



Cite this: *Med. Chem. Commun.*,
2015, 6, 477

Application of QM/MM and QM methods to investigate histone deacetylase 8†

Duangkamol Gleeson^a and M. Paul Gleeson^{*b}

Computational chemistry plays an important supporting role in the early stages of drug discovery research. Such methods are not without flaws, however they can be very useful in the development and testing of hypotheses as well as prioritizing aspects of the exploration process. In this paper we discuss some common issues with employing hybrid quantum mechanical/molecular mechanical (QM/MM) methods in certain drug discovery applications. The QM/MM method provides a means to simulate large biological systems for moderate computational cost. We use the method to assess the metalloproteins, human deacetylases (HDACs), which are targets for a variety of medical conditions including neurodegenerative diseases and HIV infection. Metalloproteins in particular are a challenge to simulate using the rapid empirical methods preferred in the pharmaceutical industry. We report the use of a QM/MM scheme of only moderate computational cost to explore the active site as well as its catalytic reaction. We also demonstrate the value of the method over smaller QM clusters and show that the method is capable of describing the kinetic differences associated with replacing Zn²⁺ with other metal co-factors.

Received 17th October 2014,
Accepted 5th January 2015

DOI: 10.1039/c4md00471j

www.rsc.org/medchemcomm

1 Introduction

Biophysical techniques such as X-ray protein crystallography, nuclear magnetic resonance (NMR) and isothermal calorimetry (ITC) provide either direct or indirect information on ligand binding to proteins which can prove extremely useful in drug discovery.^{1–4} Protein crystallography can conclusively show the binding mode of inhibitors; NMR spectroscopy can be used to screen for active site binders indirectly and determine affinities; ITC provides useful information to optimize the enthalpic efficiency of inhibitors for their target.

Theoretical methods are also commonly used to estimate ligand binding modes to proteins, rationalise structure–activity relationships (SARs), estimate simple pairwise interactions or binding free energies, predict conformational preferences of substrates, estimate the reactivity of molecules or particular functional groups, as well as having many other uses.^{5,6} The ability to generate computationally derived atomic models to understand complex biological processes is desirable as the

results can be used to post rationalize observations as well as make future prediction. While computational methods are not without flaws,^{7–9} they are useful in that they allow scientists to develop and test hypotheses as well as prioritize aspects of the exploration process. The key to this process is the application of appropriate methods in appropriate situations.

Quantum mechanical (QM) methods offer a more accurate description of atomic systems yet in an industrial setting empirical molecular mechanics (MM) force fields dominate due to their greater simplicity and low computational time costs. Nevertheless, it is widely accepted that QM methods offer a better description of non-bonded interactions, metal-ligand association, charge distributions and conformational energies, as well as being able to simulate bond breaking and formation (*i.e.* reactions). The benefits of the more widespread introduction of QM based methods in the pharmaceutical industry are increasingly being discussed with the advent of cheaper, more capable personal computers and software tools.^{10–15}

Traditional *ab initio* or DFT QM methods are typically limited to relatively small atomic models for the reason of computational cost. For systems of biological interest this means that the key active site residues can be included in a model but the effects of the extended protein or solvent must be neglected. This has prompted interest in hybrid methods such as QM/MM.^{16–23} These methods, of which a number of implementations exist, split a model system into two, with the key active site residue treated using QM and the effects of

^a Department of Chemistry, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

^b Department of Chemistry, Faculty of Science, Kasetsart University, 50 Phaholyothin Rd, Chatuchak, Bangkok 10900, Thailand.

E-mail: paul.gleeson@ku.ac.th; Fax: +66 2 5793955;

Tel: +66 2 562 5555 extn 2210

† Electronic supplementary information (ESI) available: Additional active site graphics and tables containing energies and charges are provided. See DOI: 10.1039/c4md00471j

the extended protein and solvent treated using MM methods. Crucially, this method provides a means to represent large biological systems for potentially only moderate computational cost. Given the increasing interest in covalent inhibitors of late,^{24–26} such methods could offer real potential to computationally screen compounds to maximize reactivity at pharmacological targets and potentially minimise reactivity at metabolic targets.²⁷

Histone deacetylases (HDACs) are a family of Zn metalloproteins found in eukaryotic cells that catalyze the deacetylation of acetylated lysine residues of histones.^{28,29} They play a role in important cellular processes such as apoptosis, proliferation and senescence and have therefore received attention from a pharmaceutical development perspective.^{30,31} Two inhibitors of HDACs are currently licensed as anti-cancer agents while other molecules are in clinical development as neuro-degenerative treatments.³⁰

The HDAC family can be split into four distinct classes based on phylogenetic analyses: HDAC 1–3 & 8 constitute class I; HDAC 4–7, 9 & 10 class II; sirtuins 1–7, class III; and HDAC 11 alone constituting class IV.³² Class I/II HDACs have received the largest amount of attention in the literature due to their anti-cancer properties.³³ A number of X-ray crystal structures of class I protein HDAC8 have been reported in the literature and these have helped enormously to identify the structural features that give rise to the biochemical function of the overall family.^{34–36} The active site consists of a narrow pocket found deep within the protein, and contains the Zn metal cofactor. The Zn²⁺ metal exists in a square pyramid configuration with bonds to the sidechains of His-180, Asp-178, Asp-267, and two bonds to the inhibitor. The reaction mechanism of class I/II HDACs has been postulated using crystal structure data reported for HDAC8.^{28,29}

The proteins in this family catalyse the deacetylation of substrates to give a free lysine residue and acetic acid. Catalysis is achieved *via* Zn²⁺ mediated nucleophilic attack of the carbonyl by an activated water nucleophile. This leads to the formation of a tetrahedral intermediate which can be stabilized through interaction with the side chain of Tyr-306.²⁸ The primary issue regarding the mechanism is the precise role of His-143/His-142 in the reaction, since either could act as the general base in the first step. In the second step of the reaction the tetrahedral intermediate decomposes by the reformation of the carbonyl bond, leading to the breakage of the N–C bond and the transfer of a proton to form the catalytic base.

Theoretical investigations into the reaction mechanism of HDAC8 have been reported by Zhang and co-workers.^{37,38} They used their Born–Oppenheimer *ab initio* QM/MM molecular dynamics methodology to estimate the free energy barriers rather than relying on the enthalpy alone. The inclusion of free energy effects while beneficial, comes at a cost, making the calculations lower in throughput and potentially unsuitable for use in virtual screening exercises. More recently Chen *et al.*³⁹ have employed QM/MM to investigate both HDAC8 inhibition and reactivity. Interestingly, the group noted that the wild-type reactivity profiles from different MD

snapshots gave very similar results. QM cluster calculations have been used to assess zinc binding motifs that have been the source of a number of interesting studies.^{40–42} Recent work supports the view that hydroxamic acids such as SAHA are highly effective due to their lower than expected acidic pK_a leading to their deprotonation (*via* His142) and therefore greater interaction with Zn²⁺ bound in the active site.³⁹

Novel inhibitors of HDACs are still being sought, with recent reports focusing on replacing the more common Zn²⁺ binding substituents such as carboxylates and hydroxamates with alternatives.⁴³ As such, higher throughput theoretical models to assess aspects of protein–ligand binding, and also reactivity, are desired. In this study we apply QM cluster calculations and QM/MM models to study substrate binding and reactivity at HDAC8 as the first step in our investigation of this protein. The goal of this work is to develop and validate QM and QM/MM HDAC8 models for use in a higher throughput manner. We employ a large QM active site representation, investigate the effect of the protein environment *vs.* an implicit solvent model, explore alternative mechanistic possibilities and assess the effect of ΔG correction derived from vibrational frequency calculations. We also attempt to further validate the methods by assessing how effectively the method can predict experimental rate constant differences associated with different metal co-factors. The overall aim is to identify more suitable computational models for use in analysing novel inhibitors targeting HDAC proteins, and in particular novel Zn binding moieties that are sub-optimally described by traditional MM methods.

2 Computational procedures

2.1 Model setup

The structural coordinates of HDAC8 were obtained from the RCSB protein data bank (PDB code: 2V5W, 2.0 Å).²⁹ Protein preparation was performed in the Discovery Studio (DS) 2.5 program with the CHARMM 22 force field⁴⁴ as described elsewhere. Briefly, the 623 da substrate was truncated to *n*-propyl-acetamide for the purpose of exploring the wild-type protein reactivity towards substrate deacetylation. The Y306F mutation was corrected in the model. Zn²⁺ and K⁺ counterions were retained as were X-ray water molecules.

The system was solvated in a cubic box of TIP3P water with Na⁺ and Cl[−] counterions (1.5 M) added to neutralize the system. RESP charges at the HF/6-31G(d) level were computed for the substrate.⁴⁵ Preparation was achieved using the following procedure: (1) hydrogen atoms alone were optimized in the first iteration, (2) with the backbone fixed, amino side chains, the substrate, counterions and water molecules were optimized, and (3) the backbone of the protein was harmonically restrained and the system underwent 1 ns of molecular dynamics (MD) under NPT conditions (heating 0.2 ps (0–300 K), equilibration 0.5 ps and production 0.2 ps).⁴⁶ The heavy atom RMSD between the initial X-ray structure and the protein following the MD preparation sequence was 2.1 Å.

2.1 QM/MM calculations

The MD coordinates were used to create a range of different QM and QM/MM models. In each case the QM region used to simulate the protein consisted of the side chains of key active site amino acid residues: His142, His143, Asp178, His180, Asp267, Tyr306, Zn²⁺ and the substrate (see the ESI,† Fig. S1). The valences of bonds that bridged the QM/MM interface were satisfied with hydrogen link atoms. Water molecules and Na⁺ and Cl⁻ counterions were excluded from subsequent models as the reactive center is buried within the protein. Additional QM/MM models were created by (a) adding a single proton to the system at either His142 or His143 or (b)

replacing Zn²⁺ ($S = 0/2$) with Co²⁺ ($S = 3/2$), Fe²⁺ ($S = 4/2$) and Mn²⁺ ($S = 5/2$). During the QM/MM optimization process all atoms except for the hydrogen link atoms and those beyond 15 Å of the active site were treated flexibly. The steric and electrostatic effects of the protein environment on the Zn²⁺ QM/MM model were also explored. This was achieved by extracting the active site region from the protein which was then optimized in the presence of a polarized continuum model of water. Atoms corresponding to the hydrogen link atoms were fixed in this model to limit the degrees of freedom of the system.

All QM and QM/MM calculations were performed using the Gaussian 09 program.⁴⁷ The former used a PCM solvent

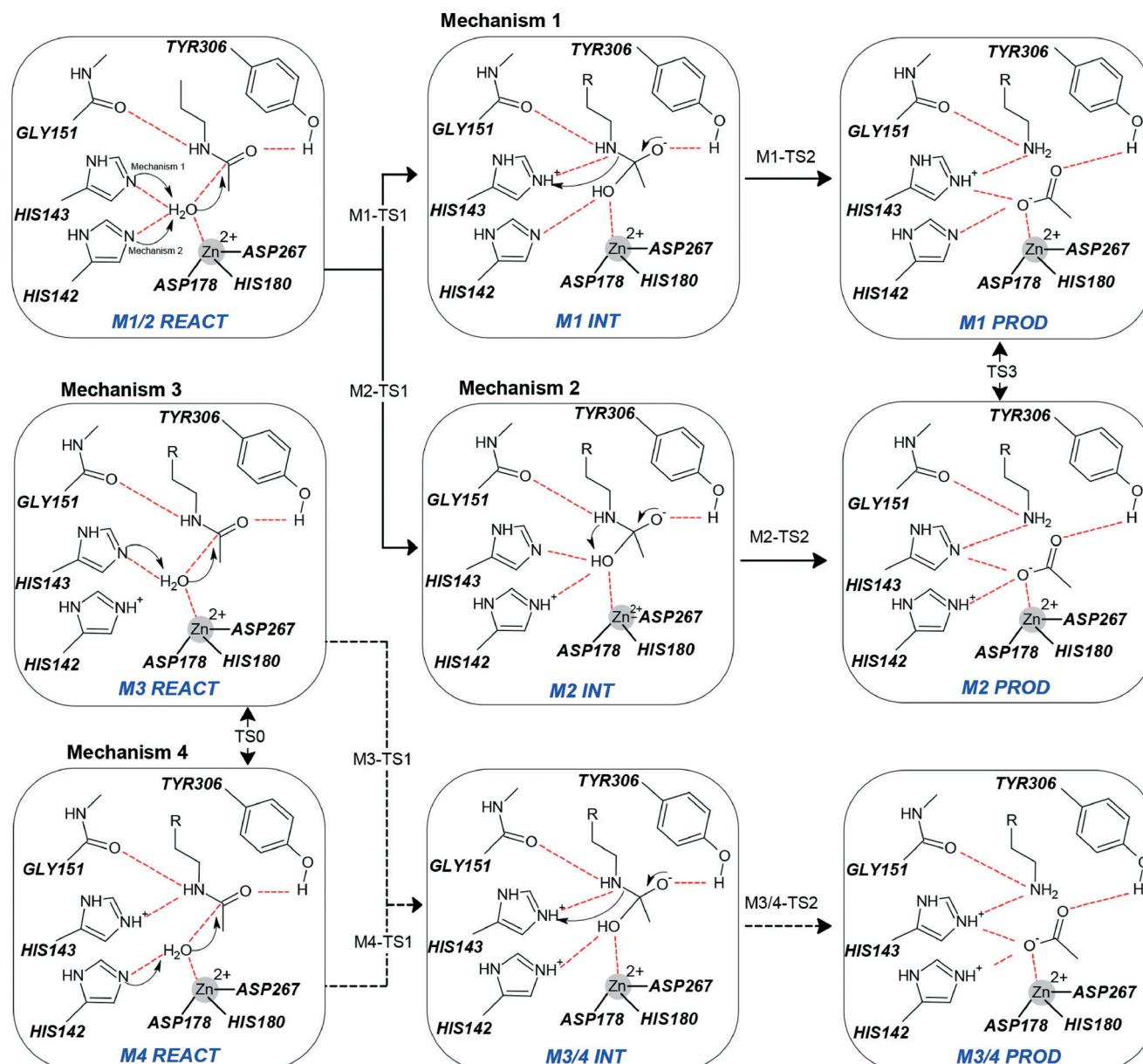


Fig. 1 Four distinct deacetylation pathways associated with HDAC8 are possible based on current proposals. Mechanisms 1 and 2 involve neutral His142 or His143 acting as the catalytic base in the initial step. Mechanisms 3 and 4 also involve either His142 or His143 acting as the general base with the other being in a protonated state.

model consisting of water while the latter employed the ONIOM methodology with the electronic embedding scheme.^{48,49} All structures were optimized at the M06/6-31G(d) for H, C, N and O, and Lanl2DZ for transition metals. Single point energies were obtained for all atoms at the M06/6-31+G(d,p) level. All structures were fully characterized using normal coordinate analysis. ΔH s were corrected for translational, electronic, rotational and vibrational motion to give the corresponding ΔG s. All calculations were performed on Intel i7 computers taking approximately 2–3 days per structure.

2 Results and discussion

We report the structures and energies for 4 possible mechanistic sequences (Fig. 1).^{38,50} To date, neither the structures nor the energetics associated with the full set of mechanistic possibilities have been reported. Mechanism 1 involves His143 acting as the general base by accepting a proton from a Zn²⁺ bound water molecule, while mechanism 2 involves His142 accepting the proton. It has also been reported that HDAC may function with an already protonated active site so we investigate mechanism 3 where His142 is protonated and His143 acts as the base and mechanism 4, where His143 is protonated and His142 acts as the base.

The QM/MM energetics associated with the 4 different mechanisms are reported in Fig. 2 (top) and the corresponding optimized structures are reported in Fig. 3. We also report the corresponding QM cluster model results in Fig. 2 (middle) as well as the residue partial charges in the ESI.† Finally, the predicted rate determining barriers for the substrate with Zn²⁺, Co²⁺, Fe²⁺ and Mn²⁺ co-factors are reported in Fig. 2 (bottom).

The M06/6-31G(d) (Lanl2dz on metals) based QM/MM model energies are found to correlate well with the corresponding single point QM/MM energies obtained at M06/6-31+G(d,p) ($r^2 = 0.94$, RMSE = 3.0 kcal mol⁻¹) and when the free energy corrections are included ($r^2 = 0.96$, RMSE = 3.3 kcal mol⁻¹). Henceforth we therefore limit our discussion of the QM/MM energies to the M06/6-31+G(d,p) single point energies, where the ΔG correction from the M06/6-31G(d) model has been included.

2.1 QM/MM protein reaction mechanism

The deacetylation reaction of HDAC8 initially requires the activation of the Zn²⁺ bound water molecule. Proton transfer can occur to either His143 (mechanism 1) or His142 (mechanism 2). In the optimized reactant complex, the Zn-bound water molecule interacts more strongly with His142 compared to His143 since the former is located closer to the metal center. The corresponding proton transfer distances in the reactant (M1/M2 REACT) are 1.70 Å and 1.91 Å, respectively. Transfer of a proton to either of the bases sees the concomitant attack of the substrate acetyl group by the resultant hydroxide anion. The QM/MM results show that a barrier of 24.1 kcal mol⁻¹ must be traversed to the intermediate (INT) with His142 acting as the base, compared to 17.5 kcal mol⁻¹ for

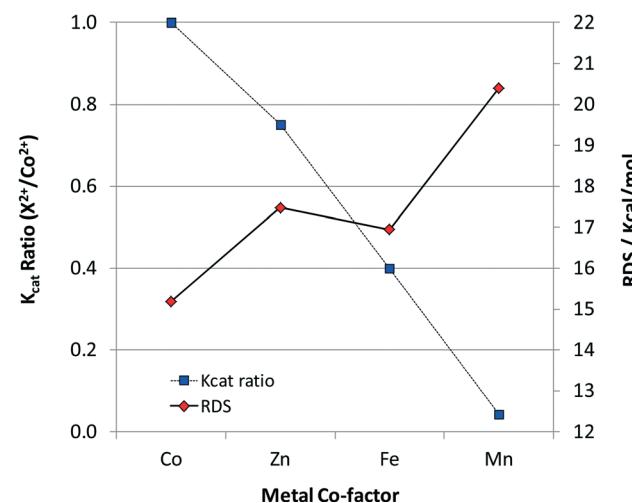
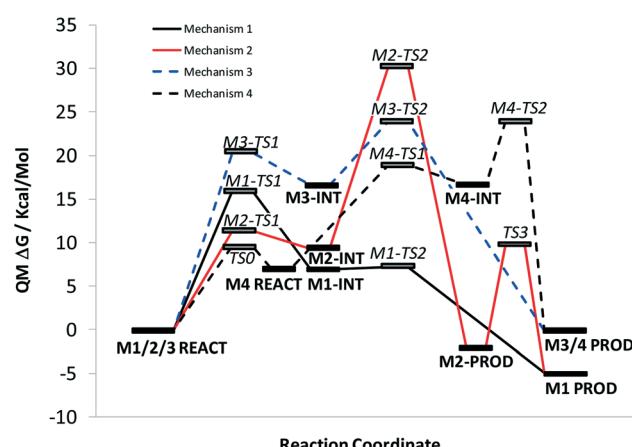
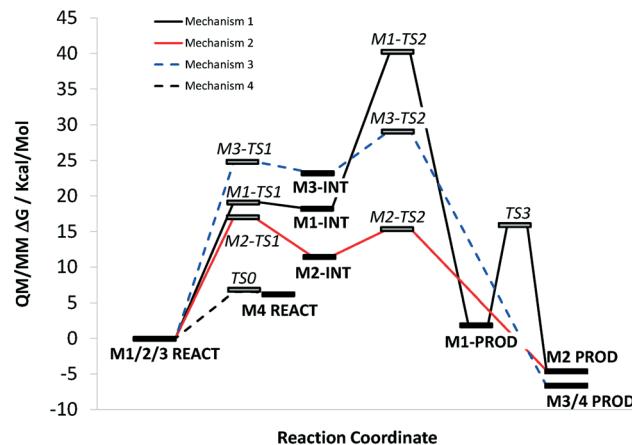


Fig. 2 Plot of the M06/6-31+G(d,p);AMBER//M06/6-31G(d):AMBER reaction profiles according to the QM/MM (top) and QM (middle) models. Also given is a plot of the predicted QM/MM rate determining step at the M06/6-31+G(d,p);AMBER//M06/6-31G(d):AMBER against the corresponding HDAC8 K_{cat} (below).

His143. The relative stability of the resulting intermediate follows the same trend of the transition states (TS) (16.3 vs.

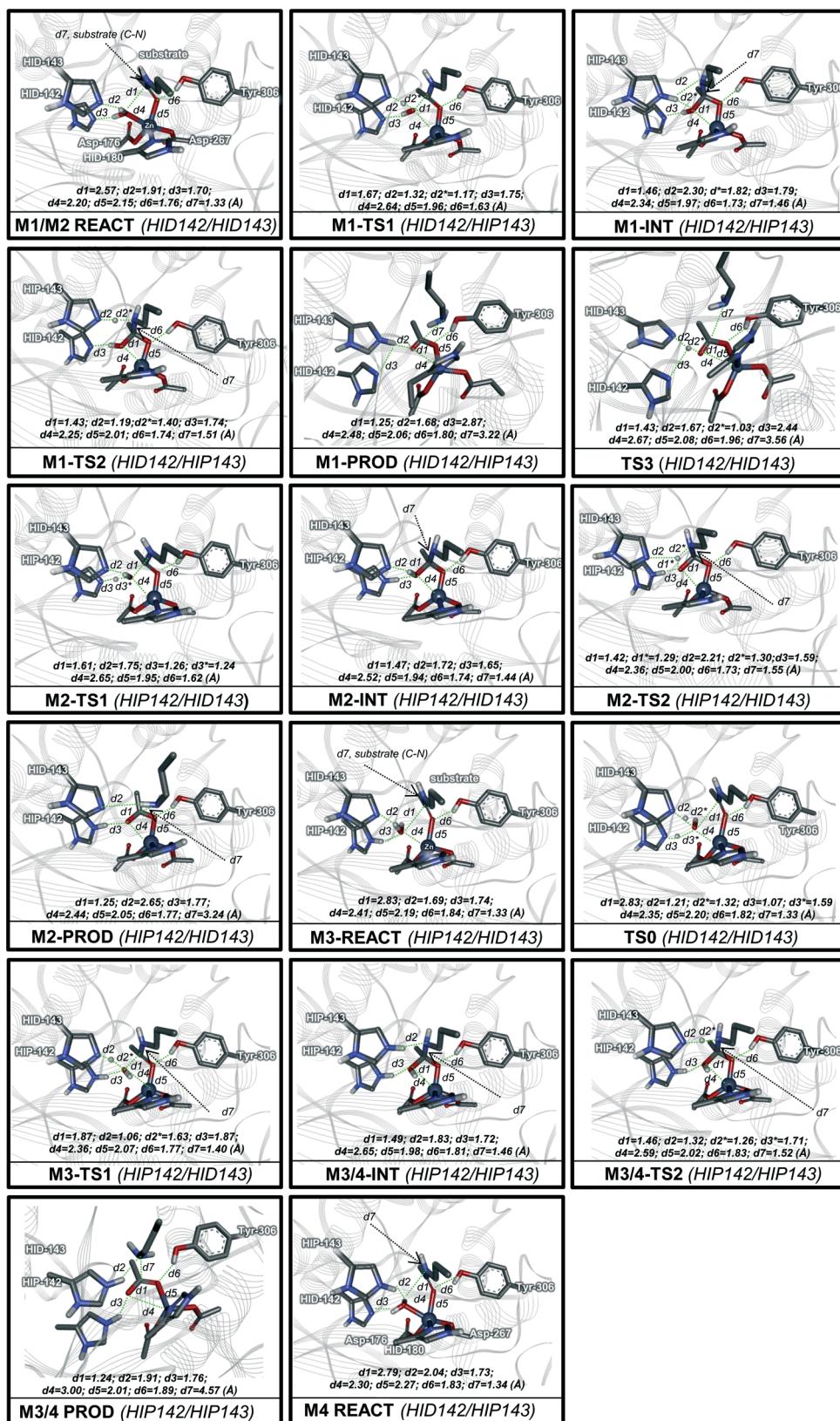


Fig. 3 Optimized QM/MM minima and transition state structures obtained at the M06/6-31G(d)//AMBER level of theory. Shown in the inset are the key distances associated with each structure.

12.2 kcal mol⁻¹). This is presumably due to the unfavourable charge–charge interaction with the Zn²⁺ center.

Formation of a tetrahedral intermediate results in a dramatic change in the partial charges on the substrate, increasing from 0.22 in the reactant to -0.48 for mechanism 1 (His143) and -0.50 for mechanism 2 (His142). Decomposition of this short lived intermediate requires proton transfer to the substrate amide nitrogen atom from the acidic His residues, along with the breaking of the amide bond. The proton transfer process associated with mechanism 1 is the more facile of the two as amide nitrogen is optimally aligned. Transition state two for mechanism 1 is 3.0 kcal mol⁻¹ compared to the dramatically larger 28.8 kcal mol⁻¹ for mechanism 2. This is because M2-TS2 involves an intramolecular 4 membered transition state with the His142 proton being shuffled to the substrate with the aid of His143. Thus His143 is found to be the most favourable base/acid, at least when both of the His residues are unprotonated.³⁷

Both mechanisms ultimately lead to the same product M1-PROD with an energy of -1.7 kcal mol⁻¹. Following the traversal of M1/M2-TS2, proton transfer from the newly formed acetic acid group to His142 is found to be spontaneous as observed by others. Mechanism 2 first sees the formation of Product 2, with a protonated His142 residue. This can form Product 1 by the shuttling of a proton with the help of the bound acetic acid residue. This requires the traversal of a barrier of 18.3 kcal mol⁻¹.

An alternative proposal was made that suggested that one of the two His residues may already be protonated (HIP) before the catalytic reaction occurs.⁵¹ Again, there are two possible mechanistic sequences that could occur (Fig. 1). Firstly, we find that the energy difference, when His142 and His143 are protonated, is 8.0 kcal mol⁻¹ in favour of the former. The barrier for proton transfer between the two configurations (TS0), with the proton being shuffled by the bound water molecule, is 4.7 kcal mol⁻¹. The barrier between M3 REACT and M3/4 INT was found to be 27.2 kcal mol⁻¹. The intermediate was found to be considerably more endothermic than those obtained for mechanisms 1 and 2 at 21.7 kcal mol⁻¹. Attempts to locate M4 TS1 were unsuccessful as proton transfer from M4 REACT was found to lead preferably to the lower energy M3-REACT rather than M3/4-INT. M3/4-TS2 was found to be 9.2 kcal mol⁻¹ higher in energy than M3/4 INT and resulted in a product -6.6 kcal mol⁻¹ lower in energy than the M3-REACT.

The systematic analysis of the 4 possible reaction pathways associated with HDAC8 reveals that the active site in the deprotonated state results in a lower barrier to the reaction. We find that the rate determining barriers determined from the QM/MM calculations follow the trend: mechanism 1 < mechanism 3 < mechanism 2. It is found that His143 acting as the general base is preferred, with His143 also acting as the general acid. These findings are in good agreement with the more computationally demanding QM/MM-MD dynamics calculations reported for mechanism 1 by Zhang and co-workers.^{37,38} The latter studies relied on a QM/MM method

consisting of the B3LYP DFT functional, an active site model that did not include Tyr306 and included free energy effects differently.³⁷

The most recent report on the HDAC8 catalytic mechanism came from Chen *et al.*³⁹ Their results suggest that the K⁺ ion bound at site 1 plays a destabilizing role, in contrast to the stabilizing effect found by Wu *et al.*³⁷ A key difference between these two models is the addition of Tyr306 and two acidic residues (Asp176 & Asp183) to the latter which will help to stabilize the protonated His residues. In the presence of a K⁺ ion in site 1, Chen *et al.* found essentially the same transition states as found by Wu *et al.* (18.3 kcal mol⁻¹),³⁷ and found here (17.9 kcal mol⁻¹), but with a somewhat lower barrier (15.3 kcal mol⁻¹). The experimental barrier for HDAC8 is reported to be ~17.7 kcal mol⁻¹ (ref. 37) based on kinetic data.⁵² In the absence of K⁺ at site 1, Chen *et al.* obtained a barrier of 8.5 kcal mol⁻¹. In this mechanism His142 acts as the general base and His143 acts as the general acid, with a proton shuttle step in between. It is worth noting that KCl is known to have an important stabilizing role in HDAC8 (increasing the thermal stability by 5 °C).⁵³ Thus it is possible that the use of the K⁺ bound state to simulate the unbound K⁺ state might be inappropriate.

2.2 QM cluster reaction mechanism

Enthalpies obtained at the M06/6-31G(d) are in good agreement with those that incorporate single point or free energy corrections. This observation is important since structure based design processes require more computationally tractable calculations. We therefore decided to investigate how simpler QM cluster models would also perform.

QM clusters allow the use of very accurate levels of theory to study protein–ligand complexes; however the effect of the surrounding protein is approximated by a continuum solvent model such as water, or neglected. This approximation is reasonable if for example the differences in ligands being explored are not influenced by the protein framework (*i.e.* if evaluating small metal binding warheads only). QM cluster calculations are nevertheless employed for many tasks due to their relative ease in terms of setup and fast computation times.^{54,55}

The QM cluster models consisted of the same QM region as used in the QM/MM model. The link atoms were frozen to restrict the conformation to one that was close to the original protein model. Nevertheless, the lack of other enzyme residues means that the conformational freedom associated with the active site atoms increases. Furthermore, the use of a solvent model as opposed to partial charges of the extended protein atoms could have an effect on the structures and energies obtained.

Fig. 2 (middle) shows that the QM derived profile for HDAC8 deacetylation is qualitatively similar to that from the QM/MM calculations. The optimized structures obtained are also found to be very similar conformationally to their QM/MM counterparts. Indeed, it is clear from a comparison

of the partial charges obtained from both models that the charge distribution within the protein is dramatically influenced by the surrounding protein. However a key observation is that mechanism 4, which was not accessible from the QM/MM model, is now possible in the QM cluster calculation. This is because protonated His143 (HIP143) is found to interact more strongly with the amide nitrogen of the substrate thereby allowing proton transfer to the substrate from the Zn^{2+} bound water molecule to occur. The rate determining barriers for mechanisms 1 and 2 are very similar at 19.7 vs. 17.9 kcal mol⁻¹, but with His142 acting as the general base now being lower in energy. The QM/MM results in contrast were 17.5 vs. 24.1 kcal mol⁻¹. The rate determining barrier for mechanism 3 is found to be 26.34 kcal mol⁻¹ according to the QM model compared to 27.2 for the QM/MM process. More importantly, the rate determining barrier for mechanism 4 is found to be only 14.9 kcal mol⁻¹. This corresponds to the barrier between M4 REACT and M3/4 INT (TS0). This result is simply an artefact of a non-ideal model cut from the original X-ray structure.

We find that the rate determining barrier determined from the QM calculations follows the trend: mechanism 4 < mechanism 2 < mechanism 1 < mechanism 3. The results are dramatically different from the QM/MM models suggesting that the effect of the protein environment plays a critical role in terms of controlling the conformation of structures formed on the reaction coordinate and in their stabilization. These results suggest that cluster calculations for HDAC8 should be considered in cases where it is clear that the effect of the extended protein environment is not so important.

2.3 Effect of the metal co-factor of reaction barriers

As an additional assessment of the QM/MM model we undertook further calculations to determine whether the method is capable of describing differences in experimental rate constants of different HDAC8 protein complexes. Catalytic rate constants have been reported for HDAC8 with 4 different transition metals Zn^{2+} , Co^{2+} , Fe^{2+} and Mn^{2+} . We therefore set out to determine whether the QM/MM model developed was sufficiently descriptive to be able to rank order the different metal co-factors in terms of their relative reactivity. The Zn^{2+} ions found in all of the mechanism 1 stationary points were replaced with Co^{2+} , Fe^{2+} and Mn^{2+} . The models were re-optimized as before and the rate determining steps were then compared to their experimental rate constants (Fig. 2 (bottom)). The QM/MM model was found to rank-order the different metals well, with Co^{2+} predicted to be the most catalytically active, Zn^{2+} and Fe^{2+} predicted to be moderately active and Mn^{2+} the least active.

The results suggest that the model is quite successful in reproducing the relative reactivity of different transition metal centres. The ability to describe transition metal-ligand interactions is highly desirable since it is a known limitation of traditional molecular modelling methods used in designing metal chelating inhibitors.⁵⁶

Conclusions

QM/MM methods can provide an effective and efficient representation of large biological systems. These methods offer considerable advantages over MM methods in that bond breaking and formation can be considered, non-standard templates do not need prior parameterization, and metals and non-bonded interactions are more accurately described.

In this study we have investigated HDAC8, a metalloprotein of widespread pharmacological interest. We were interested in developing a reasonably rapid, accurate method as part of our ultimate goal to develop a combined virtual and experimental protocol to design and assess metal chelating inhibitors. With this in mind, we have both utilized QM and QM/MM methods to explore the aspects of the ligand reactivity at HDAC8. Our calculations show that relatively modest QM/MM methods can reproduce HDAC8 results obtained from higher basis set calculations or results where free energy effects have been included.

Computational methods are of course flawed representations of real systems. Nevertheless, it is our belief that these types of models, when used in the appropriate setting, could prove useful in the design or prioritization of experimental research, including for HDACs.^{40–42} We find that gas phase cluster models do not perform so well due to the increased conformational freedom and electrostatic effects. Moreover, we have validated the accuracy of the QM/MM model in terms of the ability of the predicted rate determining step to rank-order the catalytic rate constants associated with different transition metals Zn^{2+} , Co^{2+} , Fe^{2+} and Mn^{2+} . The calculations correctly rank Mn^{2+} as the least catalytically active, Co^{2+} as the most catalytically active with Fe^{2+} and Zn^{2+} predicted correctly to have intermediate catalytic activities.

The QM/MM method could play a larger role in supporting early drug discovery research,⁵⁷ and in particular for metalloproteins such as HDACs. Selection of the active site is important, as this can have rather subtle differences in the observed results.³⁹ The key issue in the application of such methods is to know when and when not to employ them for a task under consideration. With the increasing interest in covalent inhibitors, QM/MM methods could offer real potential to computationally screen compounds to maximize reactivity at pharmacological targets and potentially minimize reactivity at metabolic targets.

Acknowledgements

MPG would like to acknowledge financial support from the Kasetsart University Research and Development Institute and the Thailand Research Fund (RSA5780068) is gratefully acknowledged. DG would like to acknowledge financial support from King Mongkut's institute of Technology Ladkrabang (2558A11802059).

Notes and references

- 1 G. Holdgate, in *Ligand-Macromolecular Interactions in Drug Discovery*, ed. A. C. A. Roque, Humana Press, 2010, ch. 7, vol. 572, pp. 101–133.
- 2 A. Ciulli, in *Protein-Ligand Interactions*, ed. M. A. Williams and T. Daviter, Humana Press, 2013, ch. 13, vol. 1008, pp. 357–388.
- 3 C. H. Reynolds and M. K. Holloway, *ACS Med. Chem. Lett.*, 2011, **2**, 433–437.
- 4 G. G. Ferenczy and G. M. Keseru, *J. Chem. Inf. Model.*, 2010, **50**, 1536–1541.
- 5 G. Sliwoski, S. Kothiwale, J. Meiler and E. W. Lowe, *Pharmacol. Rev.*, 2014, **66**, 334–395.
- 6 A. R. Leach, *Molecular Modelling. Principles and Applications*, Longman limited, Harlow, 1996.
- 7 G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff and M. S. Head, *J. Med. Chem.*, 2006, **49**, 5912–5931.
- 8 T. Scior, J. L. Medina-Franco, Q. T. Do, K. Martinez-Mayorga, J. A. Yunes Rojas and P. Bernard, *Curr. Med. Chem.*, 2009, **16**, 4297–4313.
- 9 T. Scior, A. Bender, G. Tresadern, J. L. Medina-Franco, K. Martinez-Mayorga, T. Langer, K. Cuanalo-Contreras and D. K. Agrafiotis, *J. Chem. Inf. Model.*, 2012, **52**, 867–881.
- 10 K. Raha, M. B. Peters, B. Wang, N. Yu, A. M. WollaCott, L. M. Westerhoff and K. M. Merz, *Drug Discovery Today*, 2007, **12**, 725–731.
- 11 M. B. Peters, K. Raha and K. M. Merz, *Curr. Opin. Drug Discovery Dev.*, 2006, **9**, 370–379.
- 12 K. Raha and K. M. Merz, *J. Med. Chem.*, 2005, **48**, 4558–4575.
- 13 K. E. Shaw, C. J. Woods, A. J. Mulholland and D. J. Abraham, in *Burger's Medicinal Chemistry and Drug Discovery*, John Wiley & Sons, Inc., 2003.
- 14 M. P. Gleeson and D. Gleeson, QM/MM As a Tool in Fragment Based Drug Discovery. A Cross-Docking, Rescoring Study of Kinase Inhibitors, *J. Chem. Inf. Model.*, 2009, **49**, 1437–1448.
- 15 M. P. Gleeson and D. Gleeson, QM/MM Calculations in Drug Discovery: A Useful Method for Studying Binding Phenomena?, *J. Chem. Inf. Model.*, 2009, **49**, 670–677.
- 16 M. W. van der Kamp and A. J. Mulholland, *Biochemistry*, 2013, **52**, 2708–2728.
- 17 R. Lonsdale, J. N. Harvey and A. J. Mulholland, *Chem. Soc. Rev.*, 2012, **41**, 3025–3038.
- 18 K. E. Shaw, C. J. Woods, A. J. Mulholland and D. J. Abraham, in *Burger's Medicinal Chemistry and Drug Discovery*, John Wiley & Sons, Inc., 2010.
- 19 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
- 20 H. Lin and D. G. Truhlar, *Theor. Chem. Acc.*, 2007, **117**, 185–199.
- 21 H. M. Senn and W. Thiel, *Curr. Opin. Chem. Biol.*, 2007, **11**, 182–187.
- 22 H. Senn and W. Thiel, in *Atomistic Approaches in Modern Biology*, ed. M. Reiher, Springer Berlin Heidelberg, 2007, ch. 84, vol. 268, pp. 173–290.
- 23 L. Gorb, V. Kuzmin and E. Muratov, *Application of Computational Techniques in Pharmacy and Medicine*, Springer Netherlands, Dordrecht, 2014.
- 24 R. Mah, J. R. Thomas and C. M. Shafer, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 33–39.
- 25 J. Singh, R. C. Petter, T. B. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307–317.
- 26 A. J. T. Smith, X. Zhang, A. G. Leach and K. N. Houk, *J. Med. Chem.*, 2008, **52**, 225–233.
- 27 J. Kirchmair, M. J. Williamson, J. D. Tyzack, L. Tan, P. J. Bond, A. Bender and R. C. Glen, *J. Chem. Inf. Model.*, 2012, **52**, 617–648.
- 28 M. S. Finnin, J. R. Donigian, A. Cohen, V. M. Richon, R. A. Rifkind, P. A. Marks, R. Breslow and N. P. Pavletich, *Nature*, 1999, **401**, 188–193.
- 29 A. Vannini, C. Volpari, G. Filocamo, E. C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkuhler and S. Di Marco, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15064–15069.
- 30 M. Dokmanovic, C. Clarke and P. A. Marks, *Mol. Cancer Res.*, 2007, **5**, 981–989.
- 31 C. Choudhary, C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. C. Walther, J. V. Olsen and M. Mann, *Science*, 2009, **325**, 834–840.
- 32 I. Gregoretti, Y.-M. Lee and H. V. Goodson, *J. Mol. Biol.*, 2004, **338**, 17–31.
- 33 J. Sadowski, J. Gasteiger and G. Klebe, *J. Chem. Inf. Model.*, 1994, **34**, 1000–1008.
- 34 J. R. Somoza, R. J. Skene, B. A. Katz, C. Mol, J. D. Ho, A. J. Jennings, C. Luong, A. Arvai, J. J. Buggy, E. Chi, J. Tang, B.-C. Sang, E. Verner, R. Wynands, E. M. Leahy, D. R. Dougan, G. Snell, M. Navre, M. W. Knuth, R. V. Swanson, D. E. McRee and L. W. Tari, *Structure*, 2004, **12**, 1325–1334.
- 35 D. P. Dowling, S. L. Gantt, S. G. Gattis, C. A. Fierke and D. W. Christianson, *Biochemistry*, 2008, **47**, 13554–13563.
- 36 R. Ficner, *Curr. Top. Med. Chem.*, 2009, **9**, 235–240.
- 37 R. Wu, S. Wang, N. Zhou, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2010, **132**, 9471–9479.
- 38 R. Wu, P. Hu, S. Wang, Z. Cao and Y. Zhang, *J. Chem. Theory Comput.*, 2009, **6**, 337–343.
- 39 K. Chen, X. Zhang, Y.-D. Wu and O. Wiest, *J. Am. Chem. Soc.*, 2014, **136**, 11636–11643.
- 40 K. Chen, L. Xu and O. Wiest, *J. Org. Chem.*, 2013, **78**, 5051–5055.
- 41 D. Wang, P. Helquist and O. Wiest, *J. Org. Chem.*, 2007, **72**, 5446–5449.
- 42 K. Vanommeslaeghe, S. Loverix, P. Geerlings and D. Tourwé, *Bioorg. Med. Chem.*, 2005, **13**, 6070–6082.
- 43 A. S. Madsen, H. M. E. Kristensen, G. Lanz and C. A. Olsen, *ChemMedChem*, 2014, **9**, 614–626.
- 44 *Discovery Studio*, Accelrys, Inc, San Diego, CA, USA, 2.5 edn, 2010.
- 45 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollmann, *J. Am. Chem. Soc.*, 1993, **115**, 9620–9631.
- 46 T. Somboon, J. Ochiai, W. Treesuwan, M. P. Gleeson, S. Hannongbua and S. Mori, *J. Mol. Graphics Modell.*, 2014, **52**, 20–29.

- 47 M. J. T. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian, Inc.*, Wallingford CT, 2009.
- 48 S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct.*, 1999, **1–21**, 461–462.
- 49 T. Vreven, K. S. Byun, I. Komaromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, *J. Chem. Theory Comput.*, 2006, **2**, 815–826.
- 50 D. P. Dowling, S. L. Gantt, S. G. Gattis, C. A. Fierke and D. W. Christianson, *Biochemistry*, 2008, **47**, 13554–13563.
- 51 C. Corminboeuf, P. Hu, M. E. Tuckerman and Y. Zhang, *J. Am. Chem. Soc.*, 2006, **128**, 4530–4531.
- 52 S. L. Gantt, S. G. Gattis and C. A. Fierke, *Biochemistry*, 2006, **45**, 6170–6178.
- 53 A. Vannini, C. Volpari, G. Filocamo, E. C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkühler and S. Di Marco, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15064–15069.
- 54 P. E. M. Siegbahn and F. Himo, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 323–336.
- 55 P. Siegbahn and F. Himo, *JBIC, J. Biol. Inorg. Chem.*, 2009, **14**, 643–651.
- 56 D. Santos-Martins, S. Forli, M. J. Ramos and A. J. Olson, *J. Chem. Inf. Model.*, 2014, **54**, 2371–2379.
- 57 A. J. Mulholland, *Drug Discovery Today*, 2005, **10**, 1393–1402.