A Theoretical Study of the Molecular Mechanism for the Carboxylation Chemistry in Rubisco

M. Oliva, V. S. Safont, and J. Andrés*

Departament de Ciències Experimentals, Universitat Jaume I, Box 224, 12080 Castelló, Spain

O. Tapia

Department of Physical Chemistry, Uppsala University, Box 532,S-75121 Uppsala, Sweden

Received: June 21, 1999; In Final Form: August 23, 1999

Stereochemical and structural aspects for the carbon dioxide fixation followed by hydration of D-ribulose 1,5-bisphosphate catalyzed by Rubisco are analyzed by using two model substrates, hydroxypropanone and 3,4-dihydroxy-2-pentanone. The molecular mechanisms for the carboxylation, C3-hydration and C2-inversion processes, are theoretically characterized and discussed with saddle points of index 1 representing transition structures (TS) and relevant (molded) intermediates mirroring the experimentally proposed mechanism. Ab initio SCF MO calculations at 3-21G and 6-31G** basis set levels of theory were used, while the correlation energy is included at the MP2/6-31G** level. The mapping starts from TS1, the carboxylation transition structure, which describes now the coupling of the carbon dioxide attack to the substrate C2-center in its dienol form with a synchronous interconversion of the C3 hydroxyl into a ketone group. Thereafter, water addition leads to a *gem*-diol followed by another step of intramolecular hydrogen transfer coupled with the C2−C3 bond breaking process. This TS breaks into one model of 3-D-phosphoglycerate product and an intermediate. The configuration inversion at the C2-center is found to be possible via intramolecular hydrogen transfer, as suggested by the corresponding TS relating the intermediate to the C2-inverted conformation. The complete set of steps found gives a self-contained description of the carboxylation followed by hydrolysis with proper stereochemistry. Comparisons between the transition state analogue: 2-carboxy-D-arabinitol-1,5-bisphosphate and TS1 show that both structures can be superposed with minimal root-mean-square deviation for C-atoms. All other calculated stationary TSs share the gross conformational features of TS1, and consequently, all molecular rearrangements detected in a vacuum can be accommodated without constraints into the active site of Rubisco. Transition structures invariances with respect to level of theory and molecular models are discussed.

1. Introduction

Rubisco (D-Ribulose-1,5-bisphosphate carboxylase/oxygenase, E.C. 4.1.1.39) catalyzes the initial step in Calvin's reductive pentose phosphate cycle, i.e., the photosynthetic fixation of atmospheric CO_2 to the enzyme's substrate, D-Ribulose-1,5-bisphosphate (RuBP).[1−5] Most of the carbon found in the biosphere[6] has passed at one time or another through the active site of Rubisco. Molecular mechanisms have been established over the years by using biochemical techniques. More recently, 3D data obtained from X-ray diffraction studies have been used to rationalize the molecular mechanism of carboxylation and propose alternatives to the chemically based ones.[7,8]

At the active site, there seems to be a large degree of "promiscuity" in the sense that there are a number of reactions catalyzed there. One of them is oxygenation of RuBP leading to the process of photorespiration. The dioxygen attacks the same C2-center as carbon dioxide. It is to this type of variation in the mechanism that theoretical studies from our group have focused on. The attention here is centered on the full carboxylation reaction. The oxygenation mechanism has been recently investigated by us.[9] The substrate RuBP incorporates carbon dioxide at the C2-center and a water molecule at C3, leading, via a C2−C3 bond breaking process, to two molecules of 3-phospho-D-glycerate. For this to be so, the fragment including the ingoing CO_2 must involve a configuration inversion at the C2-center. In brief, this system raises a rather complex mechanistic problem. The theoretical issue in this paper is to find whether there are minimal molecular models for which transition state structures can sustain this chemistry in a vacuum. In a preceding letter,[10] we have shown that such an objective is feasible, albeit the molecular model was minimal and no explicit enzyme elements were included. But with simple models, the role of the enzyme can be understood from a perspective different from the one obtained when all the system complexity is included in the calculations.[11]

In earlier investigations, transition structures, characterized by saddle points of index 1, have been determined for enolization, the first step in carboxylation, for oxygenation, and for self-inhibitory properties.[12−18] All the calculated transition structures (TS) have a remarkable property. Their C-frameworks have a very small root-mean-square deviation when overlapped with the crystal structure of CABP (2-carboxy-D-arabinitol 1,5-bisphosphate),[7] a transition state analogue to the six-carbon intermediate in the carboxylation reaction catalyzed by Rubisco.

* To whom mail should be addressed

10.1021/jp992052k CCC: $18.00
Published on Web 10/12/1999

The characterization of an enolization transition structure (which could also be docked at the active site), involving an intramolecular hydrogen transfer from C3 to the carbonyl oxygen at C2,[15] opened the possibility that several aspects related to the enolization reaction could be achieved via intramolecular processes. This expectation was ratified in a study made to describe self-inhibition as a set of retro-enolization processes.[18] More recently, our study of oxygenation confirms this idea.[9] In this paper, the intramolecular hydrogen transfer hypothesis is further explored and applied to study the addition of CO_2 and H_2O on five- and three-carbon molecular models representing RuBP. The promiscuity can then be understood in terms of families of transition structures having a shape complementary to the active site of Rubisco.

Here, we report on TSs and intermediate species found in a pathway describing carboxylation of the dienol form and the hydrolysis of the carboxylated substrate in a vacuum. The invariances of the transition structures to the level of electronic theory and the size of the molecular model are examined here. A discussion closes the paper.

2. Transition Structure Invariances, Methods and Models

The transition structure, obtained as a stationary geometric structure, occupies a central position in our analyses. It is then useful to give the antecedents on which we base our discussions. For simple diatom + atom reactions, besides the activation energy, Hinshelwood early noticed the need for "... some condition determining the point at which rearrangement takes place".[19] In contemporary chemistry those points are usually represented by transition structures. With the development of analytical gradients and second derivatives of the energy with respect to geometric parameters, the identification of saddle points of index 1[20−22] is becoming a tool to help check for different mechanistic proposals.[23−26]

Theoretically, there has been a shift in the perception of the nature of the transition state. This one ranges from the early concept of an unstable molecular configuration[27−29] to the present view of a molecule, albeit having finite lifetimes,[25,30] that displays a well-defined geometry (sensed as a saddle point of index 1) whose quantum vibrational states must be populated if the mechanism takes that pathway. The meaningful quantum states are those having positive eigenvalues of the Hessian. From catalytic antibodies[31−35] via the spectroscopy of the transition state[36,37] to the construction of transition state analogues,[38,39] the idea of actually existing transition structures[40−44] is becoming a natural hypothesis.[45,46]

The fact that the structure of these stationary points for Rubisco-related substrates could be overlayed with the structures of the transition state analogue CABP is important. Bearing in mind that the protein binds the substrate, which in its ground state has a totally different structure compared to the TS, one may recognize the role of productive binding: molding of the substrate so that its geometric fluctuations overlap those accessible to the TS. These configurations form sets of precursor complexes wherefrom the transition state can be populated. Surprisingly, these TSs, although calculated in the absence of the protein, have structures compatible with their active sites. And this raises an important point. Are these TSs invariant in shape to interactions with the protein at the active site? The overlapping structures strongly suggest that such might be the case, but it does not prove it. That is the reason it is important to study TS invariances under different situations. In this paper we play with the level of electronic theory and the size of the molecular model and give examples where the model complexity is increased.

The mechanistic pathway is checked by descending via the intrinsic reaction coordinate (IRC) of Fukui[47] toward lower energy stationary points. In so doing, the conformations of the corresponding intermediates (or another saddle point along an orthogonal direction to the IRC) may differ substantially from that of the parent transition structure. As the real system is, after all, in a sort of cavity (active site), the geometry change of the ingoing and outgoing channels would stop at conformations compatible with the size and shape of the active site, thereby generating not a family of transition structures but a set of activated successor and precursor complexes. For this reason, it does not make sense to speak of activation energy without defining the context.

As an extension of Pauling's lemma,[48−50] it can be stated that for a substrate system trapped at the active site of a given enzyme, and a mechanism for an elementary step described with the help of a TS, a necessary condition required to open a given interconversion pathway would be the existence of surface complementarity between the geometry associated with the TS in a vacuum and the active site. This complementarity can be checked via a geometric superposition of the calculated stationary structures with transition state analogues and slow substrates whose X-ray structure is known.

In this study, a three-carbon model (hydroxypropanone) and the five-carbon model (3,4-dihydroxy-2-pentanone) at HF-(Hartree−Fock)/3-21G and HF/6-31G** basis set levels were used. The correlation energy has been estimated at the MP2/6-31G** level of theory for the three-carbon model. All calculations were carried out using the GAUSSIAN 94 program.[51] Supplementary vibrational and spectral analyses are made with Gaussview 1.0 package.[52] The phosphate groups are not included, they are assumed to contribute in the (productive and nonproductive depending upon the protein) binding of the substrate, as documented experimentally by X-ray studies.[4,7,8,53−55]

3. TSs Mapping the Whole Carboxylation Process

The overall reaction, carbon fixation catalyzed by Rubisco leading to the addition of CO_2 and H_2O to RuBP to yield two molecules of 3-D-phosphoglycerate (PGA), involves a number of steps, some of them directly or indirectly characterized with experimental methods. After enolization, CO_2 addition at the C2-center, hydration at C3, *gem*-diol intermediate at C3, bond breaking C2−C3, and the inversion of configuration at C2 are the steps required to accomplish the chemistry.[56] The commonly accepted reaction steps for carboxylation of RuBP[4,53] are depicted in Figure 1.

Enolization is the first and, according to some authors, may be the rate-limiting step.[7,57−62] From that TS one gets the intermediate **I**, which is submitted to a C2-carboxylation via an electrophilic attack of CO_2 to form **II**. Hydration of this C3 ketone intermediate (**II**) leads to a *gem*-diol **III**. Deprotonation of the hydroxyl oxygen atom at the C3-center is assumed to initiate C2−C3 bond cleavage. Following this bond breaking process, one molecule of product, PGA, is formed together with the *aci*-acid species (**IV**). A stereospecific protonation at the C2-center is required to obtain the second PGA.

Although there is general agreement on the sequence of the major steps, at an electronic molecular level very little was known of the detailed mechanism of catalysis. The stationary points found with the three-C molecular model for the overall process are depicted in Figure 2 .[10] They give an overview of the global process, starting from the model substrate in its enediol form, and ending up in the model products. The enolization step was already studied.[15]

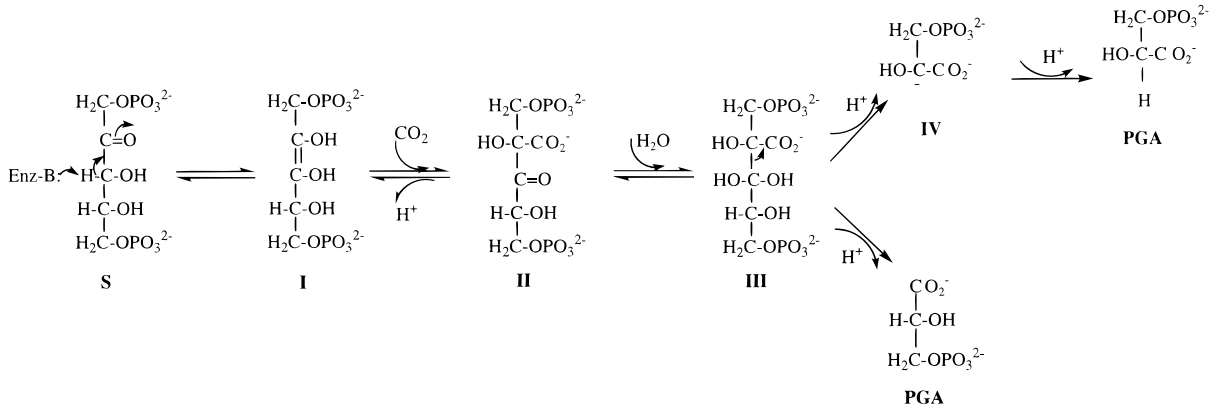

Figure 1. Proposed reaction pathway for the carboxylation of RuBP catalyzed by Rubisco extracted and modified from Hartman and Harpel.[14]

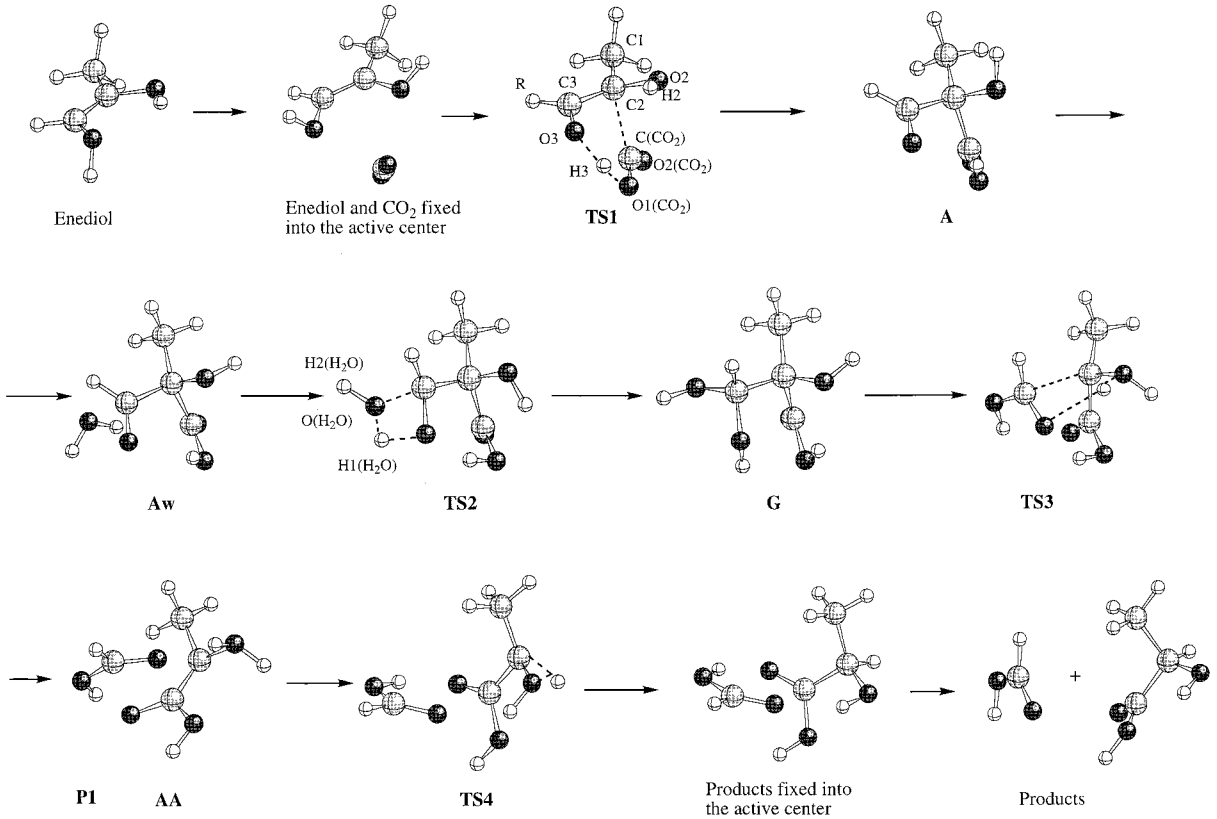

Figure 2. Schematic representation of the reaction pathways from the enediol to the final products. The geometries of stationary points along the reactive channel are depicted. The method used has been described in previous reports from our group.[15,16,18,77] The present study was done at HF/3-21G and HF/6-31G** basis set levels, while correlation energy has been included at the MP2/6-31G** level. At each point, the Berny analytical gradient optimization routines were used for the optimization.[78,79] The requested convergence on the density matrix was 10^{-9} atomic units; the threshold value of maximum displacement was 0.0018 Å and that of maximum force was 0.000 45 hartree/bohr. The nature of each stationary point was established by calculating analytically and diagonalizing the Hessian matrix (force constant matrix).

Such a mechanism in vacuo is unlikely to happen due to the enormous energy required to deform the substrate and mold it into the geometry of the TS. The study of X-ray structures[8,63−65] shows that the substrate actually is molded into a structure looking very much like the structure of the carboxylation TS in vacuo. As the model substrate contains only three carbon atoms, it is necessary to check whether a more realistic model would produce a similar mechanistic pathway. This is examined in what follows; a simultaneous comparison with the five-carbon model is given so that one can appreciate structural invariances including the transition vectors. The more complex molecular model reproduces all aspects depicted in Figure 2.

3.1. Carboxylation TS. In Figure 3, the **TS1** for carboxylation is depicted for the two molecular models used. The atom numbering is also indicated; the atoms belonging to CO_2 fragment are not numbered, and all other oxygens are recognized by their binding to a particular center. Some geometric details are given in the figure and in Table 1a.

Compared to our previously reported carboxylation TSs,[12,16,17] **TS1** presents a concomitant hydrogen transfer between the hydroxyl group at the C3-center toward one oxygen in the approaching carbon dioxide moiety. The carboxylation **TS1** couples the carbon dioxide attack to the C2-center of the substrate in its dienol form with a synchronous interconversion of the C3 hydroxyl into a ketone group. This would result in a neutral acid intermediate **A** with a carbonyl function at C3, which correlates with the 3-keto-CABP. As pointed out by Jaworowski and Rose and Pierce et al., "one of the unique

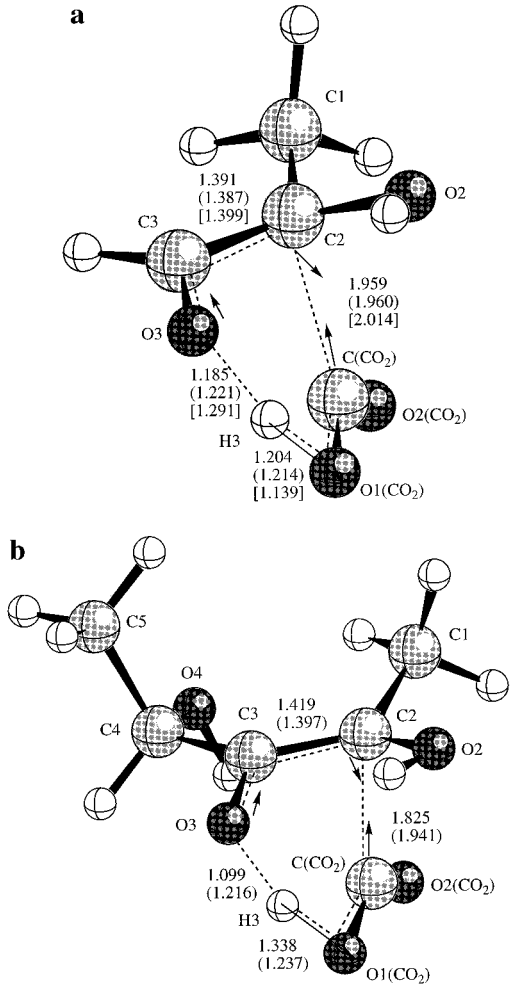

Figure 3. TS1 for the carboxylation step: (a) three-C model; (b) five-C model. The arrows represent the major components of the transition vector. Main distances are indicated, calculated at HF/6-31G**, HF/3-21G (in parentheses), and MP2/6-31G** (in brackets).

features of the carboxylation reaction is that 3-keto-CABP is an intermediate that can be isolated and that it is stable".[66,67] This intermediate has been isolated by acid quenching and stored at −80 °C. This species has a finite lifetime (1 h) in solution,[68] where it becomes decarboxylated at room temperature. Note that by deprotonating the C3-hydroxyl group during carboxylation, the **TS1** would lead directly to 3-keto-CABP if the molecular model had the phosphate groups. In our case, the carboxylate group at C2 is protonated.

The transition vector (relative atom displacement phases) of **TS1** is represented in Figure 3 as vector displacements from the stationary geometry. Inverting the global sign leads to a fluctuation that overlaps the geometry of the activated reactants. Normal mode animation shows the dominance of the hydrogen fluctuation; a carbon dioxide moiety bending; the $C(CO_2)-C2$ distance fluctuates in antiphase to the hydrogen in the O3−O bridge. The fluctuations are fully localized to this atom framework, which would explain the success of the smaller model.

Following an IRC, the passage through **TS1** would ensure the population of the acid intermediate **A**. This intermediate has an interesting feature. The carbonyl functionality at C3 would open a channel to the hydration step. As the **TS1** geometry can be overlapped with CABP, surface complementarity is ensured between the protein active site and the activated complex of the reaction catalyzed by the enzyme. This is just Pauling's lemma.[49]

TABLE 1: Main Distances (Å) Obtained by Using the Different Calculation Levels and Molecular Models, for (A) TS1, (B) TS2, (C) TS3, and (D) TS4[a]

(a) TS1

	$C2-C(CO_2)$	$H3-O1(CO_2)$	H3−O3	C2−C3
HF/3-21G 3C	1.960	1.214	1.221	1.387
HF/6-31G** 3C	1.959	1.204	1.185	1.391
MP2/6-31G** 3C	2.014	1.139	1.291	1.399
HF/3-21G 5C	1.941	1.237	1.216	1.397
HF/6-31G** 5C	1.825	1.338	1.099	1.419

(b) TS2

	$H1(H_2O)-O3$	$O(H_2O)-C3$	$O(H_2O)-H1(H_2O)$	C3−O3
HF/3-21G 3C	1.400	1.589	1.129	1.371
HF/6-31G** 3C	1.338	1.554	1.121	1.349
MP2/6-31G** 3C	1.382	1.620	1.135	1.336
HF/3-21G 5C	1.406	1.548	1.136	1.384
HF/6-31G** 5C	1.336	1.535	1.125	1.361

(c) TS3

	C2−C3	$H1(H_2O)-O2$	$H1(H_2O)-O3$	C2−O2	C3−O3
HF/3-21G 3C	2.228	0.992	1.947	1.565	1.265
HF/6-31G** 3C	2.299	0.976	1.659	1.557	1.238
MP2/6-31G** 3C	2.033	1.040	1.817	1.527	1.266
HF/3-21G 5C	2.282	1.034	1.581	1.525	1.271
HF/6-31G** 5C	2.322	0.974	1.665	1.560	1.240

(d) TS4

	C2−H	O2−H	C2−O2
HF/3-21G 3C	1.357	1.129	1.616
HF/6-31G** 3C	1.319	1.097	1.599
MP2/6-31G** 3C	1.423	1.076	1.625
HF/3-21G 5C	1.375	1.129	1.615
HF/6-31G** 5C	1.319	1.097	1.598

[a] 3C stands for the hydroxypropanone, and 5C for the 3,4-dihidroxy-2-pentanone molecular models. Atom label "H" in Table 1d corresponds to the hydrogen atom being transferred from O2 to C2.

3.2. Hydration TS. In Figure 4 the hydration **TS2** is depicted for both molecular models. In Table 1b some geometric data are given. This type of TS is rendered possible thanks to the dehydrogenation in the previous step of the hydroxyl group at C3. Note that, because the hydrogen bound to O3 has been moved to the incoming CO_2 in the previous step of our proposed alternative mechanism, there is no need for an external base to abstract now this hydrogen and prompt for hydration: hydration was made possible as soon as CO_2 addition took place. Furthermore, the substrate is not going to be overcharged during turnover.

The hydration in **Aw** (see Figure 2) at the C3-center leads to this new transition structure, **TS2** corresponding to a concerted mechanism such that oxygen addition and proton transfer to the carbonyl oxygen are accomplished in a single step. This type of TS has early been found in related hydration of carbonyl compounds.[69] The four-center character is apparent in the **TS2**. This latter is more reactant-like in the sense that the transferred hydrogen is closer to the water oxygen than the C3-carbonyl oxygen. Normal mode animation of this TS shows an almost fully localized fluctuation in the four-center species. A weak motion develops in the H−O bond of the carboxyl group. Interestingly, for most of the low frequencies (smaller than ca. 300 cm^{-1}) the local geometry of the transition structure fluctuates as a whole.

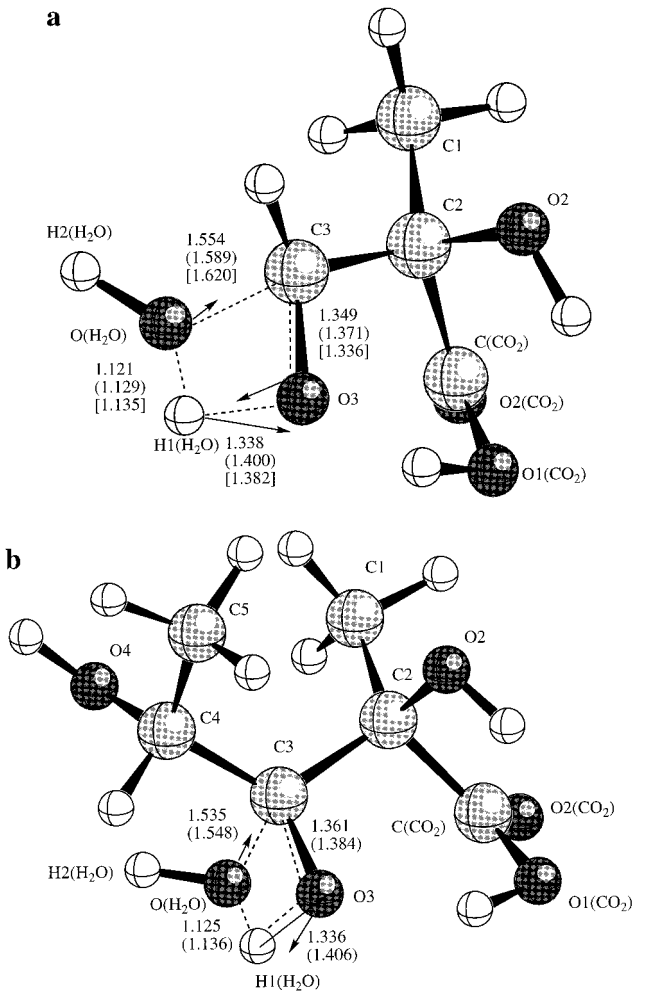

Figure 4. **TS2** for the hydration process: (a) three-C model; (b) five-C model. The arrows represent the major components of the transition vector. Main distances are indicated, calculated at HF/6-31G**, HF/3-21G (in parentheses), and MP2/6-31G** (in brackets).

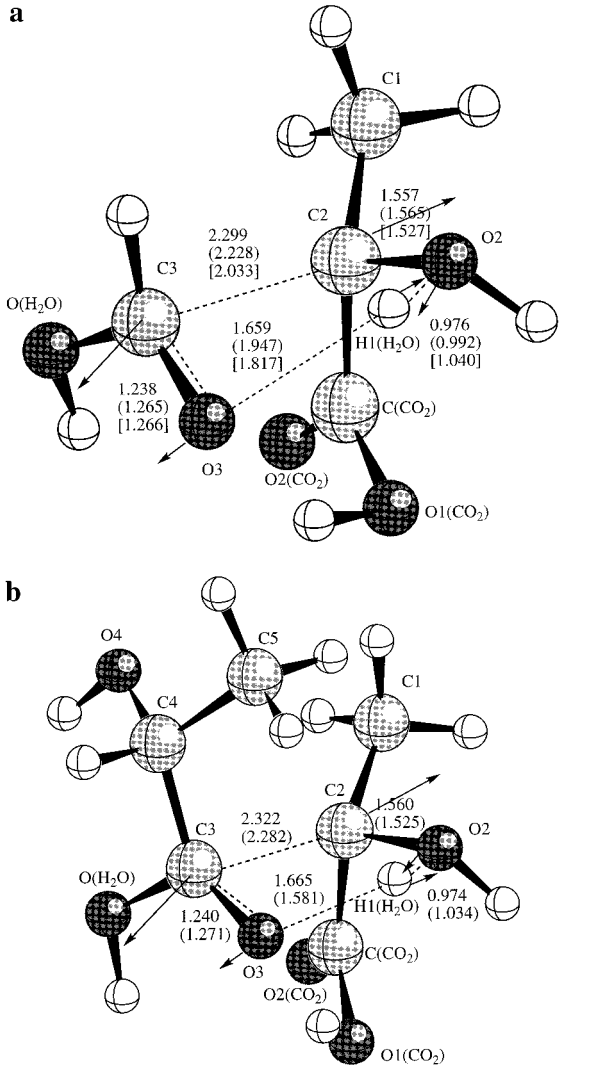

Figure 5. **TS3** for the C2−C3 bond cleavage step: (a) three-C model; (b) five-C model. The arrows represent the major components of the transition vector. Main distances are indicated, calculated at HF/6-31G**, HF/3-21G (in parentheses), and MP2/6-31G** (in brackets).

The common experimental scheme (cf. Figure 1) and our theoretical steps are molecularly identical. One of the differences resides in the previous step: instead of ressorting to an external base group to trap a proton, in the present view this could be accomplished by intramolecular hydrogen transfer.

TS2 leads directly to a *gem*-diol intermediate (G), which would correspond to the hydrated 3-keto-CABP if our model system had the phosphate groups. Experimentally, such a type of functionality has been used in connection with the mechanism of Rubisco.[3,4,53,68,70] The *gem*-diol has a strong acid character: its deprotonation will be related to the breaking of the carbon−carbon bond between C2- and C3-centers.

3.3. C2−C3 Bond Cleavage. The carbon−carbon cleavage has been a puzzling feature when coming to a theoretical explanation of the mechanism. The **TS3**, depicted in Figure 5, may offer one explanation to it. In Table 1c main distances are reported. In this model, the cleavage appears to be coupled with an intramolecular hydrogen transfer of a rather peculiar type. The hydroxyl group at C3 changes into a carbonyl and the hydrogen ends up forming a diprotonated oxygen (O2) group, thereby breaking the C2−C3 bond.

The transition vector amplitudes of this **TS3** resemble the intramolecular hydrogen transfer mechanism[18] leading from the substrate-like structure to the dienol structure. The imaginary frequency of **TS3** is much lower than the one corresponding to previous transition structures. Actually, the dominant motion corresponds to the C2−C3 fluctuation and the stationary distance is around 2.3 Å. The hydrogen is already very advanced in forming the protonated hydroxyl bound to C2. The carboxylic acid moiety presents very small fluctuations.

Following an intrinsic reaction coordinate procedure,[47] an intermediate (**AA**, see Figure 2), related to species **IV** in Figure 1, and a final product-like molecule (**P1**), are formed. In the five-carbon model, the fragment made with the C3, C4, and C5 atoms models the product, PGA, except for the fact that a phosphate group is missing there. The fragment made with C1−C2−C(CO_2) has a diprotonated oxygen bound to C2 and the stereochemistry around C2 is not yet the one found for the real product. The **AA** has a positive charge at the protonated hydroxyl group and a negative charge develops at the adjacent C2-center. Note that this species would prompt for an inversion at C2 under specific conditions by 1,2-hydrogen transfer from O2.

3.4. Configuration Inversion at the C2-Center. The **TS4** is depicted in Figure 6; some distance values are given in Table 1d. The interconversion takes place at the C1−C2−C(CO_2) fragment. It is a 1,2 hydrogen migration related to an inversion of the configuration at the C2-center. The imaginary frequency 1903i cm^{-1} corresponds to a large hydrogen fluctuation that is

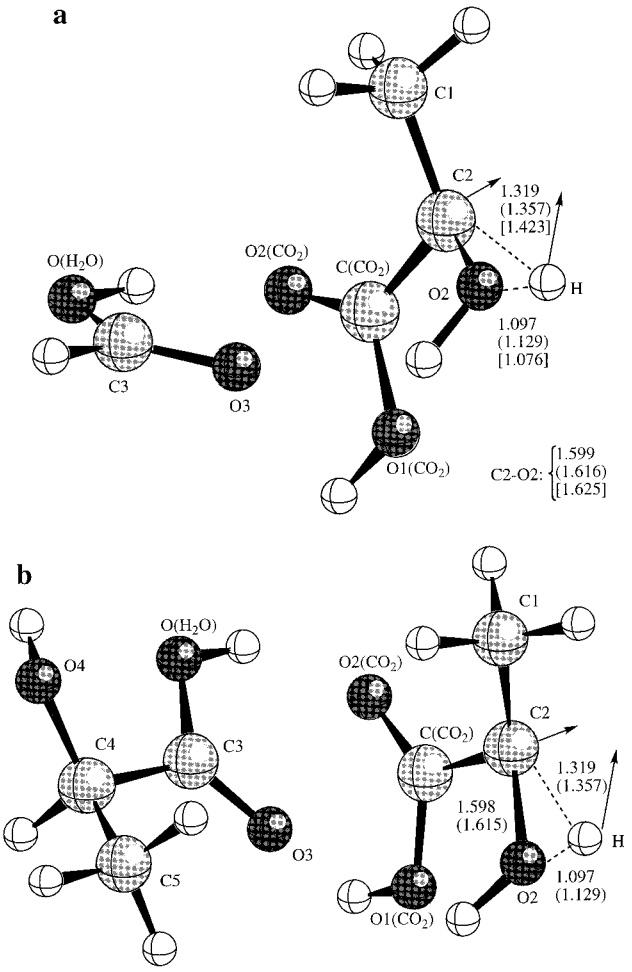

Figure 6. TS4 for the C2-inversion step: (a) three-C model; (b) five-C model. The arrows represent the major components of the transition vector. Main distances are indicated, calculated at HF/6-31G**, HF/3-21G (in parentheses), and MP2/6-31G** (in brackets).

antisymmetrically coupled to the carbon−oxygen distance fluctuation. As far as hydrogen transfer is concerned, the TS is retarded (O2−H distance around 1.1 Å and H−C2 distance 1.32 Å with a C2−O2 distance of 1.6 Å). With this transition structure the full mechanism of carboxylation−hydration of RuBP is covered.

Note that a similar TS could be obtained with no inversion of configuration. During the bond cleavage step a stereospecific protonation and inversion at C2 is required to form a stereochemically correct product. Experimentally, there is no evidence for the formation of L-PGA by Rubisco,[71] suggesting that this final step is highly specific. The present results lay down the basis to explain the proposals of Andrews and Lorimer[53] but do not solve in itself this conundrum; the experiments show that there is always inversion of configuration at C2. Thus, an explanation for this selectivity must be found in the surface complementarity between the active site of Rubisco and the corresponding TS. Further studies with enlarged molecular models are in progress in order to clarify this point. The preliminary results point toward a specific role of the carbamylated lysine 201.

4. Discussion

If the present set of transition structures were those to be found in the real system, they would provide an alternative to the standard scheme (cf. Figure 1). It yields a more simplified picture of the electronuclear events leading to the final products. It explains the nature of the "promiscuity" in this system due to the presence of transition structures having a similar geometric shapes. **TS1** by deprotonating O3 yields directly a carbonyl function on C3 and a carboxylic acid function attached to the carbon center (C2) derived from the incoming carbon dioxide. This TS avoids the creation of highly charged centers that might strongly perturb the active site. **TS2** represents the water attack leading to the *gem*-diol in **III** but with a carboxylic acid bound at C2. **TS3** produces simultaneously a C2−C3 bond breaking and a deprotonation on the oxygen O3 and a protonation of the hydroxyl group bound at C2. This transition structure is closely related to the intramolecular enolization TS reported previously.[18]

The **TS4** actually describes the inversion of configuration at C2 via an intramolecular hydrogen rearrangement. Obviously, to obtain this TS, one has selected a particular path from the precedent intermediate. If the real substrate follows the steps signaled by the TSs herein reported, there is a need for an external control of this hydrogen trafficking. Note that this study is not intended to prove that there is no need of the enzyme to carry on the chemistry. Quite the contrary, it would help to disentangle what is due to the enzyme and what might be due to the actual reacting subsystem. C2-inversion clearly depends on the enzyme surroundings. An alternative mechanism for the C2−C3 bond breaking step and configuration inversion at C2 was suggested by a referee. The cleavage would be coupled with a hydrogen transfer to $O_2(CO_2)$, leading to a product model molecule and an enetriol intermediate. Preliminary calculations suggest this hypothesis has much potential, the results will be reported in due time. However, this possibility does not eliminate the need for an enzyme action to explain the observed stereoselectivity in products formation.

The intramolecular proton reshuffling appears to be an alternative mechanism by which the carboxylation chemistry of Rubisco can be explained. As a similar result is obtained for the self-inhibitory chemistry,[18] and for the oxygenation chemistry,[9,72] one may conclude that it is the special geometry of the transition structures into which the substrate is molded that opens such pathways. They are then unavoidable to a certain extent.

A comparative normal-mode analysis between the three- and five-carbon models showed that the transition vector is mostly localized to the atoms undergoing interconversion. The imaginary frequency changes little from one model to the other. The change of basis set does not produce important geometric and fluctuation pattern variations among the models and between different basis sets.

The Born−Oppenheimer free energy values for the calculated transition structures, relative to the enediol deformed in a geometry compatible with the active site, are presented in Figure 7 for the different molecular models and theoretical levels of calculation. As one can expect, the results are dependent on these factors. Among the TSs reported, it seems that **TS3** and/or **TS4** will be those with the highest energy, if we do not take into account the possible role of the TS for the intramolecular enolization. Such high energy barriers correspond to the process in vacuo. There are many ways the protein may modulate this profile.[73,74] An enzymatically meaningful evaluation of energetics must take into account the multiple interactions between the substrate and the enzyme residues at the active center, as follows from Tommos and Babcock in a recent work on oxygen evolution chemistry in Photosystem II:[75] this is the goal of future studies using QM/MM methodology.

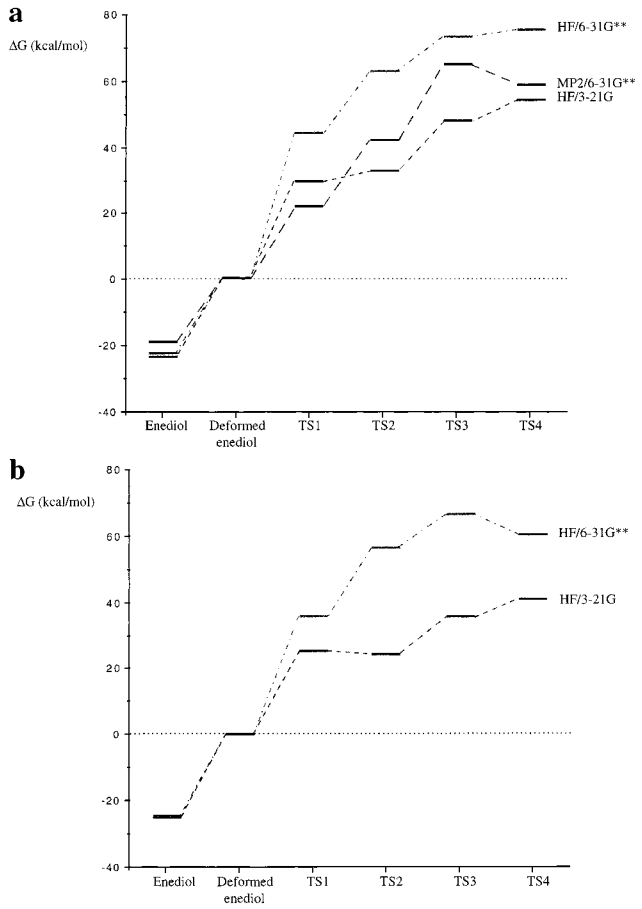

Figure 7. Free energy profile for the carboxylation pathway for model systems: (a) three-carbon model; (b) five-carbon model.

To what extent the arguments given above concerning TS invariances are wishful thinking or, on the contrary, express something meaningful is a key problem warranting the working hypothesis used in this and related works. To get a better insight on this problem of invariance, let us analyze the Hessian. In a neighborhood of the stationary geometry $\rho^o = (\rho_1^o, ..., \rho_m^o)$ for a set of m nuclei, the Hessian can be cast, for the sake of analysis, in an internal coordinate set and the system ordered so as to be possible to describe the matrix as a partitioned matrix **H**:[76]

$$\mathbf{H} = \begin{vmatrix} H_{CC} & H_{CA} & H_{CP} \\ H_{AC} & H_{AA} & H_{AP} \\ H_{PC} & H_{PA} & H_{PP} \end{vmatrix}$$

where CC stands for the degrees of freedom included in the minimal (core) molecular model, AA may be the active site elements from the protein, and PP is the symbol for the remaining enzyme. The couplings between the subspaces are represented by the rectangular matrices H_{CA}, H_{CP}, etc. In this paper, studies of different sizes of the minimal molecular model and level of theory are examined in great detail. Magnesium and residues of the coordination shell belong to the AA subspace. To substantiate the arguments concerning TS invariances, a step beyond has been taken along the computation path by using the five-C model and a model of the magnesium coordination sphere found at the active center of Rubisco. This includes the magnesium cation itself, a carbamylated ammonia that models carbamylated Lys 201, and two formiates, representing Asp 202 and Glu 204. These groups make up the AA subspace and its coupling with the CC subspace discussed here.

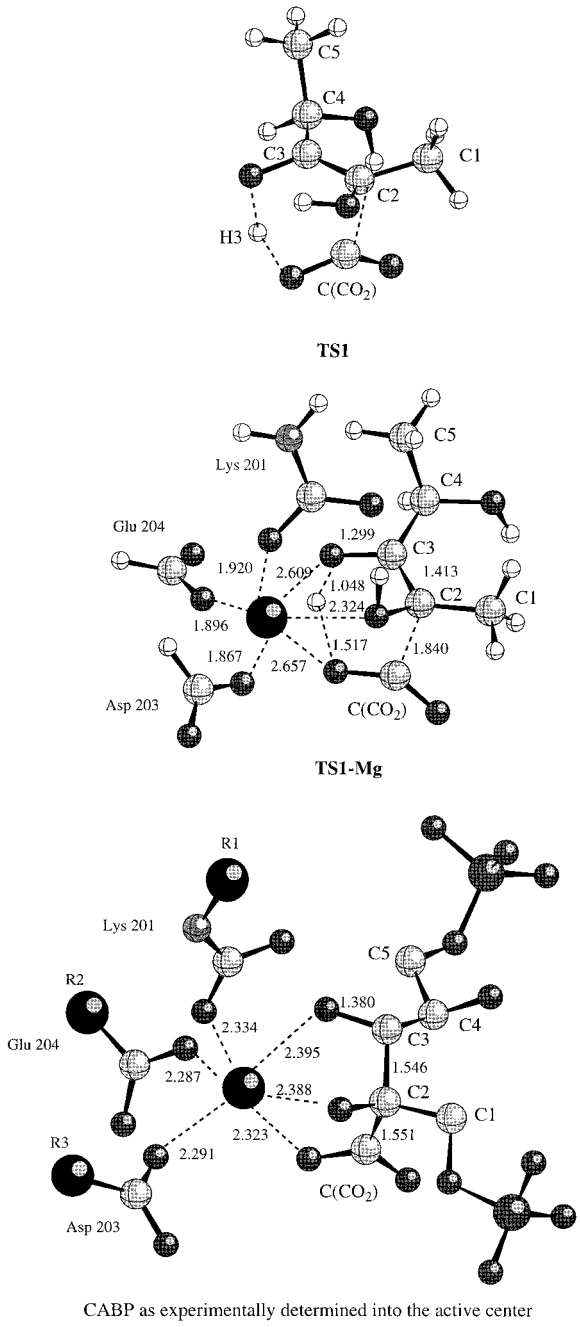

Figure 8. Comparison between five-C model **TS1**, an enlarged molecular model that includes the magnesium coordination sphere (**TS1**−**Mg**), and the experimentally determined CABP structure.[7] Main distances are indicated for the two latter. The geometric data for the third structure has been obtained from Protein Data Bank, entry number 8RUC.

The equivalent to the **TS1** was determined at a HF/3-21G level of theory as a saddle point of index 1 (**TS1**−**Mg**). This choice was made because a comparison can be made with the CABP structure at the active site.[7] In Figure 8 the **TS1** and the **TS1**−**Mg** are depicted. CABP is also included. From the reported results (compare the geometric parameters in Figure 8 with the same parameters in Figure 3b) the invariance of the geometry can be sensed. The transition vector components remain also invariant between these transition structures. Furthermore, it is apparent that the spatial thread of the C-framework of the three structures presented superpose fairly well. The idea of the invariance in transition vector amplitudes and geometry (not the energy) when a comparison is made between the minimal

molecular system in vacuo and those models incorporating the residues found at the active site is reinforced by these results. Furthermore, preliminary calculations[11] of the Hessian containing a substantial part of the enzyme using semiempirical MO methods of the kind QM/MM including the PP subspace and couplings with the AA and CC subspaces show again this remarkable property. The transition vector remains almost invariant despite the couplings.

Acknowledgment. We wish to thank Profs. G. Lorimer and I. Andersson for invaluable comments and reprints of their work on Rubisco, and we hope the theoretical study is up to the challenge conveyed by this enigmatic enzyme. This work received support from the DGICYT (Project PB93-0661) and DGES (Project PB96-0795-C02-02). The authors are grateful to the Servei d'Informàtica de la Universitat Jaume I for providing them with CPU time on two Silicon Graphics Power Challenge L. M.O. thanks the Ministerio de Educación y Ciencia and the Universitat Jaume I for FPI fellowships. O.T. thanks NFR for financial support.

References and Notes

(1) Andrews, T. J.; Hatch, M. D. *Biochem. J.* **1969**, *114*, 117.
(2) Robinson, S. P.; Portis, A. R., Jr. *Plant Physiol.* **1989**, *90*, 968.
(3) Schneider, G.; Lindqvist, Y.; Brändén, C.-I. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 119.
(4) Hartman, F. C.; Harpel, M. R. *Annu. Rev. Biochem*. **1994**, *63*, 197.
(5) Lee, E. H.; Harpel, M. R.; Chen, Y.-R.; Hartman, F. C. *J. Biol. Chem*. **1993**, *268*, 26583.
(6) Cleland, W. W.; Andrews, T. J.; Gutteridge, S.; Hartman, F. C.; Lorimer, G. H. *Chem. Rev*. **1998**, *98*, 549.
(7) Andersson, I. *J. Mol. Biol*. **1996**, *259*, 160.
(8) Taylor, T. C.; Andersson, I. *J. Mol. Biol.* **1997**, *265*, 432.
(9) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *J. Phys. Chem.* **1999**, *103*, 6009.
(10) Safont, V. S.; Oliva, M.; Andres, J.; Tapia, O. *Chem. Phys. Lett*. **1997**, *278*, 291.
(11) Moliner, V.; Andrés, J.; Oliva, M.; Safont, V. S.; Tapia, O. *Theor. Chem. Acc.* **1999**, *101*, 228.
(12) Tapia, O.; Andres, J. *Mol. Eng*. **1992**, *2*, 37.
(13) Andrés, J.; Safont, V. S.; Tapia, O. *Chem. Phys. Lett.* **1992**, *198*, 515.
(14) Andrés, J.; Safont, V. S.; Queralt, J.; Tapia, O. *J. Phys. Chem.* **1993**, *97*, 7888.
(15) Tapia, O.; Andrés, J.; Safont, V. S. *J. Phys. Chem.* **1994**, *98*, 4821.
(16) Tapia, O.; Andrés, J.; Safont, V. S. *J. Chem. Soc., Faraday Trans*. **1994**, *90*, 2365.
(17) Tapia, O.; Andrés, J.; Safont, V. S. *J. Mol. Struct. (THEOCHEM)* **1995**, *342*, 131.
(18) Tapia, O.; Andres, J.; Safont, V. S. *J. Phys. Chem.* **1996**, *100*, 8543.
(19) Hinshelwood, C. N. *Chem. Soc. Rev.* **1930**, *27*, 20.
(20) Mezey, P. G. In *Computational theoretical organic chemistry*; Csizmadia, I. G., Daudel, R., Eds.; Reidel: Dordrecht, The Netherlands, 1981; p 101.
(21) Mezey, P. G. *Theor.Chim. Acta (Berlin)* **1982**, *62*, 133.
(22) Mezey, P. G. *Potential energy hypersurfaces*; Elsevier: Amsterdam, 1987; p 538.
(23) Tantillo, D. J.; Chen, J.; Houk, K. N. *Curr. Opin. Chem. Biol*. **1998**, *2*, 743.
(24) Houk, K. N.; González, J.; Li, Y. *Acc. Chem. Res.* **1995**, *28*, 81.
(25) Williams, I. H. *Chem. Soc. Rev.* **1993**, 277.
(26) Tapia, O.; Paulino, M.; Stamato, F. M. L. G. *Mol. Eng*. **1994**, *3*, 377.
(27) Wigner, E. *Trans. Faraday Soc.* **1938**, *34*, 29.
(28) Neumark, D. M. *Acc. Chem. Res.* **1993**, *26*, 33.
(29) Zewail, A. H. *J. Phys. Chem.* **1996**, *100*, 12701.
(30) Tapia, O.; Andres, J.; Stamato, F. M. L. G. In *Solvent effects and chemical reactivity*; Tapia, O., Bertran, J., Eds.; Kluwer: Dordrecht, The Netherlands, 1996; p 283.
(31) Lerner, R. L.; Tramontano, A. *Trends Biochem. Sci*. **1987**, *12*, 427.
(32) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1988**, *234*, 1566.
(33) Schultz, P. G.; Lerner, R. A. *Acc. Chem. Res.* **1993**, *26*, 391.
(34) Reymond, J. L.; Jahangiri, G. K.; Stoudt, C.; Lerner, R. A. *J. Am. Chem. Soc.* **1993**, *115*, 3909.
(35) Gouverneur, V. E.; Houk, K. N.; Pascual-Teresa, B.; Beno, B.; Janda, K. D.; Lerner, R. A. *Science* **1993**, *262*, 204.
(36) Neumark, D. M. *Science* **1996**, *272*, 1446.
(37) Wenthold, P. G.; Hrovat, D. A.; Borden, W. T.; Lineberger, W. C. *Science* **1996**, *272*, 1456.
(38) Wolfenden, R.; Radzicka, A. *Curr. Opin. Struct. Biol*. **1991**, *1*, 780.
(39) Schröer, J.; Sanner, M.; Reymond, J.-L.; Lerner, R. A. *J. Org. Chem.* **1997**, *62*, 3220.
(40) Houk, K. N.; Rondan, N. G.; von Rague Schleyer, P.; Kaufmann, E.; Clark, T. *J. Am. Chem. Soc*. **1985**, *107*, 2821.
(41) Wu, Y.-D.; Houk, K. N. *J. Am. Chem. Soc.* **1987**, *109*, 906.
(42) Tapia, O.; Cárdenas, R.; Andrés, J.; Colonna-Cesari, F. *J. Am. Chem. Soc.* **1988**, *110*, 4046.
(43) Houk, K. N.; Gustafson, S. M.; Black, K. A. *J. Am. Chem. Soc.* **1992**, *114*, 8565.
(44) Wu, Y. D.; Lai, D. K. W.; Houk, K. N. *J. Am. Chem. Soc*. **1995**, *117*, 4100.
(45) Tapia, O.; Andres, J.; Moliner, V.; Stamato, F. M. L. G. In *Theoretical tratments of hydrogen bonding*; Hadzi, D., Ed.; Wiley & Sons: Chichester, U.K., 1997.
(46) Tapia, O. *J. Mol. Struct. (THEOCHEM)* **1998**, *433*, 95.
(47) Fukui, K. *J. Phys. Chem.* **1970**, *74*, 4161.
(48) Pauling, L. *Chem. Eng. News* **1946**, *24*, 1375.
(49) Pauling, L. *Nature* **1948**, *161*, 707.
(50) Pauling, L. *Am. Sci.* **1948**, *36*, 51.
(51) 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. *GAUSSIAN94*, Revision B.1; Gaussian, Inc.: Pittsburgh, PA, 1995
(52) *GaussView1.0*; Gaussian, Inc.: Pittsburgh, PA, 1997.
(53) Andrews, T. J.; Lorimer, G. H. In *The biochemistry of plants*; Hatch, M. D., Boardman, N. K., Eds.; Academic Press: New York, 1987; Vol. 10; p 131.
(54) Knight, S.; Andersson, I.; Brändén, C.-I. *J. Mol. Biol*. **1990**, *215*, 113.
(55) Gutteridge, S.; Rhoades, D.; Herrmann. *J. Biol. Chem*. **1993**, *268*, 7818.
(56) Lorimer, G. H.; Gutteridge, S.; Madden, M. W. In *Plant Molecular Biology*; v. Wettstein, D., Chua, N. H., Eds.; Plenum Press: New York, 1987; p 21.
(57) Saver, B. G.; Knowles, J. R. *Biochemistry* **1982**, *21*, 5398.
(58) Sue, J. M.; Knowles, J. R. *Biochemistry* **1982**, *21*, 5404.
(59) Hartman, F. C.; Soper, T. S.; Niyogi, S. K.; Mural, R. J.; Foote, R. S.; Mitra, S.; Lee, E. H.; Machanoff, R.; Larimer, W. F. *J. Biol. Chem.* **1987**, *262*, 3496.
(60) Lorimer, G. H.; Hartman, F. C. *J. Biol. Chem.* **1988**, *263*, 6468.
(61) Schloss, J. V. In *Enzymatic and model carboxylation and reduction reactions for carbon dioxide utilization*; Aresta, M., Schloss, J. V., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands; 1990; p 321.
(62) Newman, J.; Gutteridge, S. *J. Biol. Chem.* **1993**, *268*, 25876.
(63) Taylor, T. C.; Andersson, I. *Nature Struct. Biol.* **1996**, *3*, 95.
(64) Taylor, T. C.; Fothergill, D.; Andersson, I. *J. Biol. Chem.* **1996**, *271*, 32894.
(65) Taylor, T. C.; Andersson, I. *Biochemistry* **1997**, *36*, 4041.
(66) Jaworowski, A.; Rose, I. A. *J. Biol. Chem.* **1985**, *260*, 944.
(67) Pierce, J.; Andrews, T. J.; Lorimer, G. H. *J. Biol. Chem.* **1986**, *261*, 10248.
(68) Schneider, G. In *Carbon Dioxide Fixation and Reduction in Biological and Model Systems*; Brandén, C. I., Schneider, G., Eds.; Oxford University Press: Oxford, U.K., 1994.
(69) Spangler, D.; Williams, I. H.; Maggiora, G. M. *J. Comput. Chem.* **1983**, *4*, 524.
(70) Miziorko, H. M.; Lorimer, G. H. *Annu. Rev. Biochem.* **1983**, *52*, 507.
(71) Andrews, T. J.; Kane, H. J. *J. Biol. Chem.* **1991**, *266*, 9447.
(72) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *Chem. Phys. Lett.* **1998**, *294*, 87.
(73) Tapia, O.; Johannin, G. *J. Chem. Phys.* **1981**, *75*, 3624.
(74) Warshel, A. *Curr. Opin. Struct. Biol.* **1992**, *2*, 230.
(75) Tommos, C.; Babcock, G. T. *Acc. Chem. Res.* **1998**, *31*, 18.
(76) Lowdin, P. O. *J. Mol. Spectrosc.* **1963**, *10*, 12.
(77) Tapia, O.; Andres *J. Chem. Phys. Lett.* **1984**, *109*, 471.
(78) Schlegel, H. B. *J. Comput. Chem.* **1982**, *3*, 214.
(79) Schlegel, H. B. *J. Chem. Phys.* **1982**, *77*, 3676.