JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Subscriber access provided by The Libraries of the | University of North Dakota

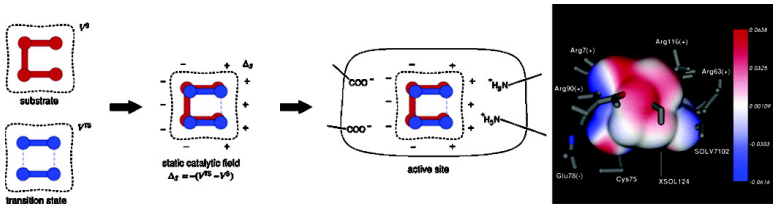
Article

Differential Transition-State Stabilization in Enzyme Catalysis: Quantum Chemical Analysis of Interactions in the Chorismate Mutase Reaction and Prediction of the Optimal Catalytic Field

Borys Szefczyk, Adrian J. Mulholland, Kara E. Ranaghan, and W. Andrzej Sokalski

J. Am. Chem. Soc., **2004**, 126 (49), 16148-16159• DOI: 10.1021/ja049376t • Publication Date (Web): 17 November 2004

Downloaded from http://pubs.acs.org on April 5, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

ACS Publications
High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
ARTICLES
Published on Web 11/17/2004

Differential Transition-State Stabilization in Enzyme Catalysis: Quantum Chemical Analysis of Interactions in the Chorismate Mutase Reaction and Prediction of the Optimal Catalytic Field

Borys Szefczyk,*,† Adrian J. Mulholland,‡ Kara E. Ranaghan,‡ and W. Andrzej Sokalski*,†

Contribution from the Institute of Physical and Theoretical Chemistry, Wrocław University of Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, and the Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

Received February 4, 2004; E-mail: borys.szefczyk@pwr.wroc.pl (B.S.); sokalski@pwr.wroc.pl (W.A.S.)

Abstract: Chorismate mutase is a key model system in the development of theories of enzyme catalysis. To analyze the physical nature of catalytic interactions within the enzyme active site and to estimate the stabilization of the transition state (TS) relative to the substrate (differential transition state stabilization, DTSS), we have carried out nonempirical variation−perturbation analysis of the electrostatic, exchange, delocalization, and correlation interactions of the enzyme-bound substrate and transition-state structures derived from ab initio QM/MM modeling of *Bacillus subtilis* chorismate mutase. Significant TS stabilization by approximately −23 kcal/mol [MP2/6-31G(d)] relative to the bound substrate is in agreement with that of previous QM/MM modeling and contrasts with suggestions that catalysis by this enzyme arises purely from conformational selection effects. The most important contributions to DTSS come from the residues, Arg90, Arg7, Glu78, a crystallographic water molecule, Arg116, and Arg63, and are dominated by electrostatic effects. Analysis of the differential electrostatic potential of the TS and substrate allows calculation of the catalytic field, predicting the optimal location of charged groups to achieve maximal DTSS. Comparison with the active site of the enzyme from those of several species shows that the positions of charged active site residues correspond closely to the optimal catalytic field, showing that the enzyme has evolved specifically to stabilize the TS relative to the substrate.

Introduction

Enzymes are remarkably efficient catalysts. The reason for their catalytic power still remains the subject of controversy and debate. For example, while conventional understanding of enzyme catalysis typically focuses on transition state stabilization,[1−4] other proposals suggest that substrate strain or conformational effects are more important.[5−7] To understand enzyme catalysis, it will be essential to analyze catalytic processes in atomic detail and to quantify the roles of various interactions in catalysis (e.g., lowering the barrier to reaction). Improved understanding of the basic physical principles involved in enzymatic catalysis should help in the design of catalysts for practical applications. For a full description of catalysis, it is essential to analyze the properties of the transition state, as well as the reactant state, for the chemical reaction.[8] Quantum chemical methods and analysis offer a good route to this goal.

Generally, the catalytic activity of an enzyme must be closely coupled with the nature of the reacting system. Therefore, it is tempting to extract information about the optimal catalytic environment from the substrate and transition state structures only. When this information is compared with the enzyme structure, it should be then possible to answer the following questions: Is the enzymatic environment the best catalyst for this reaction, or is there something to improve in this active site? Another, more challenging aim is to design a new catalytic environment based only on the information coming from the substrate and the transition state. Such information can be directly obtained within the differential transition state stabilization approach (DTSS),[9] which has been applied in this study.

One important model enzyme is chorismate mutase, which is the focus of the current debate in this area. Chorismate mutase

† Wrocław University of Technology.
‡ University of Bristol.

(1) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 10228−10237.
(2) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *J. Am. Chem. Soc.* **2004**, *126*, 311−319.
(3) Crespo, A.; Scherlis, D. A.; Martí, M. A.; Ordejón, P.; Roitberg, A. E.; Estrin, D. A. *J. Phys. Chem. B* **2003**, *107*, 13728−13736.
(4) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Org. Biomol. Chem.* **2004**, *2*, 968−980.
(5) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9032−9037.
(6) Guimarães, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6892−6899.
(7) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015−12020.
(8) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.
(9) Sokalski, W. A. *J. Mol. Catal.* **1985**, *30*, 395−410.

10.1021/ja049376t CCC: $27.50 **© 2004 American Chemical Society**

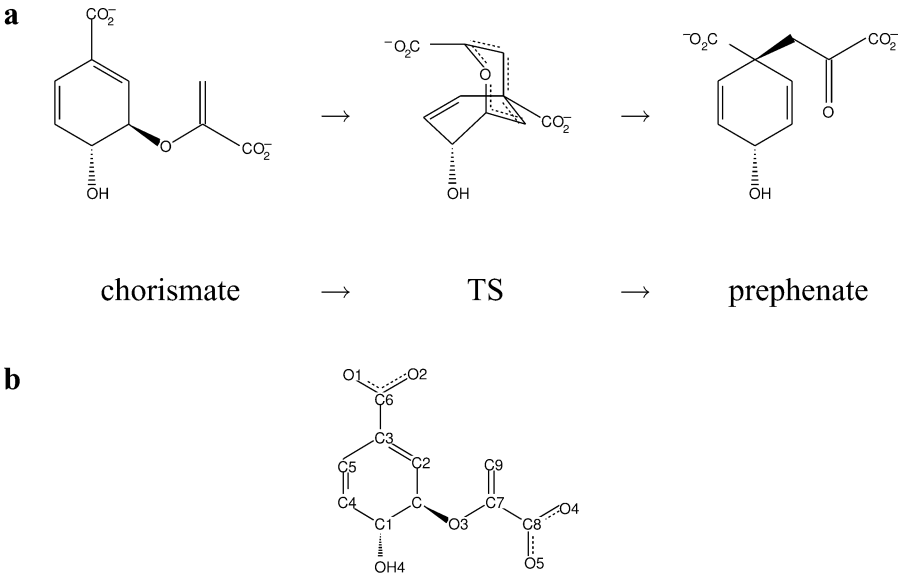


Figure 1. (a) Claisen rearrangement of chorismate to prephenate, catalyzed by chorismate mutase. (b) Numbering of atoms in chorismate.

has been studied intensively by a variety of experimental[10−15] and theoretical techniques.[1−7,16−27] Modeling of the reaction within the enzyme with combined quantum mechanics/molecular mechanics (QM/MM) methods at semiempirical[2,16,22,27,28] and ab initio[3,4,17,24] QM/MM levels shows significant TS stabilization by the enzyme.[1,3,16,19,27] In addition, several of these studies have suggested that the selection of a reactive conformer and the distortion or compression of bound chorismate also contribute to catalysis.[1,2,5,6,19] For example, some indications of TS stabilization have been given already in our previous papers[4,16,28] and in a recent paper by Crespo et al.[3] On the basis of AM1 QM/MM calculations,[16,28] it was found that the TS stabilization is provided mainly by Arg90, Arg7, and water molecules XSOL116 and XSOL124. Glu78 destabilizes the TS less than the substrate, contributing to the activation barrier lowering too. Cys75, Phe57, and water molecule SOLV7102 were found to have a minor effect on TS stabilization. These results were confirmed by HF/6-31G(d) QM/MM calculations.[4] In the current paper, TS stabilization is systematically quantified and divided into contributions from particular residues using DTSS theory. In contrast, on the basis of MM molecular dynamics simulations, Bruice and co-workers proposed that catalysis in chorismate is entirely, or largely, due to the substrate adopting a "near attack conformation" (NAC) when bound to the enzyme.[29−32] This proposal has been the subject of much debate and has been questioned.[1,2,33] While conformational effects have been acknowledged for some time as potentially important,[4−6,16,19,22,28,34] the hypothesis of Bruice et al., in its strongest form, proposes that TS stabilization is not important in chorismate mutase.[32] Ranaghan and Mulholland found that the NAC contribution is 3.8−4.6 kcal/mol[33] (5 kcal/mol according to Štrajbl et al.[35]), which is perhaps roughly half of the experimental $\Delta\Delta G^{\ddagger}$ (9.1 kcal/mol); this shows that the NAC effect by itself does not account for the observed catalysis.

This lively current debate can be found in the literature concerning other enzymes;[36−39] for example, the NAC binding mechanism was proposed to be responsible for the catalytic activity of catechol *O*-methyltransferase.[36] An NAC binding contribution was also discussed by Shurki et al. in the case of haloalkane dehalogenase and was found to play only a minor role, contrary to earlier findings.[35]

Chorismate mutase catalyzes the Claisen rearrangement of chorismate to prephenate (Figure 1), which is one of the steps in the biosynthesis of the aromatic amino acids, tyrosine and phenylalanine, in bacteria, fungi, and higher plants. Chorismate

(10) Cload, S. T.; Liu, D. R.; Pastor, R. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 1787−1788.
(11) Kast, P.; Hartgerink, J. D.; Asif-Ullah, M.; Hilvert, D. *J. Am. Chem. Soc.* **1996**, *118*, 3069−3070.
(12) Kast, P.; Asif-Ullah, M.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(13) Kast, P.; Tewari, Y. B.; Wiest, O.; Hilvert, D.; Houk, K. N.; Goldberg, R. N. *J. Phys. Chem. B* **1997**, *101*, 10976−10982.
(14) Wade, H.; Scanlan, T. S. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 461−493.
(15) Kast, P.; Grisostomi, C.; Chen, J. A.; Li, A.; Krengel, U.; Xue, Y.; Hilvert, D. *J. Biol. Chem.* **2000**, *47*, 36832−36838.
(16) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.
(17) Lee, Y. S.; Worthington, S. E.; Krauss, M.; Brooks, B. R. *J. Phys. Chem. B* **2002**, *106*, 12059−12065.
(18) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Angew. Chem., Int. Ed.* **2003**, *42*, 1508−1511.
(19) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *Chem.—Eur. J.* **2003**, *9*, 984−991.
(20) Sehgal, A.; Shao, L.; Gao, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 11337−11340.
(21) Khanjin, N. A.; Snyder, J. P.; Menger, F. M. *J. Am. Chem. Soc.* **1999**, *121*, 11831−11846.
(22) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *J. Phys. Chem. B* **2000**, *104*, 11308−11315.
(23) Hall, R. J.; Hindle, S. A.; Burton, N. A.; Hillier, I. H. *J. Comput. Chem.* **2000**, *21*, 1433−1441.
(24) Roitberg, A. E.; Worthington, S. E.; Holden, M. J.; Mayhew, M. P.; Krauss, M. *J. Am. Chem. Soc.* **2000**, *122*, 7312−7316.
(25) Worthington, S. E.; Krauss, M. *Comput. Chem.* **2000**, *24*, 275−285.
(26) Worthington, S. E.; Roitberg, A. E.; Krauss, M. *J. Phys. Chem. B* **2001**, *105*, 7087−7095.
(27) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *Theor. Chem. Acc.* **2001**, *105*, 207−212.
(28) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Mol. Phys.* **2003**, *101*, 2695−2714.
(29) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1176−1181.
(30) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 1472−1473.
(31) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 5964−5972.
(32) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 10540−10542.
(33) Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2004**, 1238−1239.
(34) Guilford, W. J.; Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5103−5109.
(35) Shurki, A.; Štrajbl, M.; Villà, J.; Warshel, A. *J. Am. Chem. Soc.* **2002**, *124*, 4097−4107.
(36) Lau, E. Y.; Bruice, T. C. *J. Am. Chem. Soc.* **2000**, *122*, 7165−7171.
(37) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 7303−7313.
(38) Schiøtt, B.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 14558−14570.
(39) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.

mutase has become a benchmark in computational chemistry,[8] mainly because the catalyzed reaction is a simple, pericyclic reaction with only one reactant and one product. The same reaction is observed in solution, making chorismate mutase an ideal case for comparative studies and for testing new methods and theories. The rate acceleration of chorismate mutase relative to aqueous solution is on the order of 10^6 (a reduction in $\Delta G^{\ddagger}$ of 9.1 kcal/mol); the activation free energy, $\Delta G^{\ddagger}$, is 15.4 kcal/mol, and the activation energy, $\Delta H^{\ddagger}$, is 12.7 kcal/mol.[12]

We have performed a nonempirical analysis to investigate the contributions of active site interactions and to examine the nature of the reaction in chorismate mutase. This analysis extends beyond previous QM/MM modeling[4,28] by quantifying (for the interactions of the reacting system with the enzyme) the contributions of effects, such as exchange, delocalization, and correlation, and by a more-sophisticated, multipole-based description of electrostatics. We have analyzed realistic structures of the enzyme-bound substrate and transition state, derived from ab initio QM/MM (RHF/6-31G(d)/CHARMM22) optimizations of *Bacillus subtilis* chorismate mutase.[4] We have analyzed the contributions of a number of active site residues, comparing their influence on the substrate with that on the TS to identify the contribution of residues to the differential transition state stabilization (i.e., stabilization of the TS relative to the substrate) within the enzyme. These results were compared to QM/MM and experimental data. We have examined some effects of conformational variability at the active site. The results demonstrate significant DTSS by the enzyme, resulting primarily from electrostatic interactions, and support findings from QM/MM modeling. Due to the importance of electrostatic effects, we have further analyzed the optimal catalytic field (by calculation of the differential electrostatic potential between the transition state and substrate). This field shows the optimal locations for unit probe charges to achieve DTSS. We have then compared the active site of chorismate mutase from *B. subtilis* and from other species with this optimal catalytic field to examine to what extent the active site of the enzyme is adapted to meet the demands of ideal DTSS. This type of analysis should be useful in the future for the design of effective catalysts. While it is unlikely that this naturally evolved enzyme can be made significantly more effective at catalyzing the overall reaction, calculation of optimal catalytic fields should assist greatly in the design of biomimetic catalysts for other reactions. We compare our findings with previous experimental and modeling work, providing new insights into the mechanism of this important enzyme.

This and previous applications of the hybrid variation−perturbation energy decomposition scheme[40] in DTSS calculations allow us to determine the theory level sufficient to describe this catalytic activity. When strongly charged systems (such as the chorismate to prephenate rearrangement) are considered, it is possible to apply simpler levels of theory (for example, the Hartree−Fock or the Heitler−London (HL) approach without the delocalization contribution, or even just the multipole electrostatic term when exchange and electrostatic−penetration effects cancel each other to a significant extent and can be neglected). This simplifies the calculation and enables larger models to be considered, but it has to be tested in a systematic way. This can be done using one of the available interaction energy decomposition schemes, and due to the considerable size of the model, the hybrid variation−perturbation energy decomposition scheme is well suited. Previous application of the DTSS analysis and the variation−perturbation scheme concerned tautomerization in models of the pyrimidine and purine base pair.[41,42] It was found that the interaction is dominated by the electrostatic (multipole) term. In such a simplified framework, the characteristics of the optimal catalytic environment for the tautomerization reaction were derived and the best location of solvent (water) molecules for enhancing this reaction was proposed.

Methods

Structures of two complexes, chorismate mutase with the substrate (chorismate) and chorismate mutase with the transition state (TS), were obtained from QM/MM calculations described in detail previously.[4] We give a brief summary of important details of the model here. Chorismate (or the TS) was treated quantum mechanically, at the HF/6-31G(d) level, with the CHARMM22 MM force field[43] to describe the enzyme environment. Calculations were performed using CHARMM (version 27a1)[44] interfaced with GAMESS-US.[45,46] The theory level and basis set [HF/6-31G(d)] are consistent with the DTSS calculations here and give a good description of TS structure.[28] The model was prepared using the active site between chains A and B in the *B. subtilis* chorismate mutase trimer taken from PDB data (PDB entry 2CHT[47]). This X-ray structure contains a transition-state analogue (TSA) bound in the active site. The TSA was replaced by the gas-phase-optimized structure of chorismate [optimized at the RHF/6-31G(d) level]. The protein structure was supplemented with hydrogen atoms and optimized by MM (hydrogens only) before solvation with TIP3P water molecules.[48] The region considered during reaction modeling was defined as a sphere centered on the C5 atom of the substrate (or TS). The sphere contained all of the residues with at least one heavy atom within 25 Å (no nonbonded cutoff was used). The structure was optimized (with the substrate fixed) by MM (1000 steps SD followed by 1143 steps ABNR), and the whole system was then optimized at the RHF/6-31G(d) QM/MM level, with chorismate as the QM part (250 steps ABNR). Atoms >16 Å away were kept fixed during the QM/MM simulation. Adiabatic mapping along a defined reaction coordinate was used to model the reaction. The reaction coordinate was defined as the difference of two bond lengths: the breaking and the forming bonds, with values ranging from −2.3 to 1.8 Å, with an interval of 0.3 Å. At each point on the reaction coordinate, 10 steps of steepest descent (SD) and 40 steps of adopted basis Newton−Raphson (ABNR) minimization were carried out. Further minimization of the model was not possible due to the high computational demands of the ab initio QM/MM method. The whole ab initio QM/MM procedure used 117 h of CPU

(40) Sokalski, W. A.; Roszak, S.; Pecul, K. *Chem. Phys. Lett.* **1988**, *153*, 153−159.
(41) Dziekoński, P.; Sokalski, W. A.; Podolyan, Y.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *367*, 367−375.
(42) Gorb, L.; Podolyan, Y.; Dziekonski, P.; Sokalski, W. A.; Leszczynski, J. *J. Am. Chem. Soc.* **2004**, *126*, 10119−10129.
(43) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fisher, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkirch, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(44) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(45) Lyne, P. D.; Hodoscek, M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 3462−3471.
(46) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(47) Chook, Y. M.; Ke, H. M.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8600−8603.
(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926−935.

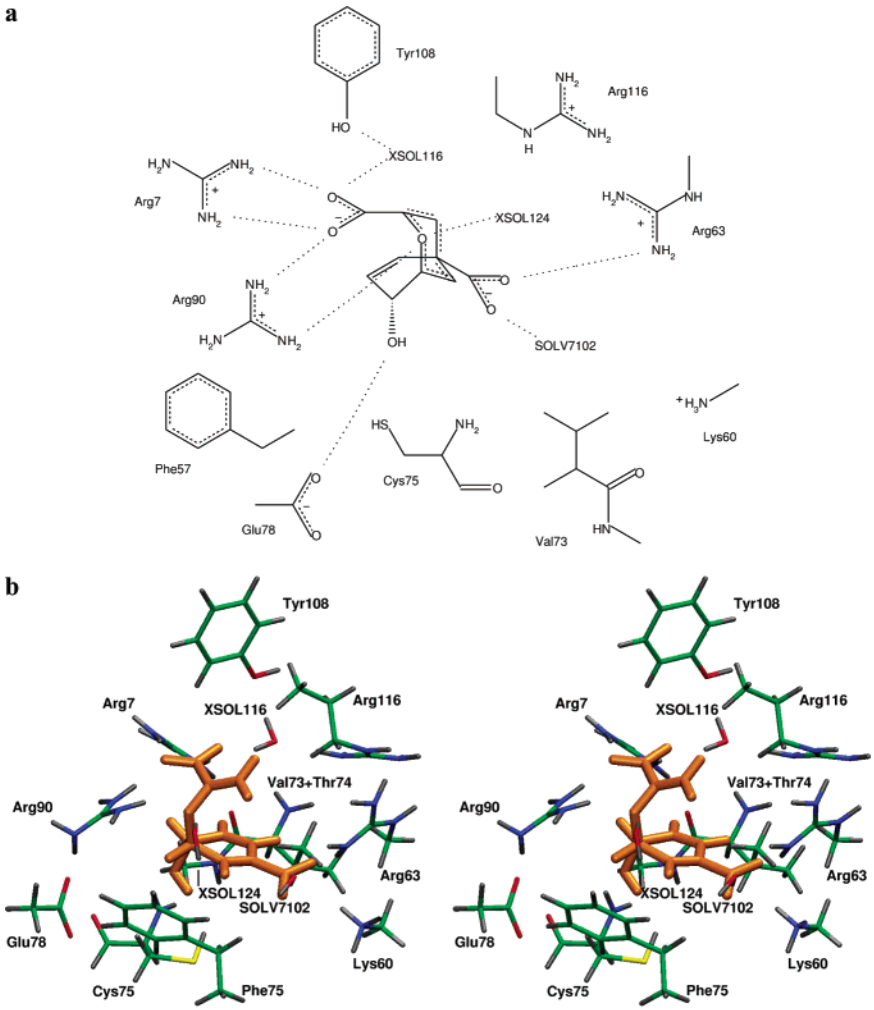


Figure 2. (a) Model of the chorismate mutase active site used in the calculations, derived from QM/MM modeling.[4] (b) Stereoview of chorismate in the active site of CM. Chorismate is shown in orange.

time on 32 processors (3744 h in total) on a Cray T3E machine.[49] Initial QM/MM minimization took 28.77 h. Pathway calculations took 82.5 h (an average of 5.5 h per point on the reaction coordinate). Averaging over many configurations was not feasible due to the computational costs outlined above; however, extensive optimization of the protein in the reactant state was performed, and therefore, any bias to stabilize the TS is excluded. Comparison with earlier semiempirical QM/MM modeling[28] shows that the adiabatic mapping procedure is reliable, and calculated barriers are consistent with those in experimental data.[11] Also, the similarity between the geometries of enzyme- and gas-phase-optimized TS structures suggests that, despite the relatively limited minimization that is possible at the ab initio QM/MM level, the optimization and reaction modeling produced properly optimized structures. For the current study, two snapshots from the QM/MM simulation were selected. The structure used here as a substrate was found at a reaction coordinate value of −2.0 Å. At this point, the C−O3 distance (bond being broken) is 1.45 Å, and the C3−C9 distance (bond being formed) is 3.44 Å. The dihedral angle of C2−C−O3−C7 (dih1 in ref 7) is 72.7°, and the dihedral angle of C−O3−C7−C9 (dih2 in ref 7) is −92.7°. As shown in Figure 2, this substrate conformation fits the definitions of NAC used by Hur and Bruice.[7] As a TS, the structure found at reaction coordinate −0.4 Å was used. This TS structure is a maximum in energy along the QM/MM path. The corresponding distances at this point are C−O3 = 2.17 Å and C3−C9 = 2.58 Å, and the dihedral angles are 63.7 and −78.9°. The simplified

(49) http://www.csar.cfs.ac.uk/.

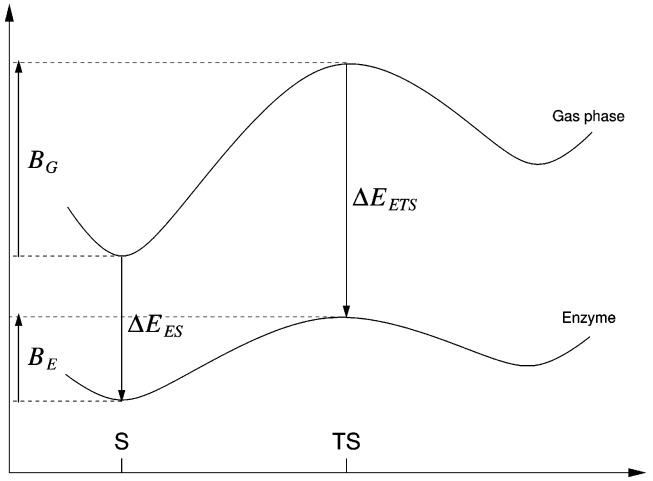


Figure 3. Schematic energy diagram showing differential transition state stabilization (DTSS). B_E denotes the barrier in the enzyme, and B_G is the barrier in the gas phase; ΔE_{ETS} is the interaction energy of the TS enzyme (or residue), and ΔE_{SS} is the interaction energy of the substrate enzyme (or residue). S and TS are substrate and transition state, respectively. On this graph, the positions of S and TS are similar for the enzyme and the gas phase, but in general, this does not need to be the case.

model used in our calculations was constructed by cutting out the side chains of the residues from the active site and terminating them with hydrogen atoms (Figure 2). All DTSS calculations were performed using

a truncated model, as shown in Figure 2, with the 6-31G(d) basis set, consistent with the QM/MM modeling. These structures were taken directly from the QM/MM model, without further optimization. The prior QM/MM optimization, in which the whole active site was optimized together with the substrate or TS, ensures that the potential catalytic residues are positioned realistically in response to the changes in electronic structure caused by the reaction.

The catalytic activity of the enzyme can be described within the differential transition state stabilization approach[9] (Figure 3), where the activation barrier changes by an amount (Δ), resulting from the presence of each individual residue or the enzyme as a whole (E), and can be expressed as the difference between the barrier in the enzyme (B_E) and in the gas phase (B_G), which both have positive values:

$$\Delta = B_E - B_G \tag{1}$$

or as

$$\Delta = \Delta E_{ETS} - \Delta E_{ES} \tag{2}$$

where ΔE_{ETS} and ΔE_{ES} denote interaction energies between the residue or enzyme (E) and transition state (TS) or substrate complex (ES), respectively. ΔE_{ETS} and ΔE_{ES} have negative values for a stabilizing effect. Δ is the amount of the DTSS or activation barrier lowering and describes how much the transition state is stabilized by a particular residue or the whole active site with respect to the reactant and is compared to the reference reaction in the gas phase. We use the term "differential transition state stabilization" (as opposed to the term "transition state stabilization" that is associated with the idea of Pauling[50]) to stress that the stabilization is relative with respect to the substrate bound to the active site. For many applications, it is useful to define the TS stabilization with respect to the reaction in water.[51] In the current work, we define the TS stabilization with respect to the gas-phase reaction, which is sufficient and convenient for analysis of the nature of interactions and for comparison of contributions to the activation barrier lowering. We have previously carried out the comparison with the reaction of chorismate in solution, modeled at the same RHF/6-31G(d) QM level (and other levels also) with the standard PCM continuum model, as well.[4] Negative Δ values indicate catalytic activity of the specific residue, E, and positive values indicate an inhibitory character (i.e., an increase in barrier). Hypothetically, the enzyme can act by destabilizing the reactant and not by stabilizing the TS[8,52]; such cases can also be treated by a DTSS analysis and are observed as positive interactions in the substrate−catalyst complex. However, while substrate destabilization has generally been found to not be the most important catalytic strategy used by enzymes,[53] some residues clearly can destabilize the substrate, whereas others stabilize. In particular, when a residue has the same charge as the reacting system (for example, the negatively charged Glu78 in the case of CM), one can expect a positive interaction energy for both the substrate and the TS complex. Such a residue can contribute to activation barrier lowering by destabilizing the TS less than the substrate, with its Δ value negative, indicating activation barrier lowering. It has to be stressed that we consider in our analysis a docked and optimized substrate structure, so the apparent substrate compression effect[6,16] does not affect the results of our analysis; for example, the reaction considered here is *from* the bound NAC (compressed form) of the substrate. Therefore, the catalytic effect observed here is due to the differential transition state stabilization only. An additional catalytic contribution is likely to arise from conformation selection−restriction by the enzyme (*preorganization* or NAC effect).[5,18,33,34]

Each interaction energy (ΔE) can be decomposed into the first-order Heitler−London (HL) term ($E^{(1)}$), the higher-order delocalization ($E^{(R)}_{DEL}$), and the correlation (E_{CORR}) components:

$$\Delta E = E^{(1)} + E^{(R)}_{DEL} + E_{CORR} \tag{3}$$

The correlation component is calculated using the MP2 level (ΔE^{MP2}) and HF level (ΔE^{SCF}) interaction energies:

$$E_{CORR} = \Delta E^{MP2} - \Delta E^{SCF} \tag{4}$$

The delocalization component is calculated according to the formula

$$\Delta E^{(R)}_{DEL} = \Delta E^{SCF} - E^{(1)} \tag{5}$$

The first-order Heitler−London term [$E^{(1)}$] is calculated according to the formula

$$E^{(1)} = \frac{\langle \hat{A}\Psi_A\Psi_B|\, \hat{H}_{AB}|\, \hat{A}\Psi_A\Psi_B\rangle}{\langle \hat{A}\Psi_A\Psi_B|\, \hat{A}\Psi_A\Psi_B\rangle} - \frac{\langle \Psi_A|\hat{H}_A|\Psi_A\rangle}{\langle \Psi_A|\Psi_A\rangle} - \frac{\langle \Psi_B|\, \hat{H}_B|\Psi_B\rangle}{\langle \Psi_B|\Psi_B\rangle} \tag{6}$$

where A and B denote monomers, AB the dimer, $\hat{H}_X$ the Hamiltonian of X (X = A, B, or AB), Ψ_X the wave function of X (X = A, B, or AB), and $\hat{A}$ the antisymmetry operator. The Heitler−London term can be decomposed into first-order electrostatic [$E^{(1)}_{EL}$] and first-order exchange [$E^{(1)}_{EX}$] terms:

$$E^{(1)} = E^{(1)}_{EL} + E^{(1)}_{EX} \tag{7}$$

where all are calculated consistently in the dimer basis set in order to eliminate basis set superposition error (BSSE), according to the counterpoise scheme (CP) proposed by Boys and Bernardi.[54] The first-order electrostatic and exchange terms are calculated as follows:

$$E^{(1)}_{EL} = \frac{\langle \Psi_A\Psi_B|\, \hat{H}_{AB} - \hat{H}_A - \hat{H}_B|\Psi_A\Psi_B\rangle}{\langle \Psi_A\Psi_B|\Psi_A\Psi_B\rangle} \tag{8}$$

$$E^{(1)}_{EX} = E^{(1)} - E^{(1)}_{EL} \tag{9}$$

The electrostatic component can be additionally decomposed into long-range multipole [$E^{(1)}_{EL\cdot MTP}$] and short-range penetration [$E^{(1)}_{EL\cdot PEN}$] terms:

$$E^{(1)}_{EL} = E^{(1)}_{EL\cdot MTP} + E^{(1)}_{EL\cdot PEN} \tag{10}$$

The multipole component can be rapidly estimated by applying the cumulative atomic multipole moment approach,[55] which could be generated using the electron density obtained from any ab initio density matrix (e.g., SCF, MP, DFT, CC, and so forth) or, as in this case, using the numerically equivalent[56] distributed multipole analysis (DMA).[57] DMA is performed in this case up to the quadrupole−quadrupole component in the monomer-centered basis set (without BSSE correction).

For all of the dimers studied here (i.e., for each residue interacting with TS and with substrate, separately), all of the described interaction energy components are calculated in the variation−perturbation energy decomposition procedure.[40]

The interaction energy components outlined above define an entire hierarchy of gradually simplified theoretical models, where a more approximate theory level is obtained by dropping the rightmost term from the higher level:

ΔE values are interaction energies between the substrate/TS and each residue; ΔE^{MP2} is the interaction energy at the MP2 level, and ΔE^{SCF} is the Hartree−Fock interaction energy; $E^{(1)}$ is the first-order HL interaction term. $E^{(1)}_{EL}$ is the first-order electrostatic interaction energy,

(50) Pauling, P. *Chem. Eng. News* **1946**, *24*, 1375−1377.
(51) Villà, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887−7907.
(52) Lewis, B. E.; Schramm, V. L. *J. Am. Chem. Soc.* **2003**, *125*, 4672−4673.
(53) Shurki, A.; Warshel, A. *Adv. Protein Chem.* **2003**, *66*, 249−313.
(54) Boys, F. S.; Bernardi, D. *Mol. Phys.* **1970**, *19*, 553−556.
(55) Sokalski, W. A.; Sawaryn, A. *J. Chem. Phys.* **1987**, *87*, 526−534.
(56) Spackman, M. A. *J. Chem. Phys.* **1986**, *85*, 6587−6601.
(57) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233−239.

$$\Delta E = E^{(1)}_{EL\cdot MTP} + E^{(1)}_{EL\cdot PEN} + E^{(1)}_{EX} + E^{(R)}_{DEL} + E_{CORR} \quad (11)$$

$O(N^5)$ $\underbrace{\qquad\qquad}_{\Delta E^{MP2}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{\Delta E^{SCF}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{E^{(1)}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{E^{(1)}_{EL}}$

$O(A^2)$ $\underbrace{\qquad}_{E^{(1)}_{EL\cdot MTP}}$

$$O(A^1) \qquad E_{EL,qV} = \sum_i V_i \cdot q_i^E \quad (12)$$

and $E^{(1)}_{\mathrm{EL}\cdot\mathrm{MTP}}$ is the electrostatic energy calculated using the multipole expansion. $E_{\mathrm{EL},qV}$ is the interaction energy calculated using the electrostatic potential generated by the substrate/TS (V_i) and charges (q_i^{E}) on active site residues. The dominant contribution to $E_{EL,q\mathrm{V}}$ usually originates from a point charge on an atom of the residue that has the shortest contact with the substrate/TS in the simplest model, the formal charge of the residue (usually $-1/0/+1$) (see, for example, ref 58). $O(X)$ in eq 11 indicates the scaling of the computational effort, where N stands for number of atomic orbitals, whereas A is the number of atoms. In particular, MP2 (the most demanding on this list) scales with the fifth power of N, and the simplest model, $E_{EL,q\mathrm{V}}$, scales linearly with A. A treatment involving partitioning of free energies would be possible with the linear response approach (LRA)[1] if the use of any empirical parameters could be avoided. Nevertheless, in the present study, nonempirical analysis is done.

Similarly, the activation barrier lowering Δ can be decomposed into components:

$$\Delta = \Delta^{(1)}_{EL\cdot MTP} + \Delta^{(1)}_{EL\cdot PEN} + \Delta^{(1)}_{EX} + \Delta^{(R)}_{DEL} + \Delta_{CORR} \quad (13)$$

$O(N^5)$ $\underbrace{\qquad\qquad}_{\Delta^{MP2}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{\Delta^{SCF}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{\Delta^{(1)}}$

$O(N^4)$ $\underbrace{\qquad\qquad}_{\Delta^{(1)}_{EL}}$

$O(A^2)$ $\underbrace{\qquad}_{\Delta^{(1)}_{EL\cdot MTP}}$

$$O(A^1) \qquad \Delta_q = \sum_i (V_i^{TS} - V_i^S) q_i^E \quad (14)$$

From the physical point of view, the Δ_{CORR} component includes in the total Δ value the effects arising from the correlation between electrons. The physical meaning of $\Delta^{(\mathrm{R})}_{\mathrm{DEL}}$ and $\Delta^{(1)}$ can be explained according to perturbation theory; $\Delta^{(1)}$ corresponds to that part of Δ which results from the interaction of "frozen", unperturbed monomers' charge distributions calculated in a perturbative manner, using eq 6, and $\Delta^{(\mathrm{R})}_{\mathrm{DEL}}$ corresponds to the change in barrier resulting from deformation of the charge distribution of one monomer induced by unperturbed charge distribution of the second monomer. $\Delta^{(1)}$ consists of two components, namely, $\Delta^{(\mathrm{R})}_{\mathrm{EX}}$, the electron exchange resulting from the Pauli exclusion principle, and also an electrostatic component, $\Delta^{(1)}_{\mathrm{EL}}$, arising from the interaction of two unperturbed charge distributions. $\Delta^{(1)}_{\mathrm{EL}\cdot\mathrm{PEN}}$ is the penetration component arising from the quantum nature of charge distribution in monomers, with the remainder, $\Delta^{(1)}_{\mathrm{EL}\cdot\mathrm{MTP}}$ or $\Delta_{\mathrm{EL},qV}$, corresponding to the interaction energy described classically (e.g., by a multipole expansion or, in the simplest case, by a single point charge representing the total charge of each residue or alternatively partial charges on atoms, making the shortest contacts with the TS/substrate).

For the calculation of the individual residue contributions to the lowering of the activation barrier, residues situated in the proximity of the substrate/TS were selected, including charged residues and residues capable of forming a hydrogen bond with the substrate/TS. These were Arg90, Arg7, Glu78, XSOL124, Arg116, Arg63, Cys75, SOLV7102, XSOL116, Tyr108, Val73, Phe57, and Lys60. Several of these residues have been suggested to be important based on QM/MM results.[4,16,28] The protonation states of these residues were chosen to be consistent with the predominant states expected for these amino acids at pH 7 (as in the QM/MM modeling).

In this study, all interaction energy components, up to the MP2 level, were calculated using a modified version of the GAMESS-US package[46] within a variation−perturbation partitioning scheme.[40] The stabilization contribution of particular residues or water molecules were computed by taking the difference between the interaction energy of two complexes: one with TS and the other with the substrate. Both complexes were treated always as dimers containing the TS or substrate and the active site residue. Interaction energies were calculated with correction for BSSE.[54] This means that the energy of a monomer (residue, reactant, or TS) is calculated in the basis set of the dimer, where empty basis functions were put on every atom of the other monomer. The total interaction energy or total stabilization effect at a specific level of theory (MP2, RHF, or Heitler−London) was calculated as a sum of the contribution from all dimers at this level of theory. Many-body interactions were not considered due to the conclusions from our earlier study, indicating their relatively minor role.[59,60] The amino acid side chains were represented by smaller molecular fragments: arginines by guanidinium or ethylguanidinium (in the case of Arg116), glutamate by acetate, lysine by protonated methylamine, tyrosine by phenol, and so forth (Figure 2).

For example, to calculate the contribution of Arg90 to the reduction of the activation barrier, the complex of chorismate with truncated Arg90 is cut out from the QM/MM model of the substrate and TS complexes and its interaction energy with the substrate and with the TS is calculated. This is done at all levels of theory (MP2, RHF, HL, multipole expansion, and molecular electrostatic potential interactions with environment point charges) with the CP correction for BSSE at MP2, RHF, and HL levels, giving all levels of interaction energy [ΔE^{MP2}, ΔE^{SCF}, $E^{(1)}$, $E^{(1)}_{\mathrm{EL}}$, $E^{(1)}_{\mathrm{EL}\cdot\mathrm{MTP}}$, and $E_{\mathrm{EL},qV}$] and the remaining interaction energy components [ΔE_{CORR}, $E^{(\mathrm{R})}_{\mathrm{DEL}}$, $E^{(1)}_{\mathrm{EX}}$, and $E^{(1)}_{\mathrm{EL}\cdot\mathrm{PEN}}$]. The difference between the interaction energy in the TS−Arg90 complex and the interaction energy in the chorismate−Arg90 complex is equal to the contribution (Δ) originating from this residue (refer to Figure 3 and eq 2), and having all of the interaction energy components, one can obtain the Δ value at various theory levels [Δ^{MP2}, Δ^{SCF}, $\Delta^{(1)}$, $\Delta^{(1)}_{\mathrm{EL}}$, $\Delta^{(1)}_{\mathrm{EL}\cdot\mathrm{MTP}}$, and $\Delta^{(1)}_{\mathrm{EL},qV}$] and the components between them [Δ_{CORR}, $\Delta^{(\mathrm{R})}_{\mathrm{DEL}}$, $\Delta^{(1)}_{\mathrm{EX}}$, and $\Delta^{(1)}_{\mathrm{EL}\cdot\mathrm{PEN}}$].

In cases where the electrostatic term [$\Delta^{(1)}_{EL}$] is dominant, one may derive the general static[9] and dynamic[9,41,61] characteristics of the molecular environment with optimal catalytic activity, just from the knowledge of the superimposed transition state and substrate (Figure 4). Fortunately, in many enzymes, this electrostatic term is dominant.[1,35] Such static and dynamic catalytic fields can be pictured as a difference map of the electrostatic potentials (Δ_s) and vector field ($-\nabla\Delta_d$),

(58) Grembecka, J.; Kedzierski, P.; Sokalski, W. A. *Chem. Phys. Lett.* **1999**, *313*, 385−392.
(59) Dziekonski, P.; Sokalski, W. A.; Leszczynski, J. *Chem. Phys.* **2001**, *272*, 37−45.
(60) Mulholland, A. J.; Lyne, P. D.; Karplus, M. *J. Am. Chem. Soc.* **2000**, *122*, 534−535.
(61) Sokalski, W. A. *J. Mol. Struct. (THEOCHEM)* **1986**, *138*, 77−87.

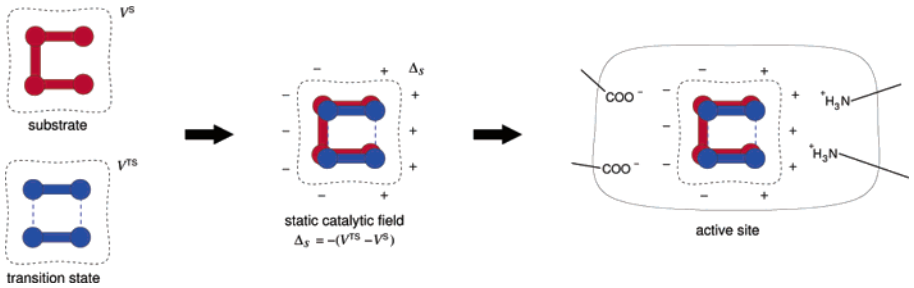


Figure 4. Idea of a static catalytic field.

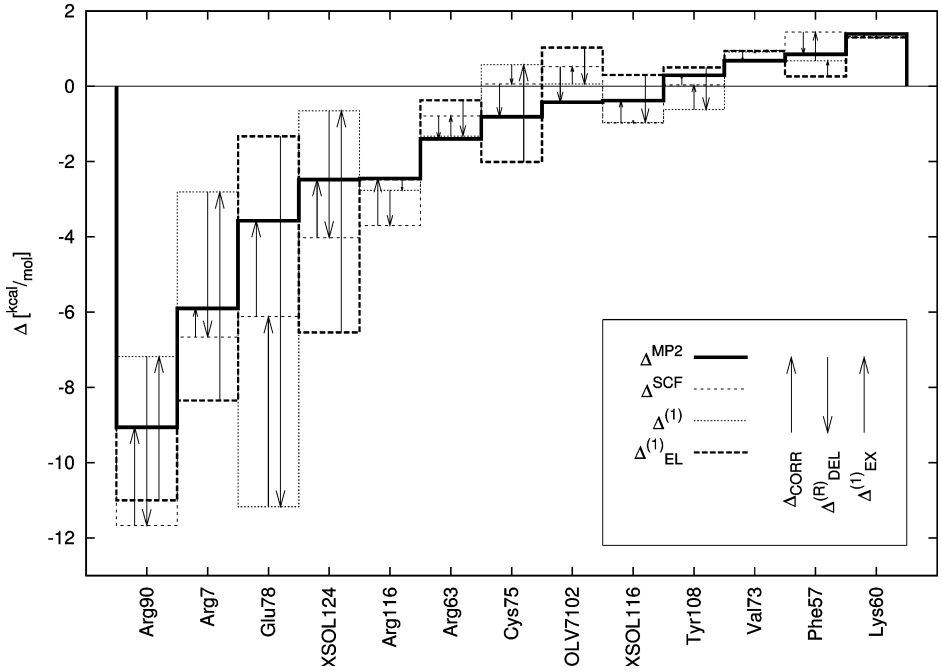


Figure 5. Value of differential transition state stabilization (Δ) for various theory levels (in kcal/mol). Arrows show cancellation of correlation, delocalization, exchange, and penetration components. The meaning of the symbols is described in the text and in the caption of Table 1.

respectively (eqs 15 and 16). This $-\nabla\Delta_d$ vector field is simply the corresponding electric field.

$$\Delta_s = -(V^{TS} - V^S) \tag{15}$$

$$-\nabla\Delta_d = \vec{E}^{TS} - \vec{E}^S \tag{16}$$

The electrostatic potential and electric field are calculated as the expectation values of the wave function obtained at the desired level of theory [HF/6-31G(d) in this case]. Calculation of these properties at higher levels of theory (i.e., MP2) is possible, but is not needed in this case, as the decomposition of Δ contributions shows that electron correlation can be neglected. The electric field can also be calculated as the gradient of the electrostatic potential using finite differences.

In the case of the maps shown in Figure 6, Δ_s is calculated on the van der Waals surface surrounding both the substrate and the TS. First, the regularly spaced points on the van der Waals surface are generated, and then the expectation values of the electrostatic potential and electric field are calculated for the substrate and TS, separately, but at the same points, using Gaussian98.[62] The van der Waals radii used here are 1.2 Å for hydrogen, 1.7 Å for carbon, and 1.4 Å for oxygen. Finally, the difference is calculated (using an external program written for this purpose), giving the differential static (or dynamic) field. The differential electrostatic potential is multiplied by −1 to reflect the environment charge. These fields are illustrated in two ways: mapped in a two-dimensional cross-section defined by the plane containing the bonds being broken (C−O3) and made (C3−C9; Figure 1). The differential electrostatic potential is visualized in Figure 6 with the use of plus and minus signs, where the size of sign corresponds to the magnitude of the catalytic effect expected from the corresponding unitary charge.

In the case of a three-dimensional catalytic field (Figure 7), four "cube" files are calculated in Gaussian98; two of them contain the electrostatic potential (one for the substrate and one for the TS), and the other two contain electron density (one for the substrate and one for the TS). The substrate and TS are aligned using a least-squares fit, and cube files are calculated in the same coordinate system. To define a common surface surrounding both the substrate and the TS, an isodensity surface is found using the condition $\rho_S + \rho_{TS} = 0.005$ au, where ρ_S and ρ_{TS} are the electron densities of substrate and TS, respectively. This condition ensures that for the substrate, as well as for the TS, the electrostatic potential is calculated on the same surface and outside the molecule. Next, the electrostatic potentials of the

(62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Rega, N.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11.3; Gaussian, Inc.: Pittsburgh, PA, 2002.

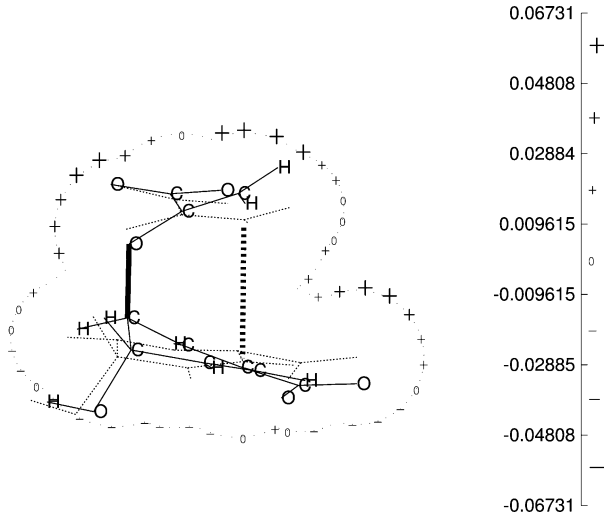


Figure 6. Static catalytic fields: the difference of the RHF/6-31G(d) electrostatic potentials (multiplied by −1) around the substrate (−) and the transition state (• • •), calculated according to eq 15 in the plane of both bonds being formed (- - -) and broken (—) during the reaction. Plus and minus signs correspond to the theoretically predicted positions where positive or negative charges accelerate the reaction. Their size illustrates the magnitude of activation barrier lowering, resulting from the presence of the corresponding single unitary charge. Signs are placed on van der Waals surface.

substrate and the TS are subtracted, and the resulting differential electrostatic potential, multiplied by −1, is mapped on the isodensity surface. The map of differential potential is visualized using a color map, where red corresponds to the maximum differential potential (positive value and TS stabilization by positively charged environment) and blue corresponds to the minimum differential potential (negative value and TS stabilization by negatively charged environment).

Such maps may be helpful in the design or redesign of active sites for new or modified reactions, as well as potentially for the design of a completely new catalytic environment for a known reaction.[41] The advantage of this approach is that the optimal catalytic environment is derived from the reacting system only. Similar approaches based on electrostatics have been used to demonstrate complementarity between the catalytic triad and the protein environment in a series of proteases and esterases[63] and to analyze the affinity for the TS analogue to bind to CM.[64,65]

Results

DTSS Analysis. All of the interaction energy components obtained within the variation−perturbation partitioning scheme[40] are presented in detail in the Supporting Information (Table SI-1). As shown in Table 1 and in Figure 5, the highest catalytic activity is exerted by the following charged amino acids: Arg7, Arg63, Arg90, Arg116 (all positively charged), and the negatively charged Glu78. The other residues, apart from a crystallographically observed water molecule (XSOL124), have no significant influence on the TS differential stabilization energy, although they are close to the substrate/TS. The total TS differential stabilization effect calculated at the MP2/6-31G(d) level is −23.3 kcal/mol (based on the sum of binary interactions with all active site residues considered in this study). The contributions of the individual residues are in the following order (all values in kcal/mol): Arg90 (−9.06), Arg7 (−5.90), Glu78 (−3.57), XSOL124 (−2.48), Arg116 (−2.45), Arg63 (−1.40), Cys75 (−0.81), SOLV7102 (−0.42), XSOL116 (−0.38), Tyr108 (0.29), Val73 (0.68), Phe57 (0.85), and Lys60 (1.39). This is consistent with results of previous QM/MM modeling, which has identified Arg90 as making the most important contribution.[4,16,28]

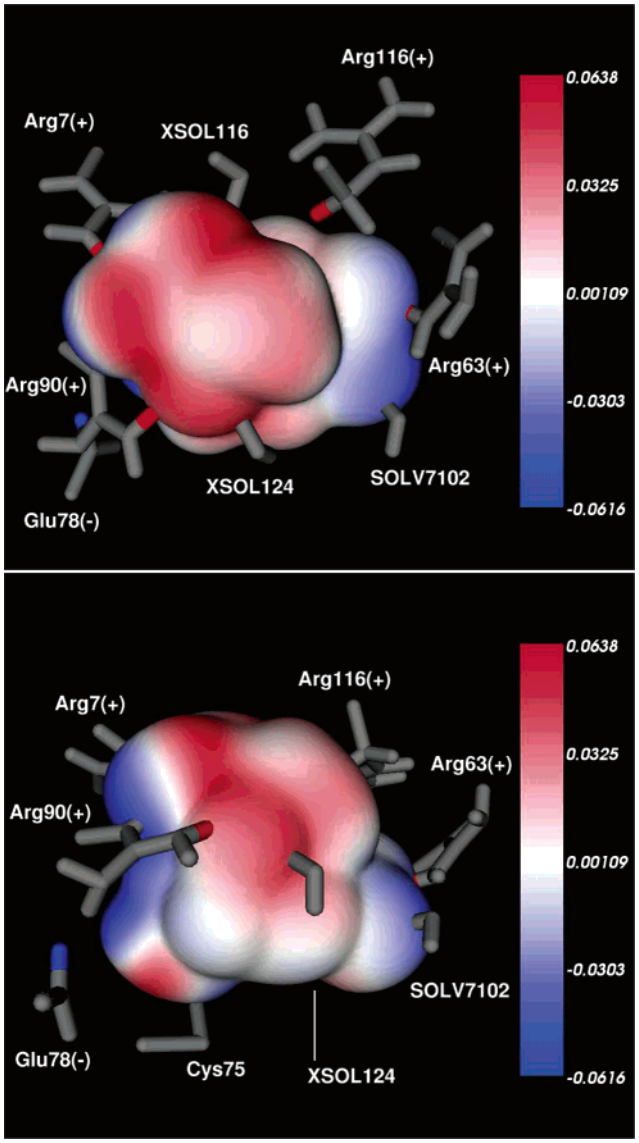


Figure 7. Static catalytic fields, for example, the difference of the RHF/6-31G(d) electrostatic potentials (multiplied by −1) of the substrate and transition state (eq 15). Red corresponds to the region where a positive charge reduces the activation barrier, and blue corresponds to regions where a negative charge exerts optimal catalytic activity. Formal charge of residues is marked with red (+) and blue (−) on the atoms forming the shortest contact with the reacting system. Color matching illustrates complementarity of charge distribution in the active site and differential catalytic field.

The total differential TS stabilization (DTSS) effect was calculated with respect to the gas phase by taking the sum of all stabilization contributions. The SCF [RHF/6-31G(d)] result is overestimated (−29.6 kcal/mol) compared to MP2, but the correlation coefficient calculated between interactions at the RHF and MP2 levels (for each complex, where MP2 is the reference level) is very high (0.98, Table 1). This means that the electron correlation components of DTSS have some significant contribution but are similar for each complex, so the electron correlation effects can be neglected when only a

(63) Gérczei, T.; Asbóth, B.; Náray-Szabo, G. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 310−315.
(64) Barbany, M.; Gutiérrez-de-Terán, H.; Sanz, F.; Villà-Freixa, J.; Warshel, A. *Chem. Biochem.* **2003**, *4*, 277−285.
(65) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880−888.

Table 1. Value of Differential Transition State Stabilization (Δ) for Various Theory Levels (in kcal/mol)[a]

residue	Δ_V	Δ_{V-HF}	$\Delta^{(1)}_{EL\text{-}MTP}$	$\Delta^{(1)}_{EL}$	$\Delta^{(1)}$	Δ^{SCF}	Δ^{MP2}	QM/MM[b]
Arg90	−7.55	−3.89 −8.95	−5.23	−11.00	−7.18	−11.67	−9.06	−4.30
Arg7	−8.18	−4.11 −5.21	−6.75	−8.35	−2.81	−6.66	−5.90	−2.72
Glu78	−3.15	−2.80	−4.67	−1.33	−11.17	−6.12	−3.57	−1.64
XSOL124	0.0	−9.28	−5.84	−6.54	−0.65	−4.02	−2.48	−2.01
Arg116	−7.32	−0.48	−2.58	−2.47	−2.76	−3.70	−2.45	
Arg63	0.76	0.32	−0.09	−0.37	−1.32	−0.79	−1.40	−2.07
Cys75	0.0	−2.69	−1.03	−2.01	0.57	0.06	−0.81	−0.08
SOLV7102	0.0	1.02	0.64	1.03	0.06	0.52	−0.42	0.10
XSOL116	0.0	−1.21	−0.11	0.30	−0.96	−0.97	−0.38	−0.59
Tyr108	0.0	0.59	0.0	0.50	−0.62	0.04	0.29	
Val73	0.0	0.53	0.72	0.94	0.91	0.95	0.68	
Phe57	0.0	−0.25	0.63	0.26	0.68	1.44	0.85	
Lys60	1.08	0.38	1.30	1.30	1.30	1.32	1.39	
sum	−24.36	−21.87 −36.03[c]	−23.01	−27.74	−23.95	−29.60	−23.26	
c.c.[d]	0.83	0.55	0.85	0.92	0.71	0.98	1.00	

[a] Δ_V is the electrostatic term calculated using the electrostatic potential. Δ_{V-HF} is the electrostatic term calculated using the electrostatic potential and Chelpg charges; in the case of Arg7 and Arg90, there are two short contacts (energies for both are shown). $\Delta^{(1)}_{EL\text{-}MTP}$ is the electrostatic term calculated using atomic multipole expansion. $\Delta^{(1)}_{EL}$ is the first-order electrostatic term. $\Delta^{(1)}$ is the first-order Heitler−London term. Δ^{SCF} is the Hartree−Fock term, and Δ^{MP2} is the differential TS stabilization energy on the MP2 level. Residues are arranged according to the contribution to the activation barrier lowering at the highest theory level. The most optimal configuration of Glu78−chorismate/TS complex was taken into account when the effects of Glu78 and Cys75 were calculated. [b] QM/MM calculations at the AM1/CHARMM22 level (model A in the reference).[28] [c] Sum including components for both short contacts of Arg7 and Arg90. [d] Correlation coefficient with respect to the MP2 results.

Table 2. Dependence of Differential Transition State Stabilization (Δ) for Glu78 (in kcal/mol), Calculated at the MP2 Level, on the Position of the Hydroxyl Group in the Chorismate/TS Complex[a]

Glu78–chorismate[b]	Glu78–TS[b]	Δ^{MP2}
250.0	250.0	0.12
80.0	80.0	−1.49
80.0	250.0	7.33
250.0	80.0	−8.70

[a] Δ for the Glu78−substrate/TS complex is shown, calculated for different values of the dihedral angle of H4−O−C1−C. The 80° angle directs the hydrogen toward Cys75, whereas the 250° angle directs the hydrogen toward Glu78. [b] Angles are given in degrees.

Table 3. Value of Differential Transition State Stabilization (Δ) for Glu78 (in kcal/mol), Calculated at the MP2 Level, in Different Configurations of the Glu78−Chorismate/TS Complex[a]

contact	Glu78–chorismate (Å)	Glu78–TS (Å)	Δ^{MP2}
Cys75, S	4.24	3.88	2.62
Glu78, O	4.24	3.88	1.29
Glu78, O	2.85	2.82	0.22
Glu78, O	4.24	2.82	5.08
Glu78, O	2.85	3.88	−3.57

[a] In the first case, the hydrogen in the hydroxyl group was directed toward the sulfur atom in the Cys75; in other cases, it was rotated in the direction of Glu78. Distance was measured between heavy atoms: O in chorismate/TS and O in Glu78.

comparative analysis is done. The DTSS values obtained from the multipole expansion are in very good agreement with the MP2 results. The stabilization energy calculated using the RHF-derived multipole expansion is −23.0 kcal/mol, which is due to the cancellation of electrostatic penetration, exchange, delocalization, and electron correlation components, showing that the DTSS is almost entirely due to the electrostatic interactions.

In our preliminary calculations, Glu78 initially indicated an increase by +2.6 kcal/mol in the activation barrier (Table 3). In this initial model, due to the limited QM/MM minimization used, Glu78 remained further away from the chorismate/TS, and the hydroxyl group in chorismate/TS rotated only after passing the barrier (on the product side, at 1.2 Å on the reaction coordinate) to form a hydrogen bond with glutamate, leading to higher stabilization of the prephenate product. To check the influence of the conformation of the hydroxyl group on the activation barrier lowering, a simple analysis of the interaction between chorismate/TS and Glu78 was performed at the MP2/6-31G(d) level (Table 2). The dihedral angle of H4−O−C1−C was set to 80 or 250° in chorismate or the TS, and Δ^{MP2} was calculated. This analysis shows that rotation of the hydroxyl group from Cys75 to Glu78, when passing from chorismate to the TS, is the most optimal scenario for the catalysis. Since this interaction has been identified as important experimentally,[11] we explored more models of the active site. Several test calculations at the MP2/6-31G(d) level were done with various distances between Glu78 and the hydroxyl group of the substrate/TS model (based on our initial QM/MM model and the model described in ref 17). The results presented in the Table 3 indicate that the agreement with experimental data is achieved (the same sign of contribution) when a shorter distance between chorismate and Glu78 and a longer distance between TS and Glu78 are used. Rotation of the hydroxyl group in both substrate and TS complexes, without changing the distance, lowers the contribution of Glu78 by 1.3 kcal/mol. Change of the distance during the reaction (increase of the distance) lowers this contribution by an additional 4.9 kcal/mol, giving a negative value (i.e., catalytic activity instead of inhibitory). In current calculations, the model with a rotated hydroxyl group (H bond present in the substrate and TS complexes) and changed substrate/TS−Glu78 distance (2.85 Å for the substrate complex and 3.88 Å for TS complex) was used. MD/QM calculations performed by Worthington et al. also indicate a large variation (almost 2 Å) in the distance between Glu78 and the substrate. Further detailed analysis of possible static and dynamic effects involved in this interaction is under investigation and will be discussed in a forthcoming study.[66]

(66) Szefczyk, B. et al. Unpublished results.

Static Catalytic Field. The results presented in the previous section indicate the dominant role of electrostatic effects in determining the catalytic activity of all of the considered active site residues, which can be approximately represented by atomic multipole electrostatic terms. As the electrostatic effects are additive, one may consider the simplest molecular environment constituted by a single unit probe charge. Its catalytic activity, compared to the gas phase, may be illustrated in a general way by the difference map of electrostatic potentials, $\Delta = V_i^{TS} - V_i^{S}$, as described in the Methods section. Static catalytic fields are presented in Figure 6 and described in the Methods section.

The signs shown in Figure 6 correspond to the charge of a particular environment fragment (a residue, for example) lowering the activation barrier. The larger the magnitude of the activation barrier lowering, resulting from the presence of this fragment, the greater is the magnitude of the + or − sign.

The various colors used in Figure 7 correspond to the optimal environment charges lowering the activation barrier. Extreme red and blue colors indicate places where positive and negative charges stabilize the TS, respectively. Inspection of the QM/MM enzyme model reveals that the charged residues, Arg90 and Glu78, are located exactly in the regions predicted by this picture (see Figure 7 for the location of the amino acids). Similarly, in Figure 6, one can observe "−" signs near the hydroxyl group (which can interact with negatively charged Glu78) and "+" signs above the O3 and C9 atoms, where arginines Arg7 and Arg90 are located.

Discussion

DTSS Analysis. The activation barrier lowering calculated using DTSS analysis at the MP2/6-31G(d) level is −23.3 kcal/mol, which is in good agreement with that in the TS stabilization (activation barrier lowering) estimated from QM/MM calculations, which is −19.7 kcal/mol.[4] Štrajbl et al. found, using empirical valence bond (EVB) and LRA methods, that TS stabilization in CM is dominated by electrostatic interactions between the reacting system and the charged environment, which helps the two negatively charged carboxylates of chorismate approach one another.[1] Similarly, DTSS analysis presented here shows that the interactions in the active site are dominated by electrostatics and can be described using the electrostatic part of the $E_{EL}^{(1)}$ term (eqs 6 and 7) or even the multipole term.

The activation barrier height calculated at the RHF/6-31G(d)/CHARMM22 QM/MM level is 11.0 kcal/mol, with single point energy corrections at the MP2/6-31+G(d) level, and 16.1 kcal/mol, with single point energy corrections at the B3LYP/6-311+G(2d,p) level.[4] The experimentally determined value of $\Delta H^{\ddagger}$ is 12.7 kcal/mol.[11] Comparison with the same reaction modeled in solution predicts rate accelerations of 1.3×10^{13} [with MP2/6-31+G(d) corrections] and 7.5×10^{6} [with B3LYP/6-311+G(2d,p) corrections], consistent with the experimental value of 4.5×10^{6}.

The reaction barrier in the gas phase is 48.4 kcal/mol at the RHF/6-31G(d) level (gas-phase optimized geometries),[28] and the barrier at the RHF/6-31G(d) level for the model used here is 49.0 kcal/mol (i.e., QM energy from the QM/MM model).[4] The barriers at the Hartree−Fock level are overestimated; better methods give significantly lower barriers, for example, 24.8 kcal/mol at the B3LYP/6-31G(d) level. Taking the Hartree−Fock barrier of the gas-phase reaction and the barrier lowering found here (−29.6 kcal/mol), the energy barrier to the reaction in the CM active site would be 19.4 kcal/mol, overestimated as well, compared to the experimental $\Delta H^{\ddagger}$. Using the B3LYP//HF barrier and Δ^{MP2} (−23.3 kcal/mol), one obtains 1.5 kcal/mol for the reaction barrier in the enzyme. However, the comparison with the experimental $\Delta H^{\ddagger}$ value should be made carefully because the total Δ includes only contributions from 13 residues and can be shifted significantly when the remaining part of enzyme is included. It also does not contain the intramolecular contributions from the protein. A rough estimate can be made by adding the MM contribution to the enzyme barrier from QM/MM calculations[4] (∼6 kcal/mol), which gives an enzyme barrier equal to 7.5 kcal/mol. Much better estimates of the reaction barrier for CM are obtained directly from QM/MM studies; the barrier calculated with corrections at the MP2/6-31+G(d) level is 11.0 kcal/mol and with corrections at the B3LYP/6-311+G(2d,p) is 16.1 kcal/mol. For comparison, the activation barrier lowering from the gas phase to the enzyme, calculated by Hall et al., in QM/MM optimizations of chorismate, TS, and prephenate is −27.7 kcal/mol,[23] with the QM part treated at the B3LYP/6-31G(d) level. This gives a very small enzyme barrier (1.4 kcal/mol); however, this was probably calculated without the MM (protein intramolecular) contribution to the barrier. Taking, once again, the MM contribution as 6 kcal/mol,[4] one obtains an enzyme barrier of 7.4 kcal/mol. Calculations of Crespo et al. at the PBE/DZVP level of theory predict activation barrier lowering, with respect to the gas phase, −27 to −28 kcal/mol.[3] From the QM/MM calculations of Lee et al., the DTSS or activation barrier lowering can be estimated as −20 kcal/mol.[17] These calculations were done at the RHF/4-31G level (QM part) and with the CHARMM force field (MM part).[17] All the QM/MM results quoted above consistently show TS stabilization in CM.

When going from the substrate to the TS, the electrostatic stabilization of the reacting system increases; simultaneously, the hydrogen bonds between chorismate/TS and the residues became shorter, as was shown using QM/MM RHF/6-31G(d) calculations.[4] This evidently indicates that the TS is better stabilized than the substrate. In the case of negatively charged Glu78, better stabilization of the TS is obtained when the hydrogen bond becomes longer at the TS[66] (see Table 3).

Arg90 has the greatest stabilization effect, which is in agreement with previous QM/MM results.[4,16,28] The role of Arg90 in catalysis has also been confirmed by mutagenesis experiments showing that mutation of this residue causes the largest decrease of k_{cat}/K_m (2.6×10^4-fold for Arg90Lys mutation; no activity observed for the Arg90Ala mutation), which was interpreted by the authors as a loss of catalytic activity.[15,67] Similar experiments indicate a 10^6-fold decrease in catalytic activity when Arg7 is substituted by alanine and a 10^3-fold decrease for an Arg7Lys mutation.[10] A similar effect is observed experimentally in the case of Arg11, the analogue of Arg7 in *Escherichia coli* chorismate mutase.[68] The minor role of Cys75 is also confirmed experimentally.[11] Generally, our findings agree qualitatively with some contributions calculated previously with semiempirical (AM1/CHARM22) QM/MM methods (Table 1), but semiempirical stabilization contri-

(67) Kienhofer, A.; Kast, P.; Hilvert, D. *J. Am. Chem. Soc.* **2003**, *125*, 3206−3207.
(68) Liu, D. R.; Cload, S. T.; Pastor, R. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 1789−1790.

butions are underestimated in a systematic way.[28] Theoretical predictions indicate also that Arg7 and Arg90 have a function in stabilizing the "chairlike" conformation of chorismate, close to the transition state conformation.[18] Such behavior is responsible for the *preorganization effects* in enzymatic catalysis[19,39,69] (this effect is called, by some researchers, *environment reorganization effect*). However, this work focuses on TS stabilization, and preorganization effects are not studied here. The substrate used in the present study is already preorganized.

We do not observe any significant contribution from Tyr108. Worthington et al. suggest that this residue is important for the formation of the prephenate.[25,26] This is probably due to the fact that different crystal structures were used for the generation of the models,[28] and in our case, a water molecule, XSOL116 (observed also in the crystal structure), forms a bridge between the substrate and Tyr108. Nevertheless, both models are reasonable.

Experiments also indicate the important role of the Glu78 residue;[11] replacing this negatively charged residue by a neutral one (glutamine) considerably reduces the catalytic activity of the enzyme (with in vitro activity $<0.05\%$, compared to the wild type measured as a reduction in k_{cat}).

The calculated differential transition state stabilization of -23.3 kcal/mol indicates that chorismate mutase acts by transition state stabilization. Bound substrate and TS have similar structures, and chorismate fits the definition of a NAC used in ref 7; however, the relative stabilization of TS is significant. This confirms the results of other simulations[1,4,6] and is in contrast with results of Bruice and co-workers.[7,29−32] The simulations of Bruice et al. did not examine the TS or its interactions. The NAC effect has recently been calculated by FEP methods to be ∼4 kcal/mol (QM/MM MD)[33] and ∼5 kcal/mol (EVB)[1], respectively. The close structural similarity of the substrate, transition state, and product bound to the enzyme active site (resembling a proposed NAC) has been observed experimentally (by means of kinetic isotope effects) in the case of purine nucleoside phosphorylase.[70] However, the NAC or preorganization effect does not necessarily mean a considerable reduction of the activation barrier, but it can contribute to the acceleration of the reaction by restriction of the conformational space or reduction of the activation barrier thickness. Our aim here is to analyze, by high level quantum chemical techniques, whether the TS is stabilized by the enzyme in representative structures. The results clearly show significant differential TS stabilization.

Static Catalytic Field. The structure of the enzyme active site corresponds closely to the theoretically derived optimal catalytic field required for this reaction (Figure 7). It is clear that the enzyme has evolved to mimic the optimal catalytic field required by the reacting system. The static catalytic fields shown here are in agreement not only with the *B. subtilis* chorismate mutase active site constituents but also with the chorismate mutase active sites from other related enzymes. For example, in *E. coli* chorismate mutase,[71] the Glu78 and Arg7 residues are conserved (corresponding to Glu52 and Arg11, respectively, in *E. coli* CM), while the positively charged Arg90 is replaced by the similarly positively charged Lys39. Also, Arg28 and Arg51 are bound in a way that is similar to Arg63 and Arg116 in BsCM. In *Saccharomyces cerevisiae* CM,[72] Glu78 and Arg7 are conserved (Glu198 and Arg16), and Arg90 is replaced by Lys168. In ScCM, Arg157 is bound in the same place as Arg63 in BsCM (Arg28 in EcCM). This electrostatic pattern is repeated through all three enzymes and is predicted by the static catalytic field. The close correspondence of the active site with the static catalytic field prediction provides further evidence of the importance of DTSS in CM catalysis and the central contribution of electrostatic interactions.

The large differential electrostatic potential in some regions demonstrates that a full study of catalysis must consider electronic differences between the substrate and the TS; geometrical/structural considerations, alone (as often employed in discussions of conformational effects, such as NACs), cannot provide a complete description of the reaction. The bound substrate used in the calculations here conforms to typical definitions of NACs (i.e., the C3−C9 distance is 3.44 Å) and is geometrically similar to TS, but it has a very different charge distribution. This means that at the last stage before passing the barrier, the geometry does not change significantly, but the charge distribution changes dramatically.

Differential fields have previously been used in a slightly different situation. Kangas and Tidor analyzed the electrostatic complementarity between the TSA and the active site of BsCM.[65] They found that the charge distribution of the TSA is compatible with the position of charged groups in the active site; however, it does not fit perfectly and could be improved. Their tests of charge optimization revealed that better binding is obtained when the total charge of the TSA is higher than the formal charge ($-2e$) by $\sim 0.5e$. They proposed that the complementarity of charge distribution could be improved by exchanging some groups, in particular, one of the carboxylate groups to nitro group. Other suggestions were given by Barbany et al., who analyzed the electrostatic similarity between the TSA and TS in order to explain low activity of catalytic antibody 1F7.[64] These authors identified the regions where the electrostatic potentials of the TS and TSA differ, in particular, the region of the carboxylate group and of the bonds being broken and formed. The lack of complementarity was explained also using a relaxed grid of Langevin dipoles. It was found that the environment preorganized to stabilize the TSA is not as proficient in stabilizing the TS because the environment preorganized specifically to stabilize the TS. The difference in solvation free energy of the TS charge distribution solvated by dipoles preorganized to solvate TS in the geometry of TS, and the geometry of TSA (3.6 kcal/mol), calculated by Barbany et al., corresponds qualitatively with the observed difference in proficiency of the enzyme and catalytic antibody (6 kcal/mol). These findings suggest that catalytic antibodies raised against the TSA may not stabilize the TS effectively. Better catalytic antibodies could be prepared, which should aim to fit the static catalytic field.

Conclusions

In the reaction catalyzed by chorismate mutase, the major contribution to differential transition state stabilization is exerted

(69) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250−5254.
(70) Fedorov, A.; Shi, W.; Kicska, G.; Fedorov, E.; Tyler, P. C.; Furneaux, R. H.; Hanson, J. C.; Gainsford, G. J.; Larese, J. Z.; Schramm, V. L.; Almo, S. C. *Biochemistry* **2001**, *40*, 853−860.
(71) Lee, A. Y.; Karplus, P. A.; Ganem, B.; Clardy, J. *J. Am. Chem. Soc.* **1995**, *117*, 3627−3628.
(72) Ma, J.; Zheng, X.; Schnappauf, G.; Braus, G.; Karplus, M.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14640−14645.

by the charged active site residues: arginines 7, 63, 90, and 116 and a water molecule, whose contribution is comparable with that of Arg116 and Arg63. Negatively charged Glu78 displays a moderate catalytic activity strongly dependent on the position of the residue and orientation of the hydroxyl group of the substrate. All of these effects are dominated by an electrostatic multipole term, which could be rapidly estimated from the atomic multipole expansion. The essential role of electrostatic and preorganization effects in chorismate mutase catalytic activity has been independently confirmed in recent QM/MM calculations.[1,3,19] The results of the DTSS analysis indicate also a significant TS stabilization. Other contributions (NAC, preorganization, and so forth) are not calculated here since they require comparison with solution reaction, but the approach used here shows clearly that TS is stabilized relative to the bound substrate. This, compared to the lack of TS stabilization found for these structures in solution,[4] indicates that DTSS is significant for CM catalysis.

The current work is, we believe, the first analysis of CM catalysis that quantifies components of interaction energy in the CM active site up to the correlation component and addresses the physical nature of contributions to the activation barrier lowering. It is shown here, by the use of first principles quantum chemical methods, that the catalytic activity and interactions in the active site of CM are dominated by electrostatic effects. Therefore, this analysis justifies the commonly used approximation of neglecting higher terms (e.g., electron correlation) in modeling the interactions by Hartree−Fock or semiempirical methods. As shown here, the higher order components (correlation, delocalization, exchange, and penetration) cancel each other for most residues. Similarly, the DTSS analysis validates the QM/MM approach for modeling CM catalysis. The high-level analysis is in good agreement with QM/MM findings, showing the QM/MM results for CM to be reliable. Current analysis was performed for two structures (one pathway), but due to considerations outlined above (comparison with experiment and other QM/MM results), this should be a representative pathway.

The positions of catalytic residues can be predicted by an analysis of the map of differential electrostatic potentials on the van der Waals envelope or isodensity surface around the reactants. This opens the way to rational design of new catalysts or redesign of existing catalysts for new reactions, which may be obtained, for example, by site directed mutation of the corresponding enzyme. Negatively or positively charged amino acids can be placed in the regions indicated by the corresponding maps of static catalytic fields to enhance differential stabilization of the transition state. These entirely theoretical predictions of the optimal catalytic environment agree with those of the observed structures of CM from several species, and mutagenesis experiments show that charged residues, such as Glu78 and Arg90, are crucial for CM catalysis.[10,11,15]

DTSS analysis allows the calculation of the amount of TS stabilization with respect to the bound (preorganized) substrate. The static catalytic field shows that a vital difference between the preorganized bound substrate (NAC) and the TS is not in the geometry, but in the charge distribution. Charged residues are positioned precisely at the active site to stabilize the TS more than the bound substrate; their positions correspond to the prediction from the static catalytic field. This explains why the strongly charged active site can stabilize the TS much more strongly than the preorganized substrate.

Acknowledgment. We thank Dr. P. Kedzierski for help with generating the catalytic field maps, Dr. R. Góra for the modified version of the GAMESS-US program with the decomposition scheme implemented, and Professor Morris Krauss for many valuable comments. A.J.M. and K.E.R. would like to thank EPSRC, BBSRC, and the IBM Life Sciences Outreach Program for funding. This work has been supported in part by KBN Grant 3T09A 056 27. The calculations were done at the Wrocław Supercomputer Center (WCSS). QM/MM calculations were carried out on a Cray T3E at the CSAR National HPC Centre and on the HPCx (www.hpcx.ac.uk).

Supporting Information Available: Table SI-1, showing the decomposition of the reactant and TS interaction energies with the active site residues. This material is available free of charge via the Internet at http://pubs.acs.org.

JA049376T