A QM/MM Study of the Conformational Equilibria in the Chorismate Mutase Active Site. The Role of the Enzymatic Deformation Energy Contribution

Sergio Martí, Juan Andrés, Vicent Moliner,* Estanislao Silla, Iñaki Tuñón,* and Juan Bertrán

Departament de Ciències Experimentals, Universitat Jaume I, Box 224, 12080 Castellón (Spain), and Departament de Química Física, Universidad de Valencia, 46100 Burjasot, Valencia (Spain), and Departament de Química, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona (Spain)

Received: May 23, 2000; In Final Form: August 24, 2000

A conformational structures search of chorismate mutase substrate has been carried out using in vacuo AM1 and MP2/6-31G* methods and by means of a hybrid quantum mechanical/molecular mechanical (QM/MM) procedure in the solvated enzyme. Apart from the pseudodiequatorial and pseudodiaxial conformers of chorismate reported in the literature, new structures have been located using both methodologies. A comparative analysis of the results reveals the importance of the mechanical and electronic constraints imposed by the enzyme in this preequilibrium. These specific interactions between the substrate and the enzyme environment change the order of stability of the different conformers obtained in vacuo, thus stabilizing those structures that can be considered as the precursor for the rearrangement of chorismate to prephenate. The deformation energy of the enzyme to mold the substrate into the active site appears as the major factor in the energetic ordering of all the studied conformers.

Introduction

The conversion of (−)-chorismate to prephenate (see Scheme 1) by chorismate mutase constitutes a key step in the shikimate pathway for biosynthesis of the phenylalanine and tyrosine aromatic amino acids in bacteria, fungi and higher plants.[1]

As pointed out by Bartlett and co-workers,[2] the chorismate mutase catalyzed reaction looks to be of special interest, since it is an intermolecular transformation that has an observable noncatalyzed solution counterpart.[3,4]

The reaction is formally a Claisen rearrangement and a rare example of an enzyme-catalyzed pericyclic process. Some kinetic studies for chorismate mutases from *Klebsiella pneumoniae* and *Streptomyces aureofaciens* suggest that the major factor in the origin of the rate acceleration must be the lowering of the entropy barrier of the reaction by restricting the conformational degrees of freedom of the flexible chorismate molecule.[5] Nevertheless, Hilvert and co-workers[6] determined activation parameters for the *Bacillus subtilus* chorismate mutase showing an entropy of activation nearly as unfavorable as the uncatalyzed reaction.[7] These results, as proposed by Hilvert et al., suggest that chorismate mutase catalysts show greater mechanistic versality than commonly believed.

As pointed out by Copley and Knowles,[8] the chorismate molecule can adopt two conformations, pseudodiequatorial and pseudodiaxial (Scheme 2), depending on the position of the ether and hydroxyl oxygen with respect to the ring. Direct NMR studies of this conformational equilibrium in solution[8] indicated that the pseudodiequatorial form was only 0.9−1.4 kcal/mol lower in energy than the pseudodiaxial one, and Campbell et al.[9] concluded that the rearrangement of chorismate to prephenate requires a conformational change in which the pseudodiequatorial conformer, which predominates in aqueous solution at physiological pH, was converted to the pseudodiaxial conformer. This is the reactive conformer capable of progressing via the pericyclic rearrangement to prephenate.

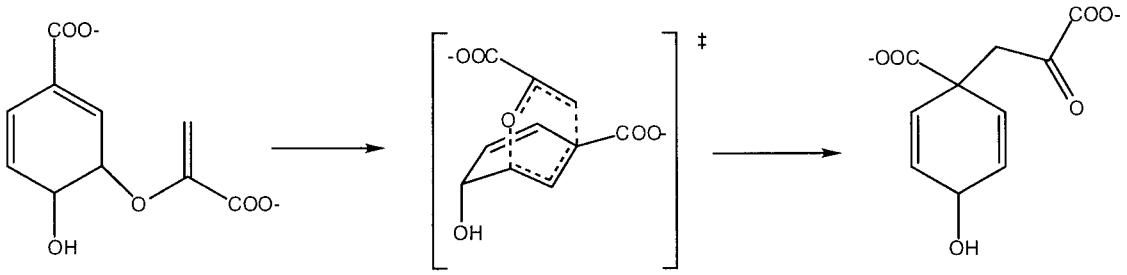
Since the early semiempirical calculations of Andrews et al.[10] on the chorismate rearrangement, several researchers have applied theoretical methods to study the conformational preequilibrium of chorismate. Wiest and Houk located and characterized as minima two chorismate conformers by Hartree−Fock and density functional theories (RHF/6-31G* and BLYP/6-31G*, respectively).[11] Hillier and co-workers obtained three chorismate conformers, two of them diaxial, by in vacuo ab initio HF/6-31G* calculations.[12] Carlson and Jorgensen[13] carried out Monte Carlo free energy perturbation FEP simulations to determine changes in free energy of solvation using the structures and charges from the ab initio calculations of Wiest and Houk.[11] The origin of the rate difference was traced solely to an enhanced population of the pseudodiaxial conformer in water. For this pseudodiequatorial−pseudodiaxial equilibrium, the gas-phase contribution to the free energy change, ΔG_0, was 10.6 kcal/mol, and 3.5 and 6.5 kcal/mol in water and methanol, respectively. Hillier et al.[14] examined the effect of water on the conformational energetics of chorismate conformers by both a continuum model (PCM) and explicit solvation within a Monte Carlo FEP treatment. Both models again described solvation as reducing the energy difference between the diequatorial and diaxial structures, but still the diequatorial form was the most stable.

The present paper describes a theoretical study of an extensive chorismate conformational searching in vacuo and in the presence of the protein environment. The first part of the work was carried out by means of ab initio and semiempirical methods, while the effect of the enzyme active site has been modeled with a hybrid QM/MM method. The analysis of the results allows us to discuss the role of the enzyme constraining the substrate in the active site and thus helping the reaction to take place. We have focused our interest in the study of the

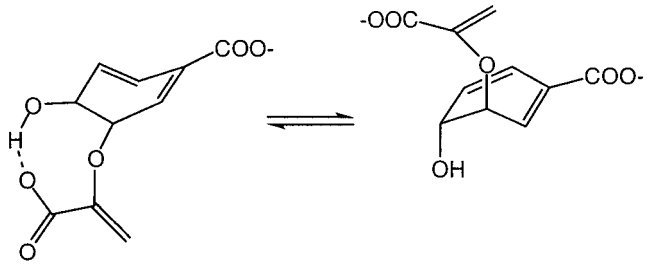
* To whom correspondence should be addressed. E-mail: Ignacio.Tunon@UV.es. Fax: +34 963864564. E-mail: moliner@exp.uji.es. Fax: +34 964728066.

10.1021/jp001888g CCC: $19.00
Published on Web 11/01/2000

SCHEME 1: Claisen Rearrangement of Chorismate to Prephenate



SCHEME 2: Pseudodiaxial-pseudodiequatorial Conformational Equilibrium of Chorismate



different energy terms that contribute to the stabilization of the enzyme−substrate Michaelis complex.

Computational Details

In vacuo calculations were carried out with the GAUSSIAN94 package of programs,[15] using the AM1 semiempirical Hamiltonian and ab initio methods including the correlation energy at second-order MP2 level with the standard 6-31G* basis set. The requested convergence on the density matrix was 10^{-9} atomic units and the threshold value of maximum displacement was 0.0018 Å and that of maximum force was 0.00045 hartree/bohr using the Berny analytical gradient optimization routine. The nature of each stationary point was checked by diagonalizing the Hessian matrix to determine the number of imaginary frequencies (zero for local minima).

The QM/MM energy hypersurface was explored using the CHARMM24 program[16] as described previously for reactions in aqueous solution[17] and in solvated enzyme active sites.[18−22] The starting geometry comes from a 2.2 Å resolution ternary complex of the prephenate in the *Bacillus subtillis* chorismate mutase (BSCM).[23] CHARMM24 was also used to add hydrogens to all titratable residues at a state complementary to pH 7. The system was placed in a cavity deleted from a preequilibrated 15 Å radius sphere of TIP3P water molecules centered on the substrate. The full system was divided into a QM region of 24 atoms, the substrate molecule, treated by the AM1 semiempirical MO methods, and a MM region, comprising the enzyme trimer plus crystallization and solvation water molecules (5573 enzyme atoms plus 224 nonrigid TIP3P water molecules). The in vacuo optimized structures of chorismate were placed in the active center of the enzyme thus being used as the substrate. These enzyme−substrate complexes were the initial structures of the QM/MM optimizations. The QM atoms and the MM atoms lying in a sphere of 18 Å of radius centered on the QM system were allowed to move (a total of 3349 atoms). The structural formula of substrate and a schematic draw of the QM/MM model are depicted in Figure 1, where the standard atom identifiers have been employed.

A partial-rational-function-operator/adopted-basis Newton−Raphson method, as implemented in CHARMM24, was employed to minimize the system until the rms residual gradient was less than 0.001 kcal/mol $Å^{-1}$; these residual gradients are lower than the commonly accepted convergence criterion for optimized geometries of small molecules in quantum chemistry.[24]

Results and Discussion

Gas Phase. The AM1 in vacuo optimized stationary point structures are shown in Figure 2, while Table 1a reports some relevant parameters (see Figure 1 for numbering of the system). A potential energy surface (PES) has been explored using the dihedral angles that define the ether and hydroxyl ring substituents. We have located five stationary points characterized as minima, which can be classified as proposed in the Introduction section. Thus, the values of the dihedral angle described by the atoms C6−C5−O7−C10 (see Table 1a) define the relative position of the ether bridge with respect to the ring. The five conformers can be classified into two types: R1 and R2 that present a value closer to 270°, and R3, R4 and R5 with a dihedral angle near to 90°. We have deliberately avoided the traditional pseudodiaxial and pseudiequatorial labels because of the ambiguity; to our concern, only R1 could be named as pseudodiaxial (see C5−O7 and C4−O8 bonds in Figure 2). On the other hand, the dihedral angle C1−C4−O8−H9 defines the hydroxyl group orientation with respect to the cyclohexadienyl ring. Its value is almost zero in R1, R4 and R5, and close to 180 in R2 and R3. These two orientations are called OH_{in} and OH_{out}. In a recent study, we have demonstrated that each one would give place to a different reaction path so leading to different prephenate conformers.[25] To distinguish between the last three conformers, the intramolecular hydrogen bond interactions can be used as a guide: R3 presents a H9···O13 interactions (O−H···O), R5 presents a weak H24···O13 interactions (C−H···O), and no hydrogen bond is presented in R4. A similar argument can be applied to distinguish between R1 and R2: a C−H···O interaction is found in the former, while the later presents a stronger O−H···O. H9···O13 and H24···O13 distances, reported in Table 1 a, reveal the magnitude of the two different hydrogen bond.

Another important internal coordinate is the C1−C14 distance. Keeping in mind that this is the bond to be formed in the chorismate to prephenate reaction, and considering the values reported in Table 1, R3, R4 and R5 structures must be the favorable conformers to lead products. R1 and R2 present a value of 5.324 and 4.961 Å, respectively; significantly larger than the values obtained for R3, R4 and R5 (4.126, 3.843, and 3.674 Å, respectively). Connecting with our previous discussion about the nomenclature, R3, R4 and R5 conformers should not be classified as pure pseudodiaxial, despite that a reactive chairlike configuration is present.

The most relevant geometrical parameters of the structures obtained with ab initio MP2 calculations, which come from the

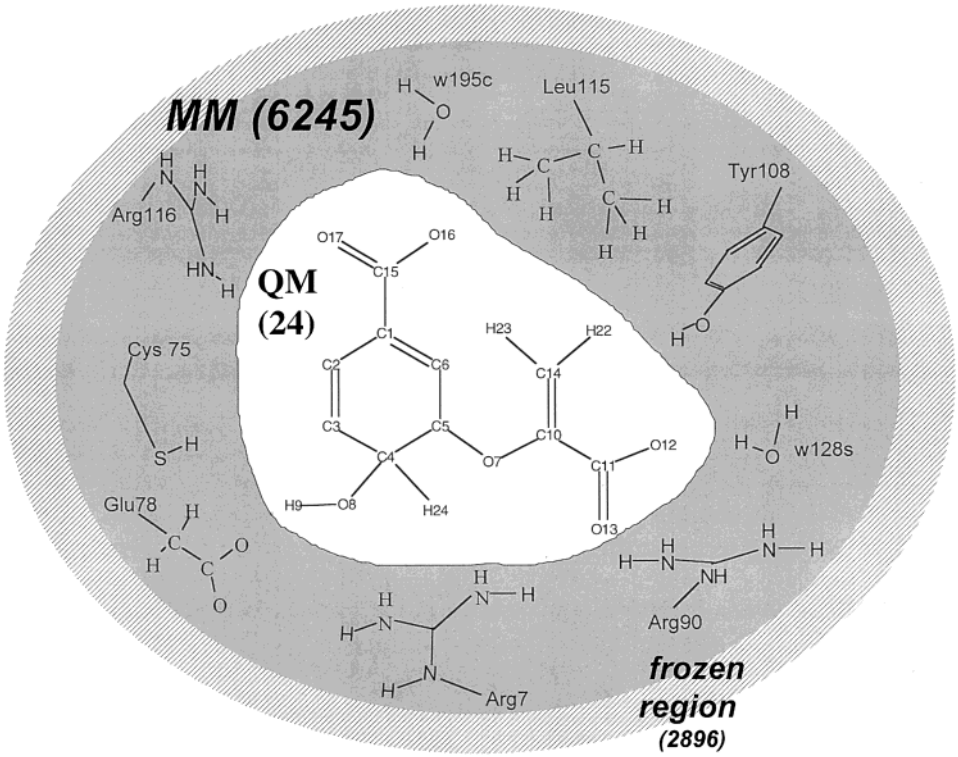


Figure 1. Covalent structure and standard atomic numbering for Chorismate Mutase active site. The full system is divided into a QM region (white area) and a MM region (gray area). The shaded area contains the MM frozen atoms.

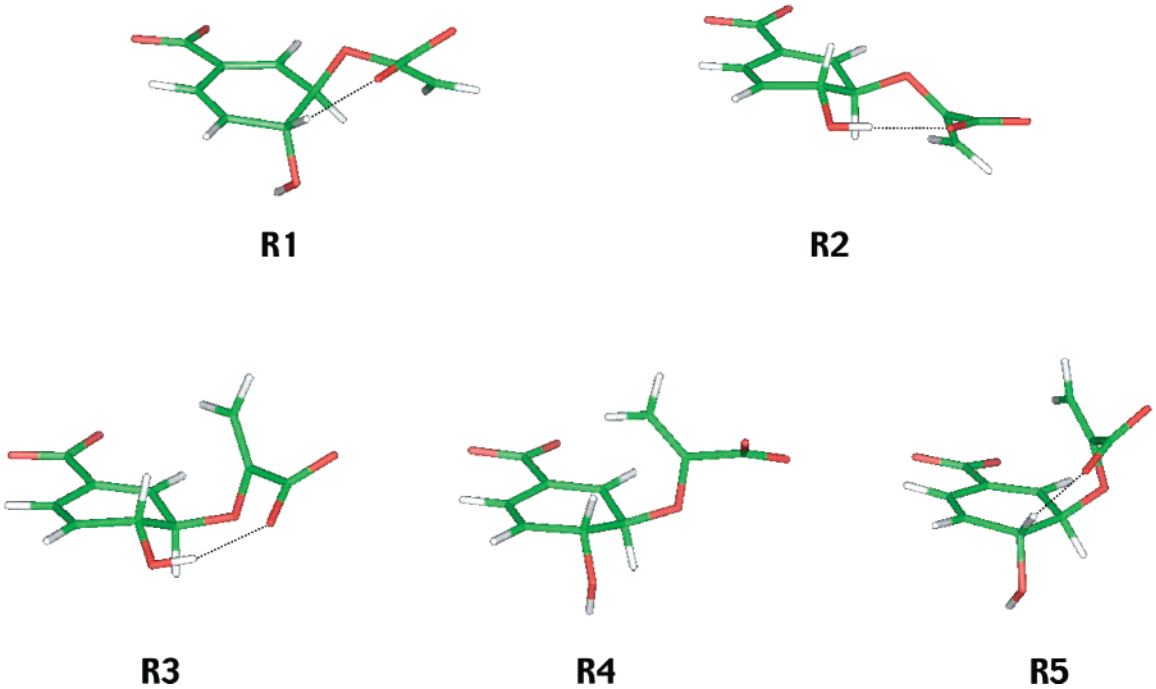


Figure 2. Stationary point structures for conformational equilibrium of chorismate obtained with the in vacuo AM1 semiempirical method.

semiempirical optimized structures, are reported in Table 1b. Comparing these results with the AM1 semiempirical ones, the main difference is that only four structures have been located on the potential energy surface. After a wide conformational structures search, R4 conformer does not appear as a minimum. Furthermore, as depicted in Figure 3, the MP2 R1′ and R5′ structures are slightly different from the AM1 R1 and R5. While the differences between R1 and R1′ can be found in the relative position of the H23−C14−H22 fragment with respect to the ring, the dihedral angle described by the O12−C11−C10−C14 atoms can be used as a guide to distinguish R5 and R5′; close to 90 and 0°, respectively. It is also important to point out that now the hydrogen bond distances are significantly shorter.

The structural changes in the different chorismate conformers have consequences on energetics. Table 2 reports absolute energies of the stationary points and relative energies to R2 obtained with in vacuo calculations. Both, AM1 and MP2 methods, describe R2 and R3 structures as the most populated structures in the gas phase, probably due to the presence of the strong O−H···O hydrogen bond. Another important feature is

TABLE 1: Selected Interatomic Distances (Å) and Dihedral Angles (deg) for the Characterized Stationary Structures Obtained with in Vacuo Calculations

	R1	R2	R3	R4	R5
		(a) AM1			
D(H9,O13)	5.014	2.070	2.062	5.789	4.948
D(H24,O13)	2.391	3.758	3.293	4.013	2.260
D(C14,C1)	5.324	4.961	4.126	3.843	3.674
T(C1,C4,O8,H9)	−3.2	118.10	125.14	9.08	−3.95
T(C6,C5,O7,C10)	234.8	239.7	81.2	67.0	73.3

MP2/6-31G*

	R1′	R2	R3	R5′
		(b) MP2/6-31G*		
D(H9,O13)	4.516	1.771	1.736	4.995
D(H24,O13)	2.427	3.642	3.181	2.237
D(C14,C1)	5.110	5.068	4.236	3.577
T(C1,C4,O8,H9)	−2.4	119.3	128.2	−7.8
T(C6,C5,O7,C10)	242.7	235.3	76.6	74.0

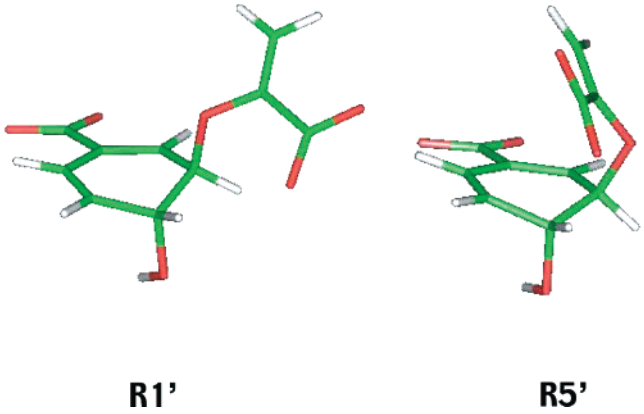


Figure 3. R1′ and R5′ chorismate conformers obtained with the in vacuo MP2/6-31G* method.

TABLE 2: Energies (Hartrees) and Relative Energies to R2 (kcal·mol^{-1}) for All Localized Species Obtained by Means of the in Vacuo AM1 and MP2/6-31G* Calculations[a]

	H (AM1)	Δ*H* (AM1)	E (MP2/6-31G*)	Δ*E* (MP2/6-31G*)
R1	−209.4	6.1	−834.772317	11.1
R2	−215.5	0.0	−834.790059	0.0
R3	−214.5	1.0	−834.786997	1.9
R4	−209.0	6.5		
R5	−208.1	7.4	−834.764696	15.9

[a] For the AM1 semiempirical procedure, heats of formation are reported while total energies are listed for ab initio MP2 results.

that the relative energies of R1′ and R5′ are almost twice the corresponding values obtained for R1 and R5. These results can be explained due to the fact that the MP2 hydrogen bond H9···O13 is shorter than the AM1 ones, as can be seen when Table 1a and 1b are compared. The AM1 and MP2 cost of OH···O hydrogen bond breaking may be, on these systems, around 5 and 10 kcal·mol^{-1}, respectively. Furthermore, as a dianion structure is being studied in a vacuum, the most populated conformers coincide with those structures that present the largest distance between the negative charged carboxylate groups (C11−C15 distance).

Environment Effect. The inclusion of the environment effect by means of the QM/MM method can render important information. Five conformers corresponding to enzyme−substrate complexes have been located as described in a previous section. According to the similitudes between the substrate structures and the previous gas-phase results (both AM1 and MP2), they have been labeled as R1, R5, R5′, R6 and R7 (see Figure 4). Structures R2, R3 and R4 obtained in the gas phase are not stable in the enzyme active site. Substrate conformation obtained in R6 is similar to R5′ but with OH_{out} orientation, as depicted in Figure 4, while in R7 it is substantially different from the rest of the described conformations. At this point it is important to underline that these structures do not have to be considered as unique; many nearly degenerated structures could be located for molecular systems containing such huge number of degrees of freedom, i.e., enzymatic systems. Furthermore, they have to be considered as relative minima in the potential energy hypersurface as the computation resulting structures are starting point geometry dependent. Nevertheless, a short 5 ps QM/MM dynamics trajectory was run using the Verlet algorithm and a constraint to fix the bonds involving hydrogen atoms,[26] as implemented in CHARMM24, starting from the most stable R6 QM/MM optimized conformer. The fact that the positions of the nuclei fluctuate around the located stationary point structure, as can be seen in Figure 5, demonstrates that this geometry can be considered as an attractor.

In Table 3 we list selected interatomic distances involving the substrate and the closest relevant enzyme active site residues and water molecules. A detail of the R6 substrate located in the active site of the enzyme is depicted in Figure 6. The first insight into these geometrical values reveals that none of the localized structures presents an intramolecular O−H···O hydrogen bond interaction; the smallest H9−O13 distance is around 4 Å. As can be seen in Table 3 and Figure 6, intermolecular hydrogen bond interactions with amino acids are preferred over the intramolecular ones obtained when the system was in the gas phase. Regarding the O13 atom, this oxygen atom is strongly interacting with Arg90, Arg7 or TIP3P water molecules, depending on the conformer. The same feature is observed on H9: a hydrogen bond interaction with Glu78B is observed in all the molecules except the R5′. In this conformer, the hydroxyl group interacts with the π electrons of the ring, thus yielding an almost perfect OH_{in} orientation. The other kind of intramolecular hydrogen bond, the weaker C−H···O (see distance H24−O13), is present in all structures but the new R7. This observation can be explained due to the arm-ring approach imposed by the active site compression, rather than because of electrostatic considerations.

An analysis of the distances between the atoms of the substrate and the protein residues shows the large number of hydrogen bond interactions that are presented in all five conformers. It is important to point out that the residues are positioned quite differently with respect to the substrate in the five structures. This suggests a considerable degree of active site flexibility, as recently observed for the LDH enzyme.[21] Thus, although there are some residues that are hydrogen bonded to the substrate in almost all the conformers (Arg90B, Glu78B, and Tyr108B), only Arg7B is making a strong interaction with chorismate in the five located structures. Furthermore, as presented in Table 3, there are some crystallographic and solvation water molecules hydrogen bonded to the substrate in all of the conformers (TIP3-195c, TIP3-41c, TIP3-51s, TIP3-70s, or TIP3-86s). This well-defined hydrogen bond network provided by the oriented water molecules and residues confirms the specific and complex character of the interactions in an enzymatic environment.

Another important geometrical feature to be analyzed is the distance between the C1 and C14 carbon atoms. On a first sight this distance obtained by means of the QM/MM method is similar to the values obtained in AM1 gas-phase calculations. Nevertheless, we must emphasize the fact that R5′ and R6 present the shortest C1−C14 distances and they are in fact the

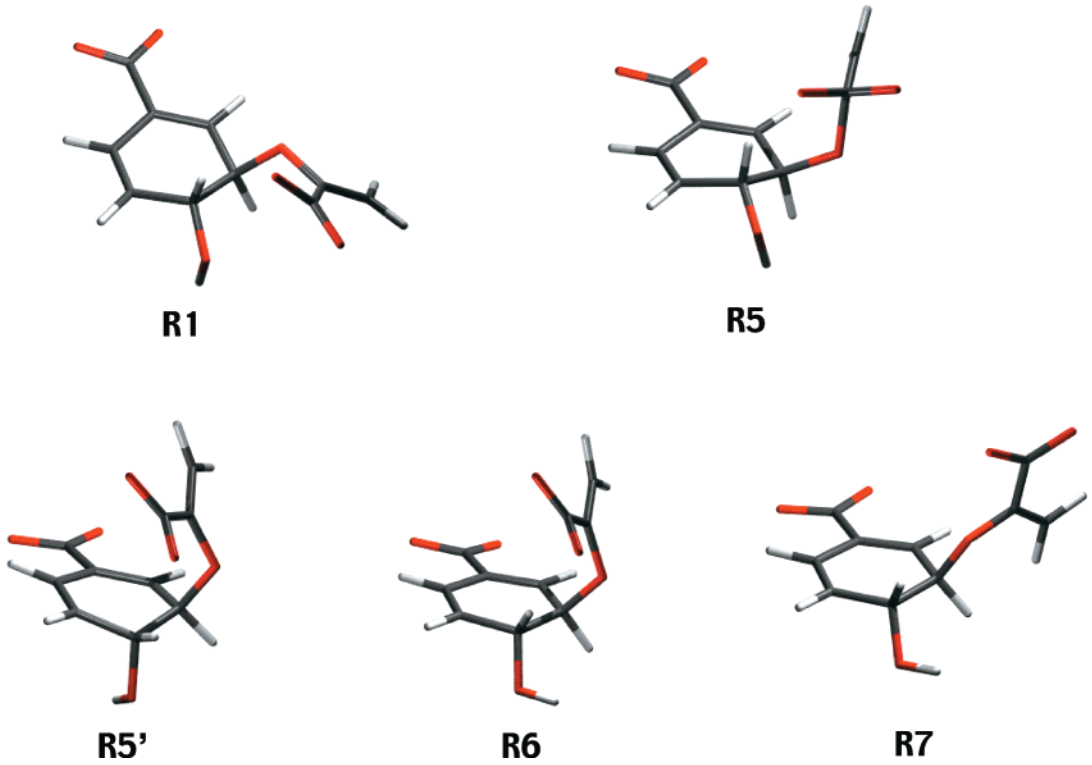


Figure 4. Detail of the different chorismate structures minimized in the active site of the BsCM enzyme by means of AM1/CHARMM/TIP3P method.

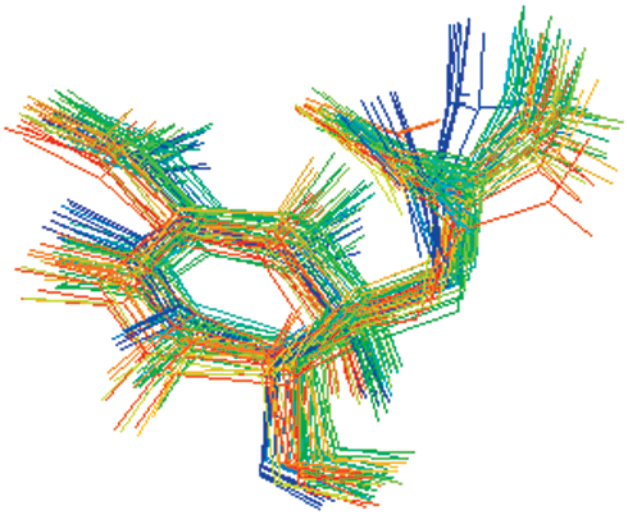


Figure 5. Superposition of structures originated from the QM/MM molecular dynamics using R6 conformer as starting geometry.

lowest energy conformers (see Table 4). This result can be interpreted as the key role of the enzyme in compressing the substrate into a configuration where the atoms to be linked are closer. In a very recent paper by Khanjin et al., the authors propose that the distance between reactive termini can be employed to analyze enzymatic reactivity, correlating this distance with the activation barrier lowering.[27] In our case, R1, R5 and R7, with C1−C14 distances of 5.568, 4.121, and 4.425 Å, respectively, cannot be the reactive structures in the enzyme reaction. The enzyme is stabilizing the structures where the C1−C14 distance is shorter, then it is helping the chorismate−prephenate reaction to take place because of a deformation of the minima toward the saddle point structure. Furthermore, looking at Table 4, the total energy difference between the five

TABLE 3: Selected Interatomic Distances (Å) and Dihedral Angles (deg) for the Characterized Stationary Structures Obtained from the QM/MM Calculations

	R1	R5	R5′	R6	R7
D(H9,O13)	4.202	4.379	5.022	3.982	5.297
D(H24,O13)	2.189	2.040	2.284	2.178	3.734
D(C14,C1)	5.568	4.121	3.666	3.777	4.425
T(C1,C4,O8,H9)	2.12	24.51	1.26	148.26	132.27
T(C11,C10,O7,C5)	264.60	102.19	88.67	89.74	162.57
T(C6,C5,O7,C10)	198.41	51.31	59.37	57.99	303.47
D(O8,HG1 Cys75A)	3.536	3.553	2.513	2.701	2.023
D(O12,HH11 Arg7B)	2.133	2.405	2.535	2.526	4.099
D(O12,HH22 Arg7B)	2.219	1.986	1.959	2.010	3.260
D(O13,HH11 Arg7B)	2.497	3.414	1.876	1.940	2.054
D(O13,HH22 Arg7B)	1.923	2.126	2.332	2.217	1.800
D(H9,OE1 Glu78B)	2.180	1.919	5.030	2.813	1.801
D(O13,HE Arg90B)	2.337	5.176	2.981	2.225	5.156
D(O13,HH12 Arg90B)	4.965	1.940	6.697	5.398	2.448
D(O7,HH22 Arg90B)	4.021	1.988	7.227	6.072	2.107
D(O13,HH21 Arg90B)	3.482	4.165	4.842	4.113	3.908
D(O13,HH22 Arg90B)	1.877	2.758	6.289	5.557	2.296
D(O13,HH Tyr108B)	3.869	3.677	3.776	3.797	1.845
D(O12,HH Tyr108B)	1.920	1.786	1.762	2.657	3.462
D(O17,HH11 Arg63A)	3.750	3.366	3.670	3.652	3.629
D(O17,H2 TIP3P-195c)	1.866	1.881	1.896	1.882	1.959
D(O7,H2 TIP3P-51s)	1.992	4.335	2.273	2.373	4.168
D(O13,H2 TIP3P-51s)	4.889	1.956	4.722	4.798	3.661
D(O12,H1 TIP3P-41c)	2.105	4.953	1.936	1.944	5.669
D(O12,H2 TIP3P-51s)	4.249	3.429	5.114	5.177	1.941
D(O16,H1 TIP3P-86s)	2.037	2.899	3.095	3.106	1.987
D(O16,H2 TIP3P-86s)	2.402	1.938	2.173	1.930	2.394
D(O17,H1 TIP3P-70s)	2.054	1.913	1.921	3.939	1.993
D(O17,H2 TIP3P-86s)	1.903	2.091	1.895	2.112	1.980

structures (ΔE_T) is very large, the structures with shorter distances between reactive centers being the most populated ones. The most stable structure is OH_{out} (R6 with a C1−C14 distance of 3.777 Å) while the second one is OH_{in} (R5′ with a C1−C14 distance of 3.666 Å). This is in agreement with our previous QM/MM study.[25]

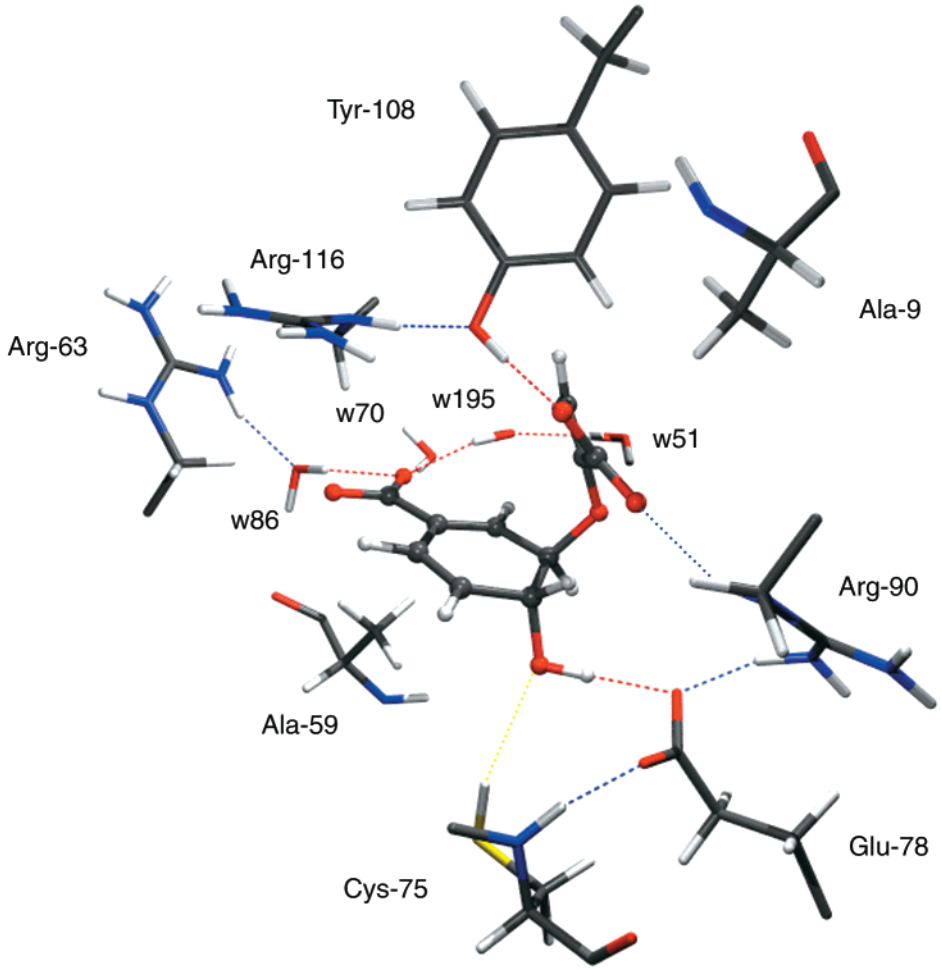


Figure 6. Detail of the R6 conformational chorismate structure minimized in the active site of the BsCM enzyme by means of AM1/CHARMM/TIP3P method. For clarity proposes, Arg7 and w41c have been removed.

TABLE 4: Total Energies (E_T) of Corresponding Structures Optimized by Means of QM/MM Methods and Relative Energies to R6 (ΔE_T)[a]

	E_T	ΔE_T	E_S^0	ΔE_S^0	ΔE_{E-S}	ΔE_{E-E}
R1	−3486.0	71.7	−199.5	−2.4	−10.5	84.6
R5	−3444.5	113.2	−199.4	−2.3	−8.5	124.1
R5′	−3550.7	7.1	−198.0	−0.9	12.9	−4.9
R6	−3557.8	0.0	−197.1	0.0	0.0	0.0
R7	−3469.4	88.3	−198.4	−1.3	−22.0	111.6

[a] E_S^0 = in vacuo energy of the QM/MM optimized structures of the substrate and relative energies to R6 (ΔE_S^0). ΔE_{E-S} = enzyme−substrate interaction energy relative to R6. ΔE_{E-E} = enzyme−enzyme interaction energy relative to R6. All values are listed in kcal/mol.

An analysis of the components of these relative energies can be carried out from the three last columns in Table 4. The sum of ΔE_S^0 (the in vacuo energy of the substrate structures, obtained in the enzyme active site, relative to the most stable conformer; R6), ΔE_{E-S} (the enzyme−substrate interaction energy relative to R6) and ΔE_{E-E} (the deformation energy of the enzyme relative to R6) gives the total relative energy, ΔE_T. As can be seen from the ΔE_S^0 column, all values are in a range of 2.5 kcal mol^{-1}, R6 being the least favorable one. The absolute values obtained using these deformed structures (E_S^0) are higher than the gas phase stationary points by ca. 15 kcal mol^{-1} (see first column of Table 2), as expected. It is amazing that R1, R7 and R5 present much higher stabilizing interaction values than R5′ and R6, the reactive conformers (see column ΔE_{E-S}). R5′ and R6, which are the least stable ones according to the substrate structure itself and the least stable ones according to the enzyme−substrate interactions (see ΔE_S^0 and ΔE_{E-S} columns, respectively), are the enzyme−substrate complexes that present the less unfavorable enzyme deformation energy term (see ΔE_{E-E} column). Furthermore, R5′ and R6 require the smallest enzyme deformation and they are, precisely, the closest structures to the saddle point geometry of the reaction that is going to be catalyzed (the chorismate to prephenate rearrangement). Although some deformation of the enzyme is required (see below), the fact that R5′ and R6 need the smallest deformation and present the highest similarity to the transition structure, is connected with the seminal Pauling's postulate:[28] the enzyme is designed to be geometrically complementary to the transition state structure of the reaction that has to catalyze.

Finally, we have optimized the structure of the enzyme without the substrate in order to obtain the deformation energy required when going from the isolated enzyme to the enzyme−substrate complex. The ΔE_{E-E} value corresponding to the isolated enzyme is −104.4 kcal mol^{-1} (using as starting geometry for the optimization the R6 structure without the substrate molecule). The energy difference between the final and initial state of the enzyme is thus very unfavorable. At this point of the discussion one could ask why we obtain a minimum energy structure if a sum of two unfavorable terms are taken into account: a deformation of the enzyme and a deformation of the substrate. The answer is that an overstabilization appears when the enzyme−substrate complex is formed and new enzyme−substrate interactions are generated (−159.3 kcal mol^{-1}

for the R6 conformer). However, the substrate structures with the largest interaction energy do not form the most stable complexes, due to the fact that interaction and deformation energies are not lineally correlated.

Conclusions

In this paper a conformational searching study has been carried out for the chorismate in the gas phase and in the active site of the chorismate mutase enzyme by means of molecular energy minimizations. The hybrid quantum-mechanical/molecular-mechanical (AM1/CHARMM24/TIP3P) technique used has allowed us to include the protein environment and to explore structural features of significance on hypersurfaces spanning several thousand degrees of freedom. We focus our interest on the comparison between the free in vacuo minimizations of the reactants-like structure and the results obtained when the enzyme constrains are taking place. The analysis of the results can be summarized as follows.

The AM1 gas phase calculations render five possible chorismate conformers, while the ab initio MP2 method allows us to locate only four; some of them not previously published. Depending on the orientation of the ring substituents, i.e., the hydroxyl group and the arm, they can be classified into OH_{in} and OH_{out}, or pseudodiaxial and pseudiequatorial, respectively. Nevertheless, while the OH_{in} OH_{out} classification has been used in this paper, we rather consider the hydrogen bond interactions criterium to distinguish them. In absence of external factors, the presence of strong O−H···O and weak C−H···O intramolecular hydrogen bonds determines the relative energy of these molecules. Thus, R2 and R3, those molecules with O−H···O interactions are the most populated conformers in vacuo using both semiempirical or ab initio methods.

When the protein environment is included, we find that not all the conformers obtained in the gas phase are minimum energy points in the QM/MM potential energy hypersurface. Some of the in vacuo structures are dramatically modified into what we consider new conformers. The substrate "prefers" the intermolecular interactions with the amino acids and water molecules of the enzyme active site rather than the intramolecular hydrogen bond interactions obtained in vacuo. These specific electrostatic enzyme−substrate interactions, as well as the compression effect of the protein that tried to avoid internal deformation itself, change the order of energies of the gas phase results and stabilizes precisely those structures capable of progressing into the chorismate-prephenate catalyzed reaction: R5′ and R6. Therefore, from a geometrical point of view, the biological catalyst selects the precursor conformer of the corresponding transition state and discards the nonreactive conformers. The comparison with the gas phase results allows us to stress the importance of including the environment effect in order to obtain a reliable picture. Important information may be lost if only in vacuo studies are carried out.

It is important to point out that the structures presented in this paper do not have to be considered as unique; many nearly degenerate structures could be located for molecular systems containing a large number of atoms (a total of 10 041 fully mobile degrees of freedom). Nevertheless, considering the flexible character of our QM/MM model the relative minima that emerge from the calculations can be used to understand the role of the enzyme. Free energy calculations can render more reliable energetics, as an average of thousands of structures are being considered, but the internal energy minimizations presented in this work allow us to carry out a deeper insight into the geometries and a detailed analysis of the different energetic terms that contribute to the catalysis. As a result of these calculations, it is important to emphasize that contrary to a standard opinion, the energy difference between several enzyme conformers comes mainly from the enzyme deformation energy term, instead of the enzyme−substrate interactions. This conclusion reminds us of similar ideas pioneered by Tapia and coworkers.[29] Interaction energy and deformation energy in enzymes are not as well correlated as in solvated systems. Although interaction energy is responsible for enzyme−substrate complex formation, the relative stability of different Michaelis complexes is not determined by this energy term; the deformation energy plays the most important role. At least part of the enhancement in rate constants is due to the preorganization effect of the enzyme, which results in the Michaelis complex containing substrates that are bound in a configuration favorable for reaction, as suggested by. Kollman et al.[30] and Bruice et al.[31]

Acknowledgment. We are indebted to DGICYT for project PB96-0795 which supported this research, and the Servei d'Informatica of the Universitat Jaume I for providing us with computer capabilities. S.M thanks Generalitat Valenciana for a doctoral fellowship.

References and Notes

(1) Haslam, E. *Shikimic Acid: Metabolism and Metabolites*; John Wiley & Sons: New York, 1993.

(2) Bartlett, P. A.; Nakagawa, Y.; Johnson, C. R.; Reich, S. H.; Luis, A. *J. Org. Chem.* **1988**, *53*, 3195−3210.

(3) Young, I. G.; Gibson, I.; MacDonald, C. G. *Biochim. Biophys. Acta* **1969**, *192*, 62−72.

(4) Gajewski, J. J.; Jurayi, J.; Kimbrough, D. R.; Gande, M. E.; Ganem, B.; Carpenter, B. K. *J. Am. Chem. Soc.* **1987**, *109*, 1170−1186.

(5) Görisch, H. *Biochemistry* **1978**, *17*, 3700−3705.

(6) Kast, P.; Asif-Ullah, M.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.

(7) Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *12*, 3492−3498.

(8) Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5008−5013.

(9) Campbell, A. P.; Tarasow, T. M.; Massefski, W.; Wright, P. E.; Hilvert, D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8663.

(10) Andrews, P. R.; Heyde, E. *J. Theor. Biol.* **1979**, *78*, 393−403.

(11) Wiest, O.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 11628−11639.

(12) Davidson, M. M.; Gould, I. R.; Hillier, I. H. *J. Chem. Soc., Perkin Trans. 2* **1996**, 525−532.

(13) Carlson, H. A.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1996**, *118*, 8475−8484.

(14) Davidson, M. M.; Guest, J. M.; Craw, J. S.; Hillier, I. H.; Vincent, M. A. *J. Chem. Soc., Perkin Trans. 2* **1997**, 1395−1399.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Peterson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian Inc.: Pittsburgh, PA, 1995.

(16) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(17) Barnes, J. A.; Williams, I. H. *J. Chem. Soc., Chem. Commun.* **1996**, 193−194.

(18) Barnes, J. A.; Williams, I. H. *Biochem. Soc. Trans.* **1996**, *24*, 263−268.

(19) Moliner, V.; Turner, A. J.; Williams, I. H. *J. Chem. Soc. Chem. Comun.* **1997**, 1271−1272.

(20) Moliner, V.; Andrés, J.; Oliva, M.; Safont, V. S.; Tapia, O. *Theor. Chem. Acc.* **1999**, *101*, 228−233.

(21) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323−1331.

(22) Castillo, R.; Andrés, J.; Moliner, V. *J. Am. Chem. Soc.* **1999**, *121*, 12140−12147.

(23) Chook, Y. M.; Ke, H.; Lipscomb, W. N. Protein Data Bank ID code 1COM.

(24) Foresman, J. B.; Frisch, A. E. *Exploring Chemistry with Electronic Structure Methods*; Gaussian, Inc.: Pittsburgh, PA, 1993.

(25) Martí, S.; Andrés, J.; Moliner, Silla, E.; Tuñón, I.; Bertrán, J. *Theor. Chem. Acc.*, in press.

(26) Van Gunsteren, W.; Berendsen, H. J. C. *J. Mol. Phys.* **1977**, *34*, 1311−1327.

(27) Khanjin, N. A.; Snyder, J. P.; Menger, F. M. *J. Am. Chem. Soc.* **1999**, *121*, 1, 11831−11846.

(28) Pauling, L. *Nature* **1948**, *161*, 707−709.

(29) (a) Tapia, O.; Cárdenas, R.; Andrés, J.; Colonna-Cesari, F. *J. Am. Chem. Soc.* **1988**, *110*, 4046−4047. (b) Tapia, O.; Andrés, J.; Safont, V. S. *J. Chem. Soc., Faraday Trans* **1994**, *90*, 2365−2374.

(30) (a) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395. (b) M. Peräkylä, M.; Kollman, P. A. *J. Phys. Chem. A* **1999**, *103*, 8067.

(31) (a) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595. (b) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1997**, *119*, 9103.