Modeling Enzyme Reaction Intermediates and Transition States: Citrate Synthase

Adrian J. Mulholland*,†,‡,§ and W. Graham Richards‡

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, and Physical and Theoretical Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, U.K.

Received: February 12, 1998

To investigate the roles of catalytic residues and the nature of the nucleophilic intermediate in citrate synthase, calculations have been carried out on models of the first stage of the reaction with ab initio (MP2/6-31+G(*d*)//6-31+G(*d*)) and semiempirical (AM1 and PM3) methods. The first stage of the reaction involves deprotonation of acetyl-CoA, with Asp-375 identified as the likely base. The resulting intermediate is believed to be stabilized by a hydrogen bond from a neutral histidine side chain (His-274), which has been suggested to be a "short, strong" or "low-barrier" hydrogen bond. Such bonds have been suggested to have exceptionally high energies and to stabilize many enzyme reaction intermediates. Transition state and stable hydrogen bonded complex geometries have been fully optimized for models of proton transfer between Asp-375 and acetyl-CoA and between the enolate of acetyl-CoA and His-274. The results support the proposal that Asp-375 is the base in the reaction and show that stabilization of the thioester enolate at the active site is required if it is to exist as a reaction intermediate. The effective basicities of methylimidazolate (representing His-274) and the thioester enolate in the hydrogen bonded complex are calculated to be clearly different. The enolate−methylimidazole complex (with the proton localized on methylimidazole) is found to be significantly lower in energy than the enol complex (in which the hydrogen bonded proton is transferred from methylimidazole), which is not stable. Unless the pK_a for deprotonation of neutral His-274 is lowered by conditions in the active site, it appears that the hydrogen bond with the enolate will not be of the low-barrier type. The highly favorable energy of the interaction between the thioester enolate and methylimidazole, as well as calculations on a larger model including all three components, indicate that a normal hydrogen bond with His-274 can make an important contribution to stabilization of the enolate intermediate.

Introduction

Many enzymes achieve remarkable efficiency and specificity for apparently challenging reactions, by mechanisms which are often not clear.[1,2] An example of such an enzyme is citrate synthase (CS), which plays an important part in biosynthetic and energy converting metabolic pathways.[3] Citrate is formed by the Claisen-type condensation of acetyl-CoA with oxaloacetate, the overall reaction being thermodynamically driven by hydrolysis of the thioester bond.[4] The enzyme is able rapidly to form (and in the reverse reaction, to cleave) a carbon−carbon bond, without a requirement for metal ions or other cofactors. Eukaryotic CS, a homodimer of molecular weight approximately 100 000, has been studied intensively by a variety of techniques.[4−7] A large conformational change from an "open" to a "closed" form, driven by substrate binding, prevents wasteful side reactions such as the hydrolysis of acetyl-CoA and is reversed to allow product release.[8] Crystallographic and site-directed mutagenesis investigations have identified Asp-375 (the numbering is for pig CS) as the base responsible for deprotonation of acetyl-CoA.[9,10] Kinetic isotope effects indicate that this step is rate-limiting.[11] His-274 is observed to bind to the carbonyl oxygen of acetyl-CoA[12] and also plays a catalytic role.[6,13,14] This mechanism appears to be conserved in *S*-citrate synthases from a wide variety of species.[15]

It is believed that deprotonation of acetyl-CoA produces the enol or enolate of the thioester which in turn acts as the nucleophile to attack oxaloacetate.[16,17] The exact nature of this nucleophilic intermediate has been uncertain, but it has been proposed that a hydrogen bond between the carbonyl oxygen of acetyl-CoA and Nδ1 of His-274 (which is neutral[7,16]) plays a central part in stabilizing it. In particular, it has been suggested that when the intermediate is formed it assumes the special character of a "short, strong", "low-barrier", or Speakman-Hadzi hydrogen bond in which the proton is effectively shared between the two heavy atoms and that the high energy of this bond provides the necessary stabilization of the "enolic" intermediate.[18−21] For such a bond to form, the pK_a's of the two hydrogen bonding partners should be approximately equal, the system should be charged, and bulk water must be excluded. This mechanism has been proposed to be of widespread importance but has been the subject of considerable debate.[22−25] Several findings indicate that, while short hydrogen bonds may exist in enzyme active sites, they do not provide significantly more stabilization than normal charged hydrogen bonds.[26−30]

To examine the mechanism of proton abstraction in CS (Figure 1) and the hydrogen bond between His-274 and the nucleophilic intermediate, we have performed calculations on active site models of CS with ab initio and semiempirical molecular orbital methods. Unstable species such as reaction intermediates and transition states (TSs) are difficult to study

* Corresponding author.
† Harvard University.
‡ Oxford University.
§ Present address: School of Chemistry, University of Bristol, Bristol, BS8 1TS, U.K. E-mail: Adrian.Mulholland@bristol.ac.uk.

S1089-5647(98)01121-3 CCC: $15.00
Published on Web 07/31/1998

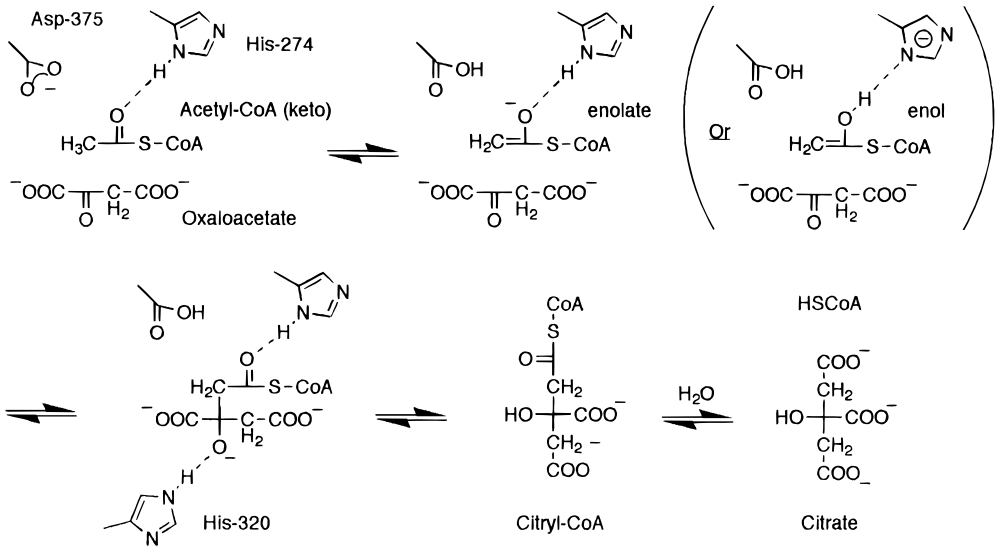

Figure 1. Mechanism of the citrate synthase reaction indicating the keto (substrate), enolate, and enol forms of acetyl-CoA.

directly, and calculations of their properties can make an important contribution to studies of enzyme catalysis.[31] TS structures and reactant and product complexes for proton abstraction by acetate from methylthioacetate (representing Asp-375 and acetyl-CoA, respectively) and for proton transfer between 5-methylimidazole and the enolate of methylthioacetate (representing His-274 and the acetyl-CoA enolate) have been optimized at the RHF/6-31+G(*d*) and RHF/6-31G(*d*) levels. These results, including MP2 correlation corrections, are compared to semiempirical calculations on the same reactions and provide a measure of the accuracy of the more approximate treatments for this system. AM1 TS structures have also been located for the same reactions in a larger model including all three components (representing His-274, Asp-375, and acetyl-CoA in their active site orientations) to calculate the contribution of His-274 to stabilizing the intermediate in the reacting system. The results show that Asp-375 is well positioned in the enzyme to deprotonate the substrate and support the conclusion drawn from solution pK_a's that significant stabilization of the resulting enolate by the enzyme is required for it to have sufficient lifetime to exist as an intermediate.[32] The basicities of the enolate and methylimidazolate are calculated to be markedly different, and consequently, unless the effective pK_a for deprotonation of neutral His-274 is lowered by interactions at the active site, it appears that the hydrogen bond formed with the enolate will not be of the low-barrier type. The calculations indicate that the increased strength of the normal hydrogen bond from His-274 to acetyl-CoA on deprotonation of the substrate may play an important part in stabilizing the intermediate. The results are compared to those from quantum mechanical/molecular mechanical (QM/MM) calculations on the reaction in CS,[7] as well as experimental results for this and related enzymes.

Methods

Three models of the enolization of acetyl-CoA in the active site of CS were studied. Each of these was based on molecular fragments taken from the crystal structure of the ternary complex of chicken heart CS with the inhibitor *R*-malate and acetyl-CoA at 1.9 Å resolution,[12] a good model of the reactive enzyme−substrate complex.[7] The three models consisted of the side chains of Asp-375 and/or His-274, as well as the thioester portion of acetyl-CoA, as follows: model A, Asp-375 and acetyl-CoA; model B, His-274 and acetyl-CoA; model C, Asp-375, His-274, and acetyl-CoA. The hydrogen bonding environment of His-274 indicates that its side chain is neutral,[7,16] protonated only on Nδ1, so it is represented here by 5-methylimidazole. The Asp-375 side chain is represented by acetate, and acetyl-CoA by methylthioacetate. Hydrogens were added in ideal geometries with the QUANTA program.[33]

Model A was used to study proton abstraction from acetyl-CoA by Asp-375, model B for proton transfer between the enolate of acetyl-CoA and His-274, and model C to study both reaction steps. Model C was too large for ab initio geometry optimizations to be feasible, due to the computational requirements for force constant calculation. The accuracy of AM1 for this model is estimated by comparison to ab initio results for models A and B. The complexes are referred to by the form of the thioester involved, namely, the keto, enolate, or enol (Figure 1). All three models bear a net negative charge. The interaction of methylthioacetate with methylimidazole was also studied semiempirically to provide an indication of the change in hydrogen bond strength on going from the substrate to the intermediate. The PM3 semiempirical parametrization was used for model B to test its performance. PM3 deals more accurately than AM1 with hydrogen bonds in some cases,[34] although an error in its nitrogen parameters makes PM3 perform poorly for N−H···O systems.[35,36] PM3 also behaves erratically for N···H, H···H, and O···H intermolecular interactions, making it more susceptible to unpredictable errors than AM1.[37] PM3 was found to provide no significant advantage over AM1 because of its poor treatment of proton transfer and charge distribution in model B, and therefore AM1 only was used for models A and C.

Semiempirical TSs were located using the SADDLE linear synchronous transit-type method[38] in MOPAC 6.0,[39] followed by unconstrained refinement by gradient norm minimization using the keyword NLLSQ or SIGMA, and finally elimination of unwanted additional imaginary frequencies by minimizing along these modes in dynamic reaction coordinate (DRC) calculations.[36] Normal-mode analysis then ascertained that the TS structures were first-order saddle points (i.e., points with one imaginary frequency). The AM1 TSs proved to be good starting points for ab initio TS optimizations for models A and B.

The AM1 intrinsic reaction coordinate (IRC)[40] (i.e., the steepest descent path in mass-weighted coordinates leading from the TS to reactants in one direction and products in the other) was calculated for each reaction. Monitoring atomic motion along this pathway indicates which coordinates are involved during the reaction. The IRC should also confirm that the TS located is that for the desired reaction by leading to the correct reactants and products. Quantum mechanical tunneling through the energy barrier can make a significant contribution for proton-transfer reactions.[41,42] Modeling such reactions by locating the TS and IRC on the potential energy surface may overestimate the effective barrier to reaction, and also predict somewhat too much heavy atom movement during the reaction, because tunneling will allow proton transfer to occur before the TS configuration is reached. However, the potential surface describes the most important features of the reaction and must be studied before reaction dynamics can be taken into account.

The geometries of the reactants and products produced by IRC calculations were optimized without constraints at the semiempirical and ab initio levels. Harmonic frequencies were calculated for all complexes to verify that they were minima or TSs. Ab initio zero-point energies (ZPE) were scaled by 0.91.[43] AM1 and PM3 are parametrized to reproduce heats of formation (ΔH_f) at 298 K, and the semiempirical results should strictly be compared with ab initio enthalpies. However, a direct comparison of reaction energies can be sufficient to demonstrate shortcomings of the semiempirical results.[44] The MOPAC program, version 6.0,[39] was used for all semiempirical geometry optimizations and Gaussian-90/92[45] for ab initio calculations. Standard Gaussian and MOPAC convergence criteria were used throughout. For these large nonbonded complexes, it was not possible to reduce the RMS gradient in the MOPAC calculations to the more stringent level requested by the "PRECISE" keyword which is desirable for small molecule calculations. Atomic charges were derived from the calculated electrostatic potential using the CHELPG facility in Gaussian,[46] or with the *Rattler* program for the semiempirical wave functions.[47,48] Comparison with charges for the isolated molecules showed that the ESP fitted charges were accurately representative of charge distribution in the complexes.

Initial ab initio optimizations used the 6-31G(*d*) basis set and were then extended to the RHF/6-31+G(*d*) level. Diffuse functions can be important for calculations on anionic systems and hydrogen bonds.[49] Electron correlation must be taken into account to treat proton transfer processes correctly,[28,42] and therefore energies corrected for correlation at the second order of perturbation theory (MP2) were calculated. It was generally not possible to optimize geometries at the correlated level due to computational limitations.

Results

Model A. Proton Abstraction from Acetyl-CoA by Asp-375. TSs for abstraction of an acetyl proton from methylthioacetate (CH_3COSCH_3, representing acetyl-CoA) by acetate (CH_3COO^-, representing Asp-375) were located (**2**, Figure 3b) and confirmed as saddle points with one imaginary frequency (Table 2). The reactant keto (CH_3COO^- and CH_3COSCH_3, **1**) and product enolate (CH_3COOH and $^-CH_2COSCH_3$, **3**) complexes produced by the IRC calculation (Figure 4a) were refined initially with AM1 and then at the RHF/6-31G(*d*) and RHF/6-31+G(*d*) levels (Figure 3). The reactant complex was also optimized at the MP2/6-31G(*d*) level (Table 1). Optimized geometrical parameters are given in Figure 3 and Table 1.

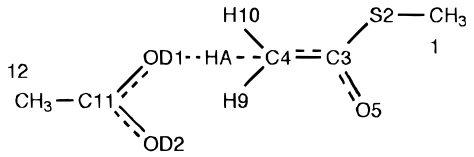

Figure 2. Atom numbering/naming scheme for the acetate (CH_3COO^-) + methylthioacetate (CH_3COSCH_3) complex **1** (keto, model A). This scheme is used for convenience and illustrative purposes only, and does not conform to standard chemical numbering schemes. Throughout, stable complexes are referred to by the form of the thioester involved (keto, enolate, or enol).

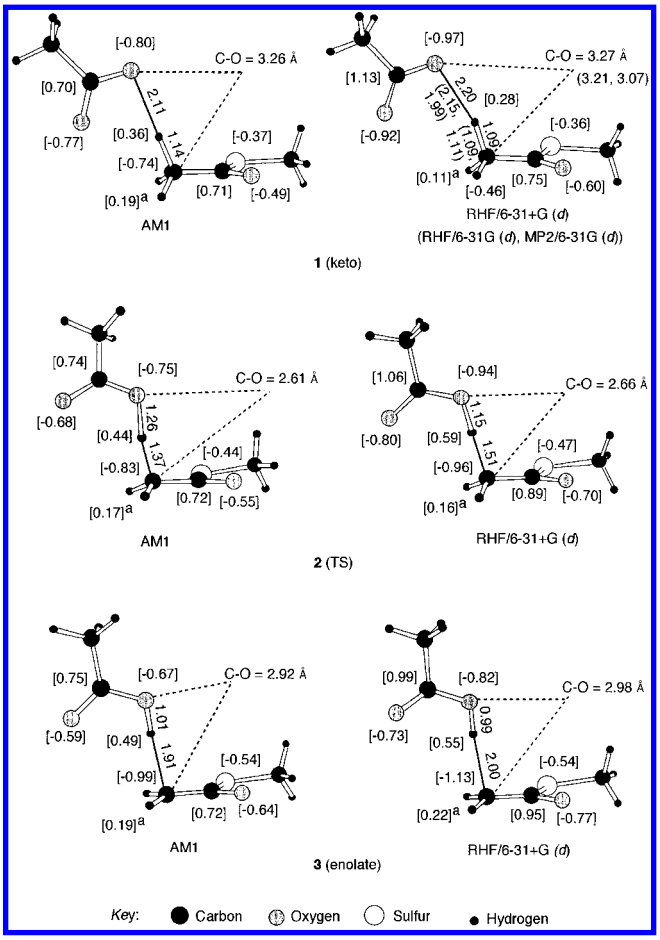

Figure 3. AM1 and RHF/6-31+G(*d*) reactant (keto, **1**), transition state **2**, and product (enolate, **3**) complexes for proton abstraction by acetate from methylthioacetate (keto → enolate reaction, model A). Interatomic separations (in Å) of the atoms involved in proton transfer are shown. RHF/6-31G(*d*) and MP2/6-31G(*d*) optimized parameters for the reactant complex (only) are also given in parentheses. Electrostatic potential fitted charges (in atomic units) of some atoms are shown in square brackets. [a] The average charge of hydrogens H9 and H10 is given.

It is immediately apparent that proton transfer from methylthioacetate to acetate is highly unfavorable (Table 3), as expected. At the RHF level, a potential barrier to the reverse reaction is predicted, but this barrier is effectively removed when ZPE is included. The TS **2** lies only 2.7 kcal/mol (RHF/6-31+G(*d*)) above the products **3**, and this difference is reduced to 0.1 kcal/mol if ZPE is included. Electron correlation significantly affects the relative energies: MP2 single point calculations give the transition state as lower in energy than the product complex. Optimization of all three stationary points at the correlated level, perhaps with a larger basis set, would be required to establish a more accurate barrier and energy change for this reaction, if the products form a stable complex and there is a TS at this level. However, it is clear that the

TABLE 1: Some Geometrical Parameters (Bond Lengths in angstroms, Angles in degrees) from Optimized Structures of Reactant (1, Keto), Product (3, Enolate) and TS 2, Complexes for the $CH_3COO^- + CH_3COSCH_3 \rightarrow CH_3COOH + {}^-CH_2COSCH_3$ Reaction (Model A)

complex	method	C4−HA	OD1−HA	C4−OD1	C4−HA−OD1
1 (keto complex	AM1	1.14	2.11	3.26	179.4
(CH_3COO^- +	RHF/6-31G(*d*)	1.09	2.15	3.21	162.6
CH_3COSCH_3)	MP2/6-31G(*d*)	1.11	1.99	3.07	164.6
	RHF/6-31+G(*d*)	1.09	2.20	3.27	164.7
2 (transition state)	AM1	1.37	1.26	2.61	168.2
	RHF/6-31G(*d*)	1.50	1.15	2.65	179.4
	RHF/6-31+G(*d*)	1.51	1.15	2.66	178.7
3 (enolate complex	AM1	1.91	1.01	2.92	172.9
(CH_3COOH +	RHF/6-31G(*d*)	1.92	1.00	2.92	177.3
${}^-CH_2COSCH_3$))	RHF/6-31+G(*d*)	2.00	0.99	2.98	175.0

complex	method	C3−C4	C3−S2	C3−O5
1 (keto complex	AM1	1.48	1.75	1.24
(CH_3COO^- +	RHF/6-31G(*d*)	1.50	1.80	1.19
CH_3COSCH_3))	MP2/6-31G(*d*)	1.49	1.81	1.22
	RHF/6-31+G(*d*)	1.50	1.80	1.19
2 transition state	AM1	1.43	1.79	1.24
	RHF/6-31G(*d*)	1.43	1.85	1.20
	RHF/6-31+G(*d*)	1.43	1.84	1.21
3 (enolate complex	AM1	1.37	1.85	1.25
(CH_3COOH +	RHF/6-31G(*d*)	1.39	1.89	1.21
${}^-CH_2COSCH_3$))	RHF/6-31+G(*d*)	1.38	1.87	1.22

TABLE 2: Imaginary Frequencies (cm^{-1}) Calculated for Transition States 2, 5, 9, and 11

	2	**5**	**9**	**11**
AM1	1089i	1367i	1204i	1291i
RHF/6-31G(*d*)	1322i	1164i		
RHF/6-31+G(*d*)	1412i	1224i		
PM3		2239i		

enolate complex **3** is very unstable relative to the reactants (**1**) and is of comparable energy to the TS **2** at all the ab initio levels. The ZPE difference between the products and the TS is comparable to the barrier to return to reactants, indicating that the product complex, if formed, will not exist for longer than approximately one molecular vibration. Only if the enolate is stabilized, and therefore the barrier to the reverse reaction raised, will it have a sufficiently long lifetime to participate as an intermediate in the reaction. Otherwise, a mechanism with deprotonation concerted with nucleophilic attack on oxaloacetate would probably be enforced.[32] A thioester enolate can exist for approximately 10^{-9}−10^{-10} s associated with a cationic acid as an ion pair in solution,[50] but the lifetime of the enolate when bound at the active site in close proximity to a carboxylic acid, in the correct orientation for the reverse reaction and without intervening solvent, would probably be smaller than this in the absence of specific stabilization.

The energy difference between the keto **1** and enolate **3** complexes may be estimated as between 12 kcal/mol (including ZPE) and 17 kcal/mol from the best ab initio results here. The energy barrier to reaction is approximately equal to this difference. AM1 calculates a comparable, somewhat smaller barrier (11.9 kcal/mol), but overestimates the stability of the products (giving an energy difference of only 7.7 kcal/mol between **1** and **3**).

The distance between the exchanging heavy atoms (C4 and OD1) is markedly reduced at the TS **2**, with the C, H, and O atoms approximately collinear, and is shorter in **3** than in **1**. The geometries produced with and without diffuse functions are similar, in particular for the TS **2**. The effect of MP2/6-31G(*d*) optimization of **1** is to reduce the C4−OD1 distance by 0.2 Å. From the AM1 IRC (Figure 4a), it is apparent that reaction occurs by initial approach of the exchanging heavy atoms toward the TS separation (approximately 2.6 Å), with the proton remaining a constant distance from its initial bonding partner. In the region of the TS, however, the carbon−oxygen distance remains almost constant, and movement along the IRC

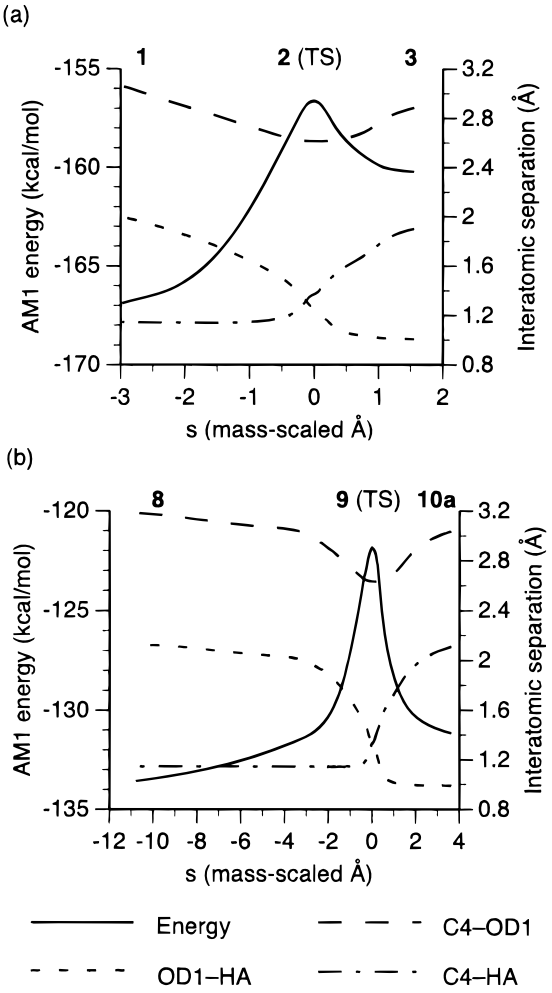

Figure 4. (a) AM1 energy (ΔH_f, kcal/mol) profile (solid line) and interatomic separations (dotted/dashed lines) along the IRC for proton abstraction from methylthioacetate by acetate (keto (**1**) → enolate (**3**) reaction, model A). The distance along the IRC *s* is given in mass-scaled units ((au)$^{1/2}$ Å). The TS is located at $s = 0$. Movement toward negative *s* leads to the reactants (CH_3COO^- + CH_3COSCH_3). **(b)** Energy profile and variation of interatomic distances along the AM1 reaction coordinate[a] for the model C keto (**8**) → enolate (**10a**) reaction. [a]In the direction of negative *s* ($s < 0$), the damped DRC is shown; for $s > 0$, the IRC is shown (see text).

TABLE 3: Energy Barriers and (Product Complex−Reactant Complex) Energy Differences (kcal/mol) for the Reaction $CH_3COO^- + CH_3COSCH_3 \rightarrow CH_3COOH + {}^-CH_2COSCH_3$ (Keto 1 → Enolate 3 Reaction, Model A)[a]

level	keto **1** → TS **2**	keto **1** → enolate **3**
AM1	11.9	7.7
RHF/6-31G (*d*)	20.0 (17.2)[b]	18.1 (17.8)[b]
MP2/6-31G(*d*)//6-31G(*d*)	12.4 (9.6)[b]	13.4 (13.1)[b]
MP2/6-31G(*d*), MP2//RHF[a]	16.2 (13.4)[b]	17.2 (16.9)[b]
RHF/6-31+G(*d*)	22.6 (19.8)[b]	19.9 (19.7)[b]
MP2/6-31+G(*d*)//6-31+G(*d*)	14.8 (12.0)[b]	17.0 (16.8)[b]

[a] The energy of the reactant (keto) complex optimized at the MP2/6-31G(*d*) level is used, with single point MP2/6-31G(*d*)//6-31G(*d*) energies used for the transition state and product (enolate) complexes. [b]Energies in parentheses are corrected for zero-point energy differences.

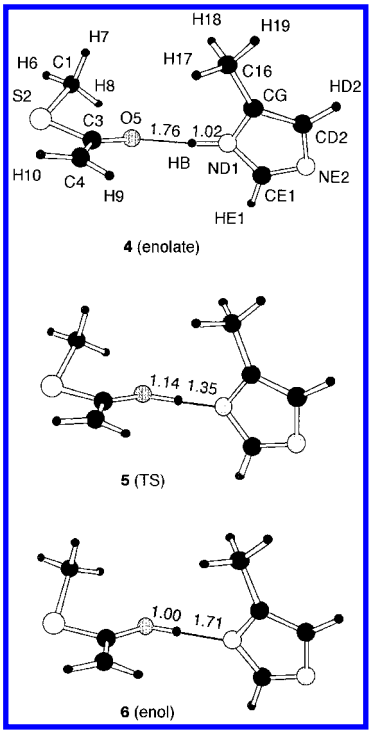

Figure 5. Model B enolate (**4**), TS (**5**), and enol (**6**) complexes for proton transfer from 5-methylimidazole (representing His-274) to the enolate of methylthioacetate ($^{-}CH_2COSCH_3$, representing the enolate of acetyl-CoA) to form the enol ($CH_2COHSCH_3$). RHF/6-31+G(*d*) geometries are shown, with O−H and N−H distances (angstroms) indicated. The numbering scheme used is shown in **4**.

corresponds to motion of the proton. Similar behavior was found for model B (below) and for proton transfer between $CH_3^-$and CH_4.[51]

AM1 and ab initio geometries for **1**, **2**, and **3** are in satisfactory agreement. The distance between the carbon and oxygen atoms which exchange the proton is predicted well by AM1. The AM1 TS is however earlier (closer to the reactants); the breaking C−H bond is shorter and the forming O−H bond too long, each by approximately 0.1 Å. This may be related to the incorrectly small energy difference between the reactant and product complexes given by AM1. All methods show the expected change in geometry at C4 from tetrahedral CH_3R toward planar CH_2R as a proton is removed. The changes in AM1 and RHF/6-31+G(*d*) charge distributions during the reaction are in good qualitative agreement (Figure 3). AM1 potential fitted charges tend to be somewhat smaller than standard ab initio results.[48] Negative charge on S2 is seen to increase in the enolate, indicating that the sulfur atom may be important in stabilizing this form by charge acceptance.

Overall, these results indicate that Asp-375 is well positioned to abstract a proton from acetyl-CoA, because the TS geometry is similar to the orientation of these groups in the enzyme complex.[12] AM1 results[7] also suggest that the TS geometry may be fairly flexible, with the reaction possible over a small range of orientations. However, the ab initio energies show that the products of this reaction are at best extremely unstable with respect to the reactants, and only if the enolate is stabilized by the enzyme can it be an intermediate.

Model B. Interaction of His-274 with Acetyl-CoA and the Enolate → Enol Reaction. The TS (**5**, Figure 5) for proton transfer from 5-methylimidazole to the thioester enolate ($^{-}CH_2COSCH_3$) to form the enol ($CH_2COHSCH_3$) was optimized at the AM1, PM3, RHF/6-31G(*d*), and RHF/6-31+G(*d*) levels, as were the reactant (enolate) and product (enol) complexes (**4** and **6**, respectively, Figure 5, Table 4). These structures are of particular interest as representing the proposed hydrogen bonded intermediate in CS.

TABLE 4: Selected Geometrical Parameters for (Model B) Complexes 4 (Enolate), 5 (TS), and 6 (Enol), Optimized at ab Initio and Semiempirical Levels[a]

complex	parameter	AM1	PM3	RHF/ 6-31G(*d*)	RHF/ 6-31+G(*d*)
enolate **4**	O5−ND1	2.93	2.75	2.77	2.78
	O5−HB	1.93	1.73	1.75	1.76
	ND1−HB	1.01	1.02	1.02	1.02
	O5−HB−ND1	170.6	174.1	175.8	175.4
	C3−O5−ND1	109.9	117.5	119.6	131.6
TS **5**	O5−ND1	2.52	2.49	2.48	2.49
	O5−HB	1.22	1.24	1.14	1.14
	ND1−HB	1.31	1.26	1.35	1.35
	O5−HB−ND1	176.3	173.6	176.5	176.0
	C3−O5−ND1	113.6	118.4	117.9	119.6
enol **6**	O5−ND1	2.86	2.68	2.69	2.71
	O5−HB	1.00	0.99	1.00	1.00
	ND1−HB	1.86	1.69	1.69	1.71
	O5−HB−ND1	178.1	172.9	178.3	177.7
	C3−O5−ND1	110.9	115.5	113.5	114.9

[a] Interatomic distances in angstroms, angles in degrees.

The ab initio calculations show the enol/methylimidazolate complex **6** to be highly unstable relative to the enolate/methylimidazole complex **4**. As for model A, the TS **5** on the ab initio RHF potential energy surface for proton transfer is actually lower in energy than the products **6** when zero-point energy is taken into account or correlation corrections at the MP2 level are applied. As is often found, the effects of electron correlation are particularly marked for the TS, reducing the **4** → **5** energy barrier by 6.1 kcal/mol to 5.0 kcal/mol (MP2/6-31+G(*d*)//6-31+G(*d*)). When zero-point energy is included, the TS **5** lies only 2.2 kcal/mol above **4** (MP2/6-31+G(*d*)//6-31+G(*d*)). This hints that, given sufficient specific stabilization of the methylimidazolate, a shared-proton hydrogen bond structure could become energetically accessible. It is known that the energy barrier to proton transfer within a hydrogen bond drops as the difference in basicity between the partners decreases and the strength of the bond increases.[52] To obtain accurate geometries and energies for proton transfer it would be necessary to optimize the complexes at higher levels, and anharmonicity corrections may be necessary.[53] Results at all levels here show, however, the conventional hydrogen bond complex with the proton localized on methylimidazole to be the most stable.

The hydrogen bond distances for the enolate **4** and enol **6** complexes here are fairly short (N−O = 2.78 and 2.71 Å, respectively, 6-31+G(*d*)), but outside the range expected for low-barrier hydrogen bonds (<∼2.6 Å[18,19]). The separation of the hydrogen bonded heavy atoms is a minimum at **5** (2.49 Å, 6-31+G(*d*)), which represents the configuration for proton transfer. Conclusions about the nature of the bond should not be drawn from the calculated geometries alone, however, as inclusion of electron correlation is generally found to shorten the distance between hydrogen bonded atoms.[34] The RHF ab initio values given here are probably somewhat too long for the stable complexes at least. Correlated optimization would be required to establish more accurate geometries and energies for proton transfer. However, it is apparent that the enol complex **6** is significantly higher in energy than the enolate **4**. The basicities of the hydrogen bonded partners are very different. A requirement for formation of a low-barrier hydrogen bond is that the pK_a's of the bonded molecules be approximately equal. In agreement with experimental measurements for model compounds (see Discussion below), these

TABLE 5: Calculated Energy Barriers and Energy Changes for the Reaction 4 → 6 (Enolate → Enol Reaction, Model B)

level	enolate **4** → TS **5**	enolate **4** → enol **6**
AM1	13.4	8.4
PM3	15.3	4.6
RHF/6-31G(*d*)	10.2 (7.3)[a]	8.6 (8.5)[a]
MP2/6-31G(*d*)//6-31G(*d*)	5.2 (2.3)[a]	6.9 (6.8)[a]
RHF/6-31+G(*d*)	11.1 (8.3)[a]	9.4 (9.5)[a]
MP2/6-31+G(*d*)//6-31+G(*d*)	5.0 (2.2)[a]	6.8 (6.9)[a]

[a] Figures in parentheses are ab initio total energy differences corrected for zero-point energy differences.

TABLE 6: Charge Transfer (au) Calculated from Atomic Charges for Model B Complexes[a]

complex	AM1	PM3	RHF/6-31G(*d*)	RHF/6-31+G(*d*)
4	0.10	0.12	0.10	0.09
6	0.13	0.25	0.18	0.18
7	0.05	0.10		

[a] In each case negative charge is transferred to the hydrogen bond donor. Charge-transfer calculated as (total charge of acceptor) − (total charge of complex) = −(total donor charge).

TABLE 7: Interaction Energies and Enthalpies, kcal/mol (Calculated as $-\Delta E$ or $-\Delta H$ of Complex Formation) for Model B Complexes[a]

	4	**6**	**7**
AM1	15.9	8.4	3.6
PM3	19.2	17.3	2.3
RHF/6-31G(*d*)	23.5 ($-\Delta H = 21.9$)		
MP2/6-31G(*d*)//6-31G(*d*)	28.4 ($-\Delta H = 26.8$)		
RHF/6-31+G(*d*)	21.6 ($-\Delta H = 20.1$)		
MP2/6-31+G(*d*)//6-31+G(*d*)	26.6 ($-\Delta H = 25.1$)		

[a] For AM1 and PM3, the semiempirical interaction enthalpy is given, and for the *ab initio* calculations the interaction energies are listed first, with approximate enthalpies (calculated using RHF harmonic frequencies by standard methods[56]) given in parentheses.

calculations indicate that, intrinsically, this is not the case for methylimidazole and a thioester enol. Therefore in CS, the hydrogen bond between His-274 and the enolate of acetyl-CoA will not be of the low-barrier type, unless the effective pK_a for deprotonation of neutral His-274 is lowered significantly at the active site.

Despite the basicity imbalance, the hydrogen bond between the enolate and methylimidazole in **4** is strong (interaction energy ($-\Delta E$ for complex formation) of 21.6 kcal/mol (6-31+G(*d*)) and 26.6 kcal/mol (MP2/6-31+G(*d*)//6-31+G(*d*)), Table 7). These values do not contain corrections for basis set superposition error (BSSE), which is not straightforward to correct for, particularly when diffuse functions are used.[49] Test calculations including the counterpoise correction[54] to estimate BSSE gave interaction energies which were smaller in magnitude by 0.03 kcal/mol (6-31+G(*d*)) to 1.65 kcal/mol (6-31G(*d*)). Ab initio interaction energies were calculated for **4** only, being the only stable model B complex at this level. The interaction energies may be considered to be due in large part to the charged hydrogen bond and are comparable to experimental results for other charged hydrogen bonded systems in the gas phase.[55] Zheng and Merz[34] found for several anionic hydrogen bonded complexes that MP2/6-31+G(*d*) and RHF/6-31+G(*d*) interaction energies bracketed the experimental results, with the MP2/6-31+G(*d*) energies too large and the RHF/6-31+G(*d*) energies smaller in magnitude. The interaction energy in the present case is likely to lie between the MP2 and RHF results. Approximate complexation enthalpies at 298 K[56] are slightly smaller ($-\Delta H = 20.1-25.1$ kcal/mol); the semiempirical results should be compared with these values.

AM1 gives an interaction energy for **4** (15.9 kcal/mol) which is too small, in line with previous findings for strongly hydrogen bonded systems,[34,57] particularly for **6**, due to excessive core−core repulsion for nitrogen as a hydrogen bond acceptor.[35] PM3 performs better, giving an interaction energy (19.2 kcal/mol) for **4** closer to the ab initio findings. However, PM3 performs incorrectly for other aspects of this model (see below). AM1 hydrogen bond lengths for the enolate **4** and enol (**6**) complexes are too long whereas the PM3 hydrogen bond lengths are smaller than ab initio values. AM1 and PM3 give O−N separations at the TS **5** which are close to the ab initio results but resemble the enolate complex **4** too closely (i.e., they are too early). The PM3 TS imaginary frequency is considerably larger than the ab initio and AM1 results (Table 2). AM1 and PM3 calculate energy barriers for the conversion of **4** to **6** which are clearly too high and predict a stable complex for the enol **6**. PM3 predicts **6** to be too stable (by 2.3 kcal/mol), whereas the AM1 energy difference between **4** and **6** is 1.5 kcal/mol larger than the best ab initio result.

Charge transfer in complexes **4**−**6** (Table 6) was calculated by summing potential-fitted atomic charges for the donor or acceptor molecules. This is of interest partly because of the suggestion that charge delocalization from the enolate of acetyl-CoA to His-274 via the hydrogen bond may be important in stabilizing the intermediate.[6] The total charge of complexes **4**−**6** is −1 au. AM1 and ab initio atomic charges for **4**−**6** are qualitatively similar. In each case, negative charge is transferred to the hydrogen bond donor. PM3 gives an incorrect pattern of charges for the hydrogen bonding N−H group in **4** (and in **7**), giving a sizable positive charge for ND1 and a zero charge for HB. Similar PM3 errors for N−H groups have been found previously.[35] An interesting observation is that, according to AM1 and the ab initio calculations, there is a balance of charges at the TS **5**. At the TS, the charges of the enolate and methylimidazolate fragments which exchange the proton are equal (excluding the charge of the exchanged proton). Summing the atomic charges for these molecules shows the same total charge (−0.61, AM1; −0.68, 6-31G(*d*); −0.71, 6-31+G(*d*)) on each. PM3 does not show this charge balance, instead it assigns a charge of zero to ND1 at the TS.

For comparison with the calculations on the largest model, model C (below), **7** (5-methylimidazole and CH_3COSCH_3, (Figure 7)) was studied as a model of the binding of the keto form of acetyl-CoA to His-274. It has been suggested that the differential binding and stabilization of unstable intermediates is a general principle of enzyme catalysis.[58] Comparing the properties of complex **7** with those of **4** and **6** gives an estimate of the difference in hydrogen bond strength between the neutral and charged complexes. Optimization of **7**, unlike the other complexes described here, resulted in some drifting of the two molecules with their final relative orientations altered noticeably. For this reason, ab initio optimizations of **7** were not performed, and the semiempirical results for **7** should only be taken as a first approximation. The trimolecular keto complex **8** is a better model of substrate binding. The interaction energy in the enolate complex **4** is considerably larger than the neutral complex **7** (by 12.3 kcal/mol, AM1; 16.9 kcal/mol, PM3; Table 7). This provides an indication that the increased strength of the hydrogen bond with His-274 caused by deprotonation of acetyl-CoA can stabilize the intermediate. However, as shown by the calculations on the largest model (model C, below), it is important to treat the reacting system as a whole, rather than just the hydrogen bonded pair, to evaluate the contribution of the hydrogen bond.

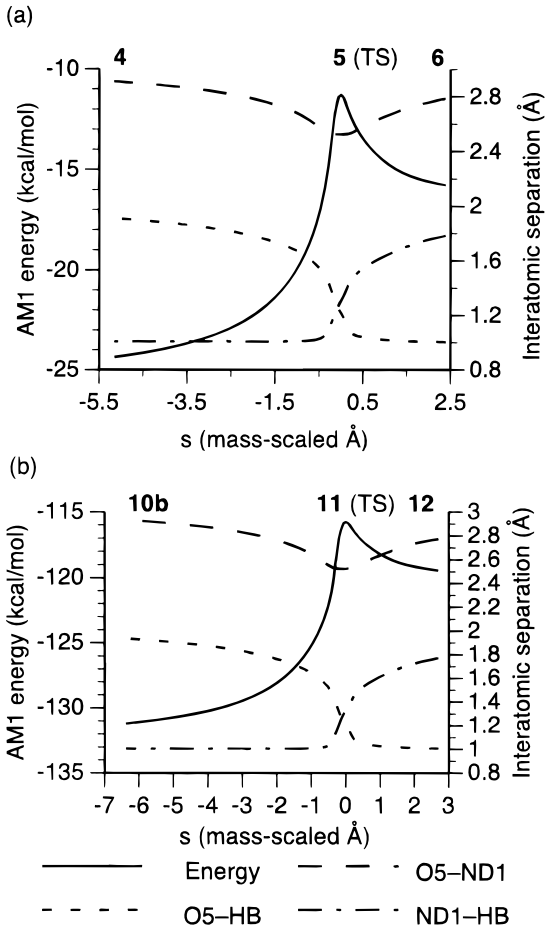

Figure 6. (a) AM1 IRC for the model B proton-transfer reaction 5-methylimidazole + $^{-}CH_2COSCH_3$ (**4**) → (4)-methylimidazolate + $CH_2COHSCH_3$ (**6**) (enolate → enol reaction). The AM1 energy (kcal/mol) is plotted on the left-hand axis, interatomic separations on the right. (b) Energy profile and variation in interatomic distances along the AM1 IRC for the enolate **10b** → enol **12** reaction (model C).

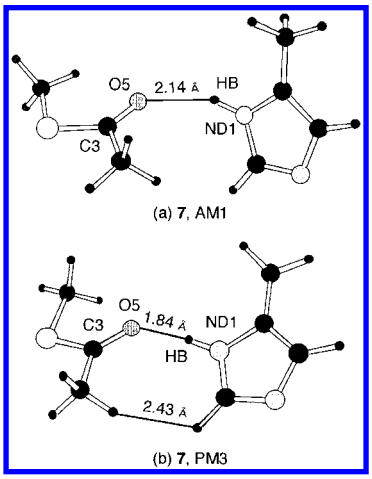

Figure 7. (a) AM1 and (b) PM3 optimized geometries of complex **7** (5-methylimidazole and methylthioacetate), representing the interaction of acetyl-CoA (keto form) with His-274.

To summarize, AM1 calculates structural, energetic and electronic properties of complexes **4**−**6** in reasonable agreement with ab initio results, although the barrier to the reaction is too high. It must also be remembered that the interaction energy for the enolate with methylimidazole in **4** given by AM1 is 4.2−9.2 kcal/mol too small compared to (RHF or MP2) 6-31+G(*d*) findings. PM3 gives a better hydrogen bond energy for the enolate complex **4** and appears more reasonable for **6**, but calculates a small interaction energy for the keto complex **7**. The PM3 charge distribution of the model B complexes is incorrect, and the TS, barrier, and energy change for proton transfer are relatively inaccurate. For this system PM3 does not provide a distinct advantage over AM1 and in some respects performs more poorly, and so AM1 was selected for calculations on model C.

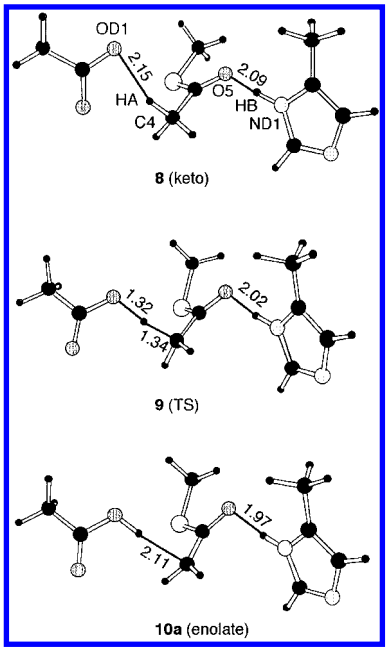

Figure 8. AM1 optimized geometries of reactant (keto, **8**), TS **9**, and product (enolate, **10a**) model C complexes for proton abstraction from methylthioacetate (CH_3COSCH_3) by acetate, to form the enolate $^{-}CH_2COSCH_3$ (keto → enolate reaction). Some interatomic distances (angstroms), and atom names (using the same naming scheme as for models A and B) are indicated.

Model C. Asp-375, His-274, and Acetyl-CoA. Model C (Figure 8) consists of acetate, 5-methylimidazole, and methylthioacetate, representing Asp-375, His-274, and acetyl-CoA, respectively, and was used to study the entire enolization reaction. The large size of this model made ab initio optimizations impractical, and so it was studied at the AM1 level only. The accuracy of the results can be judged by comparison with the ab initio results for the smaller models above. The keto **8** → enolate **10** reaction should be compared to model A (**1** → **3**) and the enolate **10** → enol **12** reaction with model B (**4** → **6**).

Attempts to locate a TS for concerted enol formation produced instead complex **11** (Figure 9), the TS for the enolate → enol reaction. The TS for the keto → enolate reaction, **9** (Figure 8), was also located, independently. The reaction coordinate calculated from **9**, the TS for proton transfer between acetate and the thioester, is plotted in Figure 4b. In the direction of positive *s* in this diagram, a standard IRC calculation led to **10a**, the enolate complex. It proved necessary to use a slightly different method to follow the reaction coordinate in the opposite direction (i.e., toward negative *s* in Figure 4b). The initial displacement in the IRC calculation in this direction was too small to prevent subsequent optimization taking the system toward the same product complex **10a**. A dynamic reaction coordinate calculation (DRC) was performed instead,[36] damped (2 fs half-life for kinetic energy) to ensure that this pathway was as similar as possible to the IRC steepest descent path. This produced **8**, the complex of the keto form of the thioester. The geometries of **8** and **10a** were optimized using AM1, and frequency calculations showed that they were minima. That **9**

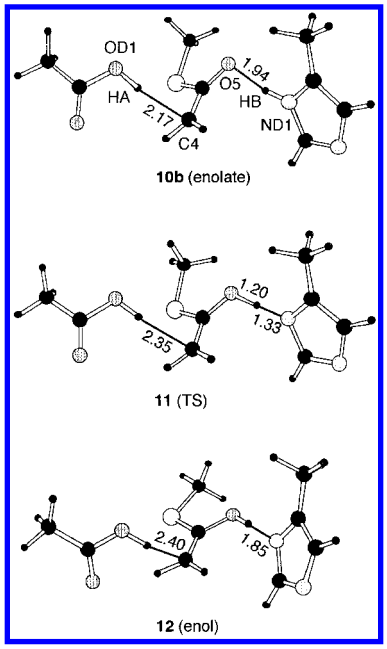

Figure 9. AM1 optimized geometries of reactant (enolate, **10b**), TS **11**, and product (enol, **12**) model C complexes for proton transfer from methylimidazole to the enolate of methylthioacetate (enolate → enol reaction). Some interatomic distances (angstroms) are shown.

TABLE 8: Selected Geometrical Parameters (Interatomic Distances in angstroms) for AM1 Optimized Model C Complexes[a]

complex	OD1−HA	C4−HA	C4−OD1	O5−HB	ND1−HB	ND1−O5
8 (keto)	2.15	1.14	3.22	2.09	0.99	3.01
9 (TS)	1.32	1.34	2.63	2.02	1.00	3.00
10a (enolate)	0.99	2.11	3.04	1.97	1.00	2.96
10b (enolate)	0.99	2.17	3.09	1.94	1.01	2.93
11 (TS)	0.98	2.35	3.20	1.20	1.33	2.53
12 (enol)	0.98	2.40	3.23	1.00	1.85	2.85

[a] The form of the thioester (methylthioacetate, CH_3COSCH_3) in each complex is indicated.

is the TS associated with the keto−enolate reaction was verified by inspecting the eigenvector for movement along the reaction coordinate; the correct atoms made the dominant contributions. IRC calculations (Figure 6(b)) on TS **11** followed by refinement (and characterization) produced the enolate complex **10b** and the enol complex **12**.

The structures of the model C complexes are generally similar to those of the smaller models A and B. For example, the model C keto **8** and enolate **10a** complexes show interatomic separations for the reacting fragments close to those found for the equivalent AM1 model A structures **1** and **3** (Tables 1 and 8). The thioester oxygen−imidazole nitrogen hydrogen bond (O5−ND1) length in **10a** is 0.12 Å shorter than in **8**, suggesting that the enolate forms a stronger hydrogen bond with 5-methylimidazole than does the keto form. The effect of the hydrogen bond on the TS **9** for the keto → enolate reaction is to make it earlier (i.e., more similar to the acetate + thioester reactants). 5-Methylimidazole stabilizes the enolate (product) relative to the keto form, and so the observed change in TS structure is in accord with the Hammond postulate.[59] For the enolate→ enol reaction, the distances between the reacting atoms (Table 8) are the same (in **10b**, **11**, and **12**), to within 0.02 Å, as the equivalent distances in the model B complexes **4**−**6**. The structure of the TS for this process is almost unchanged.

The enolate complexes **10a** and **10b**, produced by independent IRC calculations, have very similar structures and virtually identical energies. This shows that the model C enolate complex may be converted into either the keto complex **8** or to the enol complex **12**, and therefore the energies of all these complexes may be compared directly. The molecules in each of the complexes remain close to the crystallographic orientations of the groups they represent and do not interact in ways which would be impossible within the confines of the active site, so may be taken as relevant to the reaction occurring in the enzyme. The keto complex **8** is the most stable, with the enolate **10** 2.4 kcal/mol higher and the enol complex **12** considerably (12.5 kcal/mol) higher in energy. Significant barriers are found for both reaction steps. On the basis of the comparisons to the ab initio calculations on the smaller models above, the accuracy of these AM1 results can be estimated roughly, to the extent that nonspecific interactions and three-body effects are treated accurately. AM1 overestimates the stability of the acetic acid and thioester enolate products of the keto → enolate reaction (by approximately 9.1 kcal/mol, Table 3). A fortuitous compensating error for model C is the low AM1 hydrogen bond energy for the enolate and methylimidazole (too small by 4.2−9.2 kcal/mol, Table 7). The strength of the hydrogen bond between the neutral thioester and methylimidazole given by AM1 (3.6 kcal/mol) is probably too small by about 3 kcal/mol, based on results for neutral N−H···O hydrogen bonded complexes.[34,60,61] The error in the AM1 energy change for the keto **8** → enolate **10** reaction can therefore be estimated as 2.9−7.9 kcal/mol, with the enolate complex too stable. For the enolate → enol reaction, the presence of acetic acid does not stabilize the model C enol complex **12** relative to the enolate **10** and in fact disfavors enol formation. The energy difference between the enolate and enol complexes is increased from 8.4 kcal/mol (model B) to 10.1 kcal/mol for model C, and may be somewhat too large (by approximately 1.5 kcal/mol, Table 5). The interactions of acetic acid have not been studied separately here, but given the likelihood that the energy of its interaction with the enolate is underestimated more by AM1 than that with the neutral enol, it is probable that more sophisticated calculations on model C would also predict the enolate complex to be more stable than the enol. Taking into account all these factors and those noted above, the AM1 energy order of keto (most stable) < enolate < enol (least stable) appears qualitatively correct, a conclusion supported by comparison of AM1 and ab initio QM/MM calculations on the reaction.[7,61]

The AM1 barrier to the model C keto → enolate reaction is 11.7 kcal/mol and that for the enolate → enol reaction 15.4 kcal/mol. This latter barrier is certainly too high, as shown by the calculations on model B (Table 5). It is difficult to assess the

TABLE 9: AM1 Energy Changes and Barriers for Model C Keto 8 → Enolate 10a and Enolate 10b → Enol 12 Reactions

process	description	energy (kcal/mol)
8 → 9	energy barrier to keto → enolate reaction	11.7
8 → 10a	energy difference between enolate and keto complexes	2.4
10b → 11	energy barrier to enolate → enol reaction	15.4
10b → 12	energy difference between enol and enolate complexes	10.1
8 → 12	energy difference between enol and keto complexes	12.5

accuracy of the barrier for the keto → enolate reaction because a barrier for this reaction could not be clearly differentiated from the energy change for the reaction for model A. The model C AM1 result of 11.7 kcal/mol for **8** → **9** is scarcely less than that for model A (**1**→ **2**) (11.9 kcal/mol). No definite conclusion can be drawn about the barrier to reaction because of the unreliability of semiempirical methods for activation barriers and because AM1 gives too low a hydrogen bond energy in the enolate complex. The AM1 barrier is comparable to the large "intrinsic" barriers for deprotonation of carbon acids in solution of 10−13 kcal/mol[19] and is consistent with the apparent experimental activation energy ($\Delta G^{\ddagger}$) for the CS reaction of 14.9 kcal/mol,[9] for which deprotonation of acetyl-CoA is thought to be rate-limiting.[11] In this context the AM1 result appears reasonable. The AM1 imaginary frequency for the keto **8** → enolate **10** reaction, 1204i cm^{-1}, is large enough to suggest that tunneling could be important,[62] but a more accurate value must await correlated optimization.

The stabilization of the enolate by the hydrogen bond with methylimidazole can be seen by comparing the reaction coordinate plots (Figure 4a and b) for models A and C. The contribution of this interaction to stabilizing the enolate (relative to the keto form) calculated by AM1 is 5.3 kcal/mol (from the change in the enolate−keto energy difference between model A (**3**−**1** = 7.7 kcal/mol) and model C (**10**-**8** =2.4 kcal/mol)). This is an underestimate because of the erroneously small hydrogen AM1 bond energy for the enolate. Two conclusions can be drawn, however. First, the enolate is stabilized because it forms a stronger hydrogen bond with 5-methylimidazole than does the keto form, as shown by the interaction energies calculated for these molecules in model B. Similarly, the increased strength of the hydrogen bond between His-274 and the enolate of acetyl-CoA compared to the bond formed with the neutral substrate is likely to play an important part in stabilizing the intermediate in CS. Hydrogen bond energies are strongly affected by environmental effects,[19,29,60] and for comparison with experimental findings for mutant enzymes lacking His-274,[6,9] it would be necessary to evaluate the free energy changes on binding and reaction caused by the mutation. As discussed below, however, it appears that the hydrogen bond can undergo a large increase in strength on deprotonation of acetyl-CoA at the buried active site and is a good means of stabilizing the enolate. Second, it is important to note that the stabilization provided by this hydrogen bond in the reacting system (i.e., in model C, including acetate/acetic acid) is less than the increase in hydrogen bond strength calculated for bimolecular complexes of the thioester with methylimidazole (**7** and **4**) on going from the neutral keto to the charged enolate form. That is, the actual stabilization of the enolate relative to the keto form is affected by interactions with the third component of the reacting system. For this reason, the stabilization of the enolate relative to the keto by methylimidazole in the reacting system (5.3 kcal/mol, AM1) is significantly less than the difference in hydrogen bond strength between the keto and enolate complexes **7** and **4** (12.3 kcal/mol, AM1). This also affects higher level calculations,[61] and it must be stressed that contributions to stabilizing the enolate in the reaction cannot be deduced from bimolecular complexation energies alone. The whole reacting system should be considered.

Discussion

A crucial problem in the study of enzyme catalysis is the explanation of how unstable species are formed and stabilized. From observed activation energies, it is clear that many enzymes are capable of significantly stabilizing intermediates and TSs relative to bound substrates, compared to equivalent processes in solution,[1,63] and yet in many cases the processes responsible remain uncertain. The CS reaction proceeds rapidly, using only amino acid side chains to catalyze what is an apparently difficult reaction. The enzyme must deprotonate a weakly acidic thioester, form a carbon−carbon bond and finally hydrolyze the thioester bond, avoiding premature hydrolysis. This is achieved through a conformational change which places the two substrates in close proximity to each other, orienting binding and catalytic residues around them and excluding bulk solvent from the buried active site, as crystallographic studies have shown.[64] The details of the chemical mechanism of CS have proved difficult to establish, however, with several possibilities put forward.

The identity of the base involved in deprotonation of acetyl-CoA, for example, has been somewhat uncertain. Asp-375 has been considered the most likely base,[9,10] but it has been proposed that His-274 or another group performs this role.[5,65] The present results indicate that Asp-375 can act as the base. It is unlikely that His-274 could function in this way and there are no other obvious candidates.[7] The results also show the necessity of stabilizing the intermediate produced by deprotonation of acetyl-CoA: the enolate complex is otherwise highly unstable relative to the reactants, with a vanishingly small lifetime. How this stabilization is achieved, and the exact nature of the nucleophilic intermediate, are central to the mechanism. Given that deprotonation of acetyl-CoA is probably not concerted with nucleophilic attack on oxaloacetate,[17,66] a crucial question has been whether the enolate or the enol is the likely nucleophilic intermediate. The difficulty with the enolate of acetyl-CoA as an intermediate is its apparent instability. Consideration of the pK_a's of acetyl-CoA (approximately 20.4−21.5 in aqueous solution[50]) and Asp-375 (approximately 6.5[67]) shows that the products of the proton abstraction reaction (the thioester enolate and protonated Asp-375) are highly unstable relative to the reactants ($\Delta G^{\circ} = 20$ kcal/mol). The observed activation energy of the reaction is only 14.9 kcal/mol ($k_{cat} = 96\ s^{-1}$ for pig CS[9]), and so the enolate must be stabilized significantly at the active site if it is to be an intermediate in the reaction. However, the finding that His-274 is neutral[7,16] (and so a weak acid, pK_a = 14[68]) means that formation of the enol offers no thermodynamic advantage.[7]

It has been proposed that the requisite stabilization is achieved through a low-barrier hydrogen bond between His-274 and the intermediate,[18,19] with the pK_a of His-274 equal to that of the enol due to conditions in the active site, and the proton shared between them. Such a bond is suggested to be of very much higher energy than a normal hydrogen bond and therefore to stabilize the intermediate (by up to 10−20 kcal/mol[18]). For the models of the intermediate hydrogen bonded to His-274 studied here, only the enolate is stable (at the ab initio level). It is highly energetically unfavorable to transfer the proton from the imidazole ring to form the thioester enol. The deprotonation energies of the enol and methylimidazole are clearly different. Interactions at the active site could lower the effective pK_a of the His-274 side chain (for example a conserved serine residue, Ser-244, donates a hydrogen bond to the imidazole[7,69]). However, ab initio[61] and semiempirical[7] QM/MM calculations which include the effects of the protein environment show that proton transfer from His-274 is highly disfavored and that the enolate is a more likely reaction intermediate than the enol or a proton-sharing enolic form. It therefore appears that the effective pK_a's of the intermediate and His-274 are not equal, and so the

hydrogen bond between them will not be of the low-barrier type. In a strongly hydrogen bonded complex, some degree of proton delocalization and transfer will occur, and the basicities of the bonded partners may be made more equal by active site conditions, but the indications are that the complex with the proton localized on His-274 and the enolate acting as the hydrogen bond acceptor is favored. Analysis of the related triosephosphate isomerase reaction has shown that a neutral histidine residue plays a similar role in stabilizing a charged intermediate via a conventional hydrogen bond.[69,70,71]

It has been suggested that the intermediate in CS has enol character compared to malate synthase for which the enolate predominates (based on gas phase conformational energies and stereochemical preferences for reaction of the fluoro substituted species).[72] The intermediate cannot be regarded simply as the enol, as no group is in a position to act as an acid and protonate the carbonyl oxygen completely. The stereochemical behavior may instead be accounted for by the hydrogen bonded enolate (in CS) exhibiting conformational preferences similar to the enol, whereas in malate synthase electrostatic stabilization of the enolate by a metal ion[73] (malate synthase requires Mg^{2+} [66]) may mean that hydrogen bonding is not as important, consistent with a lower degree of protonation of the intermediate. Alternatively other interactions with substrate or active site residues could affect the conformational behavior of fluoroacetyl-CoA. In any case, for both enzymes the enolate character of the intermediate is likely to be important for the subsequent nucleophilic attack to proceed efficiently.[74]

The findings here indicate that the hydrogen bond with His-274 can make an important contribution to stabilizing the enolate intermediate. In isolation, methylimidazole interacts strongly with the thioester enolate ($\Delta E = -21.6$ kcal/mol, 6-31+G(*d*); -26.6 kcal/mol, MP2/6-31+G(*d*)//6-31+G(*d*) for **4**) in a conventionally hydrogen bonded complex with the proton localized on the methylimidazole. The hydrogen bond does not need to be of the low-barrier type for the interaction to be highly energetically favorable. Calculations and experimental measurements have found similarly large hydrogen bond energies for other charged complexes without pK_a balance,[34,55] and although the strongest interactions have small or nonexistent proton-transfer barriers, there is no special stabilization associated with disappearance of the barrier or at $\Delta pK_a = 0$.[27,28] Unusually short hydrogen bonds have been observed in CS/inhibitor complexes, but they do not lead to exceptionally high binding affinities.[26] It appears, therefore, that a short hydrogen bond could form with acetyl-CoA within CS, but that it is not of vital importance for the pK_a of His-274 to match that of the intermediate to stabilize it significantly. QM/MM calculations on the reaction in the enzyme indicate that other interactions also significantly affect the reaction energetics. For example, a water molecule may play an important part in stabilizing the enolate by hydrogen bonding.[7,61]

Based on calculations treating CS as a low-dielectric continuum, it has been proposed that the enolate and TS can be stabilized relative to the ground state by charge delocalization on to His-274.[6] In effect, a more delocalized charge distribution is suggested to be stabilized relative to the reactants (in which negative charge is localized on Asp-375) by a low dielectric environment. For the small models studied here, a transfer of approximately -0.1 au to the imidazole ring is indeed found in the hydrogen bonded enolate complex. However, Warshel and co-workers have argued persuasively that the catalytic power of enzymes in general arises from enhanced solvation of TSs and intermediates in polar active sites compared to the reaction in solution, and that a uniformly nonpolar active site would destabilize a TS relative to its energy in water.[29,75] Charge distributions representing the ground, TS, and intermediate states are all destabilized significantly (by 13−20 kcal/mol) in a low dielectric ($\epsilon = 4$) continuum compared to aqueous solution.[6] To make binding polar and charged molecules favorable, the enzyme must contain polar hydrogen bonding groups at the active site to provide sufficient binding energy. Assessing the electrostatic energies of the reacting species will require the structure of the enzyme surrounding the active site, and the solvent, to be taken into account. However, it seems that stabilization of the enolate intermediate by hydrogen bonding is the most important contribution to deprotonation of acetyl-CoA in CS.[7,61] Charge transfer from acetyl-CoA or its enolate to His-274 is a result of the hydrogen bond between them, and is likely to increase as the strength of the hydrogen bond increases, but is not the root of the enzyme's catalytic power. It could however assist in stabilization of the intermediate because of enhanced interactions of His-274 with Ser-244[7] as negative charge increases on the imidazole ring.

An outline of the mechanism of the CS reaction can be built up from experimental and theoretical investigations as follows. Binding of the substrates (acetyl-CoA and oxaloacetate) causes a conformational change to a closed form, burying the active site.[8,76] Asp-375 removes a proton from acetyl-CoA, resulting in the enolate which is stabilized by hydrogen bonds from His-274 and a water molecule.[7,69] These bonds stabilize the enolate because they are stronger than the equivalent bonds with the neutral substrate. Both the larger interaction energies of charged hydrogen bonded species in the active site compared to solution and the lower entropic cost of forming hydrogen bonds in the more ordered protein environment contribute. The enolate is the nucleophile for attack on oxaloacetate,[74] aided by polarization of the target carbonyl,[77] and stabilization of the tetrahedral product by a hydrogen bond from His-320.[16] The remaining chemical steps of the reaction are less clear. The thioester bond is hydrolyzed, a step which may involve Asp-375 in its protonated form.[8,9,65] This triggers a conformational change back to the open form from which the products are released. The proposed mechanism of enolate stabilization, namely the increased strength of the hydrogen bond with His-274 (and with a bound water molecule[7,61,69]), has been shown to be operative in experimental studies of model systems: hydrogen bond strength undergoes a much larger increase in low-dielectric media (such as a buried active site) upon increase of negative charge on the acceptor atom than in aqueous solution.[78] Stabilization of charge by multiple interactions appears to be a feature of many enzymes.[1,2] It will be of interest to study the possible catalytic contribution of the bound water molecule, relative to that of His-274, the effects of replacing Ser-244, and other aspects of this complex enzyme, for which simulation approaches may be useful.[7,76,79,80]

Conclusions

These results should provide valuable models of the TS for proton abstraction from acetyl-CoA and of the hydrogen bonded enolate intermediate in citrate synthase. They support the assignment[9,10] of Asp-375 as the catalytic base, and the ab initio calculations show that the resulting acetyl-CoA enolate must be stabilized at the active site if it is to exist as a reaction intermediate. The hydrogen bond between Nδ1 of His-274 and the carbonyl oxygen of bound acetyl-CoA is likely to be important for stabilizing this intermediate. In the enolate

complex this interaction is highly favorable energetically, and the increased strength of this hydrogen bond on deprotonation of acetyl-CoA will stabilize the intermediate relative to bound substrate. This differentially stabilizes[58] the enolate, prolonging the lifetime of this reactive species, allowing it to act as the nucleophile in the condensation step of the reaction.[74] The stabilities of the hydrogen bonded enolate and enol complexes are significantly different. For these models, the enolate complex, with the hydrogen bonding proton localized on the imidazole ring, is preferred. On the basis of these results and findings for the reaction in the enzyme,[7,61] the hydrogen bond between His-274 and the intermediate is not expected to be of the low-barrier type. Interactions at the active site (for example with the side chain of Ser-244[7,69]) could increase the effective acidity of His-274 and lower the barrier to proton transfer, but would not be expected to cause a dramatic increase in bond strength. It appears that the best description of the nucleophilic intermediate in citrate synthase is the hydrogen bonded enolate of acetyl-CoA. AM1 calculations on a larger model show that the stabilization of the enolate provided by the hydrogen bond with the imidazole is affected by interaction with the third component of the reacting system (acetate/acetic acid). Due to approximate cancelation of errors, the AM1 energy ordering of enol (least stable) > enolate > keto (most stable) appears to be correct for this model. Quantum mechanical/molecular mechanical calculations on this reaction employing the AM1 method may therefore be expected to provide useful results.[7,69] The results show that the proposed mechanism for the first stage of the reaction is both structurally and energetically plausible. These findings will assist in further studies of the reaction, aimed at determining the mechanism of the progress of the reaction and the sources of the enzyme's catalytic power.

Acknowledgment. A.J.M. is a Wellcome Trust International Prize Travelling Research Fellow (Grant 041229). We thank Dr. Guy Grant for fruitful discussions.

Supporting Information Available: Ab initio and semiempirical energies for all complexes discussed and electrostatic potential derived atomic charges (5 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; Freeman: New York, 1985.
(2) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover: New York, 1987.
(3) Srere, P. A. *Curr. Top. Cell. Regul.* **1992**, *33*, 261−275.
(4) Srere, P. In *Advances in Enzymology*; Meister, A., Ed.; Wiley: New York, 1975; Vol. 43; pp 57−102.
(5) Wiegand, G.; Remington, S. J. *Annu. Rev. Biophys. Biophys. Chem.* **1986**, *15*, 97−117.
(6) Evans, C. T.; Kurz, L. C.; Remington, S. J.; Srere, P. A. *Biochemistry* **1996**, *35*, 10661−10672.
(7) Mulholland, A. J.; Richards, W. G. *Proteins: Struct., Funct., Genet.* **1997**, *27*, 9−25.
(8) Remington, S. J. *Curr. Top. Cell. Regul.* **1992**, *33*, 209−229.
(9) Alter, G. M.; Casazza, J. P.; Zhi, W.; Nemeth, P.; Srere, P. A.; Evans, C. T. *Biochemistry* **1990**, *29*, 7557−7563.
(10) Karpusas, M.; Branchaud, B.; Remington, S. J. *Biochemistry* **1990**, *29*, 2213−2219.
(11) Lenz, H.; Buckel, W.; Wunderwald, P.; Biederman, G.; Buschmeier, V.; Eggerer, H.; Cornforth, J. W.; Redmond, J. W.; Mallaby, R. *Eur. J. Biochem.* **1971**, *24*, 207−215.
(12) Karpusas, M.; Holland, D.; Remington, S. J. *Biochemistry* **1991**, *30*, 6024−6031.
(13) Zhi, W.; Srere, P. A.; Evans, C. T. *Biochemistry* **1991**, *30*, 9281−9286.
(14) Kurz, L. C.; Drysdale, G. R.; Riley, M. C.; Evans, C. T.; Srere, P. A. *Biochemistry* **1992**, *31*, 7908−7914.
(15) Russell, R. J. M.; Hough, D. W.; Danson, M. J.; Taylor, G. L. *Structure* **1994**, *2*, 1157−1167.
(16) Remington, S. J. *Curr. Opin. Struct. Biol.* **1992**, *2*, 730−735.
(17) Wlassics, I. D.; Anderson, V. E. *Biochemistry* **1989**, *28*, 1627−1633.
(18) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.
(19) Gerlt, J. A.; Gassman, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 11552−11568.
(20) Gerlt, J. A.; Gassman, P. G. *Biochemistry* **1993**, *32*, 11943−11952.
(21) Marimanikkuppam, S. S.; Han Lee, I.-S.; Binder, D. A.; Young, V. G.; Kreevoy, M. M. *Croat. Chem. Acta* **1997**, *69*, 1661−1674.
(22) Warshel, A.; Papazyan, A.; Kollman, P. A. *Science* **1995**, *269*, 102−103.
(23) Cleland, W. W.; Kreevoy, M. M. *Science* **1995**, *269*, 104.
(24) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(25) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163−170.
(26) Usher, K. C.; Remington, S. J.; Martin, D. P.; Drueckhammer, D. G. *Biochemistry* **1994**, *33*, 7753−7759.
(27) Shan, S.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97−101.
(28) Scheiner, S.; Kar, T. *J. Am. Chem. Soc.* **1995**, *117*, 6970−6975.
(29) Warshel, A.; Papazyan, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13665−13670.
(30) Wang, Z.; Luecke, H.; Yao, N.; Quiocho, F. *Nat. Struct. Biol.* **1997**, *4*, 519−522.
(31) Mulholland, A. J.; Grant, G. H.; Richards, W. G. *Protein Eng.* **1993**, *6*, 133−147.
(32) Thibblin, A.; Jencks, W. P. *J. Am. Chem. Soc.* **1979**, *101*, 4963−4973.
(33) Quanta 3.3; Molecular Simulations, Inc., Waltham, MA, 1992.
(34) Zheng, Y.-J.; Merz, K. M., Jr. *J. Comput. Chem.* **1992**, *13*, 1151−1169.
(35) Rzepa, H. S.; Yi, M. *J. Chem. Soc., Perkin Trans. 2* **1990**, 943−951.
(36) Stewart, J. J. P. *MOPAC 93.00 Manual*; Fujitsu Ltd.: Tokyo, 1993.
(37) Csonka, G. I.; Ángyan, J. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *393*, 31−38.
(38) Dewar, M. J. S.; Healy, E. F.; Stewart, J. J. P. *J. Chem. Soc., Faraday Trans. 2* **1984**, *80*, 227−233.
(39) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.
(40) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363−368.
(41) Bruno, W. J.; Bialek, W. *Biophys. J.* **1992**, *63*, 689−699.
(42) Liedl, K.; Sekusak, S.; Kroemer, R. T.; Rode, B. M. *J. Phys. Chem. A* **1997**, *101*, 4707−4716.
(43) Grev, R. S.; Janssen, C. L.; Schaefer, H. F., III *J. Chem. Phys.* **1991**, *95*, 5128−5132.
(44) Mulholland, A. J.; Richards, W. G. *Int. J. Quantum Chem.* **1994**, *51*, 161−172.
(45) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. Gaussian-92; Gaussian, Inc.: Pittsburgh, PA, 1993.
(46) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894−905.
(47) Ferenczy, G. G.; Reynolds, C. A.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 159−169.
(48) *Rattler*, Version 1.01; Oxford Molecular Ltd., Oxford, 1992.
(49) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104−6114.
(50) Amyes, T. L.; Richard, J. P. *J. Am. Chem. Soc.* **1992**, *114*, 10297−10302.
(51) Isaacson, A. D.; Wang, L.; Scheiner, S. *J. Phys. Chem.* **1993**, *97*, 1765−1769.
(52) Scheiner, S. *J. Mol. Struct. (THEOCHEM)* **1994**, *307*, 65−71.
(53) Del Bene, J. E.; Szczepaniak, K.; Chabrier, P.; Person, W. B. *J. Phys. Chem. A* **1997**, *101*, 4481−4483.
(54) Clark, T. *A Handbook of Computational Chemistry*; Wiley-Interscience: New York, 1985.
(55) Meot-Ner, M.; Sieck, L. W. *J. Am. Chem. Soc.* **1986**, *108*, 7525−7529.
(56) Del Bene, J. E.; Mettee, H. D.; Frisch, M. J.; Luke, B. T.; Pople, J. A. *J. Phys. Chem.* **1983**, *87*, 3279−3282.
(57) Kolb, M.; Thiel, W. *J. Comput. Chem.* **1993**, *14*, 775−789.
(58) Albery, W. J.; Knowles, J. R. *Biochemistry* **1976**, *15*, 5631−5640.
(59) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.
(60) Ben-Tal, N.; Sitkoff, D.; Topol, I. A.; Yang, A.-S.; Burt, S. K.; Honig, B. *J. Phys. Chem. B* **1997**, *101*, 450−457.
(61) Mulholland, A. J.; Lyne, P. D.; Karplus, M. In preparation.
(62) Saunders, W. H., Jr. *J. Am. Chem. Soc.* **1994**, *116*, 5400−5404.
(63) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105−6109.

(64) Remington, S.; Wiegand, G.; Huber, R. *J. Mol. Biol.* **1982**, *158*, 111−152.
(65) Man, W.-J.; Li, Y.; O'Connor, D.; Wilton, D. C. *Biochem. J.* **1991**, *280*, 521−526.
(66) Clark, J. D.; O'Keefe, S. J.; Knowles, J. R. *Biochemistry* **1988**, *27*, 5961−5971.
(67) Kosicki, G. W.; Srere, P. A. *J. Biol. Chem.* **1961**, *236*, 2560−2565.
(68) Karplus, M.; Evanseck, J. D.; Joseph, D.; Bash, P. A.; Field, M. J. *Faraday Discuss.* **1992**, *93*, 239−248.
(69) Mulholland, A. J.; Karplus, M. *Biochem. Soc. Trans.* **1996**, *24*, 247−254.
(70) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 9855−9862.
(71) Åqvist, J.; Fothergill, M. *J. Biol. Chem.* **1996**, *271*, 10010−10016.
(72) O'Hagan, D.; Rzepa, H. S. *J. Chem. Soc., Chem. Commun.* **1994**, 2029−2030.
(73) Guthrie, J. P.; Kluger, R. *J. Am. Chem. Soc.* **1993**, *115*, 11569−11572.
(74) Mulholland, A. J.; Richards, W. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *427*, 175−184.
(75) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523−2544.
(76) Ech-Cherif El-Kettani, M. A.; Zakrzewska, K.; Durup, J.; Lavery, R. *Proteins* **1993**, *16*, 393−407.
(77) Kurz, L. C.; Shah, S.; Frieden, C.; Nakra, T.; Stein, R. E.; Drysdale, G. R.; Evans, C. T.; Srere, P. A. *Biochemistry* **1995**, *34*, 13278−13288.
(78) Shan, S.-O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474−14479.
(79) Elcock, A. H.; McCammon, J. A. *Biochemistry* **1996**, *35*, 12652−12658.
(80) Mulholland, A. J.; Richards, W. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *429*, 13-21.