The effect of MM polarization on the QM/MM transition state stabilization: application to chorismate mutase — Source link ↗

Christopher J. R. Illingworth, Kevin E. B. Parkes, Christopher R. Snell, Sergio Martí ...+2 more authors

Institutions: University of Essex, James I University

Topics: Chorismate mutase and QM/MM

Related papers:

- Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme

- Classical Polarization in Hybrid QM/MM Methods

- QM/MM Methods for Biomolecular Systems

- Development of an Accurate and Robust Polarizable Molecular Mechanics Force Field from ab Initio Quantum Chemistry

- CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model.

The effect of MM polarization on the QM/MM transition state stabilization: application to chorismate mutase

Christopher Reynolds, Christopher J Illingworth, Kevin E Parkes, Chris R Snell, Vicent Moliner, Sergio Marti

The effect of MM polarization on the QM/MM transition state stabilization: application to chorismate mutase

The effect of MM polarization on the QM/MM transition state stabilization: application to chorismate mutase

C.J. R. ILLINGWORTH[1], K. E. PARKES[2], C. R. SNELL[2], S. MARTI[3], V. MOLINER[3], C. A. REYNOLDS[1,*]

[1]Department of Biological Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom

[2]Medivir UK Ltd., Chesterford Research Park, Little Chesterford, Essex CB10 1XL, United Kingdom

[3]Departament de Quimica Fisica i Analitica, Universitat Jaume I, Box 224, Castellón, Spain.

Abstract

Hybrid quantum mechanics / molecular mechanics (QM/MM) calculations provide a mechanism for studying enzyme catalysed reactions at the molecular level. Here, through applications on the chorismate to prephenate rearrangement within the enzyme chorismate mutase, we have demonstrated the feasibility of including MM polarization into these calculations using the method of induced charges. MM polarization is shown to be a short-range effect, such that 80% of the energy of MM polarization occurs within a 5 Å residue-based cut-off of the substrate. MM polarization was shown to have a greater magnitude within the enzyme catalysed reaction than in the aqueous reaction, indicating that MM polarization may in principle have a significant effect on enzyme rate enhancement and mechanism. In both the enzyme and the aqueous case, the percentage contribution of MM polarization to the total stabilization energy was towards the upper end of the expected value. For the specific structures studied here, MM polarization lowered the energy barrier for the aqueous reaction, but the calculated contribution of MM polarization to both the reactant and transition structure stability were similar.

Keywords: QM/MM, induced charge, polarization, chorismate mutase, transition state stabilization

1. Introduction

Transition state stabilization has been proposed as a major driving force behind enzyme rate enhancement [1] and since the transition state is believed to bind more tightly than the substrate, the use of transition state analogues has become a common strategy in rational drug design [2–7]. Here we propose to investigate the role of enzyme polarization on this transition state stabilization using hybrid quantum mechanical / molecular mechanics (QM/MM) methods with polarizable MM charges [8].

For particular reactions, the origin of enzyme rate enhancement may still be a matter of debate, and this is the case for the Claisen rearrangement of chorismate to prephanate. This reaction within the enzyme chorismate mutase has become the focus of a large number of hybrid quantum mechanical/molecular mechanics (QM/MM) computational studies, partly as an exemplar system for methodology development, because the reaction does not involve any covalent bonding between the enzyme and the substrate, thus avoiding the difficult issue of link atoms. Moreover, the mechanism is believed to occur with the same mechanism in solution, and as such the enzyme has played a central role in recent debates about theories of enzyme catalysis. Here we propose to investigate whether MM polarization, which is currently ignored in most QM/MM studies, can play a role in transition state stabilization.

Early studies of the system indicated, in line with the ideas of Pauling [1,9], that the catalysis proceeds by a mechanism of transition state stabilisation [10–12], in which the transition state is bound more strongly by the enzyme than the chorismate reactant. More recent controversy has focused around the argument that catalysis is instead due to the preferential binding by the enzyme of a reactive conformation of chorismate, known as a near-attack conformation.[13–15] Through improved computational techniques, more accurate simulations of the Claisen rearrangement have been carried out [16–18], leading to estimates

of the reaction energy barriers that are in close correspondence with experimentally derived

values [16,19], and seeming to confirm the importance of transition state stabilisation.

However, in these models electronic polarization is included in the QM model of the ligand

but it is not incorporated into the MM model of the enzyme. Noting the importance of

electrostatic effects in this system, [16,20] here we apply the method of induced charges

[8,21−23] (IC) to incorporate polarization into a QM/MM(IC) model of the chorismate

mutase reaction, with the aim of demonstrating the feasibility of studying the effect of

enzyme polarization on transition state stabilization.

2. Methods

Calculations were performed on reactant, product, and transition state structures of *Bacillus*

subtilis chorismate mutase (BsCM), and of the ligand in aqueous solution, taken from the

OH_{in} pathway calculated by Marti et al. [24]. In each case, the QM region was defined to be

the substrate. The enzymatic system was modelled using an MM region consisting of all

residues with at least one atom within 10 Å of the substrate, in at least one of the three

structures, making a total of 1438 MM points in each case. The aqueous solvent was

modelled using an MM region with a cut-off chosen to include the 364 water molecules

closest to the substrate, this being equivalent to a 10 Å cut-off in the reactant system. This

MM region is smaller than in other models of this system. In order to verify that this was

sufficient to capture the effects of polarization in the MM region, a series of calculations

were carried out measuring the MM polarization in the aqueous reactant system with cut-offs

varying from 3 to 11 Å, which were extrapolated to give an estimate of the total MM

polarization in an aqueous solvent with no cut-off.

Initial charges for residues in the MM region were derived from the Amber03 [25] forcefield.

Gaussian03 [26] was used to carry out a QM optimisation on a water monomer at the

B3LYP/6-31G(d) level of theory [27,28], and Stone's GDMA 1.3 program [29] was used to convert the charge distribution into a multipole series, which was then converted into potential derived point charges using the mulfit methodology [21,22,30−32]. Both the protein and the water charges are derived from correlated wavefunctions and therefore suitable for polarization calculations [32]. These are potential derived charges in which the potential is determined from a distributed multipole analysis up to hexadecapole in a spherical shell around the multipole centre. The optimal charges that reproduce this potential on the multipole centre and the atoms bonded to it are determined using an analytical procedure that avoids the use of a numerical grid [30].

For each QM/MM system, the wavefunction was calculated using Gaussian03 [26] and expressed as a distributed multipole analysis using GDMA 1.3 [29]. The field was then calculated at the MM (enzyme and/or water) atoms from the multipole series using ORIENT 4.5 [33]. Induced dipoles, $\underline{\mu}$, were calculated for each atom in the MM region, according to the equation

$$\underline{\mu} = \alpha \underline{E}$$

where $\underline{E}$ is the field, and α is the isotropic atomic polarizability, taken from Miller and Savchik [34]. These dipoles were then expressed as induced charges using the mulfit technology, described above [21,22,30−32]. These new induced charges were added to the permanent charges to generate a new QM/MM system, and in accordance with previous work [8] this process was iterated four times, bringing the system very close to convergence. Examining the individual point charges, the largest change in the charge on an MM point as a result of the final iteration of the method was of magnitude 0.000011. In each case, calculations were performed at the B3LYP/6-31G(d) level of theory. Five energy values were calculated, labelled E_1 to E_5 (all of which are determined at the geometry of the complex). E_1 was defined as the energy of the monomer in vacuo. E_2 was defined to be the first SCF energy from the first SCF cycle of a QM/MM run, namely the energy of the

monomer wavefunction in the presence of the point charges. Thus E_2-E_1 gives the energy of the electrostatic interaction between the unperturbed wavefunction and the point charges. E_3 was defined as the converged SCF energy for the initial QM/MM run, such that E_3-E_2 is equal to the polarization energy of the QM system in response to the MM charges. E_4 was defined as the energy of the converged QM/MM run after one iteration of the classical polarization process, while E_5 was defined as the energy of the final QM/MM run, after four iterations of the classical polarization process. Hence the total MM polarization energy is derived from E_5-E_3. The induced charge polarization energies included the self energy correction [8].

Barrier heights for the reaction were calculated by comparing the energy values at different stages in the QM/MM calculation. For example, the height of the barrier to transition state, incorporating MM polarization, was given by subtracting the value of E_5 from the QM/MM run on the reactant from the value of E_5 from the QM/MM run on the transition state. The energy values E_1 to E_5 are tabulated in supporting information. However, the key focus of this article is not so much the magnitude of the energy barriers but the extent to which they are effected by MM polarization, which is essentially given by E_4-E_3 (one iteration) and E_5-E_3 (4 iterations).

3. Results and Discussion
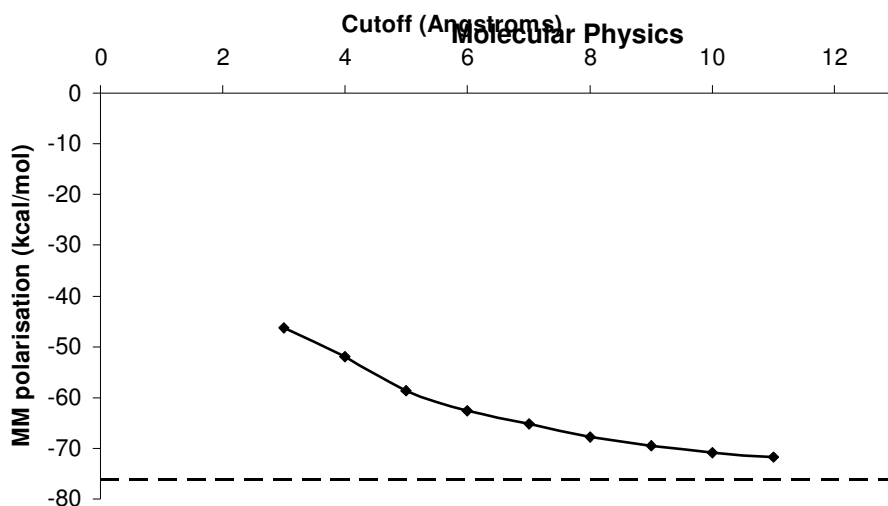
Figure 1 MM polarisation energy ($E_5 - E_3$) as a function of cut-off for the chorismate reactant in an aqueous solution, modelled using explicit waters. The calculated values are shown by the solid diamonds. The dotted line shows the asymptote of the extrapolated curve, that is, the estimated limit of the polarisation energy as the cutoff tends to infinity.

The MM polarisation energies for the reactant in aqueous solution as a function of the cutoff are shown in Figure 1. The MM energy of polarisation with a 10 Å cutoff was calculated to be -70.9 kcal mol^{-1}. The curve of the graph appears to increase asymptotically. Using the assumption that the polarisation energy in a shell of distance r from the ligand is proportional to r^{-3} (see supporting information), an estimate for the asymptote was calculated at -76.2 kcal mol^{-1}. The figure suggests that MM polarisation is primarily a short-range effect, with 77% of the MM polarisation energy occurring on molecules within 5 Å of the ligand, and 93% of the MM polarisation energy being captured by a simulation with a cutoff of 10 Å. This high percentage justifies the use of the 10 Å cutoff. Where different structures for the enzyme at different stages of the reaction are compared, the differences are likely to be close to the site of ligand binding. Thus, the effect on the energy of MM polarization of residues more than 10 Å from the binding site is likely not only to be small, but very similar between the structures. Because of this, the resultant error in the energy of transition state stabilisation caused by neglecting MM polarization of residues more than 10 Å from the ligand is likely to be negligible.

Table 1. Polarization energy values for the enzyme and aqueous systems, at different stages in the reaction pathway. The final column gives the MM polarization energy as a percentage of sum of the electrostatic and polarization energies.

	Energy (kcal mol^{-1})		
	QM polarization E_3-E_2	MM polarization E_5-E_3	% MM polarization
Enzyme system			
Reactant	-15.4	-94.9	14.9
Transition state	-14.4	-94.9	17.4
Product	-14.1	-91.2	16.2
Aqueous system			
Reactant	-10.2	-70.9	18.1
Transition state	-9.4	-72.3	19.1
Product	-9.4	-72.8	19.1

Table 1 shows QM and MM polarization energies for the reactants, transition structure and products in both the enzyme and water systems. The total MM polarization energy is relatively large, at about 17±2% of the total electrostatic and polarization interaction, which places this towards the upper end of the expected percentage value [32]. This high percentage is to be expected as the chorismate and prephanate are both highly charged and so will exhert a large electric field. The Arg counterions will also exert a large field on their neighbours within the enzyme. Because the numerical MM polarization energy values are indeed large, MM polarization clearly has the potential to influence the rate and mechanism of enzyme catalysed reactions determined using QM/MM methods. Examination of the induced charges

in the MM system from the enzyme system showed that the largest changes occur on residues and water molecules close to the carboxylate groups in the ligand, e.g. Arg 90.

Figure 2 shows a graphical representation of the changes in MM charges for the three systems. Comparison of the figures for reactant, transition state and product systems show that the changes in the MM charges are roughly constant across the three systems. Residues and water molecules containing atoms that have large induced charges (i.e. greater than 0.1) are listed in supporting information.

Table 1 shows that the MM polarization of the enzyme in the reactant and transition structure configurations is also very similar (the MM polarization energies do not differ by more than 4%). The effect of this in this case, is that inclusion of MM polarization has a small effect on the energy barrier for the conversion of chorismate to prephenate since polarization increases the stability of the ligand in both the reactant and transition structure configurations by a similar amount of the energy. A similar conclusion was reached by Szefczyk et al. on the basis of quantum mechanical energy decomposition calculations on a limited number of active site residues [35]. Thus the increase in the barrier from reactant to transition state when MM polarization is incorporated into the model is only 0.1 kcal mol^{-1} (once the full QM energies have been included, see supporting information).

MM polarization in the product configuration is -91.2 kcal mol^{-1} and this is smaller in magnitude than the -94.9 kcal mol^{-1} for the transition structure. As a result of this, MM polarization lowers the
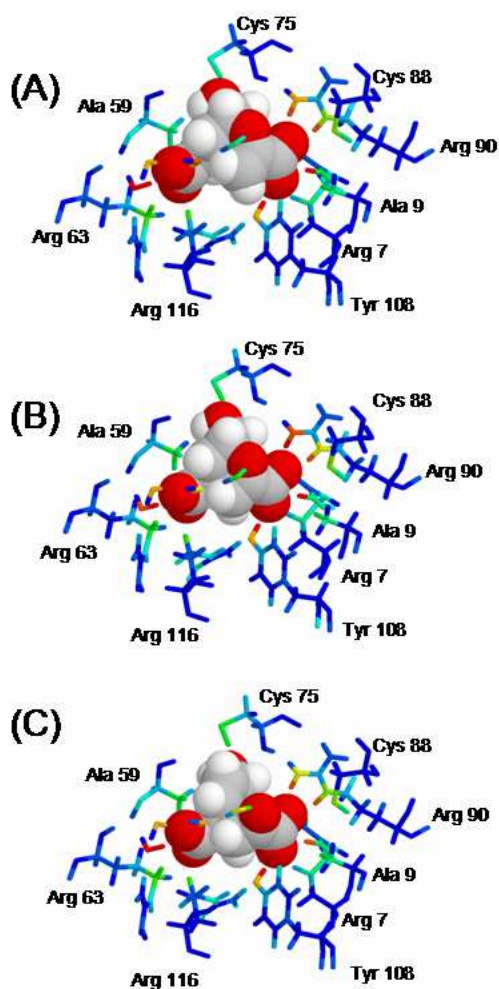
Figure 2. A graphical representation of the induced charges on the enzyme and key water

molecules in (A) the reactant configuration, (B) the transition structure and (C) the product

configuration. Atoms shown in red have the largest induced charges; atoms shown in blue

have the smallest induced charges. The magnitude of the induced charge is equal to the

change made to the original charge in the model. The chorismate ligand is shown in a space-

filling representation in CPK colours. In each of the 3 cases, the induced charges are shown

mapped onto the transition state configuration.

barrier for the reverse reaction by 7.5 kcal mol[-1] (once the full QM energies have been

included)

For the reaction in aqueous solution, the numerical values of the MM polarization of the water are interesting for two reasons. Firstly, the electronic polarization effects are greater in the enzyme environment than in the water and this suggests that MM polarization could in general have more influence on reactions carried out within enzymes than on reactions carried out in aqueous solution. Secondly, the effect of MM polarization on the environment (i.e. the water) has a greater effect in the transition state (-72.3 kcal mol^{-1}) than in the reactant configuration (-70.9 kcal mol^{-1}). These results mean that polarization lowers the energy barrier for the forward reaction by 2.7 kcal mol^{-1} (once the full QM energies have been included). In this sense then, the calculations predict that MM polarization has an unexpected indirect adverse effect on the enzyme rate enhancement.

In contrast to the reaction in the enzyme, in aqueous reaction, MM polarization has essentially no effect on the reverse reaction.

Here we have demonstrated the feasibility of including classical polarization of MM charges [8] in QM/MM studies of enzyme mechanisms. When considering the enzyme alone, our calculations suggest that MM polarization does not affect the barrier height to the transition state in the enzyme system. An implication of the lower predicted contribution of MM polarization to transition state stabilization in the enzyme compared to water is that if MM polarization plays a lower role than expected, then reactive conformation selection may play a higher role than expected. However, a definitive answer on the impact of MM polarization on the chorismate mutase reaction would require a more careful treatment of the enzyme system, possibly by investigating multiple pathways [16].

References

1. L.Pauling, Chem. Eng. News **24**, 1375 (1946).

2. L.Gan, M. R. Seyedsayamdost, S. Shuto, A. Matsuda, G. A. Petsko, and L. Hedstrom, Biochemistry **42**, 857 (2003).

3. M.J.Costanzo, S. C. Yabut, H. R. Almond, Jr., P. Andrade-Gordon, T. W. Corcoran, L. De Garavilla, J. A. Kauffman, W. M. Abraham, R. Recacha, D. Chattopadhyay, and B. E. Maryanoff, J Med. Chem. **46**, 3865 (2003).

4. S.P.Dharmasena, D. S. Wimalasena, and K. Wimalasena, Biochemistry **41**, 12414 (2002).

5. G.A.Kicska, P. C. Tyler, G. B. Evans, R. H. Furneaux, V. L. Schramm, and K. Kim, J Biol. Chem. **277**, 3226 (2002).

6. I.Basu, G. Cordovano, I. Das, T. J. Belbin, C. Guha, and V. L. Schramm, J Biol. Chem. **282**, 21477 (2007).

7. R.E.Rigsby, C. L. Rife, K. L. Fillgrove, M. E. Newcomer, and R. N. Armstrong, Biochemistry **43**, 13666 (2004).

8. C.J.Illingworth, S. R. Gooding, P. J. Winn, G. A. Jones, G. G. Ferenczy, and C. A. Reynolds, J. Phys. Chem. A **110**, 6487 (2006).

9. L.Pauling, Nature **161**, 707 (1948).

10. S.Marti, J. Andres, V. Moliner, E. Silla, I. Tunon, J. Bertran, and M. J. Field, J. Amer. Chem. Soc. **123**, 1709 (2001).

11. P.D.Lyne, A. J. Mulholland, and W. G. Richards, J. Amer. Chem. Soc. **117**, 11345 (1995).

12. K.E.Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann, and A. J. Mulholland, Mol. Phys. **101**, 2695 (2003).

13. S.Hur, T. C. Bruice, Proc. Natl. Acad. Sci. U. S. A **100**, 12015 (2003).

14. S.Hur, T. C. Bruice, J. Amer. Chem. Soc. **125**, 10540 (2003).

15. S.Hur, T. C. Bruice, J. Amer. Chem. Soc. **125**, 5964 (2003).

16. F.Claeyssens, K. E. Ranaghan, F. R. Manby, J. N. Harvey, and A. J. Mulholland, Chem Commun. 5068 (2005).

17. K.E.Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann, and A. J. Mulholland, Org. Biomol. Chem **2**, 968 (2004).

18. F.Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schutz, S. Thiel, W. Thiel, and H. J. Werner, Angew. Chem. Int. Ed. **45**, 6856 (2006).

Deleted: G.

19. P.Kast, M. AsifUllah, and D. Hilvert, Tetrahedron Letters **37**, 2691 (1996).

20. M.Barbany, H. Gutierrez-de-Teran, F. Sanz, J. Villa-Freixa, and A. Warshel, Chembiochem **4**, 277 (2003).

21. S.R.Gooding, P. J. Winn, R. I. Maurer, G. G. Ferenczy, J. R. Miller, J. E. Harris, D. V. Griffiths, and C. A. Reynolds, J. Comput. Chem. **21**, 478 (2000).

22. P.J.Winn, G. G. Ferenczy, and C. A. Reynolds, J. Comput. Chem. **20**, 704 (1999).

23. J.H.Wu, P. J. Winn, G. G. Ferenczy, and C. A. Reynolds, Int. J. Quant. Chem. **73**, 229 (1999).

24. S.Marti, J. Andres, V. Moliner, E. Silla, I. Tunon, and J. Bertran, J. Phy. Chem. B **104**, 11308 (2000).

25. Y.Duan, C. Wu, S. Chowdhury, M. C. Lee, G. M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. M. Wang, and P. Kollman, J. Comput. Chem. **24**, 1999 (2003).

26. Frisch MJ, Trucks GW, Shlegel HB, Scuseria GE, Robb MA et al. (2004) Gaussian 03, Revision C.02, version Wallingford CT:

27. A.D.Becke, J. Chem. Phys. **98**, 5648 (1993).

28. W.J.Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

29. A.J.Stone, J. Chem. Theory Comput. **1**, 1128 (2005).

30. G.G.Ferenczy, J. Comput. Chem. **12**, 913 (1991).

31. G.G.Ferenczy, P. J. Winn, and C. A. Reynolds, J. Phys. Chem. A **101**, 5446 (1997).

32. P.J.Winn, G. G. Ferenczy, and C. A. Reynolds, J. Phys. Chem. A **101**, 5437 (1997).

33. Stone AJ, Dullweber A, Engkvist O, Fraschini E, Modges MP et al. (2002) Orient: a program for studying interactions between molecules, version 4.5, version University of Cambridge, Enquiries to AJ Stone, ajs1@cam.ac.uk. 2006.:

34. K.J.Miller, J. A. Savchik, J. Amer. Chem. Soc. **101**, 7206 (1979).

35. B.Szefczyk, F. Claeyssens, A. J. Mulholland, and W. A. Sokalski, Int. J. Quant, Chem. **107**, 2274 (2007).

Deleted: -&

Deleted: u

The effect of MM polarization on the QM/MM transition state stabilization: application to chorismate mutase

Christopher J. R. Illingworth[1], Kevin E. Parkes[2], Christopher R. Snell[2], S. Marti[3], Vincent Moliner[3], Christopher A. Reynolds[1,*]

[1]Department of Biological Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom, [2]Medivir UK Ltd., Chesterford Research Park, Little Chesterford, Essex CB10 1XL, United Kingdom [3]Departament de Ciències Experimentals, Universitat Jaume I, Box 224, Castellón, Spain,.

Table S1 Energy values (E1 to E5) for the different systems tested. The measured data gives the raw energy values, from the QM/MM run. The 'point charges' terms are the energy of the MM point charges, considered in isolation. These are subtracted from the measured energy values to give the corrected energy terms, which give the energy of the QM region. All values are in Hartrees.

Enzyme system

Reactant

	E1	E2	E3	E4	E5
Measured	-837.116367	-890.763672	-890.788264	-891.809700	-891.858188
Point charges		-52.561503	-52.561503	-53.288037	-53.328917
Corrected	-837.116367	-838.202169	-838.226761	-838.521662	-838.529271

Transition state

	E1	E2	E3	E4	E5
Measured	-837.077742	-891.310861	-891.333852	-892.422144	-892.462504
Point charges		-53.291836	-53.291836	-54.085155	-54.118100
Corrected	-837.077742	-838.019025	-838.042016	-838.336988	-838.344403

Product

	E1	E2	E3	E4	E5
Measured	-837.154240	-890.767666	-890.790086	-891.782741	-891.813676
Point charges		-52.538422	-52.538422	-53.246409	-53.271563
Corrected	-837.154240	-838.229243	-838.251663	-838.536332	-838.542113

Water system

Reactant

	E1	E2	E3	E4	E5
Measured	-837.118475	-934.850326	-934.866611	-937.768985	-937.836082
Point charges		-97.237494	-97.237494	-99.919034	-99.981136
Corrected	-837.118475	-837.612832	-837.629117	-837.849952	-837.854946

Transition state	E1	E2	E3	E4	E5
Measured	-837.067469	-934.813571	-934.828516	-937.601088	-937.667587
Point charges		-97.274160	-97.274160	-99.821504	-99.883029
Corrected	-837.067469	-837.539411	-837.554356	-837.779584	-837.784558

Formatted Table

Product	E1	E2	E3	E4	E5
Measured	-837.154179	-934.850215	-934.865217	-937.656686	-937.725032
Point charges		-97.218959	-97.218959	-99.783683	-99.846802
Corrected	-837.154179	-837.631257	-837.646258	-837.873003	-837.878230

Formatted Table

Table S2 Largest changes in point charges in MM region of reactant system

Residue	Atom	Initial charge	Final charge	Change
Tyr 108	HH	0.358	0.555	0.197
TIP3 86	H	0.372	0.567	0.195
Arg 7	2HH2	0.421	0.609	0.188
TIP3 86	O	-0.744	-0.931	-0.187
Arg 90	2HH1	0.421	0.599	0.179
Arg 90	HE	0.323	0.499	0.176
HOH 105	O	-0.744	-0.918	-0.174
Tyr 108	OU	-0.476	-0.642	-0.166
HOH 105	H	0.372	0.536	0.164
Arg 90	N2	-0.835	-0.998	-0.163
TIP3 70	H	0.372	0.531	0.159
TIP3 70	O	-0.744	-0.903	-0.158
HOH 41	H	0.372	0.524	0.152
HOH 41	O	-0.744	-0.885	-0.141
Arg 7	1HH1	0.421	0.557	0.137
Arg 90	NE	-0.511	-0.638	-0.127
Cys 88	SG	-0.243	-0.119	0.124
Arg 116	1HD	0.058	0.170	0.112
Arg 63	1HD	0.058	0.168	0.111
Arg 63	CD	0.101	-0.008	-0.109

Table S3 Largest changes in point charges in MM region of transition state system

Residue	Atom	Initial charge	Final charge	Change
Tyr 108	HH	0.358	0.552	0.193
TIP3 86	H	0.372	0.561	0.189
Arg 90	2HH1	0.421	0.607	0.187
Arg 7	2HH2	0.421	0.603	0.183
TIP3 86	O	-0.744	-0.922	-0.178
Arg 90	HE	0.323	0.499	0.176
Arg 90	N2	-0.835	-1.002	-0.167
Tyr 108	OU	-0.476	-0.638	-0.162

Residue	Atom			
HOH 105	H	0.372	0.531	0.159
HOH 41	H	0.372	0.530	0.158
TIP3 70	H	0.372	0.525	0.152
TIP3 70	O	-0.744	-0.894	-0.150
HOH 41	O	-0.744	-0.894	-0.150
HOH 105	O	-0.744	-0.888	-0.144
ARG7	1HH1	0.421	0.562	0.141
ARG90	NE	-0.511	-0.643	-0.131
CYS88	SG	-0.243	-0.121	0.122
ARG116	1HD	0.058	0.160	0.102
TIP3 51	O	-0.744	-0.845	-0.101
ARG63	1HD	0.058	0.158	0.101
TYR108	HH	0.358	0.552	0.193
TIP3 86	H	0.372	0.561	0.189

Table S4 Largest changes in point charges in MM region of product system

Residue	Atom	Initial charge	Final charge	Change
TIP3 86	H	0.372	0.558	0.185
TYR108	HH	0.358	0.539	0.180
TIP3 86	O	-0.744	-0.918	-0.173
ARG7	2HH2	0.421	0.592	0.171
ARG90	2HH1	0.421	0.588	0.168
ARG90	HE	0.323	0.484	0.160
HOH 105	H	0.372	0.529	0.157
TIP3 70	H	0.372	0.523	0.151
TYR108	OU	-0.476	-0.626	-0.150
HOH 41	H	0.372	0.518	0.146
TIP3 70	O	-0.744	-0.889	-0.144
HOH 105	O	-0.744	-0.885	-0.141
ARG90	N2	-0.835	-0.972	-0.138
ARG7	1HH1	0.421	0.556	0.136
HOH 41	O	-0.744	-0.880	-0.136
TIP3 51	O	-0.744	-0.870	-0.126
CYS88	SG	-0.243	-0.124	0.119
ARG90	NE	-0.511	-0.630	-0.119
ARG116	1HD	0.058	0.158	0.100
ARG63	1HD	0.058	0.157	0.100

Curve estimation

In the reactant water system, an estimate is given for the total MM polarisation energy, if the reactant was in an infinite amount of water. This is given by the following method:

Suppose that the total magnitude of the MM polarisation energy is equal to some limit, L. Dividing space according to its distance from the ligand into regions of thickness 1 Å, and defining the polarisation energy of the waters in the shell of distance a to a+1 Å from the ligand to be e(a), we have:

$$L = \sum_{a=0}^{\infty} e(a)$$

If the polarization energy e(a) of each shell is defined primarily in terms of charge-dipole interactions, the average strength of an individual interaction will be roughly proportional to a^{-4}, while the volume of the shell is roughly proportional to a^2. Thus it roughly holds that

$$e(a) \propto a^{-4} * a^2 = a^{-2}$$

If we define

$$E(r) = \sum_{a=0}^{r-1} e(a)$$

we can then say that

$$L = E(r) + \sum_{a=r}^{\infty} e(a)$$

Given the proportionality relation defined for e(a), it then holds roughly that

$$\sum_{a=r}^{\infty} e(a) \propto \sum_{a=r}^{\infty} \frac{1}{a^2}$$

Noting that

$$\frac{1}{a^2} - \frac{1}{a(a+1)} = \frac{1}{a^2(a+1)}$$

we then have the result

$$\sum_{a=r}^{\infty} e(a) \propto \sum_{a=r}^{\infty} \frac{1}{a^2} = \sum_{a=r}^{\infty} \left[\frac{1}{a(a+1)} + \frac{1}{a^2(a+1)} \right] = \sum_{a=r}^{\infty} \frac{1}{a(a+1)} + \sum_{a=r}^{\infty} \frac{1}{a^2(a+1)}$$

Looking at the first sum, we have

$$\sum_{a=r}^{\infty} \frac{1}{a(a+1)} = \sum_{a=r}^{\infty} \left[\frac{1}{a} - \frac{1}{a+1} \right] = \frac{1}{r}$$

since successive terms of the sum cancel out. For the second sum, we have

$$\sum_{a=r}^{\infty} \frac{1}{a^2(a+1)} = \sum_{a=r}^{\infty} \frac{1}{a} \left[\frac{1}{a} - \frac{1}{a+1} \right] < \sum_{a=r}^{\infty} \frac{1}{r} \left[\frac{1}{a} - \frac{1}{a+1} \right] = \frac{1}{r^2}$$

As r increases, this second sum becomes small, so we can say approximately that

$$\sum_{a=r}^{\infty} e(a) \propto \frac{1}{r}$$

Hence, for some constant c_1, we have the model

$$L = E(r) + \frac{c_1}{r} .$$

If, however, the energy e(a) of each shell is defined primarily in terms of dipole-dipole interactions, the average strength of an individual interaction will be roughly proportional to a^{-6}. It would then follow, again roughly, that

$$e(a) \propto a^{-6} * a^2 = a^{-4}$$

$$\sum_{a=r}^{\infty} e(a) \propto \frac{1}{r^3} \text{ and}$$

Deleted: for some constant c_1.

Formatted: Lowered by 14 pt

Deleted: $\sum_{a=r}^{\infty} e(a) \propto \sum_{a=}^{\infty}$

Formatted: Lowered by 14 pt

Formatted: Lowered by 15 pt

Formatted: Lowered by 14 pt

Formatted: Lowered by 14 pt

Formatted: Lowered by 14 pt

Deleted: and so just above we showed that e(a) prop to a^{-2} so why do we have 1/r here and not $1/r^2$?

Formatted: Superscript

Formatted: Superscript

Deleted: for some constant c_2

$L = E(r) + \dfrac{c_2}{r^3}$ for some constant c_2.

Therefore, suppose that for some constant values k and c, the amount of polarisation energy in the volume up to r Ångstroms away from the ligand obeys the function

$$E(r) = L - \frac{c}{r^k} \qquad\qquad \text{Eq.1}$$

This would imply that

$$E(r+1) - E(r) = \left(L - \frac{c}{(r+1)^k} \right) - \left(L - \frac{c}{r^k} \right) = c\left(\left(\frac{1}{r^k}\right) - \left(\frac{1}{(r+1)^k}\right) \right) \quad \text{Eq.2}$$

By looking at the calculated data, we then attempt to deduce the values of k and c, in order to create a model for the polarisation energy.

Taking values of r from 8 to 11 Ångstroms gives the following observed differences:

r	E(r)	E(r)-E(r-1)
8	-67.73682	-2.571943
9	-69.48578	-1.748963
10	-70.85499	-1.36921
11	-71.73236	-0.877372

These differences can be compared to model difference values computed using different values of k, for example where k is equal to 1, as follows:

r	E(r)	E(r)-E(r-1)	Predicted difference	c(r)
8	-67.73682	-2.571943	0.017857143	-144.0288241
9	-69.48578	-1.748963	0.013888889	-125.9253718
10	-70.85499	-1.36921	0.011111111	-123.2288624
11	-71.73236	-0.877372	0.009090909	-96.51094997

Here the predicted difference is equal to E(r)-E(r-1) as given from Eq. 2 above, under the assumption that c=1, while c(r) is the observed difference divided by the predicted difference at each r, that is, the value of c in Eq.2 needed to give a correct answer at that r.

By repeating this with different values of k, different sets of values of c(r) can similarly be computed. The goodness of fit of the model is indicated by the consistency of the values c(r) at each value of r – if c(r) was constant that would indicate a perfect fit to the model, while a large variation in c(r) would indicate a poor model fit.

The consistency of the value of c(r) was measured at different values of k by calculating the standard deviation of the c(r) values divided by the mean of the c(r) values at that value of k. For simplicity, k was assumed to be an integer. This gave the following results:

k	stdev[c(r)] / mean[c(r)]
1	0.160008
2	0.059987
3	0.14485
4	0.273271

As can be seen from the table, k=2 gives the best fit for this model. Given this result, the constant c was defined as the mean of the values c(8) to c(11) already calculated, giving the model:

$$E(r+1) - E(r) = -540.1 \left(\left(\frac{1}{r^2} \right) - \left(\frac{1}{(r+1)^2} \right) \right)$$

$$E(11) = -71.732$$

The limit of this sequence was calculated computationally up to r=10000, giving a limiting value for the energy of -76.2 kcal mol^{-1}.

This model does not fit directly to either of the two scenarios presented above, but a number of factors need to be taken into account. Firstly, it is primarily a numerical procedure and there is not much data to extrapolate from. Secondly, it ignores higher order terms as they rapidly become small. For example if the energy varied according to c1/r + 0.3 c2/r3 then at r = 11 Å, the first term is 50 times bigger than the second. It was argued above that the charge - (induced) dipole correction would be proportional to r-2 but this would only be true if the magnitude of the changes in the charges was constant – but this is unlikely to be true given the charged chorismate involved and this decrease in induced charge with distance from the chorismate may explain the preference for k=2 rather than k=1 in our model.