



Published in final edited form as:

J Phys Chem B. 2005 August 18; 109(32): 15645–15650.

On Possible Pitfalls in *ab initio* QM/MM Minimization Approaches For Studies of Enzymatic Reactions

Marco Klähn, Sonja Braun-Sand, Edina Rosta, and Arieh Warshel*

Department of Chemistry, University of Southern California, 3620 S. McClintock Ave., Los Angeles, California, 90089-1062

Abstract

Reliable studies of enzymatic reactions by combined quantum mechanics /molecular mechanics (QM/MM) approaches, with an *ab initio* description of the quantum region, presents a major challenge to computational chemists. The main problem is the need for a very large computer time for the evaluation of the QM energy, which in turn makes it extremely challenging to perform proper configurational sampling. A seemingly reasonable alternative is to perform energy minimization studies of the type used in gas phase *ab initio* studies. However, it is hard to see why such an approach should give reliable results in protein active sites. In order to examine the problems with energy minimization QM/MM approaches we chose the hypothetical reaction of a metaphosphate ion with water in the Ras•GAP complex. This hypothetical reaction served as a simple benchmark reaction. The possible problems with the QM/MM minimization were explored by generating several protein configurations from long MD simulations and using energy minimization and scanning of the reaction coordinates to evaluate the corresponding potential energy surfaces of the reaction for each of these different protein configurations. Comparing these potential energy surfaces, we found major variations of the minima of the different total potential energy surfaces. Furthermore, the reaction energies and activation energies also varied significantly even for similar protein configurations. The specific coordination of a magnesium ion, present in the active center of the protein complex, turned out to influence the energetics of the reaction in a major way and a direct coordination to the reactant leads to an increase of the activation energy by 17 kcal/mol. This study demonstrates that energy minimizations starting from a single protein structure could lead to major errors in calculations of activation free energies and binding free energies. Thus we believe that extensive samplings of the configurational space of the protein are essential for meaningful determination of the energetics of enzymatic reactions. The possible relevance of our conclusion with regard to a recent study of the RasGAP reaction is discussed.

I. Introduction

QM/MM approaches have provided a general scheme for studies of chemical processes in proteins^{1–11}. A significant progress has been made with calibrated semiempirical QM/MM approaches^{2,7,10,11}. However, one would like to move to an *ab initio* representation with a QM/MM treatment since such QM representations have been shown to provide “chemical accuracy” in studies of gas phase reactions of small molecules¹². In principle, when one uses a reliable large *ab initio* QM region we can expect to obtain a reliable description of the potential surface of the reaction region. Unfortunately, this does not mean that the actual free energy barriers are estimated correctly, and in fact we may have many serious problems. First, it is essential to provide a proper long-range treatment and effective boundary conditions for the protein+solvent environments¹³. Second, it is important to use a polarizable force field for the

* Corresponding author. E-mail: warshel@usc.edu.

MM region and finally it is important to perform sufficient sampling to obtain the actual free energy^{2,14}.

The present work will not focus on the sampling problem, which was discussed recently^{2,15,16} in terms of possible simulation strategies. Here we focus, however, on the issue of using energy minimization for a few protein configurations^{17,45}. Obviously this provides an advance over gas phase or single point calculations. However, this approximated strategy was not subjected to a systematic scrutiny.

When one deals with enzyme active sites the challenge of finding the transition state region and evaluating the activation barrier is quite different than in gas phase energy search, since the protein landscape is very complex and different configurations cannot be found by standard energy minimization approaches. Furthermore, results obtained from different minima can be very different, as already observed¹⁷.

In order to illustrate the above problems we take a relatively simple case, a segment of a possible path in the reaction of Ras, where an H₂O attacks a metaphosphate which can in principal be formed by a dissociative mechanism. This type of reaction has been considered in a recent QM/MM energy minimization study where it was concluded that the barrier for the dissociative bond breaking is very low and the rate determining step is a proton transfer to the metaphosphate through a concerted path that involves both Gln61 and the catalytic water. It has also been concluded that a direct proton transfer from the catalytic water to the metaphosphate involves a very high barrier⁴⁵. The present paper does not attempt to explore the possible involvement of such a reaction path in the reaction of the Ras•GAP system, since this should involve a very careful evaluation of the free energy surface for the associative and dissociative mechanisms, a careful calibration and a systematic comparison to relevant experimental results. Instead we merely chose the hypothetical water-metaphosphate system as a simple benchmark for QM/MM calculations, and we will only make some general comments about the new proposal of ref⁴⁵.

Section II describes our benchmark and outlines the computational methods. Section III describes the results of the calculations and demonstrates their sensitivity to the protein conformational states and the position of the Mg²⁺ ion. Finally, we discuss in section IV the implications of our findings and emphasize the importance of a proper average over the protein configurations.

II. Methods and systems

This work considers some aspects of the rather complex mechanism of the Ras•GAP system, whose active site and a model of the bound GTP is depicted in Fig 1.

However, instead of focusing on the actual reaction and its biological implications we focus on a relatively simple model reaction, namely the attack of H₂O on a metaphosphate formed in the active site of the Ras•GAP system, following an hypothetical dissociative cleavage of the GTP substrate (Fig 1). The metaphosphate water system was used as a benchmark in exploring the validity of QM/MM transition state searches in condensed phases in general and in proteins in particular.

The potential energy surface (PES) of our reaction was explored by considering the two reaction coordinates described in Fig 2, the distance of the water oxygen to the phosphorous, R, and the distance of the transferred proton to proton acceptor oxygen of the metaphosphate, r.

We also examined the possibility that a possible third reaction coordinate (i.e. the distance of the transferred proton to the water oxygen) can be neglected and its inclusion does not change our results quantitatively.

In order to evaluate the potential surface for the reaction in solution we started by scanning the surface in the gas phase as a function of R and r, where at each point we constrained the system to the specific R and r and minimized the energy with regards to all other coordinates. Next we considered the effect of the solvent by using the conductor reaction field approach COSMO²⁰ and performed a single point COSMO calculation at each of the scanning points. The resulting PES served as an approximation for the energetics of the reaction in solution. In doing so we assumed implicitly that minimization with respect to the coordinates orthogonal to R and r will give similar results in the gas phase and in solution. This assumption was verified in some specific cases²¹. It seems to us that the above scanning procedure is at present more effective than alternative options of performing a transition state search on the solution surface using analytical derivatives. That is, in our experience there are problems with the analytical derivatives of most current *ab initio* solvation models, at least with respect to the dependence of the cavity surface and the reaction field on the solute coordinate. Here the use of a systematic scanning procedure is the simplest way of obtaining a reliable mapping of the transition state region.

In addition to the COSMO calculations we also considered the previously reported results obtained by the Langevin dipole model²². The gas phase optimizations were performed using the Hartree Fock method and a 6-31 G(d) basis set, whereas the single point calculations that considered the solvent effect used the DFT method with the hybrid functional B3LYP and an extensive 6-311++G(d,p) basis set. The results of this minimization/scanning approach will be described in the next section.

The next step of our study involved a QM/MM evaluation of the potential surface of our system in the protein active site. Before describing the actual calculations we will provide below a brief description of our QM/MM approach emphasizing the specific technical points of the current implementation. The starting point in QM/MM models is the separation of the system into QM and MM regions¹. In systems where the quantum region is composed of a molecule or molecules that are not bound to the classical region, as is often the case in solution-phase reactions, separation of the two regions is rather trivial. The separation into quantum and classical regions is not as straightforward in enzyme reactions, because the quantum region is frequently bonded to the classical region, and it is not clear how to define the boundary conditions for the electronic structure calculations of the quantum region, nor how to incorporate the electrostatic and van der Waals effects of the classical region into the quantum region energy expression. An effective way of connecting the two regions can be provided using hybrid orbital¹ and related techniques^{14,23}. Unfortunately, these approaches are somewhat difficult to implement, and because the current QM(ai)/MM method aims to be an efficient method of linking standard programs (in this case GAUSSIAN²⁴ and MOLARIS²⁵), we chose the simpler method of using link atoms. A link atom (LA)²⁶ is an atom inserted along the bond between the quantum and classical regions. Such an approach was introduced¹, in addition to the introduction of the use of the hybrid orbital approach. Using the nomenclature of ref²⁷, the quantum mechanical atom to which the link atom is bonded is referred to as the link atom bond partner (LABP), and the classical atom replaced by the link atom is referred to as the link atom host (LAH). Warshel and Levitt¹ introduced the link atom treatment in their AMI/MM model of the catalytic reaction of lysozyme, while using the more reliable hybrid orbital treatment¹ in their QCFF/all study of lysozyme. Although the name LA was not introduced the method was incorporated into the ERFN program²⁹ and thus constitutes the first LA QM/MM treatment. The LA treatment was also introduced in QM/MM studies of π -electron systems.⁴⁶ Warshel and Levitt¹ and Ostlund³⁰ developed methods that adjust the

ionization potential of the LAs to reproduce properties of the given molecule in the presence of the actual LAHs. In our and previous studies^{30,27}, the LA is a hydrogen atom. Figure 3 illustrates the process of defining the quantum region and inserting LAs.

This figure shows the reacting system separated into the quantum region and a small part of the MM region, which is covalently linked to the QM region. The structure below the quantum fragment shows this fragment capped with LAs to produce the quantum region. The LAs are inserted along the LAH-LABP bond at a distance determined by gas-phase *ab initio* energy minimization. Despite the intense current interest in proper treatment of link atoms and the availability of more effective hybrid orbital and related localized bond approaches^{1,23}, we feel that this problem is not so crucial in studies of enzymatic reactions. That is, errors introduced by using a small fragment are generally similar in solution phase and protein simulations. Thus, such errors are largely canceled when one is concerned with enzyme catalysis, which reflects the difference between the activation energy in the enzyme and in the reference solution reaction. In the present case we treated the metaphosphate and the water quantum mechanically using the Hartree-Fock approach with a 6-31G(d) basis set, as was done for the reaction in solution. We also estimated a correction for the free reaction energies and free activation energies by calculating the energies at the reactant, product and transition states using the B3LYP functionals with a 6-311++G(d,p) basis set.

Before applying the above QM/MM approach we generated starting protein configurations. This was done by constructing an EVB potential for the QM system and used this potential to propagate trajectories in the reactant state. The trajectories started from a model structure of the Ras•GAP complex derived from the crystal structure of Scheffzek et al.¹⁹. Because we have been only interested here in the hypothetical reaction of metaphosphate (that is already dissociated from the GDP moiety) with a water molecule, we separated the metaphosphate from GDP by applying a soft harmonic potential with a minimum at 3 Å distance between the two compounds. This model system was first equilibrated by an MD run of 500 ps and then the simulation continued for another 500 ps, generating five representative protein configurations. These configurations were taken as the starting points for the QM/MM studies described in the next section.

III. Results and discussion

Although it is convenient to assume that *ab initio* QM/MM approaches should give reliable results, this is far from being certain. Thus before performing QM/MM calculations of chemical processes in enzymes it is important to validate the method used by calculating the energetics of the given reaction in solution. That is, since QM/MM calculations in enzymes do not yet give chemical accuracy it is essential to calibrate (or validate) the specific method used by examining the agreement between the calculated and observed energetics in solution reactions (see discussion in II). Such calculations are also needed in order to assess the catalytic effect of the enzyme. At any rate, we established the validity of the quantum mechanical model by evaluating the free energy surface for the solution reaction using the COSMO solvent and the mapping procedure described in the method section. The corresponding results are summarized in Table 1 (together with the previously reported results with LD solvent model) and Fig 4.

Fig. 4 depicts the equipotential lines of the PES, the transition state and the corresponding reaction path. The figure shows that the PES in the reactant region is shallow and that the water molecule can reach a very close distance ($R=2.1$ Å) to the metaphosphate before the proton is transferred. The transition state is found at $R=1.9$ Å and $r=1.3$ Å. The reaction free energy, ΔG_0 , is -23.2 kcal/mol and the corresponding activation free energy is $\Delta G^\ddagger=13.3$ kcal/mol. These energy differences are in a very good agreement with the $\Delta G_0=-24.0$ and $\Delta G^\ddagger=11.0$

kcal/mol values obtained by Florian and Warshel²¹ from *ab initio*/LD calculations and from analysis of relevant experimental data.

We also examined the influence of electron correlation and the basis set on the PES. That is, in addition to our chosen B3LYP/6-311++G(d,p) calculations we also evaluated the activation barrier and reaction energy with the HF/LanL2DZ and B3LYP/LanL2DZ methods. The results of this study are also summarized in Table 1.

From this comparison we conclude that the neglect of Coulomb electron correlation in the Hartree Fock approach leads to rather inaccurate estimates for both the reaction and the activation free energy, as expected. Both energy differences deviate around 50% from our original results. The deviations of the smaller basis set, LanL2DZ, compared to the triple zeta basis set are smaller but using this basis set somewhat underestimates the reaction free energy.

After establishing the reliability of the potential surfaces for the solution reaction we explored the reaction in the active site of the Ras•GAP complex. In this case we performed the transition state search with several initial protein configurations. The potential surface for each protein configuration was evaluated by the same mapping procedure described in the discussion of the solvent surface and the method section, but this time using the QM/MM description of the protein. It is important to notice that even in the case of a single protein configuration there is a risk of obtaining an incorrect transition state. This issue is considered in Fig. 5, where we depict an example of the PES for a specific protein configuration. As seen from the figure, starting from different initial reactant coordinates and using an incomplete configurational search can lead us to different transition states. This problem, that can also occur in gas phase studies, is rather well known^{31,32} and is not the topic of this work.

The actual serious problem addressed by this paper starts to become apparent when we consider the effect of the protein configurations. This issue is demonstrated in Fig. 6 and Table 2, where we examine the dependence of the lowest energy profile on the protein configuration used.

As seen from the figure the absolute energies of the ground state can easily vary by 30 kcal/mol between different protein configurations. This finding makes it rather clear that calculations of binding free energies, which reflect the energy of the reactant state, cannot be performed by QM/MM approaches without a very extensive averaging. The situation seems to be less catastrophic when one considers the activation barriers relative to the given ground state energy. Here the variation is about 6 kcal/mol, which is large but can be used in some qualitative considerations. Unfortunately this is only the tip of the iceberg. That is, while the error in ΔU^\ddagger obtained in closely related configurations may be tolerable, some configurational changes might lead to much larger effects.

A detailed analysis of the effect of the protein fluctuations on the QM/MM surfaces is usually rather complicated. However, one can focus on some key elements. Here we explored the effect of changes in the position of the Mg^{2+} ion, considering two different coordinations that were both found to be stable (assuming that the Mg^{2+} ion is fixed at the single position found in the transition state analog is apparently unjustified). The results of this study are summarized in Fig 7 where we compared the reaction profile associated with the two initial configurations depicted on the right hand side of the figure. As seen there, increasing the distance between the Mg^{2+} ion and the metaphosphate oxygen reduces the barrier by 17 kcal/mol.

The results presented in Fig. 7 clarify that the situation in QM/MM calculations in proteins are quite different than those obtained in gas phase calculations. That is, while in gas phase calculations it is reasonable to assume that a convergent minimization approach will lead automatically to a displacement of the magnesium ion, it is unlikely that this will happen in an energy minimization in the protein active site. These results underscore the importance of a

proper conformational search in QM/MM evaluations of activation energies. Here one of the best options is to use EVB as a reference potential for QM/MM free energy calculations³³.

The determination of the relative role of the two magnesium ion configurations of Fig. 7 is far from being simple. In principle one can start by evaluating the electrostatic free energy of the Mg^{2+} ion in the two configurations by a FEP approach. Then by assuming that the energy of moving the Mg^{2+} ion between the two sites is small, it would be possible to estimate the relative energy of the corresponding configurations. This treatment can be improved by the linear response approximation (LRA) treatment, which is quite effective in capturing the protein reorganization energy³⁴. However, such a study is out of the scope of the present work, since it requires one to resolve the issue of the optimal representation of the Mg^{2+} ion. That is, an MM representation of the magnesium ion considers charge transfer to the substrate by the empirical van der Waals parameters. On the other hand, the seemingly superior treatment that includes the Mg^{2+} ion in the QM region may be problematic if the ligands of this ion are treated classically (in this case one neglects the charge transfer from the ligands to the metal and the inclusion of the ligands in the QM region is likely to be too expensive^{35,36}). At present we are focusing on the assumption that both configurations are accessible and thus can be used as an instructive example of the risk in QM/MM energy minimization approaches. In view of the problems with simple QM/MM minimizations or scanning procedures, it seems obvious that one should attempt to average such calculations over the relevant protein configurations. The most rigorous way of obtaining such averages is by free energy calculations but such calculations can be too expensive when one uses a high level *ab initio* treatment for the QM region and regular free energy perturbation treatments. This problem can be drastically reduced by using the EVB or other simple potentials as a reference for the *ab initio* QM/MM free energy calculations^{37,38}.

IV. Concluding Remarks

The evaluation of the energetics of chemical processes in proteins by classical force fields requires usually very extensive averaging over the configurational space of the protein. Thus it is quite likely that the same requirement will hold for QM/MM calculations. This means that simple minimization approaches of the type used in gas phase QM calculations might not be effective in evaluations of activation energies of chemical reactions in proteins. The present work examines this issue and demonstrates that QM/MM energy minimization approaches can lead to significant errors in evaluation of activation barriers of enzymatic reactions. The difficulties in using minimization approaches can be particularly serious when the enzyme substrate complex involves a rugged landscape with many local minima. An instructive example of the coupling between the protein coordinate and the solute activation barrier is provided here where the position of the Mg^{2+} ion determines the transition state energy.

While the afore mentioned problems with regards to the evaluation of activation barriers might not occur in some cases, the related problems associated with calculations of binding free energies are much more serious. That is, as demonstrated in Fig 6, calculations of absolute ground state energies and thus also binding energies by QM/MM minimization are extremely problematic³⁹. Here one expects to obtain very different results in different local minima. Obviously the only way to obtain reasonable binding energies is to perform very extensive averaging of the type performed in QM/MM redox calculations⁴⁰.

The present test case involves the attack of a water molecule on a hypothetical metaphosphate intermediate in the Ras•GAP complex. This intermediate has been proposed recently in the QM/MM energy minimization study of Grigorenko et al.⁴⁵. As stated in the introduction we did not examine yet the proposal of a very low barrier for the dissociative formation of the metaphosphate, but we find it to be rather unlikely. Furthermore, the idea that a concerted H₂O-

Gln61 attack on the metaphosphate as the rate limiting step is problematic. For example, this concerted reaction cannot occur in the elongation factor Tu (EF-Tu) which has histidine instead of Gln61, but this system hydrolyses GTP quite effectively.⁴⁷ Thus we believe that despite the intriguing insight provided by the studies of Grigorenko et al. it is crucial to reexamine the corresponding conclusions by a well calibrated approach (with a careful attention to a more complete sampling). It would also be crucial to examine the corresponding result in related systems such as Ras alone as was done in EVB studies.⁴³

Finally, perhaps the simplest message that emerges from this study is the idea that QM/MM minimization and scanning approaches should involve averaging (and validation) over several protein configurations generated by long MD simulations. This idea has been found to be very useful for theoretical IR spectroscopy of the ground state of GTP in Ras^{41,42} and in statistical mechanically more rigorous treatments such as FEP calculations of electrostatic energies¹³ and in EVB calculations of enzymatic reactions^{43,44}.

Acknowledgements

This work was supported by NSF grant MCB-0342276 and NCI Grant 1 U19 CA105010-01. The computational work was supported by University of Southern California High Performance Computing and Communication Center (HPCC).

References

1. Warshel A, Levitt M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 1976;103:227. [PubMed: 985660]
2. Shurki A, Warshel A. Structure/Function Correlations of Proteins using MM, QM/MM and Related Approaches; Methods, Concepts, Pitfalls and Current Progress. *Advances in Protein Chemistry* 2003;66:249. [PubMed: 14631821]
3. Gao J. Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc Chem Res* 1996;29:298.
4. Bakowies D, Thiel W. Hybrid Models for Combined Quantum Mechanical and Molecular Approaches. *J Phys Chem* 1996;100:10580.
5. Field MJ, Bash PA, Karplus M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J Comp Chem* 1990;11:700.
6. Friesner R, Beachy MD. Quantum Mechanical Calculations on Biological Systems. *Curr Op Struct Biol* 1998;8:257.
7. Monard G, Merz KM. Combined Quantum Mechanical/Molecular Mechanical Methodologies Applied to Biomolecular Systems. *Acc Chem Res* 1999;32:904.
8. Garcia-Viloca M, Gonzalez-Lafont A, Lluch JM. A QM/MM study of the Racemization of Vinylglycolate Catalysis by Mandelate Racemase Enzyme. *J Am Chem Soc* 2001;123:709. [PubMed: 11456585]
9. Marti S, Andres J, Moliner V, Silla E, Tunon I, Bertran J. Transition structure selectivity in enzyme catalysis: a QM/MM study of chorismate mutase. *Theor Chem Acc* 2001;105:207.
10. Field M. Stimulating Enzyme Reactions: Challenges and Perspectives. *J Comp Chem* 2002;23:48. [PubMed: 11913389]
11. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J Phys Chem B* 2001;105:569.
12. Pople JA. Quantum chemical models (Nobel lecture). *Angew Chem Int Ed Engl* 1999;38:1894.
13. Lee FS, Warshel A. A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *J Chem Phys* 1992;97:3100.
14. Zhang Y, Liu H, Yang W. Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface. *J Chem Phys* 2000;112:3483.

15. Strajbl M, Hong G, Warshel A. Ab-initio QM/MM Simulation with Proper Sampling: "First Principle" Calculations of the Free Energy of the Auto-dissociation of Water in Aqueous Solution. *J Phys Chem B* 2002;106:13333.
16. Bentzien, J.; Florián, J.; Glennon, T.; Warshel, A. QM/MM approaches for studying chemical reactions in proteins and solution. In *Combined Quantum Mechanical & Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; American Chemical Society: Washington, DC, 1998; Vol. 712; pp 16
17. Nemukhin AV, Grigorenko BL, Rogov AV, Topol IA, Burt SK. Modeling of serine protease prototype reactions with the flexible effective fragment potential quantum mechanical/molecular mechanical method. *Theor Chem Acc* 2004;111:36.
18. Topol IA, Cachau RE, Nemukhin AV, Grigorenko BL, Burt SK. Quantum chemical modeling of the GTP hydrolysis by the RAS-GAP protein complex. *Biochimica Et Biophysica Acta-Proteins and Proteomics* 2004;1700:125.
19. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, Schmitz F, Wittinghofer A. The Ras-RasGAP Complex - Structural Basis for GTPase Activation and its Loss in Oncogenic Ras Mutants. *Science* 1997;277:333. [PubMed: 9219684]
20. Eckert F, Klamt A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *Aiche Journal* 2002;48:369.
21. Florián J, Warshel A. Phosphate Ester Hydrolysis in Aqueous Solution: Associative Versus Dissociative Mechanisms. *J Phys Chem B* 1998;102:719.
22. Florián J, Warshel A. Langevin Dipoles Model for Ab Initio Calculations of Chemical Processes in Solution: Parameterization and Application to Hydration Free Energies of Neutral and Ionic Solutes, and Conformational Analysis in Aqueous Solution. *J Phys Chem B* 1997;101:5583.
23. Théry V, Rinaldi D, Rivail JL, Maigret B, Ferenczy GG. Quantum Mechanical Computations on Very Large Molecular Systems: The Local Self-Consistent Field Method. *J Comp Chem* 1994;15:269.
24. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, Ö. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komáromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
- 25a. Lee FS, Chu ZT, Warshel A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by POLARIS and ENZYMIK Programs. *J Comp Chem* 1993;14:161.
- 25b. Chu, Z. T.; Villa, J.; Strajbl, M.; Schutz, C. N.; Shurki, A.; Warshel, A. "MOLARIS version beta9.05". 2004.
26. Singh UC, Kollman PAA. Combined Ab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH₃Cl + Cl⁻ Exchange Reaction and Gas Phase Protonation of Polyethers. *J Comput Chem* 1986;7:718.
27. Eurenium KP, Chatfield DC, Brooks BR, Hodoscek M. Enzyme Mechanisms with Hybrid Quantum and Molecular Mechanical Potentials. 1. Theoretical Considerations. *Int J Quant Chem* 1996;60:1189.
28. Sharon, R.; Warshel, A. QCFF/ALL; Weizmann Institute of Science: Israel, 1973.
29. Levitt, M.; Warshel, A. ERFN/QM; Weizmann Institute of Science: Israel, 1974.
30. Ostlund, N. S. Hyperchem; Hypercube, Inc.: Cambridge, Ontario, 1985.
31. Gonzales C, Schlegel BH. Reaction path following in mass-weighted internal coordinates. *J Phys Chem* 1990;94:5523.
32. Ayala PY, Schlegel HB. A combined method for determining reaction paths, minima, and transition state geometries. *J Chem Phys* 1997;107:375.
33. Bentzien J, Muller RP, Florián J, Warshel A. Hybrid ab Initio Quantum Mechanics/Molecular Mechanics Calculations of Free Energy Surfaces for Enzymatic Reactions: The Nucleophilic Attack in Subtilisin. *J Phys Chem B* 1998;102:2293.

34. Strajbl M, Shurki A, Warshel A. Converting conformational changes to electrostatic energy in molecular motors: The energetics of ATP synthase. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100:14834. [PubMed: 14657336]
35. Warshel, A. Computer Simulations of the Action of Metalloenzymes. In *NATO: ASI Series, Molecular Modeling and Dynamics of Bioinorganic Systems*; Banci, L., Comba, P., Eds.; Kluwer Academic Publishers, 1997; pp 343.
36. Hong G, Strajbl M, Wesolowski TA, Warshel A. Constraining the electron densities in DFT method as an effective way for ab initio studies of metal catalyzed reactions. *J Comp Chem* 2000;21:1554.
37. Muller RP, Warshel A. Ab Initio Calculations of Free Energy Barriers for Chemical Reactions in Solution. *J Phys Chem* 1995;99:17516.
38. Bentzien J, Muller RP, Florian J, Warshel A. Hybrid ab initio Quantum Mechanics/Molecular Mechanics (QM(ai)MM) Calculations of Free Energy Surfaces For Enzymatic Reactions: The Nucleophilic Attack in Subtilisin. *J Phys Chem B* 1998;102:2293.
39. Olsen L, Rasmussen T, Hemmingsen L, Ryde U. Binding of benzylpenicillin to metallo-beta-lactamase: A QM/MM study. *J Phys Chem B* 2004;108:17639.
40. Olsson MHM, Hong G, Warshel A. Frozen Density Functional Free Energy Simulations of Redox Proteins: Computational Studies of the Reduction Potential of Plastocyanin and Rusticyanin. *J Am Chem Soc* 2003;125:5025. [PubMed: 12708852]
41. Klähn, M.; Schlitter, J.; Gerwert, K. Theoretical IR spectroscopy based on QM/MM calculations provides changes in charge distribution, bond lengths and bond angles of the GTP ligand induced by the Ras-protein, *Biophys. J.* 2005, 88, in press.
42. Klähn M, Mathias G, Kottling C, Nonella M, Schlitter J, Gerwert K, Tavan P. IR spectra of phosphate ions in aqueous solution: Predictions of a DFT/MM approach compared with observations. *J Phys Chem A* 2004;108:6186.
43. Shurki A, Warshel A. Why does the Ras switch “break” by oncogenic mutations? *Proteins-Structure Function and Bioinformatics* 2004;55:1.
44. Florian, J.; Goodman, M. F.; Warshel, A. Computer simulations of protein functions: Searching for the molecular origin of the replication fidelity of DNA polymerase, *Proceedings of the National Academy of Sciences of the United States of America* 2005, in press.
45. Grigorenko, B. L.; Nemukhin, A. V.; Topol, I. A.; Cachau, R. E.; Burt, S. K. QM/MM modeling the Ras-GAP catalyzed hydrolysis of guanosine triphosphate, *Proteins: Structure, Function, and Genetics* 2005, in press.
46. Warshel A, Lippicirella A. Calculations of Ground-and Excited-State Potential Surfaces for Conjugated Heteroatomic Molecules. *J Am Chem Soc* 1981;103:4664.
47. Daviter T, Wieden HJ, Rodnina MV. Essential Role of Histidine 84 in Elongation Factor Tu for the Chemical Step of GTP Hydrolysis on the Ribosome. *J Mol Biol* 2003;332:689. [PubMed: 12963376]

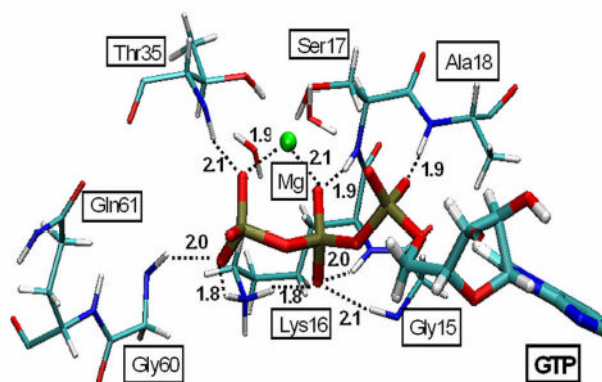


Figure 1. Showing the substrate GTP and its hydrogen bonds to residues of the Ras•GAP complex as derived from the crystal structure of the hydrolysis transition state analogue of Scheffzek et al. (PDB code: 1WQ1¹⁹).

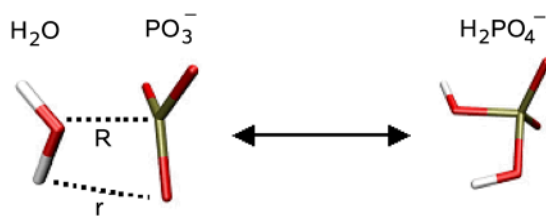


Figure 2.
Showing the two reaction coordinates R and r used in the present study.

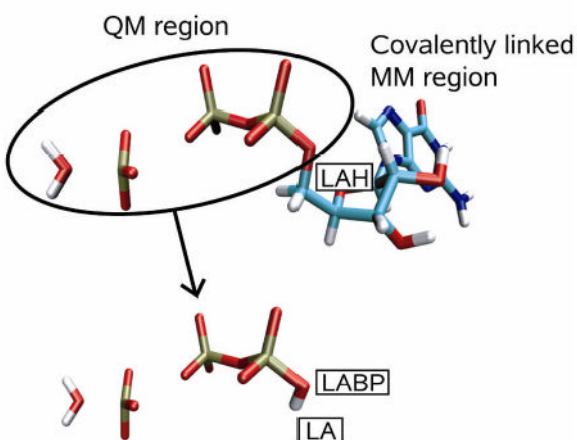


Figure 3. Division of the simulated system into QM and MM regions. The upper part shows the QM region and a part of the MM region that contains the guanosine moiety. The rest of the MM region, which is not shown in the figure, contains the surrounding protein and solvent molecules. The lower part shows the QM fragment capped with a hydrogen link atom. The notations LABP and LAH designate, respectively, the link atom bond partner and the link atom host. The link atom host is replaced here by the hydrogen link atom (LA).

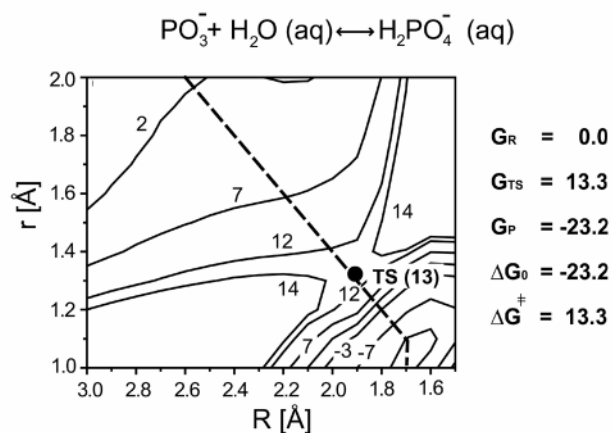


Figure 4.

Showing the free energy surface for the reference reaction of metaphosphate with water in solution and dependency on the two reaction coordinates R and r . The solid lines are the equipotential lines of the surface with the corresponding free energy values given next to them. The dashed line represents the reaction pathway connecting the shallow reactant region (upper left corner) to the product valley (lower right corner) through the transition state (marked with a dot). The free energies of the ground, transition and product as well as the reaction and activation free energy are given on the right side of the figure. All energies are given in kcal/mol.

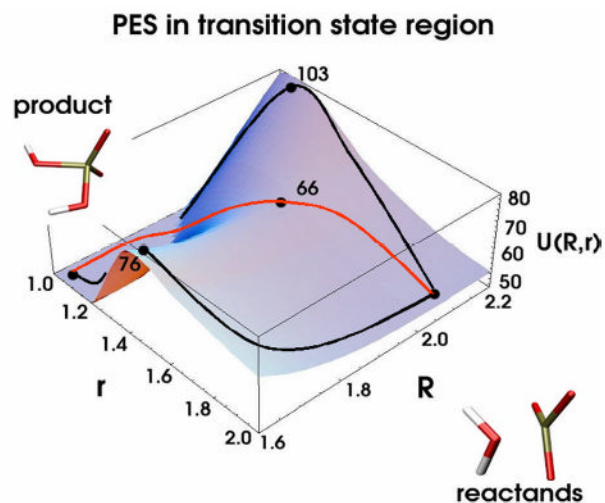


Figure 5. Showing the potential energy surface of the reaction of metaphosphate with water in the active site of Ras•GAP for one representative protein structure. The red line represents the reaction pathway connecting the reactant (on the left side) with the product (right side)

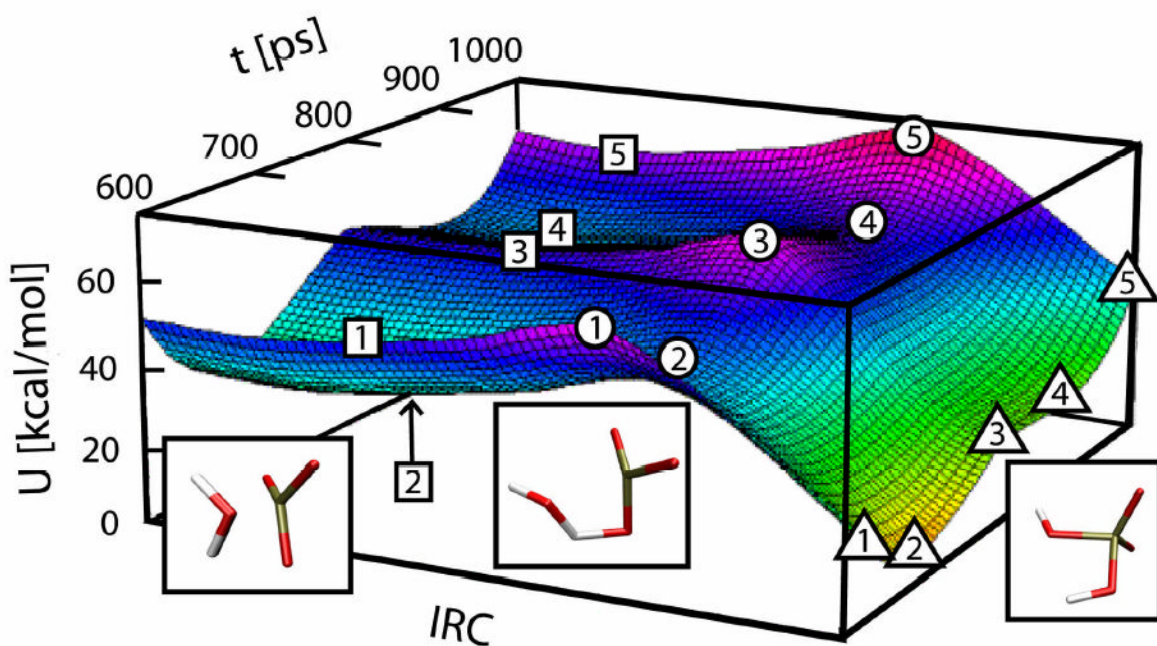


Figure 6. Showing the temporal development of the potential energy surface of the reaction metaphosphate with water in the active site of Ras•GAP. The generic surface is based on the two dimensional PES's, like the one shown in Fig. 5, projected on an intrinsic reaction coordinate (IRC), determined at five different protein configurations taken from the MD trajectory after 600, 700, 800, 900 and 1000 ps. The reactant states are marked with squares, the product states with triangles and the transition states with circles. The inscribed numbers correspond to one of the five protein structures.

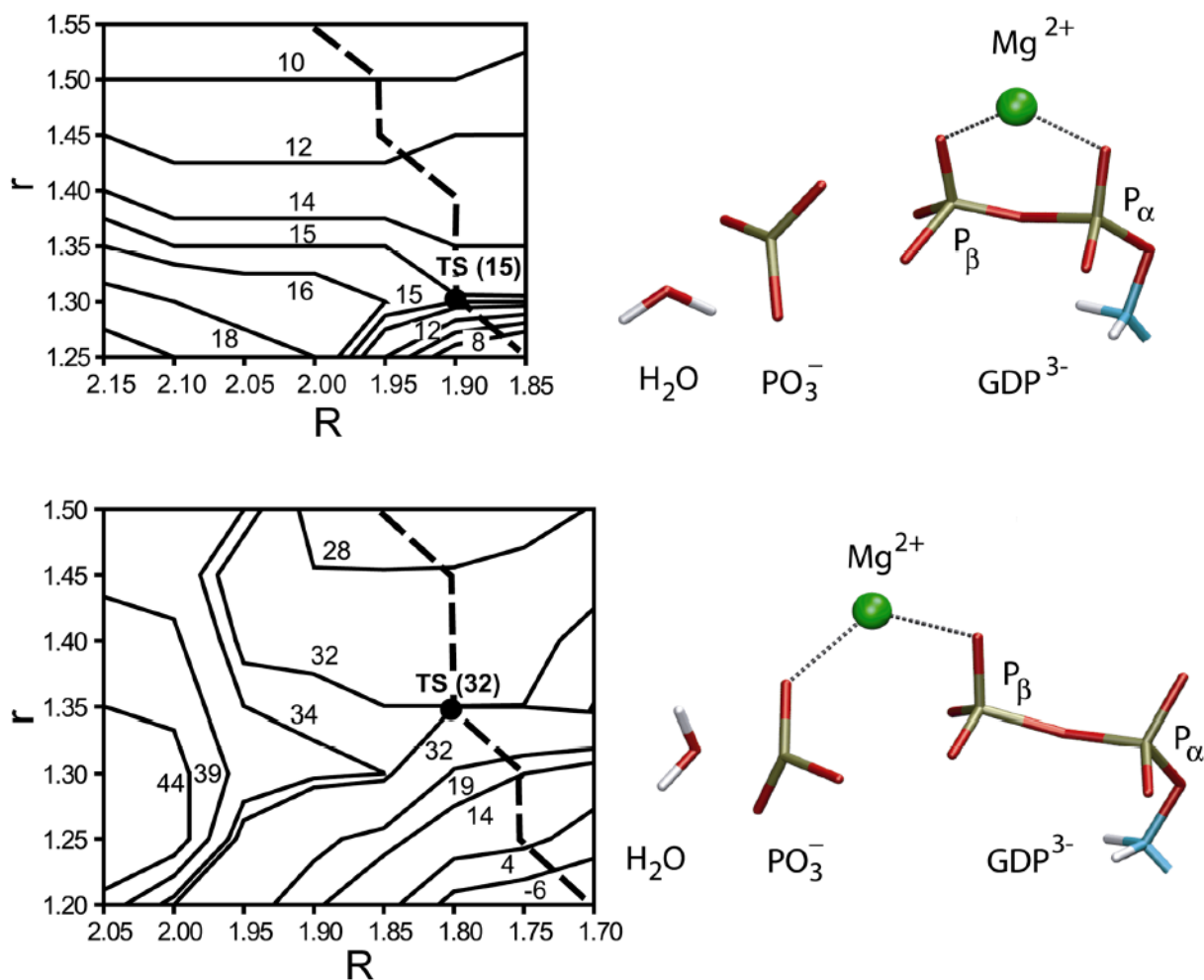


Figure 7. Showing the potential energy surfaces in the transition state region for the reaction of metaphosphate with water in Ras analogous to Fig. 4. The lower half displays the case where the Mg^{2+} ion is coordinated between the metaphosphate reactant and GDP. The upper half shows the PES for the case where the Mg^{2+} ion is coordinated to two GDP oxygens and not to the metaphosphate.

Table 1

Comparison of the free reaction energy and free activation energy in kcal/mol for the reaction of metaphosphate with water in solution for different methods*

	B3LYP/ 6-311 ++G (d,p) COSMO	Florian and Warshel ²¹ MP2/ 6-31+G(d,p) Langevin Dipoles	Florian and Warshel ²¹ Experimentally estimates of the actual free energy	HF/ based LanL2DZ	B3LYP/ LanL2DZ
ΔG_0	-23	-21	-24	-37	-27
ΔG^\ddagger	13	16	11	20	11

* All energies are given in kcal/mol. The Langevin dipole results and the estimates of the actual experimental free energies are taken from Florian and Warshel²¹.

Table 2

The dependence of the energetics of the reacting system on the protein configurations

protein structure	reactant	transition state (R/r)	product	ΔU_0	ΔU^\ddagger
600 (1)	57	75 (1.85/1.31)	36	-21	18
700 (2)	39	52 (1.89/1.33)	16	-23	13
800 (3)	52	67 (1.84/1.29)	18	-34	15
900 (4)	27	46 (1.85/1.32)	4	-23	19
1000 (5)	46	65 (1.85/1.31)	25	-21	19

The energies of the reactants, products and transition states as well as the corresponding reaction free energies and activation free energies in kcal/mol. Each row contains the values for one protein structure derived from the MD trajectories after 600, 700, 800, 900 and 1000 ps. The number in brackets for the protein structure corresponds to the label used in Fig. 6. For the transition states the values for R and r are given in brackets in Å.