# Multiple high-level QM/MM reaction paths demonstrate transition-state stabilization in chorismate mutase: correlation of barrier height with transition-state stabilization

**Frederik Claeyssens, Kara E. Ranaghan, Frederick R. Manby, Jeremy N. Harvey and Adrian J. Mulholland***

*Received (in Cambridge, UK) 10th June 2005, Accepted 8th August 2005*
*First published as an Advance Article on the web 7th September 2005*
**DOI: 10.1039/b508181e**

**Multiple profiles for the reaction from chorismate to prephenate in the enzyme chorismate mutase calculated with hybrid density functional combined quantum mechanics/molecular mechanics methods (B3LYP/6-31G(d)–CHARMM27) agree well with experiment, and provide direct evidence of transition-state stabilization by this important enzyme, which is at the centre of current debates about the nature of enzyme catalysis.**

The fundamental nature of the catalytic power of chorismate mutase (CM) has been the subject of much recent controversy. The central question is: does the enzyme stabilize the transition state (TS) relative to the bound substrate? Some recent proposals suggest that TS stabilization is not involved in catalysis by this key model enzyme. This controversy has implications for enzyme catalysis in general. We have investigated this crucial question through high level quantum mechanics/molecular mechanics (QM/MM) modelling of multiple reaction pathways in the protein. A strong correlation is found between TS stabilization and the reaction barrier in the enzyme. The results are in excellent agreement with experimental data and demonstrate the importance of TS stabilization in CM.

CM catalyses the Claisen rearrangement of chorismate to prephenate. The widely studied *Bacillus subtilis* enzyme BsCM (PDB 2CHT)[1] was chosen here. This is an excellent system for testing theories of catalysis, because the reaction does not involve any covalent bonding between the enzyme and the substrate, and because the same reaction occurs in solution with the same reaction mechanism. The activation free energy $\Delta^{\ddagger}G$ = 15.4 kcal mol$^{-1}$ ($\Delta^{\ddagger}H$ = 12.7 kcal mol$^{-1}$) in the (BsCM) enzyme is much lower than that found for the uncatalysed reaction in aqueous solution ($\Delta^{\ddagger}G$ = 24.5 kcal mol$^{-1}$, $\Delta^{\ddagger}H$ = 20.7 kcal mol$^{-1}$).[2] This translates to a rate acceleration of $10^6$ by the enzyme ($\Delta\Delta^{\ddagger}G$ = 9.1 kcal mol$^{-1}$). Earlier studies, *e.g.* with lower-level semiempirical QM/MM methods, indicated TS stabilization by the enzyme.[3–8] This and much subsequent work showed that the enzyme-bound conformation is significantly different from that in solution, and more closely resembles the TS.[9–13] This conformational selection is thought to contribute to catalysis,[10] although it may be a consequence of TS stabilization.[8,9,13] However Bruice *et al.* have controversially argued that TS stabilization is not involved, and that catalysis is instead almost entirely due to the selection of a reactive conformation, described as a near-attack conformation (NAC).[14] This conformation is far less likely in solution than in the enzyme. Estimates of the free energy cost of NAC formation (*e.g.* from MM molecular dynamics (MD) simulations (with or without restraints on the substrate)) led to the proposal that catalysis in CM is due entirely to its ability to maintain high populations of NACs.[14] However, extensive free energy perturbation molecular dynamics methods give a free energy cost of 3.8–4.6, or 5 kcal mol$^{-1}$ (by semiempirical QM/MM or empirical valence bond methods, respectively[9,13]), *i.e.* only accounting for 40–55% of the total $\Delta\Delta^{\ddagger}G$ between enzyme and solvent.

This indicates that catalysis involves TS stabilization relative to the bound conformation, in agreement with semiempirical QM/MM results for the enzyme-catalysed reaction. The reliability of these lower-level methods has been questioned, however. The central issue (related to Pauling's seminal hypothesis) of whether the TS is stabilized relative to the bound substrate, therefore remains highly controversial.

In order to study the effect of the enzyme on the reaction, it is important to sample many different reaction pathways corresponding to different conformational substates of the enzyme-substrate complex. We have calculated 16 different adiabatic reaction pathways using high-level QM/MM methods. Jaguar[15] and TINKER,[16] linked by in-house routines,[17] were used for QM and MM calculations. The electronic coupling between the two regions was treated by including MM charges in the QM Hamiltonian. Standard CHARMM Lennard-Jones parameters were used to describe QM/MM van der Waals interactions.[6] Full QM/MM energy minimizations were performed for the whole system to give RMS gradients below 0.1 kcal mol$^{-1}$ Å$^{-1}$ and 0.002 hartree Å$^{-1}$ on the MM and QM atoms, respectively.

The substrate does not form any covalent bonds to the enzyme and hence it is the natural choice, and the one used here, for the QM region. It was treated at the hybrid density functional B3LYP/6-31G(d) level of theory, which is known to give a good description of this reaction.[5,6] It has been shown previously that the effects of including some amino acid sidechains (*e.g.* Glu78 and Arg90) in the QM region are small.[18,19] In particular, the barrier height has been found to be insensitive to the size of the QM region.[19] The QM/MM treatment of the important interactions at the active site has been found to be accurate for chorismate mutase: this is because electrostatic interactions dominate.[8]

The MM region comprised a 25 Å radius sphere of protein and solvent, treated with the CHARMM27 force field.[20] The outer 5 Å was fixed, with all other atoms free to move. The set-up of the model is described in detail elsewhere.[5,6,13] Starting structures for calculating reaction pathways were derived from snapshots taken

*School of Chemistry, University of Bristol, Bristol, UK BS8 1TS. E-mail: Adrian.Mulholland@bris.ac.uk*

 

from two different 30 ps QM/MM molecular dynamics runs[13] (using two distinct semi-empirical methods to model the QM region, AM1 and PM3, respectively) for the BsCM complex, with chorismate restrained to be close to the TS. It is important to note that no major differences were observed between the AM1/CHARMM and PM3/CHARMM molecular dynamics trajectories. Structures were saved at regular intervals from 5 to 30 ps for both trajectories, giving 16 different TS complexes. As there is no evidence for large-scale conformational changes during the reaction,[1,5] this approach should give a representative sample of reactive conformations in the enzyme.

The difference in length between the forming C–C and breaking C–O bonds has been shown to be a good reaction coordinate, $r$.[5,6] Each structure was fully optimized at the B3LYP/6-31G(d)-CHARMM27 QM/MM level, while restraining the reaction coordinate, $r$, to $-0.3$ Å with a harmonic force constant of 500 kcal mol$^{-1}$ Å$^{-2}$. Reaction pathways were generated by restrained optimizations in both directions along the reaction coordinate, towards the reactant and the product, in steps of 0.2 Å (0.1 Å around the TS), with both the MM and QM systems fully and consistently optimized at each step. The end points were $r = -1.8$ Å (reactant) and 1.8 Å (product). Reoptimization of the reactant without constraints gave very similar structures (and energy differences below 1 kcal mol$^{-1}$).

The 16 different reaction pathways are shown in Fig. 1. The average reaction barrier is 12.0 kcal mol$^{-1}$ (standard deviation, $\sigma = 1.7$ kcal mol$^{-1}$) in excellent agreement with the experimental activation enthalpy (12.7 kcal mol$^{-1}$).[2] 13 of the 16 reaction barriers lie within $1\sigma$ of the average value. The energy profiles all have a similar shape, with the highest point, the approximate TS structure, at $r$ values between $-0.5$ and $-0.7$ Å. This is slightly earlier than in semiempirical QM/MM studies.[5–7] The TS structures from the various pathways here are similar to one another. For example, at the highest point along the profiles (at $r = -0.6$ Å), which is close to the TS, the average length of the breaking C–O bond is 2.02 Å, with a standard deviation, $\sigma$, of only 0.03 Å. The length of the forming C–C bond is also consistent from one structure to another, at 2.63 ($\sigma = 0.03$) Å. The substrate geometries are also all similar to one another. There is, however, a large spread in the calculated energy barriers (9 to 15 kcal mol$^{-1}$). As shown below, this variation is due to differences in the protein environment.

To examine the stabilization provided by the enzyme, single-point calculations on the isolated QM region were carried out at the QM/MM optimized geometries along the different reaction paths. All these *in vacuo* curves are very similar, with an average barrier height of 16.2 kcal mol$^{-1}$ ($\sigma = 0.3$ kcal mol$^{-1}$). This shows that the differences in barrier height between the 16 QM/MM reaction profiles are due to differences in the microstructure of the active site.

Discussion of catalytic effects ideally requires comparison of energy profiles in the enzyme and in solution. The barrier in solution, relative to the enzyme-bound conformation, has been found in previous work to be similar to that in the gas phase.[6,8,9] As all the pathways are structurally similar, solvent effects will be similar for them all. We use the gas-phase profiles as a convenient and meaningful reference.

The difference between the gas-phase and QM/MM energy gives the stabilization of the reacting system by the protein environment.

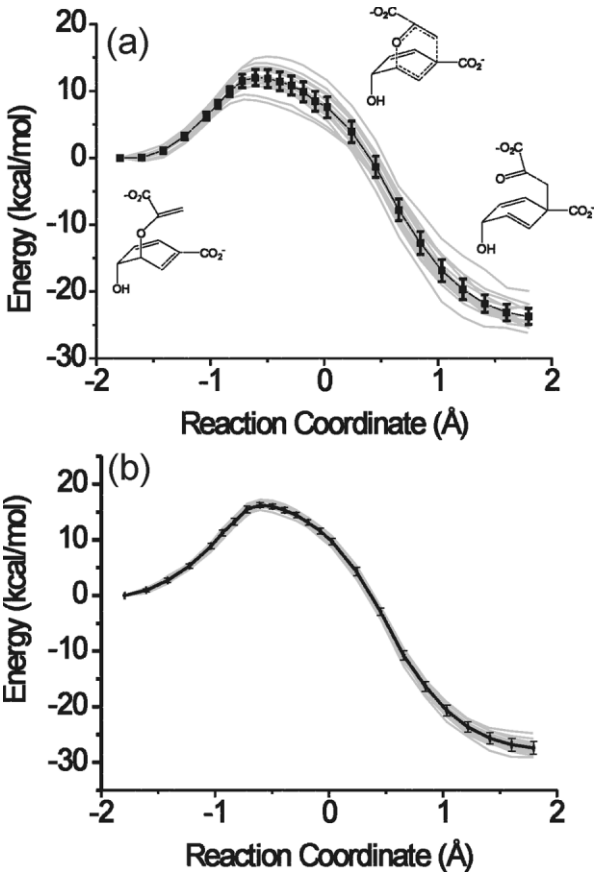


**Fig. 1** Energy profiles for reaction (a) in the enzyme and (b) *in vacuo* (B3LYP/6-31G(d)–CHARMM27). Grey lines depict individual paths, black points show the average values of the paths (all energies shown relative to the substrate), and the error bars show one standard deviation. Structures of chorismate, the TS and prephenate are shown.

This term is large and negative along the whole reaction coordinate, corresponding to favourable Coulombic interactions between the dianionic substrate and the positively charged side-chains in the reaction site. Most importantly, there is a systematic variation in the stabilization energy relative to that of the reactant complex along the reaction coordinate, as shown in Fig. 2(a), for the different optimized pathways. In all cases, the TS is stabilized more than the reactant. In all but one case, the product is destabilized (relative to the reactant). The stabilization of the TS is quite variable, and is plotted against the computed barrier height in Fig. 2(b). As can be seen, there is an excellent linear correlation between the two: higher TS stabilization equates to a lower barrier. On average, the enzyme stabilizes the TS by 4.2 kcal mol$^{-1}$ more than it stabilizes the reactant. The gradient is 0.95 and the intercept is 16.0 kcal mol$^{-1}$; this latter value is the predicted 'intrinsic' barrier from the bound substrate conformation in the absence of TS stabilization.

The stabilization energy can be divided into an electrostatic term due to the interaction of the enzyme MM partial charges with the QM wavefunction, which includes polarization of the latter, QM/MM van der Waals interactions, and a term corresponding to changes within the MM environment. The first of these terms yields an average stabilization of 4.7 kcal mol$^{-1}$, showing that the

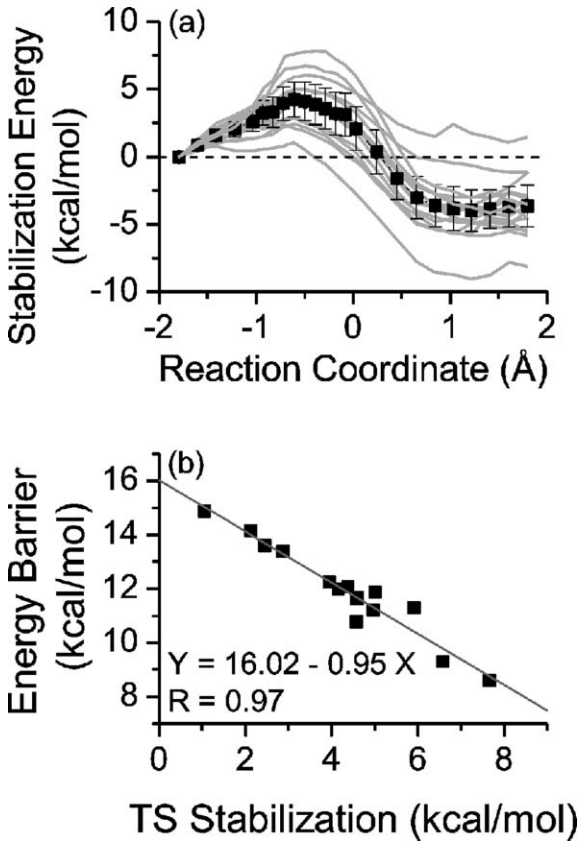


**Fig. 2** (a) Relative stabilization energy along the reaction coordinate, (grey) for individual reaction profiles and (black) for the average of the 16 profiles. Error bars show the standard deviation. (b) Energy barrier *vs.* TS stabilization energy.

TS stabilization is overwhelmingly electrostatic in nature, in agreement with previous findings.[3,8,9]

To conclude, multiple high-level QM/MM reaction pathways in this crucial enzyme give an average barrier (12.0 kcal $mol^{-1}$) in good agreement with experiment[1] ($\Delta^{\ddagger}H$ = 12.7 kcal $mol^{-1}$). They show significant TS stabilization by the enzyme. CM stabilizes the TS on average by 4.2 kcal $mol^{-1}$ more than it stabilizes the reactant, by electrostatic interactions. This study provides the most accurate estimate of TS stabilization by the enzyme to date, by applying high level calculations, and including multiple reaction pathways. The correlation in Fig. 2 shows that reactivity along given pathways is determined by TS stabilization by the enzyme: *i.e.*, the correlation of barrier height with TS stabilization shows that the efficiency of reaction in the enzyme is determined by the degree of TS stabilization in the enzyme. It appears that conformational effects (*i.e.* binding of a reactive conformation),[11,13] and TS stabilization (relative to the bound substrate) contribute roughly equally to catalysis in CM. We note that the calculated average TS stabilization here (4.2 kcal $mol^{-1}$) and the previously calculated cost of forming a reactive conformation in the enzyme, compared to solution[9,13] (3.8–5 kcal $mol^{-1}$) sum to give a value very close to the experimentally observed catalytic effect of CM (a lowering of the barrier of $\Delta\Delta^{\ddagger}G$ = 9.1 kcal $mol^{-1}$), and thus can together account for catalysis by the enzyme. The ability to bind a reactive conformation, resembling the TS, is also probably due to the enzyme's adaptation to bind the TS.[9,13] CM is thus a good example of an enzyme for which TS stabilization is central to catalysis.

A. J. M. thanks the IBM High Performance Computing Life Sciences Outreach Programme and BBSRC (with F. C.), and EPSRC (with K. E. R.) for support. J. N. H. is an EPSRC Advanced Research Fellow. F. R. M. thanks the Royal Society.

## Notes and references

1 Y. M. Chook, H. Ke and W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 8600.
2 P. Kast, M. Asif-Ullah and D. Hilvert, *Tetrahedron Lett.*, 1996, **37**, 2691.
3 P. D. Lyne, A. J. Mulholland and W. G. Richards, *J. Am. Chem. Soc.*, 1995, **117**, 345.
4 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *J. Phys. Chem. B*, 2000, **104**, 11308.
5 K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Mol. Phys.*, 2003, **101**, 2695.
6 K. E. Ranaghan, L. Ridder, B. Szefczyk, W.A. Sokalski, J. C. Hermann and A. J. Mulholland, *Org. Biomol. Chem.*, 2004, **2**, 968.
7 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Theor. Chem. Acc.*, 2001, **105**, 207.
8 B. Szefczyk, A. J. Mulholland, K. E. Ranaghan and W. A. Sokalski, *J. Am. Chem. Soc.*, 2004, **126**, 16148.
9 M. Strajbl, A. Shurki, M. Kato and A. Warshel, *J. Am. Chem. Soc.*, 2003, **125**, 10228.
10 H. Guo, Q. Cui, W. N. Lipscomb and M. Karplus, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 9032.
11 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Chem. Eur. J.*, 2003, **9**, 984.
12 C. R. W. Guimarães, M. P. Repasky, J. Chandrasekhar, J. Tirado-Rives and W. L. Jorgensen, *J. Am. Chem. Soc.*, 2003, **125**, 6892.
13 K. E. Ranaghan and A. J. Mulholland, *Chem. Commun.*, 2004, 1238.
14 S. Hur and T. C. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 1472; S. Hur and T. C. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 5964; S. Hur and T. C. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 10540.
15 *Jaguar, 4.0*, Schrödinger, Inc., Portland, OR, 1996–2001.
16 J. W. Ponder, TINKER: *Software Tools for Molecular Design, v4.0*, Saint Louis, MO, 2003.
17 J. N. Harvey, *Faraday Discuss.*, 2004, **127**, 165.
18 Y. S. Lee, S. E. Worthington, M. Krauss and B. R. Brooks, *J. Phys. Chem. B*, 2002, **106**, 12059.
19 A. Crespo, D. A. Scherlis, M. A. Marti, P. Ordejon, A. E. Roitberg and D. A. Estrin, *J. Phys. Chem. B*, 2003, **107**, 13728.
20 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586.

 