MP2 method stands for the spin-component scaled MP2 method of Grimme [36, 37], which has shown to give results close to those from coupled cluster methods for other enzyme catalysed reactions [6]. Single point energy

calculations were also carried out at selected points along the SCS-MP2 path (R, TS1, TS2 and P) at the LCCSD(t)/aVDZ/CHARMM27 level of theory to confirm their agreement with a more accurate ab initio method. The LCCSD(T) acronym indicates that local approximations were used within the coupled cluster framework including single, double and pertubative triple excitations.

Coordinate driving approaches such as adiabatic mapping are a useful solution to the complex problem of identifying transition states in large systems. However, energy maxima on the potential energy surface produced in this way are only approximations to the true saddle point structure. Structures from the pathway generated by adiabatic mapping were used in nudged elastic band (NEB) calculations [38] in an attempt to refine the reaction path. NEB calculations do not require the definition of any reaction coordinate; instead, they involve the simultaneous optimization of a chain of structures towards the minimum energy path. A NEB path was generated at the B3LYP/6-31G(d)/CHARMM27 level by optimizing seven structures along the path linking the reactant, TS1 and the product. Intermediate structures were generated by linear interpolation between the reactant and TS1 or TS1 and product structures. Attempts were made to include additional structures, but these pathways would not converge. Single point energy calculations at the MP2 and SCS-MP2 levels of theory were also carried out on the structures generated by the NEB calculation. Climbing image NEB techniques [39] were then attempted to further refine the path and locate the transition state.

Results

The two-dimensional contour maps of the potential energy for the reaction indicate that the reaction proceeds in an asynchronous but concerted manner at all levels of QM/MM theory tested here (see Figure 1), in agreement with the results of kinetics experiments (~15 kcal/mol) [9]. The energy barriers for the two steps (relative to the substrate) are summarised in Table 1.The best estimate of the potential barrier to reaction is 16.9 kcal/mol at the SCS-MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 level of theory where proton transfer precedes phosphate elimination, while the potential barrier is 19.9 kcal/mol for phosphate elimination as the first step. This is in good agreement with the barrier deduced (using transition state theory) from experiment of ~15 kcal/mol. Note that the experimentally derived barrier is a free energy and is not directly comparable to the potential energy barrier, because the effects of zero-point energy and entropy are not included, but this does not affect the comparison of results from different levels of theory to the experimental data. The MP2 method

also predicts that the reaction proceeds with proton transfer as the first step, but the predicted barriers are lower (13.5 kcal/mol and 16.9 kcal/mol for proton transfer and phosphate elimination as the first step of the reaction, respectively). LCCSD(T)/aVDZ/CHARMM27//B3LYP/6-31G(d)/CHARMM27 single point energy calculations predict a barrier of 16.7 kcal/mol for the proton transfer step in excellent agreement with the 16.9 kcal/mol barrier at the SCS-MP2 level of theory. Coupled cluster theory can predict energies to within chemical accuracy [5] and this result and those reported for other enzymes [6] shows that the SCS-MP2 method gives results very close to this 'gold standard' ab initio method.

FMN acts as a cofactor for many different types of reaction. CS falls into a more unusual class of flavoproteins where there is no net redox change in the reaction. In this class of flavoproteins, FMN can act in several ways: through 2 electron chemistry, through free radical chemistry or even with no clearly defined or direct role in catalysis [40]. It has previously been suggested that the FMN cofactor in CS is involved in radical chemistry but results in good agreement with experimental data can be obtained without invoking a radical mechanism for the reaction. Further modelling of a radical pathway would be required to rule it out completely from catalysis in CS.

In contrast to the ab initio QM/MM methods, B3LYP predicts that phosphate elimination is the first step of the reaction. The total potential energy barrier (relative to the substrate) is 10.1 kcal/mol when phosphate elimination is the first step and 20.1 kcal/mol if proton transfer is first. The difference of 10 kcal/mol in the energy barriers at the B3LYP/6-31G(d)/CHARMM27 level is significantly larger than the ~3 kcal/mol difference found with the ab initio QM/MM methods. The MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 and B3LYP/6-31G(d)/CHARMM27 results are lower than the experimental value. SCS-MP2 calculations give better results than B3LYP calculations when compared to more accurate coupled cluster methods [36]. The B3LYP method has also been shown to perform poorly for the condensation reaction in citrate synthase, where there is also quite a high concentration of charge in the reaction region [6]. Due to their computational efficiency and general accuracy, DFT-based methods, particularly the B3LYP functional, are widely used in the calculation of enzyme reactions. However, these results highlight the fact that careful testing against methods such as SCS-MP2 must be carried out to ensure that the correct conclusions are drawn from computational modelling.

Figure 2 shows graphically the overall reaction mechanism for the formation of chorismate in CS from adiabatic mapping. The first step of the reaction is proton transfer from the C2 position of EPSP to N35 of $FMNH^-$ (see Figure 2, label 1, corresponding to reactant state (1) in Figure 1). The donor to acceptor distance is 3.50 Å and the proton to acceptor distance is 2.40 Å. At the SCS-MP2/CHARMM27 transition state for this step the distance between the atoms is reduced to 1.28 Å and in the intermediate (2) the atoms are 1.03 Å. The phosphate group of EPSP is eliminated in the second step (see Figure 2, label 2, corresponding to intermediate state (2) in Figure 1). The bond between C4 and O26 of EPSP 1.44 Å in (2) and the separation between C4 and O26 is 3.09 Å in the product (3) (see Figure 2, label 3, corresponding to product state (3) in Figure 1). $FMNH^-$ is stabilized in the active site of the enzyme by hydrogen bonds with Arg45. These hydrogen bonds are most stabilizing for the reduced form $FMNH^-$, $FMNH^-$ can exist in the active site without a discrete reaction step for its formation, the protonation of Asp339 is not required to stabilize it [13]. Arg45 may be a more suitable candidate for a base to abstract the proton from $FMNH_2$. Hydrogen bonds involving the carboyxlate oxygens of EPSP shorten slightly in the intermediate (2) e.g. the hydrogen bond between O23 and Arg45 has an initial length of 1.80 Å, which shortens to 1.75 Å in the intermediate (2), before lengthening again to a value of 1.77 Å in the product (3). This stabilization of EPSP in the intermediate is further evidence that the reaction can proceed with proton transfer as the first step rather than phosphate elimination. The phosphate anion released from EPSP in the second step is stabilized by hydrogen bonds with 3 positively charged residues: Arg48, Arg337 and the protonated His10 residue. Arg48 interacts with O26, the reacting centre of the elimination, through both the NE and NH2 positions. Both hydrogen bonds are significantly shorter in the product (3) [*d*(O26 – Arg48 HE) = 2.37 Å in (1) and 1.80 Å in (3) and *d*(O26 – Arg48 HH22) = 1.96 Å in (1) and 1.75 Å in (3)]. The hydrogen bond between O30 and Arg337 decreases in length from 1.97 Å in (1) to 1.76 Å in (3), and the O29 – His10 HD2 hydrogen bond shortens from 1.64 Å in (1) to 1.54 Å in (3).These two positively charged residues could provide significant stabilization to the second step of the reaction, which could be probed further through *in silico* and experimental mutagenesis studies.

NEB calculations at the B3LYP/6-31G(d)/CHARMM27 level of theory predict a reaction where phosphate elimination precedes proton transfer, in agreement with the potential energy surface generated by adiabatic mapping (Table 2). However, analysis of the distances involved in the reaction shows that the *d*(H8–N35) shortens significantly between the first and second images without a

significant increase in energy: from 2.40 Å to 2.22 Å with the relative energy increasing by 0.16 kcal/mol. The distance involved in phosphate elimination increases significantly (from 1.50 Å to 2.3 Å) and *d*(H8–N35) also shortens by ~0.1 Å between image 2 and image 3. Image 3 is the highest point along the path which corresponds to a reaction barrier of 8.51 kcal/mol, lower than the 10.1 kcal/mol barrier predicted by adiabatic mapping techniques. The reaction coordinates used in the adiabatic mapping approach provide a reasonable description of the reaction, but the differences in energy barrier predicted by the two techniques can be accounted for when the geometry of the maxima are considered. Figure 3 shows a comparison of the geometry of the QM atoms at the energy maximum in the adiabatic mapping and NEB calculations. There is very good agreement between the two methods for the position of the atoms of the FMN cofactor and EPSP. However, both reactions are more advanced at the maximum from adiabatic mapping, with reaction coordinate values of $R_{PO\ elimination}$ = 2.9 Å and $R_{H+\ transfer}$ = −0.5 Å for phosphate elimination and proton transfer reactions, respectively. The corresponding values in the maximum from the NEB pathway are $R_{PO\ elimination}$ = 2.3 Å and $R_{H+\ transfer}$ = −1.0 Å, respectively. MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 and SCS-MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 single point energies again predict very different energetic for the reaction compared to B3LYP (Table 2 and Figure 4). The ab initio methods predict that image 2 is ~ 1 kcal/mol lower in energy than image 1. This is in agreement with the adiabatic mapping methods which gave a reactant minimum at $R_{H+\ transfer}$ = −1.1 Å rather than the B3LYP value of $R_{H+\ transfer}$ = −1.3 Å. The MP2 and SCS-MP2 energies for image 3 are more than 10 kcal/mol higher than the B3LYP prediction, showing that the predictions made by the B3LYP method are qualitatively incorrect, irrespective of how the pathway was generated.

Conclusions

A reaction mechanism for the formation of chorismate from EPSP in CS has been modelled using high-level QM/MM methods. Adiabatic mapping techniques were used to generate a 2D potential energy surface for reaction coordinates involving phosphate elimination and proton transfer to the FMN cofactor. The best estimate of the potential energy barrier from LCCSD(T)/aVDZ//B3LYP/6-31G(d)/CHARMM27 calculations is 16.7 kcal/mol for a pathway in which proton transfer precedes phosphate elimination, giving a potential energy barrier in agreement with the experimental value of

15 kcal/mol. The enzyme stabilizes the negative charge of the flavin cofactor and EPSP substrate via hydrogen bonds with positively charged arginine residues (Arg45, Arg48 and Arg337).

The SCS-MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 energy is in excellent agreement with the LCCSD(T)/aVDZ//B3LYP/6-31G(d)/CHARMM27 value, showing that reliable energies can be obtained by just applying the spin component scaling treatment to the MP2 method without the need to move on to a more expensive ab initio method such as coupled cluster. Comparison of the DFT results with those from ab initio methods shows that the B3LYP functional performs poorly for this system, predicting a different order of the steps in the reaction mechanism to ab initio methods. High-level calculations (e.g. SCS-MP2) are apparently necessary to give reliable predictions in good agreement with the experimental data for this enzyme. This highlights the careful testing of QM/MM methods that is required before reliable conclusions can be drawn. Other mechanisms for the reaction in CS and protonation states for EPSP cannot be ruled out, but this study indicates that any further modelling of the reaction will have to be carried out at the SCS-MP2 level of QM theory.

Acknowledgements

AJM is an EPSRC Leadership Fellow and with KER thanks EPSRC for support. NL thanks the Thai Government for Higher Educational Strategic Scholarships for Frontier Research Network (SFR) scholarship.

References

1. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/advanced-chemistryprize2013.pdf. 2013.
2. Ranaghan, K.E. and Mulholland, A.J., *Investigations of enzyme-catalysed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods.* Int. Rev. Phys. Chem., 2010. **29**(1): p. 65-133.
3. Senn, H.M. and Thiel, W., *QM/MM Methods for Biomolecular Systems.* Angew. Chem. Int. Edit., 2009. **48**(7): p. 1198-1229.
4. Van Der Kamp, M.W. and Mulholland, A.J., *Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology.* Biochemistry-US, 2013. **52**(16): p. 2708-2728.
5. Claeyssens, F., Harvey, J.N., Manby, F.R., et al., *High-Accuracy Computation of Reaction Barriers in Enzymes.* Angew. Chem. Int. Ed., 2006. **45**: p. 6856 -6859.
6. Van Der Kamp, M.W., ZUrek, J., Manby, F.R., et al., *Testing High-Level QM/MM Methods for Modeling Enzyme Reactions: Acetyl-CoA Deprotonation in Citrate Synthase.* J. Phys. Chem. B, 2010. **114**(34): p. 11303-11314.

7. Rauch, G., Ehammer, H., Bornemann, S., et al., *Replacement of two invariant serine residues in chorismate synthase provides evidence that a proton relay system is essential for intermediate formation and catalytic activity.* FEBS Journal, 2008. **275**(7): p. 1464-1473.
8. Maclean, J. and Ali, S., *The Structure of Chorismate Synthase Reveals a Novel Flavin Binding Site Fundamental to a Unique Chemical Reaction.* Structure, 2003. **11**(12): p. 1499-1511.
9. Bornemann, S., Lowe, D.J., and Thorneley, R.N.F., *The Transient Kinetics of Escherichia coli Chorismate Synthase: Substrate Consumption, Product Formation, Phosphate Dissociation, and Characterization of a Flavin Intermediate.* Biochemistry, 1996. **35**(30): p. 9907–9916.
10. Osborne, A., Thorneley, R.N.F., Abell, C., et al., *Studies with substrate and cofactor analogues provide evidence for a radical mechanism in the chorismate synthase reaction.* J. Biol. Chem., 2000. **275**(46): p. 35825-35830.
11. Rauch, G., Ehammer, H., Bornemann, S., et al., *Mutagenic analysis of an invariant aspartate residue in chorismate synthase supports its role as an active site base.* Biochemistry, 2007. **46**(12): p. 3768-3774.
12. Macheroux, P., Bornemann, S., Ghisla, S., et al., *Studies with flavin analogs provide evidence that a protonated reduced FMN is the substrate-induced transient intermediate in the reaction of Escherichia coli chorismate synthase.* J. Biol. Chem., 1996. **271**(42): p. 25850-25858.
13. Lederer, F., *Extreme pKa displacements at the active sites of FMN-dependent alpha-hydroxy acid-oxidizing enzymes.* Protein Science, 1992. **1**: p. 540-548.
14. Vriend, G., *WHAT IF: A Molecular Modeling and Drug Design Program.* J. Mol. Graphics., 1990. **8**: p. 52-56.
15. Bas, D.C., Rogers, D.M., and Jensen, J.H., *Very fast prediction and rationalization of pKa values for protein-ligand complexes.* Proteins: Structure, Function, and Bioinformatics, 2008. **73**(3): p. 765-783.
16. Schlippe, Y.V.G. and Hedstrom, L., *Is Arg418 the catalytic base required for the hydrolysis step of the IMP dehydrogenase reaction?* Biochemistry, 2005. **44**(35): p. 11700-11707.
17. Schlippe, Y.V.G. and Hedstrom, L., *A twisted base? The role of arginine in enzyme-catalyzed proton abstractions.* Archives of Biochemistry and Biophysics, 2005. **433**(1): p. 266-278.
18. Mackerell, A.D., Bashford, D., Bellott, M., et al., *All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins.* J. Phys. Chem. B, 1998. **102**(18): p. 3586-3616.
19. Ranaghan, K.E., Ridder, L., Szefczyk, B., et al., *Insights into Enzyme Catalysis from QM/MM Modelling: Transition State Stabilization in Chorismate Mutase* Mol. Phys., 2003. **101**(17): p. 2695-2714.
20. Ranaghan, K.E., Ridder, L., Szefczyk, B., et al., *Transition State Stabilization and Substrate Strain in Enzyme Catalysis: Ab Initio QM/MM Modelling of the Chorismate Mutase Reaction.* Org. Biomol. Chem., 2004. **2**: p. 968 - 980.
21. Lawan, N., *QM/MM modelling of the reaction mechanism of chorismate synthase and chorismate mutase.* 2010. Ph.D thesis, School of Chemistry, University of Bristol.
22. Brunger, A.T. and Karplus, M., *Polar Hydrogen Positions in Proteins: Empirical Energy Placement and Neutron Diffraction Comparison.* Proteins, 1988. **4**(2): p. 148 - 156.
23. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., et al., *Comparison of Simple Potential Functions for Simulating Liquid Water.* J. Chem. Phys., 1983. **79**(2): p. 926-935.
24. Neria, E., Fischer, S., and Karplus, M., *Simulation of activation free energies in molecular systems.* J. Chem. Phys., 1996. **105**(5): p. 1902-1921.

25. Masgrau, L., Roujeinikova, A., Johannissen, L.O., et al., *Atomic Description of an Enzyme Reaction Dominated by Proton Tunneling.* Science, 2006. **312**(5771): p. 237-241.
26. Ridder, L., Mulholland, A.J., Rietjens, I.M.C.M., et al., *A quantum mechanical/molecular mechanical study of the hydroxylation of phenol and halogenated derivatives by phenol hydroxylase.* J. Am. Chem. Soc., 2000. **122**(36): p. 8728-8738.
27. Reuter, N., Dejaegere, A., Maigret, B., et al., *Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches.* J. Phys. Chem. A, 2000. **104**(8): p. 1720-1735.
28. Harvey, J.N., *Spin-forbidden CO ligand recombination in myoglobin.* Faraday Discuss., 2004. **127**: p. 165 - 177.
29. *Jaguar, 4.0.* Schrödinger Inc., Portland, Oregon, 1996-2001.
30. Werner, H.J., Knowles, P.J., Lindh, R., et al.,Molpro, 2006, http://www.molpro.net.
31. Ponder, J.W.,TINKER: Software Tools for Molecular Design, v4.0, 2003,
32. Claeyssens, F., Ranaghan, K.E., Lawan, N., et al., *Analysis of chorismate mutase catalysis by QM/MM modelling of enzyme-catalysed and uncatalysed reactions.* Org. Biomol. Chem., 2011. **9**: p. 1578-1590.
33. Lonsdale, R., Harvey, J.N., and Mulholland, A.J., *Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions.* J. Chem. Theo. Comput., 2012. **8**(11): p. 4637-4645.
34. Peterson, K.A., Wilson, A.K., Woon, D.E., et al., *Benchmark calculations with correlated molecular wave functions .12. Core correlation effects on the homonuclear diatomic molecules B-2-F-2.* Theo. Chem. Acc., 1997. **97**(1-4): p. 251-259.
35. Woon, D.E. and Dunning, T.H., *Gaussian basis sets for use in correlated molecular calculations 5. Core-valence basis sets for Boron through Neon.* J. Chem. Phys., 1995. **103**(11): p. 4572-4585.
36. Grimme, S., *Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies.* J. Chem. Phys., 2003. **118**: p. 9095.
37. Antony, J. and Grimme, S., *Is Spin-Component Scaled Second-Order Møller−Plesset Perturbation Theory an Appropriate Method for the Study of Noncovalent Interactions in Molecules?* J. Phys. Chem. A, 2007. **111**(22): p. 4862-4868.
38. Henkelman, G. and Jonsson, H., *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points.* Journal of Chemical Physics, 2000. **113**(22): p. 9978-9985.
39. Henkelman, G., Uberuaga, B.P., and Jonsson, H., *A climbing image nudged elastic band method for finding saddle points and minimum energy paths.* Journal of Chemical Physics, 2000. **113**(22): p. 9901-9904.
40. Bornemann, S., *Flavoenzymes that catalyse reactions with no net redox change.* Nat. Prod. Rep., 2002. **19**: p. 761 - 772.

Table 1 Total potential energy barriers for the reaction in CS with either proton transfer or phosphate elimination as the first step

	$\Delta V^{\ddagger}$ / kcal/mol	
	Proton transfer step as the first step	**Phosphate elimination as the first step**
B3LYP/6-31G(d)/ CHARMM27	20.1	10.1
MP2/aVDZ/ CHARMM27[a]	13.6	16.9
SCS-MP2/aVDZ/ CHARMM27[a]	16.9	19.9
LCCSD(T)/aVDZ/ CHARMM27[a]	16.7	–

[a] Single point energies based on B3LYP/6-31G(d)/CHARMM27 geometries with MM charges included in the QM Hamiltonian (see text).

Table 2 Distances (in Å) involved in the reaction in structures generated by nudged elastic band (NEB) calculations at the B3LYP/6-31G(d)/CHARMM27 level of theory.

	***d*(C4 – O26)**	***d*(C2–H8)**	***d*(H8–N35)**	**$R_{H+ transfer}$**	**B3LYP/ 6-31G(d)/ CHARMM27**	**MP2/aVDZ/ CHARMM27[a]**	**SCS-MP2/aVDZ/ CHARMM27[a]**
1	1.50	1.10	2.40	-1.3	0.00	0.00	0.00
2	1.51	1.11	2.22	-1.1	0.16	-1.34	-0.72
3	2.28	1.12	2.14	-1.0	8.51	17.57	20.75
4	2.62	1.12	2.07	-1.0	6.30	3.56	8.15
5	2.73	1.12	1.95	-0.8	6.62	-4.01	1.78
6	2.73	1.95	1.09	0.9	-2.73	-12.47	-12.19
7	2.70	2.32	1.02	1.3	-7.36	-14.73	-15.63

[a] Single point energies based on B3LYP/6-31G(d)/CHARMM27 geometries with MM charges included in the QM Hamiltonian (see text).

Scheme 1 The reaction of $FMNH^-$ and EPSP catalysed by CS. Red and blue arrows indicate the phosphate elimination and H^+ transfer reaction coordinates modelled simultaneously to generate 2D potential energy surfaces for the reaction. Atoms shown in black are included in the QM region for the QM/MM calculations, whereas those in grey form part of the MM region with a link atom to terminate the boundary. Hydrogen bonds between QM atoms and MM residues are indicated by green dotted lines.

Fig.1 Two dimensional potential energy surfaces (shown as contour maps) for proton transfer and phosphate elimination steps involved in the conversion of EPSP to chorismate in CS (a) at the B3LYP/6-31G(d)/CHARMM27 level of theory (b) at the MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 level of theory and (c) at the SCS-MP2/aVDZ//B3LYP/6-31G(d)/CHARMM27 level of theory. The reaction proceeds from reactant state (1) to intermediate state (2) then to product state (3).

Fig. 2 Chorismate synthase reaction mechanism, with proton transfer as the first step and phosphate elimination as the second step as predicted by SCS-MP2/aVDZ//B3LYP/6-31G(d)-CHARMM27 calculations, the arrows show the atom movements.

Fig. 3 A comparison of the geometry of the QM atoms at energy maxima from adiabatic mapping (magenta) and nudged elastic band (NEB) calculations (atom colours) from B3LYP/6-31G(d)/CHARMM27 calculations.

Fig. 4 Energy profiles for a reaction pathway generated using nudged elastic band (NEB) techniques at the B3LYP/6-31G(d)/CHARMM27 level of theory with single point energies calculated at the MP2/aVDZ// B3LYP/6-31G(d)/CHARMM27 and MP2/aVDZ// B3LYP/6-31G(d)/CHARMM27 levels of theory.

Scheme 1

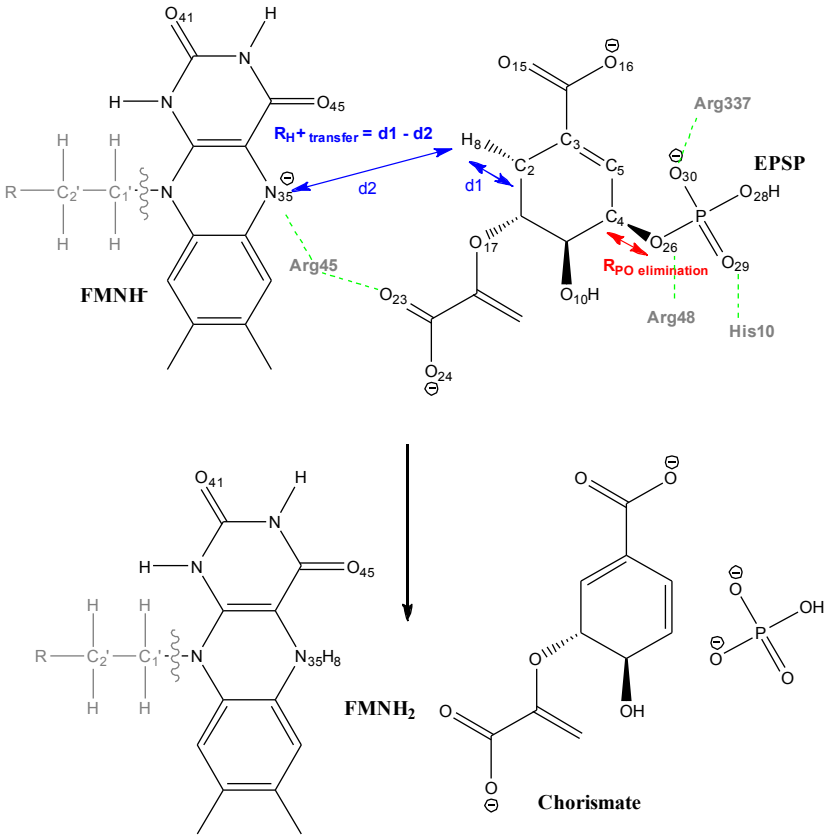
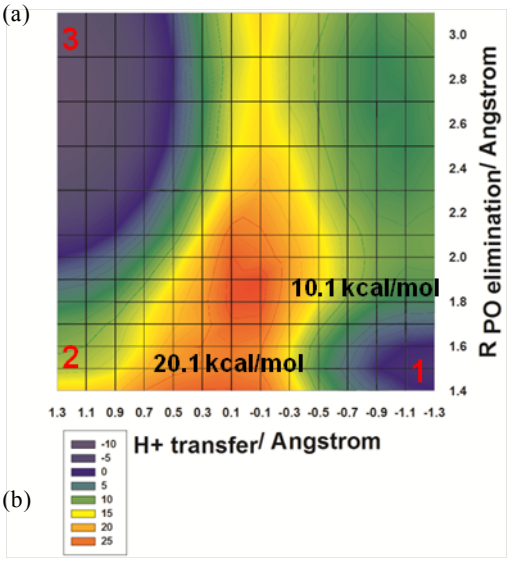

FMNH⁻
EPSP
Arg337
Arg45
Arg48
His10
R_H+ transfer = d1 - d2
d1
d2
R_{PO} elimination
FMNH$_2$
Chorismate

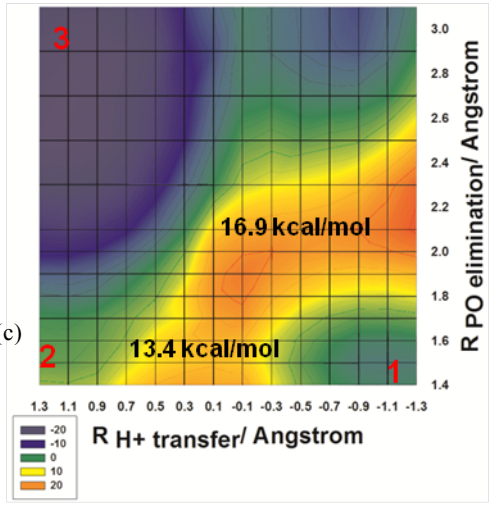
Fig. 1

(a)


3
2
1
10.1 kcal/mol
20.1 kcal/mol
3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4
R PO elimination/ Angstrom
1.3 1.1 0.9 0.7 0.5 0.3 0.1 -0.1 -0.3 -0.5 -0.7 -0.9 -1.1 -1.3
H+ transfer/ Angstrom
-10 -5 0 5 10 15 20 25

(b)

(c)


3
2
1
16.9 kcal/mol
13.4 kcal/mol
3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4
R PO elimination/ Angstrom
1.3 1.1 0.9 0.7 0.5 0.3 0.1 -0.1 -0.3 -0.5 -0.7 -0.9 -1.1 -1.3
R H+ transfer/ Angstrom
-20 -10 0 10 20

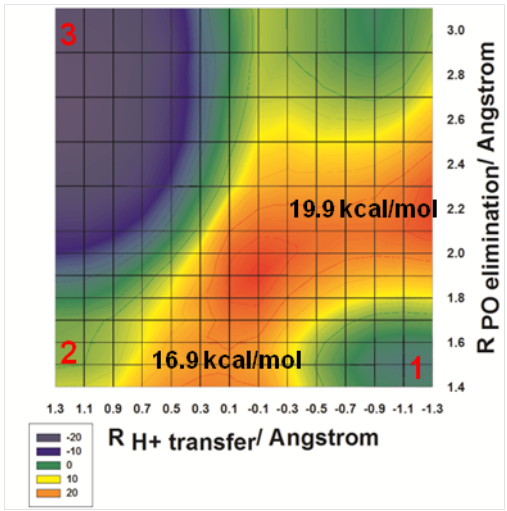

3
2
1
19.9 kcal/mol
16.9 kcal/mol
3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4
R PO elimination/ Angstrom
1.3 1.1 0.9 0.7 0.5 0.3 0.1 -0.1 -0.3 -0.5 -0.7 -0.9 -1.1 -1.3
R H+ transfer/ Angstrom
-20 -10 0 10 20

Fig. 2

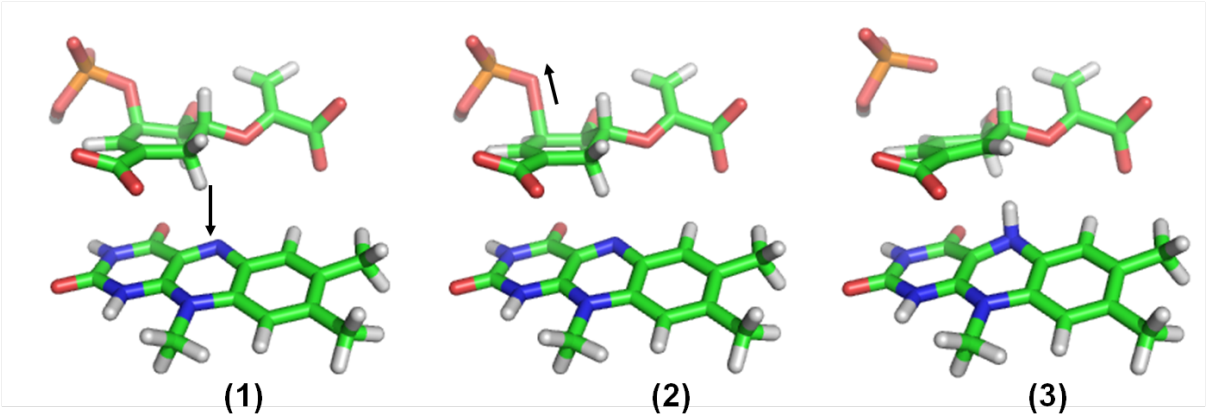


Fig. 3

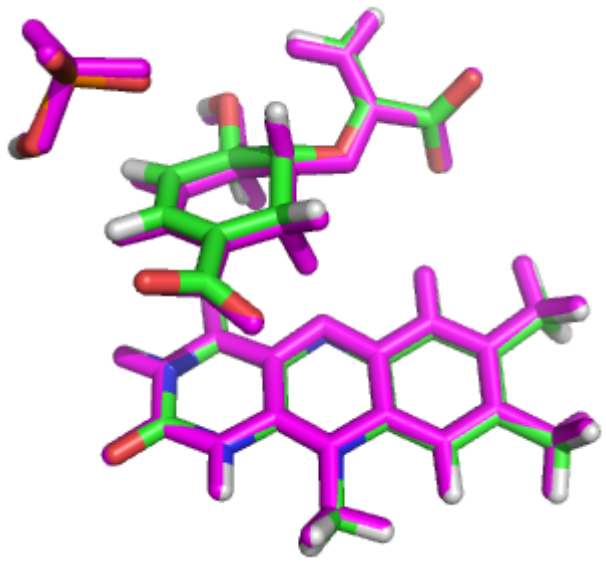
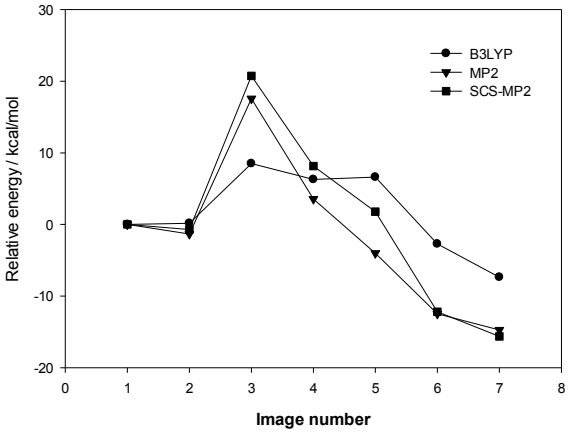
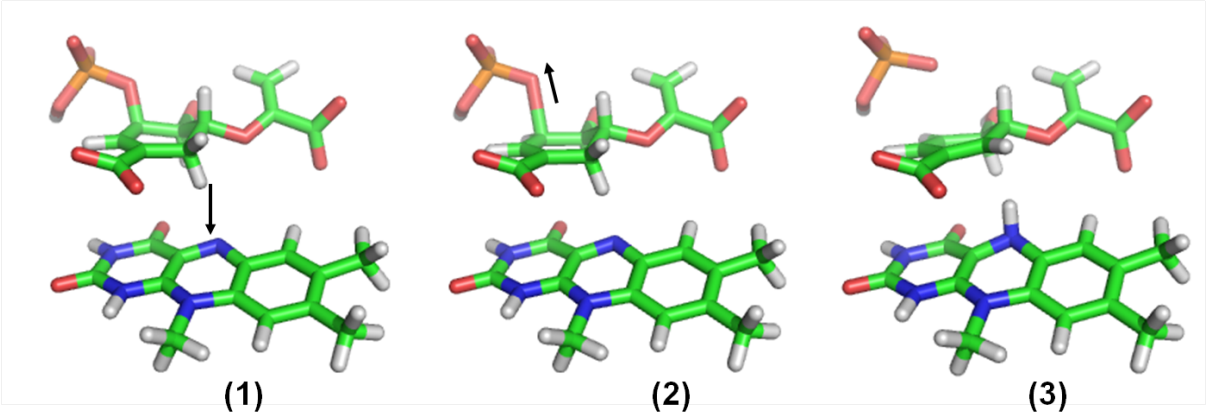


Fig. 4


B3LYP
MP2
SCS-MP2
Relative energy / kcal/mol
Image number
30
20
10
0
-10
-20
0
1
2
3
4
5
6
7
8

Graphical abstract



DFT/MM fails to accurately model the formation of chorismate in chorismate synthase