Uppsala University

This is a published version of a paper published in *Thermodynamics and Catalysis*.

Citation for the published paper:

Prasad, R., Kamerlin, L., Plotnikov, N., Warshel, A. (2012)

"Studying catalysis by QM/MM approaches should not be a black box process."

Thermodynamics and Catalysis, 3: 4

Access to the published version may require subscription.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-190390>

<http://uu.diva-portal.org>

Studying Catalysis by QM/MM Approaches should not be a Black Box Process

B. Ram Prasad¹, S. C. L. Kamerlin², Nikolay V. Plotnikov¹ and Arieh Warshel^{1*}

¹Department of Chemistry, University of Southern California, SGM 418, 3620 McClintock Avenue, Los Angeles, California 90089, USA

²Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, S-751 24 Uppsala, Sweden

The increasing usage of combined quantum mechanical/molecular mechanical (QM/MM) models, which started in 1976 [1], has paved the way for realistic studies of enzymatic reactions and other reactions in condensed phases. Over the years, we have witnessed an explosion in the field, with the emergence of the different versions of the QM/MM embedding idea, ranging from the empirical valence bond approach [2-7], through semi-empirical molecular orbital approaches [8-11] and *ab initio* QM/MM approaches [12,13], to frozen DFT approaches [14-17]. Additionally, increase in computational power have led to progress also with cluster models [18-21] that are clearly good substitute for some problematic QM/MM studies (see below) but are unable to explore the effect of the protein environment relative to the aqueous solvent environment. At any rate, recent years have led to increasing awareness of the fact that simple energy minimization with QM/MM approaches can lead to major problems and artifacts [22]. However, accompanying awareness of the fact that one needs very extensive sampling and free energy calculations has not reached universal recognition. That is, despite the development of powerful approaches for performing *ab initio* QM/MM (QM(ai)/MM) studies, which are based on the use of the EVB or other reference potentials [23-26] as well as closely related approaches such as the powerful metadynamics (MTD) method [27,28], the risk of obtaining problematic results still exists. Here, one of the biggest risks is associated with the belief that using a black box computational approach one can automatically get the atomic level insights about the reaction mechanisms and the factors responsible for the observed catalytic effect. As we have previously written very extensively about the risks of using completely inappropriate approaches, i.e. models such as gas-phase models or quantum mechanical cluster models [29] in exploring enzymatic reactions, here, we will focus on the traps associated with much more sophisticated QM/MM approaches, focusing on a single case, namely computational studies of phosphate hydrolysis, which serves to highlight a more general problem.

The basic issues pertaining to the mechanism of the hydrolysis of phosphate esters (and related systems) can be clearly and uniquely formulated by considering the reaction paths in the landscape presented in figure 1, where R_1 represents the nuclear reaction coordinate between the phosphorous and the oxygen atom of the departing leaving group, and R_2 describes the nuclear reaction coordinate between the phosphorous and the oxygen atom of the attacking nucleophilic water, in the example of the hydrolysis of the monomethyl pyrophosphate trianion. Computational studies on the hydrolysis of this compound have illustrated that both associative and dissociative concerted pathways are equally viable for this system [30], and that the reaction in aqueous solution is likely to proceed with proton transfer (PT) from the attacking nucleophile onto one of the non-bridging oxygen of the γ -phosphate, in line with studies on related systems (see e.g. [31-37], amongst others). However, the nature of this PT [38-40], and whether it even occurs [41,42], have been the subject of controversy. Now, the PT can occur along either of the associative (compact) or more dissociative-like (expansive) pathways. Additionally, this PT can occur

in either a direct manner (1W), as was observed in the prior works mentioned above, or with the assistance of a second water molecule (2W) (this case is also taken as a representative example for potential PT pathways involving even more water molecules). The key mechanistic issues can be examined by figure 1, which is highly relevant to recent studies involving QM/MM energy minimization [38,43] where it was found that the PT to the phosphate occurs through several intervening water molecules, rather than occurring directly from the nucleophile to the substrate. A similar scenario involving intervening water molecules was also recently suggested for phosphate triesters [44], although this work was performed using only DFT and not QM/MM calculations. A more recent study of the hydrolysis of triphosphates has also been reported by Marx and coworkers [40], who used the MTD approach in order to construct the relevant free energy surfaces, although this work is particularly problematic in light of the fact that the authors call a fairly compact “associative –like” transition state obtained in their work “dissociative”, and argue against the substrate-as-base mechanism although the phosphate is still the ultimate proton acceptor. In any case, our recent work [45], which involved very systematic mapping (one of the resulting surfaces is depicted in Figure 1), compared the 2W mechanism in the path explored by refs. [38,40,43] to the corresponding 1W mechanism that was overlooked by those studies. As can be seen from the figure, the 1W and 2W mechanisms have similar activation barriers, and, therefore, it is unjustified to assume that the 2W (or even 3W) mechanisms would be favored in aqueous solution. The reason why the 1W mechanism might have been overlooked in other studies will be clarified further below.

Specifically, an important point to take into account is that many studies that are supporting the idea of proton transfer through multiple water molecules tend to (with the exception of e.g. [44]) then overlook the comparable pathway involving a single water molecule. In fact, such studies as [40], also try to obtain a proton transfer that can occur “spontaneously” in a black box approach. What such works overlook is the fact that, as the bond to the incoming nucleophile is formed, the pK_a of the attacking water molecule is reduced drastically (from about 16 in the ground state to a negative value). As this water molecule becomes increasingly acidic, it will, at some stage, transfer its proton to the best available base. However, here there is a major difference in

***Corresponding author:** Arieh Warshel, Department of Chemistry, University of Southern California, SGM 418, 3620 McClintock Avenue, Los Angeles, California 90089, USA, E-mail: warshel@usc.edu

Received September 17, 2012; **Accepted** September 18, 2012; **Published** September 21, 2012

Citation: Prasad BR, Kamerlin SCL, Plotnikov NV, Warshel A (2012) Studying Catalysis by QM/MM Approaches should not be a Black Box Process. J Thermodynam Cat 3: e111. doi:10.4172/2157-7544.1000e111

Copyright: © 2012 Prasad BR. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Figure 1: The calculated free energy surfaces for the hydrolysis of monomethyl pyrophosphate trianion in solution for the path that starts with an associative and dissociative paths that follows by a late PT to the phosphate oxygen. As seen from the figure the rate determining TSs for the 2W and 1W have similar energies. Note that the activation entropy of the 2W path (which is not included) should be larger than that of the 1W barrier. The surfaces are drawn as functions of R_1 and R_2 for several values of the proton transfer coordinate X , and focuses on the comparing the PT barriers that starts from the intermediate where R_2 is around 2 Å and R_1 is partially broken. The transition states on the surface were explored by starting with locating corresponding gas phase transition state, solvating it by IEF-PCM, and then reevaluating the relevant energy in solution by constrained Cosmo minimization.

the energetics of the 1W and 2W pathways. That is, the 2W mechanism obtains an artificial advantage in a black box approach, because a second water molecule that could act as a base can be arranged at an optimal distance, such that compression of the P-O bond to the nucleophile results in a spontaneous PT to this water molecule (in this scenario, the exothermicity and TS energy are almost equal but the energy of the P-O bond compression is very large). On the other hand, in the case of the 1W mechanism, the energy will always go up along the PT coordinate, since it is necessary to bend the O-P-O angle to obtain optimal PT although the PT is very exothermic and the compression energy is smaller than that needed in the 2W path. The barrier for the PT in the case of the 1W model is what we believe is almost certainly the reason this path has been overlooked by standardized energy minimization (or MTD) studies. That is, to actually obtain this path, one requires a more complex reaction coordinate that either takes into account the uphill proton transfer as a distinct extra dimension to the surface, or scans in both directions (as is done in figure 1) such that one can obtain a More O'Ferrall-Jencks plot that is an intersection of two surfaces (dependent on proton position), thus compressing the PT coordinate into a 2-D projection. In any case, as seen by the examples above, such a complex scenario will almost certainly be missed by an automated approach that do not search in the PT direction, and it is worth highlighting here that

high level QM/MM studies using explicit water molecules that correctly define the reaction coordinate obtain a similar 1W mechanism to that in our previous work [46] (although these authors have not explored the alternate, 2W pathway).

We believe this should serve as both a guide and a warning to those workers who are performing cutting-edge studies of complex biologically relevant systems, and, in fact, this warning holds not only for energy minimization approaches, but also for the far more sophisticated MTD approach (and other improved sampling strategies). Additionally, the same points highlighted here are also relevant to phosphate hydrolysis in proteins. Here, the chance of getting problematic results is much larger, and the confusion is likely to persist for a longer time. This is why we view EVB studies as the most reliable current approach for exploring enzymatic reactions. This means, however, that one must start with a very careful search of all key mechanistic options in solution, and then ask (without prior bias) what happens to these mechanisms when we move to the protein. Of course, with increasing computational power, we can start to explore the effect of moving from the EVB to QM(ai)/MM surfaces by the Para dynamics (PD) approach. However, this approach is likely to give reasonable results only when performing extensive EVB sampling.

Acknowledgments

This work was supported by NSF grant MCB-08364000 and NCI (1U19CA105010) to AW, as well as Swedish Research Council (VR) grant 2012-5026, Carl Tryggers grant CT 11:226 and ERC starting grant 306474 to SCLK. We also acknowledge the University of Southern California's High Performance Computing and Communications Center for computer time.

References

- Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J Mol Biol 103: 227-249.
- Warshel A, Weiss RM (1980) An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. J Am Chem Soc 102: 6218-6226.
- Warshel A, Weiss RM (1981) Empirical valence bond calculations of enzyme catalysis. Ann N Y Acad Sci 367: 370-382.
- Åqvist J, Warshel A (1993) Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. Chem Rev 93: 2523-2544.
- Warshel A (1997) Computer Modeling of Chemical Reactions in Enzymes and Solutions. John Wiley & Sons.
- Kamerlin SCL, Warshel A (2011) The empirical valence bond model: theory and applications. Wiley Interdiscip Rev Comput Mol Sci 1: 30-45.
- Kamerlin SCL, Warshel A (2010) The EVB as a quantitative tool for formulating simulations and analyzing biological and chemical reactions. Faraday Discuss 145:71-106.
- Bakowies D, Thiel W (1996) Hybrid Models for Combined Quantum Mechanical and Molecular Approaches. J Phys Chem 100: 10580-10594.
- Bakowies D, Thiel W (1996) Semiempirical treatment of electrostatic potentials and partial charges in combined quantum mechanical and molecular mechanical approaches. J Comput Chem 17: 87-108.
- Cummins PL, Gready JE (2003) Computational methods for the study of enzymic reaction mechanisms. II. An overlapping mechanically embedded method for hybrid semi-empirical-QM/MM calculations. J Mol Struct 632: 247-257.
- Devi-Kesavan LS, Garcia-Viloca M, Gao J (2003) Semiempirical QM/MM potential with simple valence bond (SVB) for enzyme reactions. Application to the nucleophilic addition reaction in haloalkane dehalogenase. Theor Chem Acc 109: 133.
- Dapprich S, Komaromi I, Byun KS, Morokuma K, Frisch MJ (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. J. Mol Struct (THEOCHEM) 462: 1-21.
- Maseras F, Morokuma K (1995) IMOMM: A new integrated *ab initio* + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. J Comput Chem 16 :1170-1179.
- Wesolowski TA, Warshel A (1993) Frozen density functional approach for *ab initio* calculations of solvated molecules. J Phys Chem 97: 8050-8053.
- Wesolowski T, Warshel A (1994) *Ab Initio* Free Energy Perturbation Calculations of Solvation Free Energy Using the Frozen Density Functional Approach. J Phys Chem 98: 5183-5187.
- Wesolowski T, Muller RP, Warshel A (1996) *Ab Initio* Frozen Density Functional Calculations of Proton Transfer Reactions in Solution. J Phys Chem 100: 15444-15449.
- Olsson MH, Hong G, Warshel A (2003) Frozen density functional free energy simulations of redox proteins: computational studies of the reduction potential of plastocyanin and rusticyanin. J Am Chem Soc 125: 5025-5039.
- Blomberg MR, Siegbahn PE (2006) Quantum chemistry applied to the mechanisms of transition metal containing enzymes -- cytochrome c oxidase, a particularly challenging case. J Comput Chem 27: 1373-1384.
- Bloomberg MRA, Siegbahn PEM (2001) A Quantum Mechanical Approach to the Study of Reaction Mechanisms of Redox Active Metalloenzymes. J Phys Chem 105: 9376-9386.
- Siegbahn PEM (2001) Modeling aspects of mechanisms for reactions catalyzed by metalloenzymes. J Comput Chem 22: 1634-1645.
- Siegbahn PE, Blomberg MR (1999) Density functional theory of biologically relevant metal centers. Annu Rev Phys Chem 50: 221-249.
- Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in *ab initio* quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. J Phys Chem B 109: 15645-15650.
- Muller RP, Warshel A (1996) *Ab initio* calculations of free energy barriers for chemical reactions in solution: proton transfer in [FHF]⁻. Pac Symp Biocomput 524-538.
- Kamerlin SC, Haranczyk M, Warshel A (2009) Progress in *ab initio* QM/MM free-energy simulations of electrostatic energies in proteins: accelerated QM/MM studies of pKa, redox reactions and solvation free energies. J Phys Chem B 113: 1253-1272.
- Plotnikov NV, Kamerlin SC, Warshel A (2011) Paradynamics: an effective and reliable model for *ab initio* QM/MM free-energy calculations and related tasks. J Phys Chem B 115: 7950-7962.
- Plotnikov NV, Warshel A (2012) Exploring, Refining, and Validating the Paradynamics QM/MM Sampling. J Phys Chem B 116: 10342-10356.
- Laio A, Parrinello M (2002) Escaping free-energy minima. Proc Natl Acad Sci U S A 99: 12562-12566.
- Ensing B, De Vivo M, Liu Z, Moore P, Klein ML (2006) Metadynamics as a tool for exploring free energy landscapes of chemical reactions. Acc Chem Res 39: 73-81.
- Warshel A (2003) Computer simulations of enzyme catalysis: methods, progress, and insights. Annu Rev Biophys Biomol Struct 32: 425-443.
- Kamerlin SC, Florian J, Warshel A (2008) Associative versus dissociative mechanisms of phosphate monoester hydrolysis: on the interpretation of activation entropies. Chemphyschem 9: 1767-1773.
- Langen R, Schweins T, Warshel A (1992) On the mechanism of guanosine triphosphate hydrolysis in ras p21 proteins. Biochemistry 31: 8691-8696.
- Florián J, Warshel A (1997) A Fundamental Assumption about OH⁻ Attack in Phosphate Hydrolysis is not Fully Justified. J Am Chem Soc 119: 5473-5474.
- Kamerlin SC (2011) Theoretical comparison of p-nitrophenyl phosphate and sulfate hydrolysis in aqueous solution: implications for enzyme-catalyzed sulfuryl transfer. J Org Chem 76: 9228-9238.
- Kamerlin SC, McKenna CE, Goodman MF, Warshel A (2009) A computational study of the hydrolysis of dGTP analogues with halomethylene-modified leaving groups in solution: implications for the mechanism of DNA polymerases. Biochemistry 48: 5963-5971.
- Florián J, Warshel A (1998) Phosphate Ester Hydrolysis in Aqueous Solution: Associative versus Dissociative Mechanisms. J Phys Chem B 102: 719-734.
- Klahn M, Rosta E, Warshel A (2006) On the mechanism of hydrolysis of phosphate monoesters dianions in solutions and proteins. J Am Chem Soc 128: 15310-15323.
- Kamerlin SCL, Sharma PK, Prasad BR, Warshel A (2012) Why nature really chose Phosphate? Q Rev Biophys (In Press).
- Grigorenko BL, Nemukhin AV, Shadrina MS, Topol IA, Burt SK (2007) Mechanisms of guanosine triphosphate hydrolysis by Ras and Ras-GAP proteins as rationalized by *ab initio* QM/MM simulations. Proteins Struct Funct Bioinf 66: 456-466.
- Grigorenko BL, Nemukhin AV, Topol IA, Cachau RE, Burt SK (2005) QM/MM modeling the Ras-GAP catalyzed hydrolysis of guanosine triphosphate. Proteins Struct Funct Bioinf 60: 495-503.
- Glaves R, Mathias G, Marx D (2012) Mechanistic insights into the hydrolysis of a nucleoside triphosphate model in neutral and acidic solution. J Am Chem Soc 134: 6995-7000.

41. Admiraal SJ, Herschlag D (2000) The Substrate-Assisted General Base Catalysis Model for Phosphate Monoester Hydrolysis: Evaluation Using Reactivity Comparisons. *J Am Chem Soc* 122: 2145-2148.
42. Lassila JK, Zalatan JG, Herschlag D (2011) Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu Rev Biochem* 80: 669-702.
43. Grigorenko BL, Rogov AV, Nemukhin AV (2006) Mechanism of triphosphate hydrolysis in aqueous solution: QM/MM simulations in water clusters. *J Phys Chem B* 110: 4407-4412.
44. Mora JR, Kirby AJ, Nome F (2012) Theoretical study of the importance of the spectator groups on the hydrolysis of phosphate triesters. *J Org Chem* 77: 7061-7070.
45. Ram Prasad B, Plotnikov N, Warshel A (2012) Resolving uncertainties about phosphate hydrolysis pathways by careful free energy mapping. *J Phys Chem B* (submitted).
46. Li W, Rudack T, Gerwert K, Gräter F, Schlitter J (2012) Exploring the multidimensional free energy surface of phosphoester hydrolysis with constrained QM/MM dynamics. *J Chem Theory Comput* (DOI: 10.1021/ct300022m).