CO_2 Fixation by Rubisco: Computational Dissection of the Key Steps of Carboxylation, Hydration, and C−C Bond Cleavage

Harald Mauser,[‡] William A. King,[‡] Jill E. Gready,*[,‡] and T. John Andrews[§]

Contribution from the Computational Molecular Biology and Drug Design Group, John Curtin School of Medical Research, and Molecular Plant Physiology Group, Research School of Biological Sciences, Australian National University, Canberra ACT 0200, Australia

Received June 4, 2001

Abstract: Despite intensive experimental and computational studies, some important features of the mechanism of the photosynthetic CO_2-fixing enzyme, Rubisco, are still not understood. To complement our previous investigation of the first catalytic step, the enolization of D-ribulose-1,5-bisphosphate (King et al., *Biochemistry* **1998**, *44*, 15414−15422), we present the first complete computational dissection of subsequent steps of the carboxylation reaction that includes the roles of the central magnesium ion and modeled residues of the active site. We investigated carboxylation, hydration, and C−C bond cleavage using the density functional method and the B3LYP/6-31G(d) level to perform geometry optimizations. The energies were determined by B3LYP/6-311+G(2d,p) single-point calculations. We modeled a fragment of the active site and substrate, taking into account experimental findings that the residues coordinated to the Mg ion, especially the carbamylated Lys-201, play critical roles in this reaction sequence. The carbamate appears to act as a general base, not only for enolization but also for hydration of the β ketoacid formed by addition of CO_2 and, as well, cleavage of the C2−C3 bond of the hydrate. We show that CO_2 is added directly, without assistance of a Michaelis complex, and that hydration of the resultant β ketoacid occurs in a separate subsequent step with a discrete transition state. We suggest that two conformations of the hydrate (*gem*-diol), with different metal coordination, are possible. The step with the highest activation energy during the carboxylation cycle is the C−C bond cleavage. Depending on the conformations of the *gem*-diol, different pathways are possible for this step. In either case, special arrangements of the metal coordination result in bond breaking occurring at remarkably low activation energies (between 28 and 37 kcal mol^{-1}) which might be reduced further in the enzyme environment.

Introduction

D-Ribulose-1,5-bisphosphate (RuBP) carboxylase-oxygenase (Rubisco, EC 4.1.1.39) catalyzes the fixation of CO_2 in photosynthesis. The minimum functional unit of the enzyme is a dimer of 50−55 kDa large subunits which bears two identical active sites at the interface between the subunits with residues of both subunits contributing to each active site. Some bacterial and dinoflagellate Rubiscos (form II) consist of this dimer alone, often assembled into higher oligomers. The common form (form I), encountered in plastids, cyanobacteria, and some bacteria, is a cubical tetramer of these dimers held together with tetrameric sets of 12−18 kDa small subunits on two opposing faces of the cube.[1−5] Rubisco is activated by carbamylation of a specific Lys residue in the active site by a CO_2 molecule distinct from the substrate CO_2 that is fixed during carboxylation.[6] The carbamylated Lys residue not only completes the binding site for the other cofactor, a Mg ion,[7] but also appears to play a critical role as a proton-abstracting base in several steps of the catalytic cycle.[8,9]

Despite its unique position in the evolution of life, Rubisco seems inefficient because its catalytic rate is very slow and, in a competitive oxygenation reaction, it wastes carbon and energy.[1−4] Many aspects of the enzyme's complex structure and function suggest it is a compromise solution that enables quite difficult chemistry. The multistep reaction sequence (Scheme 1) involves as many as four enzyme-bound intermediates whose instability gives rise to multiple side reactions that further compromise efficiency. Therefore, one of the aims of current research is to enhance the performance of Rubisco by genetic engineering, leading to more resource-efficient crops.[10]

Despite intensive experimental studies, key features of the mechanism are still not entirely understood. For example, there is uncertainty about whether the critical CO_2 addition to C2 of the enediol of RuBP occurs as a discrete step or is concerted with the hydration of the C3 carbonyl so formed (Scheme 1).[9,11] The subsequent step, in which the C−C bond is cleaved, has also been difficult to investigate experimentally.

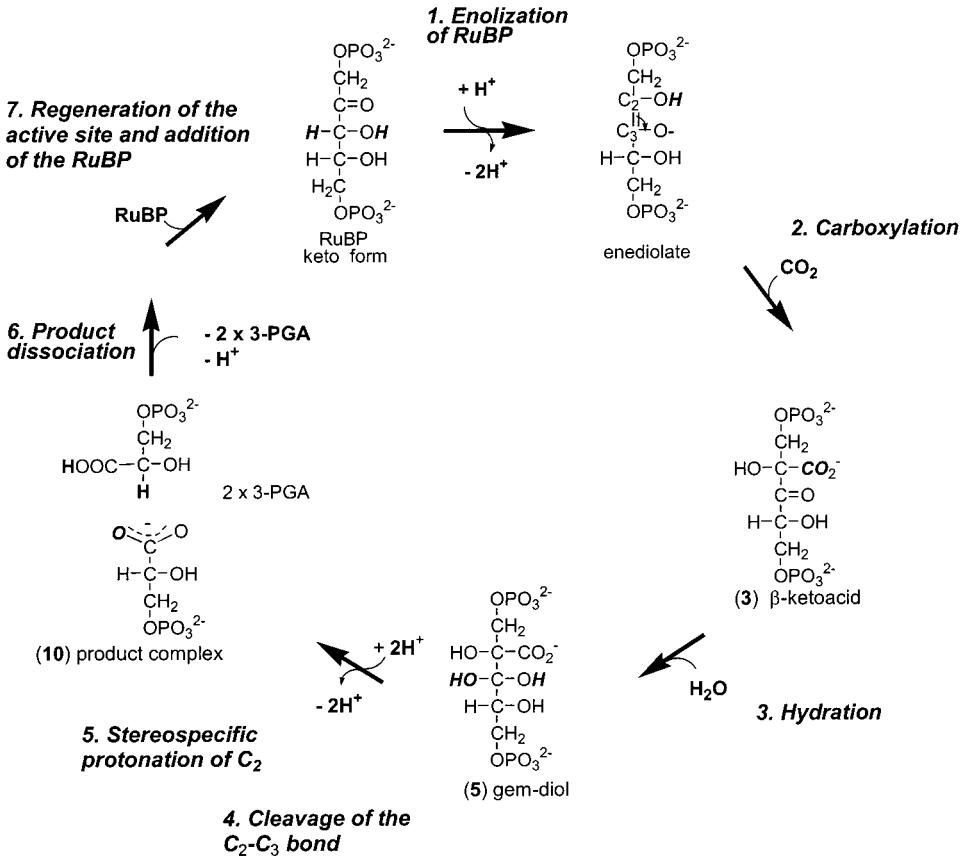
To explore these and other aspects of the multistep catalytic cycle (Scheme 1) and to interpret experimental findings, we

* Corresponding author. E-mail: Jill.Gready@anu.edu.au.
‡ John Curtin School of Medical Research.
§ Research School of Biological Sciences.

(1) Andrews, T. J.; Lorimer, G. H. In *The Biochemistry of Plants: A Comprehensive Treatise, Vol. 10, Photosynthesis*; Hatch, M. D., Boardman, N. K., Eds.; Academic Press: New York, 1987; pp 131−218.
(2) Hartman, F. C.; Harpel, M. R. *Annu. Rev. Biochem.* **1994**, *63*, 197−234.
(3) Gutteridge, S.; Gatenby, A. A. *Plant Cell* **1995**, *7*, 809−819.
(4) Roy, H.; Andrews, T. J. In *Photosynthesis: Physiology and Metabolism*; Leegood, R. C., Sharkey, T. D., von Caemmerer, S., Eds.; Kluwer Academic Publishers: Dordrecht, 2000; pp 53−83.
(5) Delwiche, C. F.; Palmer, J. D. *Mol. Biol. Evol.* **1996**, *13*, 873−882.
(6) Lorimer, G. H. *Biochemistry* **1981**, *20*, 1236−1240.
(7) Lorimer, G. H.; Badger, M. R.; Andrews, T. J. *Biochemistry* **1976**, *15*, 529−536.
(8) Taylor, T. C.; Andersson, I. *J. Mol. Biol.* **1997**, *265*, 432−444.
(9) Cleland, W. W.; Andrews, T. J.; Gutteridge, S.; Hartman, F. C.; Lorimer, G. H. *Chem. Rev.* **1998**, *98*, 549−561.
(10) Mann, C. C. *Science* **1999**, *283*, 314−316.
(11) Cleland, W. W. *Biochemistry* **1990**, *29*, 3194−3197.

10.1021/ja011362p CCC: $20.00
Published on Web 10/12/2001

Scheme 1. Multistep Mechanism for the Carboxylation Reaction Catalyzed by Rubiso[a]



[a] Shown are the key steps of carboxylation, hydration (see Scheme 2), and C−C cleavage (see Schemes 3 and 4) leading from the 2,3-enediolate of RuBP to two molecules of 3-phospho-D-glycerate.

undertook computational studies on model systems for the addition of CO_2 to enolized RuBP and the subsequent hydration and cleavage of the resulting β ketoacid intermediate. In contrast to recent studies of the "nonenzymatic" [12] reaction pathway,[12−19] we included the central metal ion and some residues of the active site in our calculations to achieve a more realistic representation.

Methods

Calculations were performed using the Gaussian 98[20] program package. We used the density functional theory (DFT) approach to include the effects of electron correlation. The B3LYP hybrid functional[21] generally yields accurate properties,[22−24] and we found in our previous study[25] that the DFT calculations yielded results in close agreement with those of the MP2 method. Especially in the large-scale calculations required for this study, DFT calculations are far more efficient than conventional ab initio correlated methods.

All stationary points were optimized with the B3LYP functional and the 6-31G(d) basis set. A vibrational analysis was performed on all optimized structures. The imaginary frequency modes of transition structures were visualized to check that they connected reactant and product states. In complex cases, where several protonation and deprotonation steps were involved, we calculated intrinsic reaction coordinates. Following literature suggestions,[26] the 6-311G+(2d,p) basis set was used to calculate reliable energies.

All stationary points on the reaction path were optimized without constraints to meet the convergence criteria of a maximum step size of 0.0018 au and an RMS force of 0.0003 au. Unless indicated otherwise, Mg-coordinated residues were fixed, and the substrate was optimized to more relaxed convergence criteria (maximum step size = 0.01 au, RMS force = 0.0017 au) during scans of the potential energy surfaces (Figure 4).

(12) Zhan, C. G.; Niu, S. Q.; Ornstein, R. L. *J. Chem. Soc., Perkin Trans. 2* **2001**, 23−29.
(13) Safont, V. S.; Oliva, M.; Andrés, J.; Tapia, O. *Chem. Phys. Lett.* **2000**, *318*, 361−369.
(14) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *J. Phys. Chem. A* **1999**, *103*, 8725−8732.
(15) Andrés, J.; Oliva, M.; Safont, V. S.; Moliner, V.; Tapia, O. *Theor. Chem. Acc.* **1999**, *101*, 234−240.
(16) Safont, V. S.; Oliva, M.; Andrés, J.; Tapia, O. *Chem. Phys. Lett.* **1997**, *278*, 291−296.
(17) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *Chem. Phys. Lett.* **1998**, *294*, 87−94.
(18) Tapia, O.; Andrés, J.; Safont, V. S. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 2365−2374.
(19) Tapia, O.; Andrés, J.; Safont, V. S. *J. Phys. Chem.* **1994**, *98*, 4821−4830.
(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(22) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000.
(23) Miani, A.; Cane, E.; Palmieri, P.; Trombetti, A.; Handy, N. C. *J. Chem. Phys.* **2000**, *112*, 248−259.
(24) Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2000**, *316*, 160−166.
(25) King, W. A.; Gready, J. E.; Andrews, T. J. *Biochemistry* **1998**, *37*, 15414−15422.
(26) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian, Inc.: Pittsburgh, PA, 1996.

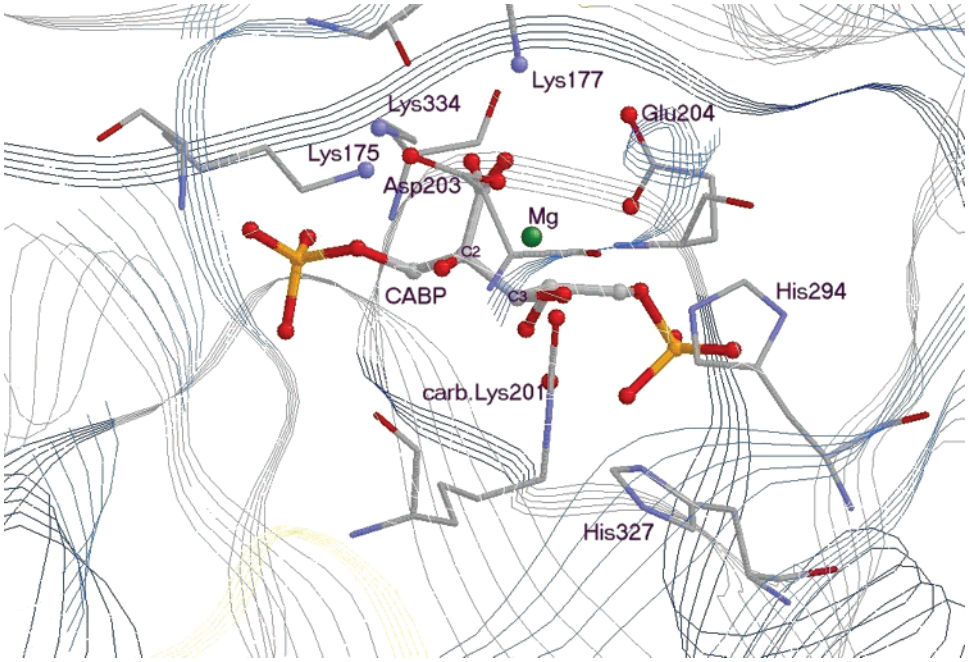
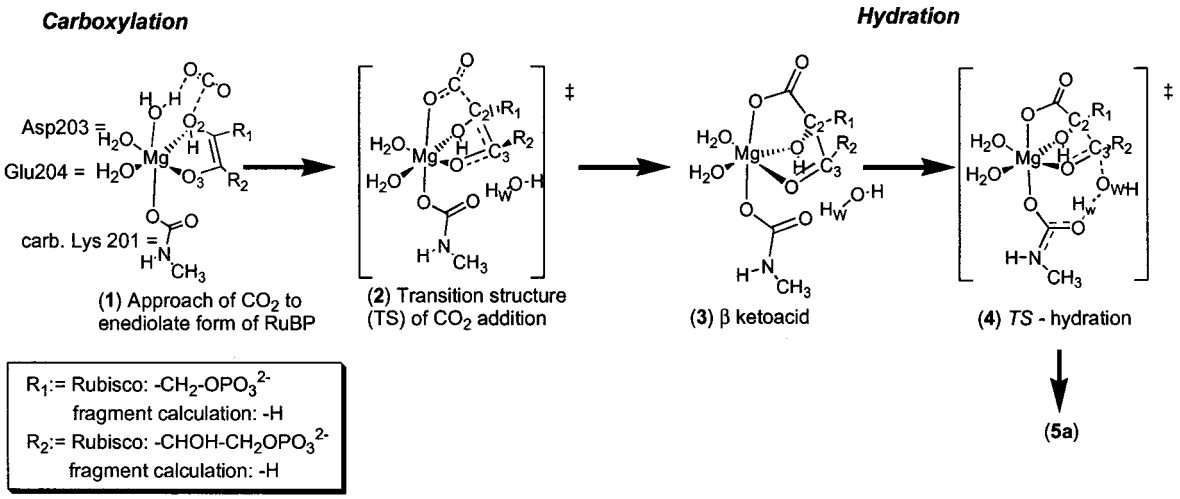


Figure 1. Active site of Rubisco according to the X-ray structure of activated spinach Rubisco.[27] The residues close to the Mg ion are shown explicitly; for other residues only the backbone is drawn.

Scheme 2. Carboxylation and Hydration Reactions Starting at the Approach of CO_2 to the Enediolate Form of RuBP[a]



[a] Schematic representations for the structures obtained with the active site fragment model are shown (R_1 = R_2 = H). "‡" indicates a first-order transition state.

The active site fragment model was constructed in a manner analogous to that described in our previous study (see ref 25 for a detailed description). The starting coordinates were taken from the 1.6 Å crystal structure of spinach Rubisco[27] (see Figure 1) and used for a constrained molecular dynamics (MD) simulation. The cocrystallized 2′-carboxyarabinitol 1,5-bisphosphate (CABP) ligand, an analogue of the β ketoacid intermediate (**3**, Scheme 2), was replaced by RuBP and water to reproduce the initial complex. We included the complete environment within 22 Å of the magnesium ion in the simulation. This included residues of two large and two small subunits of the enzyme and 279 water molecules present in the crystal structure. We added a further 238 water molecules within an approximately 28 Å sphere. We constrained the geometry of the magnesium coordination sphere and the binding orientation of the two phosphate groups of RuBP but allowed other atoms in the sphere to be flexible during the MD simulations (simulation time 1 ns, initial temperature 10 K, simulation temperature 300 K, coupled to a heat bath with the time constant of 0.2 ps). The starting coordinates for the 29-fragment model were then constructed out of six sample sets generated at 5 ps simulation intervals.

To achieve a suitable representation of the electrostatics within the active site and to reduce the number of heavy atoms in the calculation, we replaced Asp-203 and Glu-204, whose carboxylate groups are coordinated to the magnesium ion, with water molecules (see Figure 1 and structure **1** in Scheme 2). As pointed out by Siegbahn and Blomberg,[28] it is crucial to model the charge distribution within the active site by finding a balance between charged groups included and excluded from the fragment considered in the calculation. In test calculations, we found that the repulsion due to close contact of two negatively charged ligands overcompensated the attraction of the magnesium ion. There is structural evidence that, even if not neutralized by addition of protons, the overall charge on Asp-203 and Glu-204 is reduced by the interaction with the Lys-175, Lys-177, and His-294 residues.[27] Considering the partial covalent character of the bonds (see population analysis in Results) between the Mg ion and its ligands and the neglected interactions with the second coordination shell, we found an overall neutral complex to be the most suitable description for the active site of Rubisco. We would not expect structures or energies

(27) Andersson, I. *J. Mol. Biol.* **1996**, *259*, 160−174.

(28) Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221−249.

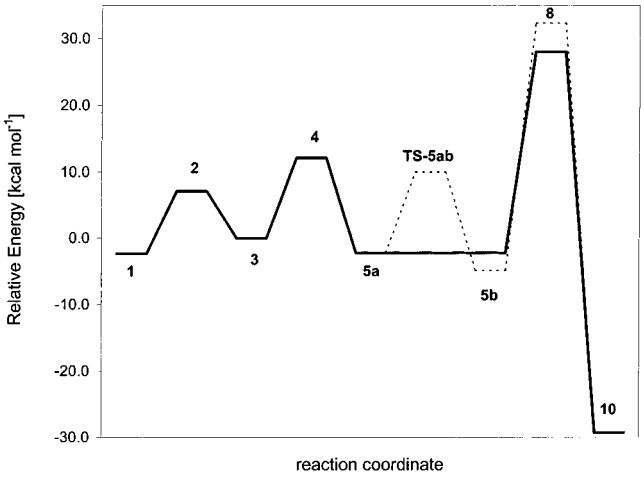


Figure 2. Energy profile of the Rubisco-catalyzed carboxylation of the 2,3-enediolate of RuBP. The dotted line represents an isomerization of *gem*-diol form I (**5a**) to form II (**5b**) via the estimated transition state TS-**5ab** and subsequent C−C cleavage (pathway II in Scheme 4). The energy according to the B3LYP/6-311+G(2d,p)//6-31G(d) results is relative to the β ketoacid (**3**) and corresponds to column 3 in Table 1.

Table 1. Energy Differences (in kcal mol^{-1}) and Bond Distances for the Reaction Steps of the Rubisco-Catalyzed Carboxylation of the 2,3-Enediolate of RuBP

complex[a]	B3LYP/ 6-31(d)[b]	B3LYP/ 6-311+G(2d,p)// 6-31(d)	distance C2−C3 (Å)
1	−1.9	−2.3	1.35
2	5.0	7.1	1.38
3[c]	0.0	0.0	1.51
4	9.3	12.1	1.54
5a	−3.4	−2.2	1.54
5b	−4.4	−4.8	1.53
8a	24.8	28.0	2.40
8b	28.0	32.3	2.00
10	−30.9	−29.2	4.67

[a] The complex numbering refers to Schemes 2 and 3. [b] Zero-point energies are included. [c] Reference energy: −1130.420284 au and −1130.800571 au for the 6-31(d) and 6-311+G(2d,p) basis sets, respectively.

calculated with this complex to differ significantly from those that might be obtained from calculations with a larger complex that included the outer coordination sphere of the Mg ion and carboxylate groups rather than water molecules.

The steric restraints of the environment were checked by modeling the resulting structures into the active site by replacing the analogous atoms of the CABP ligand in the 1.6 Å X-ray crystal structure of activated Rubisco.[27]

The starting point of our present calculations is the enediolate form of RuBP. As previously,[25] only the C2 and C3 carbon atoms of the substrate are included; therefore, our starting species is ethene-1,2-diol-2-ate.

Results

Addition of CO_2. Structures on the reaction pathway during the carboxylation and hydration of RuBP are shown in Scheme 2; the corresponding energies are given in Figure 2 and Table 1. The first step after enolization is the approach of carbon dioxide to the enediolate form of RuBP (**1**). In this very weakly bound complex, CO_2 is located in the outer coordination sphere with a water molecule still remaining in the upper vertical position of the octahedral magnesium complex. Following the normal mode, the carbon atom of CO_2 moves toward the C2 carbon of RuBP to form the transition structure (**2**), in which the Mg-coordinated water is replaced by CO_2. Consistent with the experimental observations,[4] a stable noncovalent Michaelis complex where the coordinated water molecule is supplanted by carbon dioxide could not be calculated. Rather, in the simulation, the reverse substitution (water for carbon dioxide) was favored. After overcoming an activation barrier of 9 (7) kcal mol^{-1} (result for the 6-311+G(2d,p) basis set; result for the 6-31G(d) basis set is given in parentheses, see Table 1), the β ketoacid (**3**) is formed as a stable intermediate, corresponding to the 3-keto-2′-carboxy-arabinitol-P_2 in the enzymatic reaction.

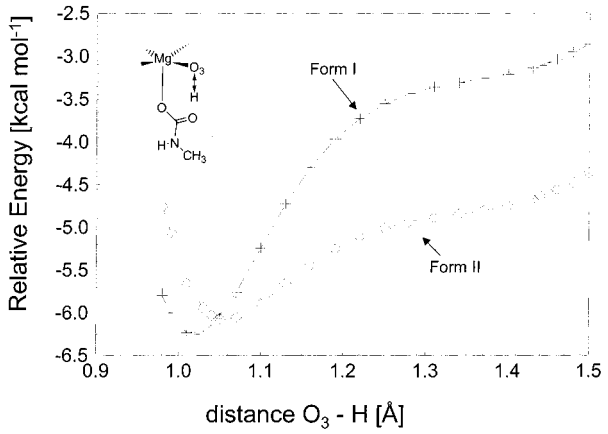
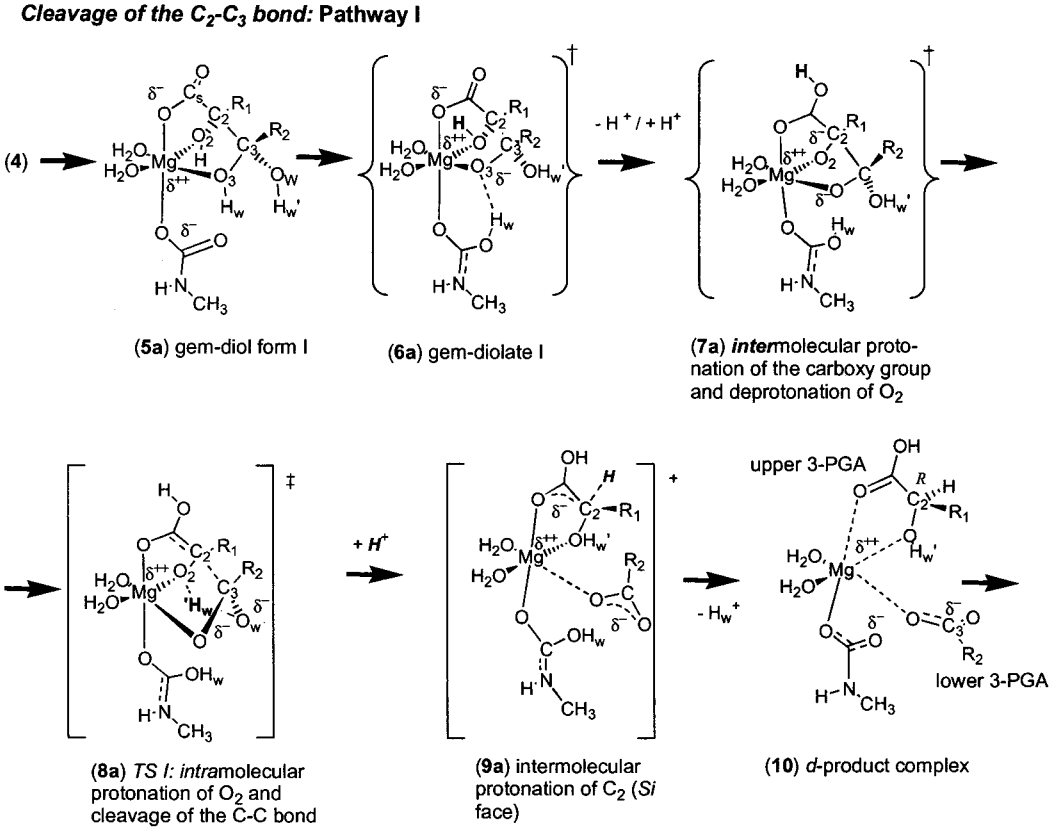


Figure 3. Reaction pathways of the proton transfer from O-3 of (a) the *gem*-diol **I** (+; dashed line) and (b) the *gem*-diol **II** (◇; dotted line) to the carbamate group of Lys 201 (structures **5a**/**6a** and **5b**/**6b** in Scheme 3 and 4). The energies are relative to the β ketoacid (**3**), and all atoms were flexible during the surface scan.

Addition of H_2O. In the subsequent hydration step, the O atom of the hydrogen-bonded water molecule of the outer coordination sphere is added to the positively polarized C3 atom of RuBP, forming a second transition complex (**4**) which is 12 (9) kcal mol^{-1} (Table 1) higher in energy than the β ketoacid. The hydration results in the *gem*-diol (**5a**). The attack of the water molecule on C3 is supported by the carbamylated Lys-201, which facilitates the abstraction of one proton (H_w, Scheme 2). Figure 3 (form I) shows that the *gem*-diolate structure with the proton remaining at the carbamate (corresponding to an O3−H distance of ca. 1.4 Å) is not stable; the proton transfer to the negatively charged O3 occurs spontaneously without activation energy. The coordination to the magnesium ion and the close proximity of the neutralized carbamate avoids an overcharged substrate in the transition structure **4**. An analysis of the Mulliken charges (according to the 6-311+G(2d,p) basis set) shows that the formal negative charge on O3 is reduced to −0.7 and is almost identical with the other coordinated oxygen atoms (Mg, +1.3; Lys-201 oxygens, −0.7; Asp-203/ Glu-204-coordinated oxygens, −0.8; the latter are modeled as water; the total charge on the fragment is zero).

Alternative Conformations of the *gem*-Diol. Starting from the *gem*-diol (**5a**), there are two possible pathways to continue the reaction sequence (Schemes 3 and 4): (1) the carboxyl group attached to C2 is protonated by the enzyme environment or (2) the intermediate changes its conformation, rearranging the coordination to the magnesium ion. First, we will describe the latter isomerization of the *gem*-diol. Starting from **5a**, as indicated in Scheme 4 by the bidirectional arrows, the C2−C3 bond rotates so that O_w approaches the Mg ion. This results in both oxygen atoms bound to C3 becoming coordinated to the metal. Concomitantly, the CS−C2 bond rotates in the opposite

Scheme 3. Pathway I for the Cleavage of the C2−C3 Bond Starting from the *gem*-Diol (Form **I**) Resulting from the Hydration Reaction[a]



[a] All structures marked with "†" are not local minima but are intermediately formed on the reaction path. The stereospecific protonation (**9a**) of C2 was not part of the calculations. However, the direction of the proton attack could be deduced from the last optimized structure (**8a**) of the bond cleavage leading to the product complex **10**.

direction, freeing O2 from its metal coordination. The result is the *gem*-diol form II (**5b**), which the active site of the enzyme should be able to accommodate without difficulties. The two *gem*-diol forms are approximately equal in energy, form II being slightly more stable, 2 (1) kcal mol^{-1} (Table 1). We estimated the activation energy of the transition from **5a** to **5b** to be ca. 10 kcal mol^{-1} (O2 and O3 in syn position). Although, in the enzyme, the energy for the isomerization might be higher due to the rearrangement of the hydrogen-bond network required, it probably should still be significantly lower than the energy necessary for the cleavage of the C−C bond. Therefore, we propose that both conformations of the *gem*-diol **5a** and **5b** are in equilibrium and suggest alternative routes for the C−C bond cleavage starting from either **5a** or **5b**.

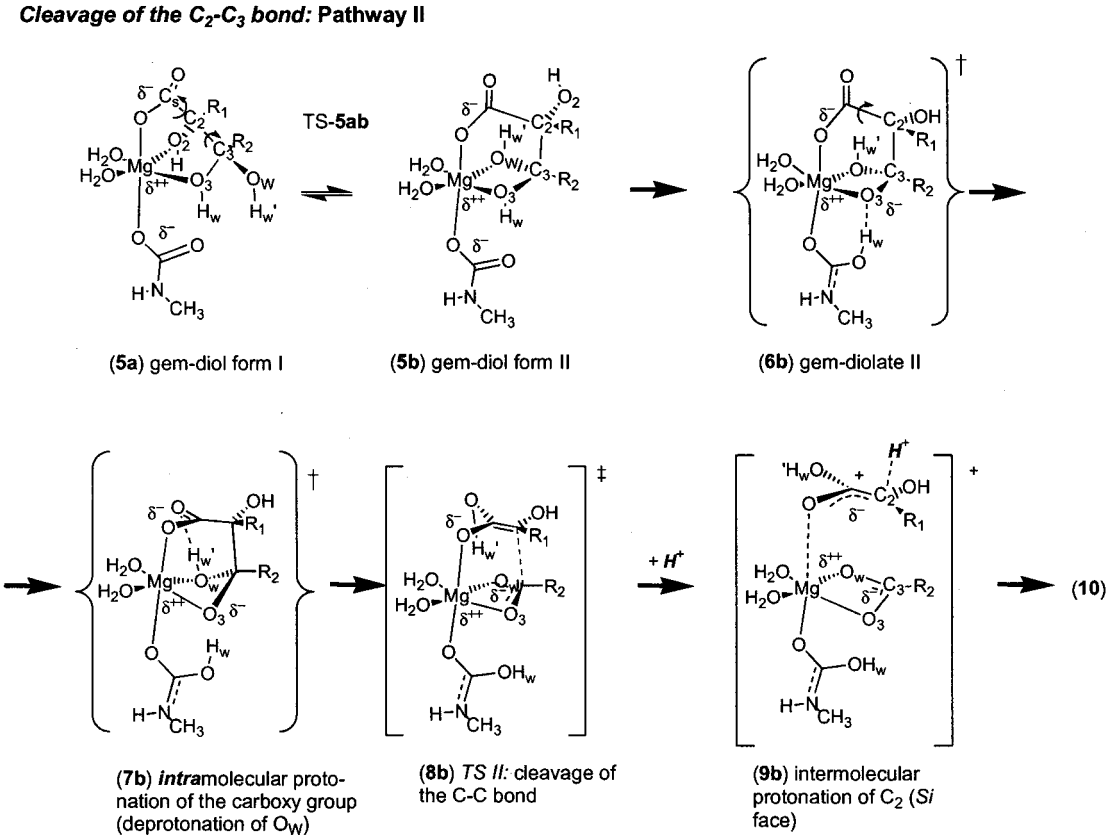
Cleavage of the C2−C3 Bond—Pathway I. The steps from both *gem*-diol forms to the cleavage of the C2−C3 bond are shown in Schemes 3 and 4; the corresponding energy profiles are plotted in Figure 4. In reaction pathway I (Scheme 3), the reaction sequence starts from the *gem*-diol **5a** as a local minimum at a C2−C3 distance of 1.54 Å (Table 1, Figure 4a). When the C−C bond elongates to a distance between 1.8 and 2.0 Å, the *gem*-diolate **6a** becomes energetically more favorable and is intermediately formed without activation energy, the H_w proton migrating again to the carbamate. However, this structure is not a local minimum on the energy surface, and relaxation of the C2−C3 bond leads back to the *gem*-diol **5a** (compare Figure 3). In contrast, further elongation of the C−C bond results in a steady increase of the energy (reaction coordinate α) unless general acid/base catalysis intervenes to assist C−C cleavage (reaction coordinate β).

The hydrogen-bond network within the active site may provide appropriate acidic and basic groups (Figure 1). Structural evidence is consistent with hydrogen bonding between the side chains of Lys-334 (with the carboxyl group of CABP) and Lys-175 (with O2).[8,27,29] We postulate a proton shift from protonated Lys-334 to the noncoordinated carboxylate O atom and another proton shift from O2 to Lys-175. The latter reverses the migration of a proton from Lys-175 to O2 that occurred during the enolization step.[25] This leads to **7a** (Scheme 3). Now the requirements for a low-energy cleavage of the C2−C3 bond are accomplished. Moving along the reaction coordinate β (Figure 4a, Scheme 3) leads to the transition structure **8a** of the bond cleavage which is further stabilized by the intramolecular transfer of the O_w proton to O2. We calculated the activation energy for the cleavage of the C−C bond in this manner to be 30 (28) kcal mol^{-1}.

Cleavage of the C2−C3 Bond—Pathway II. The reaction sequence is different for the second pathway for the cleavage of the C2−C3 bond that involves the isomerized form II of the *gem*-diol (Scheme 4). Analogously to pathway I, the O3 proton of *gem*-diol II is first transferred to the carbamate, forming the diolate **6b (**Figure 3, form II**)**. In contrast to pathway I, the carboxyl group is protonated intramolecularly by the hydrogen bound to O_w (**7b**). Both of these proton shifts occur when the C2−C3 bond distance is between 1.8 and 1.9 Å on the reaction coordinate (Figure 4b) where the energy of the system increases rapidly. The length of the C2−C3 bond in **7b** (2.0 Å) is close

(29) Knight, S.; Andersson, I.; Brändén, C.-I. *J. Mol. Biol.* **1990**, *215*, 113−160.

Scheme 4. Pathway II for the Cleavage of the C2−C3 Bond Starting from the Isomerization of *gem*-Diol Form **I** to Form **II**[a]



[a] As in Scheme 3, **9b** is shown to demonstrate the only possible direction for the stereospecific protonation of C2.

to that in the transition structure of the bond cleavage **8b**. We calculated the activation energy for pathway II to be 37 (32) kcal mol^{-1}.

Transition Structures for C−C Cleavage. A comparison of the highest occupied molecular orbitals (HOMOs) given in Figure 5 shows the similarities between transition structures of the two pathways. The highest MO coefficients are located between the carbon atoms C2, C3, and CS, indicating that the breaking of the C2−C3 bond is concerted with the formation of a C2−CS double bond. Hereby, the CS−OS bond is weakened, as indicated by the opposite sign of the wave function. The system is stabilized by the binding orbitals between the hydrogen atoms bound to O_w and OS and the oxygen atoms O2 and O_w, respectively. In reaction pathway I (Figure 4a), there is a sudden decrease of the energy on the reaction coordinate between 2.4 and 2.5 Å, shortly after passing the transition structure. This is due to completion of the proton transfer from O_w to O2.

Protonation of C2. The final step of the formation of the two molecules of 3-phospho-D-glycerate (3-PGA) is the protonation of C2 leading to complex **9a** (Scheme 3). Structurally, protonated Lys-175 is ideally positioned to donate this proton, and experimental evidence shows that deletion of its side chain cripples this step.[30] As the number of atoms in our current model was restricted, we were not able to calculate the energy for this protonation. However, calculations on an extended system showed that the activation energy for this step is likely to be negligible compared with that of carboxylation, hydration, or cleavage of the C−C bond.

In the case of pathway II, the intramolecular proton transfer is accomplished before the transition structure, and the system energy drops as the upper 3-PGA molecule moves away from the magnesium ion. In the enzyme environment (Figure 1), this movement would lead directly to Lys-175. This residue is therefore most likely to be responsible for the protonation of C2 in pathway II, just as it is in pathway I. However, to react as a proton donor, Lys-175 must be protonated, and pathway II lacks an explicit means by which this might be accomplished. Therefore, in pathway II, Lys-175 would need to accept a proton from the hydrogen-bond network (e.g., from Lys-334).

Dissociation of Products. The catalytic cycle is completed with the dissociation of the two molecules 3-PGA to regenerate the active site. Although the X-ray structure of the product complex[31] indicates that only one 3-PGA molecule remains coordinated to the metal, we modeled a product complex with two molecules of 3-PGA coordinated to the Mg ion to compare the energies. We assume that structure **10**, which has both product molecules coordinated to the metal (Scheme 3), is formed intermediately and provides an upper limit for the energy of the reaction products (see Table 1 and Figure 2).

Discussion

Proton Abstractions and Possible Reuse of the Carbamylated Lys-201 as a Base. To produce our model for the enediolate of RuBP, ethene-1,2-diolate, from ethene-1,2-diol, the finishing point of our previous study of the enolization of RuBP,[25] O3 must be deprotonated. This deprotonation is required to direct attack of CO_2 to C2; otherwise, C3 would be equally vulnerable,[4,9,25,32] an issue that has not been addressed in other calculations of CO_2 addition to the enediol.[15,17,33] His-

(30) Harpel, M. R.; Hartman, F. C. *Biochemistry* **1996**, *35*, 13865−13870.

(31) Taylor, T. C.; Andersson, I. *Biochemistry* **1997**, *36*, 4041−4046.

(32) Harpel, M. R.; Larimer, F. W.; Hartman, F. C. *Protein Sci.* **1998**, *7*, 730−738.

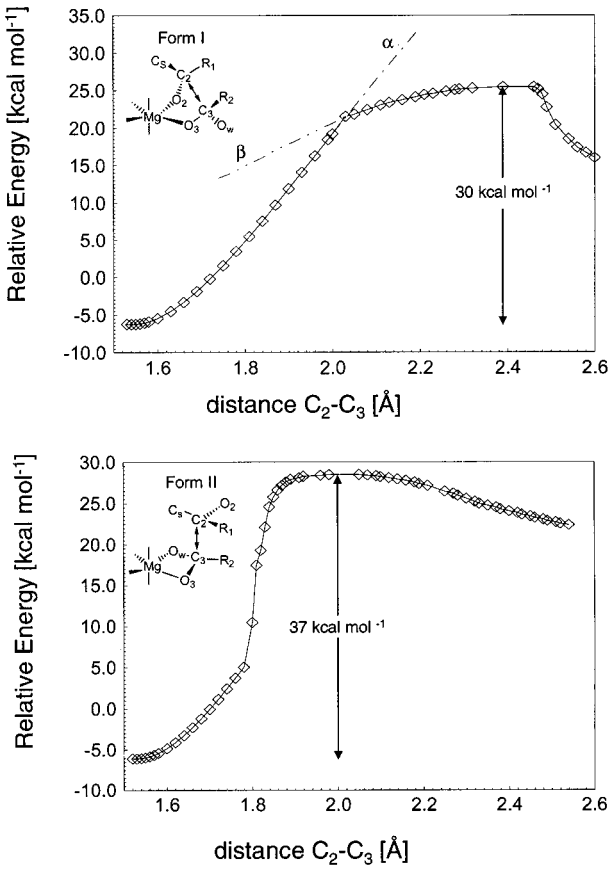


Figure 4. Reaction pathways for the cleavage of the C_2-C_3 bond of (a) the *gem*-diol **I** and (b) the *gem*-diol **II**. In the case of form **I**, the sharp bend results from a crossing of the energy surfaces with a deprotonated carboxyl group and a protonated O_3 atom (α) and the inverse arrangement (β). The energies are relative to the β ketoacid **3**.

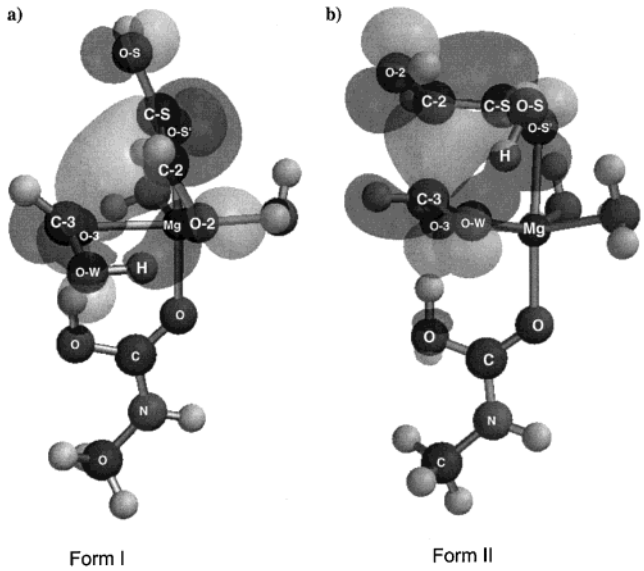


Figure 5. Highest occupied molecular orbitals of the transition structures for the C2−C3 bond cleavage: (a) structure **8a** and (b) structure **8b**. The positive wave function is represented by light gray and the negative wave function by dark gray. In both cases a new double bond is formed between C2 and CS, while the former CS−OS π-interaction becomes antibinding.

294 might abstract the O3 proton, as suggested previously,[25,32] or the carbamate might be reused for this purpose.[32] As our calculations show, there is no need for the 2,3-enediol to be twisted to achieve a low-energy addition of CO_2. We found the dihedral O2−C2−C3−O2 in the transition structure **2** to be −9.6°. This distortion can be explained by the rehybridization of C2 from sp^2 to sp^3.

We find that the carbamylated Lys-201 again plays a crucial role as a base during the subsequent reaction steps of hydration and C−C cleavage. Therefore, after each proton abstraction step, the enzyme environment must remove the proton from the carbamate to recycle it back into its basic, unprotonated form before the next proton abstraction can occur. Preliminary combined quantum mechanical and molecular mechanical (QM/MM) calculations indicate that the phosphate groups of RuBP could act as sinks for protons. Both phosphates are fully ionized in solution at physiological pH, but they may be at least partly protonated when desolvated in the enzyme environment. The phosphate oxygen atoms are connected to the carbamate by an extensive network of hydrogen bonds[8,27,29] which could provide a conduit for transporting protons within the active site. More detailed investigations of this possible channeling of protons are in progress.

Active Site Features Lower Activation Energies. The results of our calculations show that the residues within the active site, and especially the central metal, have an important impact on the activation energies of the transition structures involved in the carboxylation reaction sequence (**2**, **4**, **8a**/**8b**). Stabilization of the transition structures by these residues becomes particularly evident when the energies we calculate with our model are compared with those calculated in recent studies of the analogous nonenzymatic reaction pathway. For example, the activation energies we predict for the addition of CO_2 and the subsequent hydration (9 (7) and 12 (9) kcal mol^{-1}, respectively) are less than half the analogous values calculated for nonenzymatic systems [39 and 40 kcal mol^{-1},[15] 56 and 66 kcal mol^{-1},[16] or 20 and 60 kcal mol^{-1} (20 kcal mol^{-1} corresponds to free energy calculations on nonenzymatic model systems; 60 kcal mol^{-1} is estimated from the graphical representation)].[14] Even when solvent effects were included in a nonenzymatic model, the calculated activation energies for these steps were still very large (25 and 42 kcal mol^{-1}, respectively).[12]

CO_2 and H_2O Add Sequentially. Our proposed mechanism has separate transition states for CO_2 addition and H_2O addition, separated by a β ketoacid intermediate. We could not locate a transition structure in which the carboxylation and hydration reactions were concerted. This contradicts suggestions that carboxylation and hydration might occur in the same transition state[9,11] but is in agreement with all other quantum mechanical simulations of carboxylation.[12,14−16,18,34] Simulations of the analogous oxygenation reaction also have O_2 and H_2O adding sequentially.[17,18,33,35,36] The *gem*-diol **5a** has no normal frequency modes in which the C2−CS bond extension leading to decarboxylation and the C3−O_w bond extension of dehydration are coupled. The conjecture that CO_2 addition and hydration might be concerted rests on the observation that, when the isolated β ketoacid (which is predominantly or exclusively in the unhydrated ketone form[37]) is used as substrate, it reacts only slowly

(33) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *J. Phys. Chem. A* **2001**, *105*, 4726−4736.
(34) Moliner, V.; Andrés, J.; Oliva, M.; Safont, V. S.; Tapia, O. *Theor. Chem. Acc.* **1999**, *101*, 228−233.
(35) Oliva, M.; Safont, V. S.; Andrés, J.; Tapia, O. *J. Phys. Chem. A* **1999**, *103*, 6009−6016.
(36) Tapia, O.; Oliva, M.; Safont, V. S.; Andrés, J. *Chem. Phys. Lett.* **2000**, *323*, 29−34.
(37) Pierce, J.; Andrews, T. J.; Lorimer, G. H. *J. Biol. Chem.* **1986**, *261*, 10248−10256.

(≈2% of the k_{cat} with CO_2 and RuBP as substrates),[37] indicating that its rate of reaction is limited by a process not on the catalytic pathway. This rate limitation might be imposed by the need to catalyze the hydration of the β ketoacid, a process that the active site would not normally be required to catalyze if CO_2 addition and hydration were concerted. However, the sluggish rate with which β ketoacid reacts is matched by an equally slow rate of inhibition by its reduced analogue, 2′-carboxyarabinitol-P_2,[38] suggesting alternatively that the rate limitation results from slow binding of these cumbersome molecules. Therefore, slow processing of the β ketoacid does not establish that it is not an intermediate on the catalytic pathway.

When used as substrate, the β ketoacid partitions nearly exclusively to products. Unless nonphysiological metals are used, little or no release of CO_2 from it can be detected.[37] The enzymatic CO_2 addition is thus effectively irreversible. Reasons for this are not apparent in our present calculations, nor in previous calculations by others,[14−16] which suggest that subsequent steps have similar or higher activation energies. Unmodeled features of the active site must promote hydration and C2−C3 bond cleavage so strongly that the β ketoacid is fully committed kinetically to the forward pathway. Perhaps conformational rearrangements in the protein, which "close" and "open" the active site during each catalytic cycle,[4,39] might favor such commitment. It is not certain when closure occurs, but it necessarily must happen between the binding of RuBP and the formation of the β ketoacid (represented by **3**).

Calculations for the nonenzymatic reaction suggested that hydration of the β ketoacid (**3**) could involve a four-center transition structure where one proton of the substrate water is directly transferred to O3.[15] By contrast, we found that the carbamylated Lys-201 is better positioned to abstract the water proton and then transfer it to the deprotonated O3, which is the stronger base (see energy plot of form I in Figure 2). Therefore, transfer of the proton to O3 via the carbamate (**4)** is more favorable energetically.

Alternative Pathways for C2−C3 Bond Cleavage. Consistent with predictions for the nonenzymatic reaction,[15,17,40] we found that transition structures for cleavage of the C2−C3 bond are coupled with an intramolecular proton transfer from O_w. However, the destination of the transferred proton differs according to the conformation of the *gem*-diol (**5a** and **5b**).

For the transition structure **8a** arising from the form I *gem*-diol **5a** where O3 and O2 are metal coordinated (pathway I, Scheme 3), the C2−C3 bond distance we calculate with the B3LYP/6-31G(d) method (2.4 Å, Table 1) is similar to that estimated for the nonenzymatic transition structure with the HF/6-31G(d,p) method (2.3 Å).[14] However, a detailed analysis of transition structure **8a** shows that the pathway for transfer of the proton is hindered by the presence of an existing proton on O2. Coordination of O2 to the Mg ion makes it unlikely that this O atom could become hypervalent in the transition structure, transiently accommodating two protons, as suggested for the nonenzymatic reaction. As an alternative, we suggest a preliminary abstraction of the O2 proton by the enzyme. Lys-175 is appropriately positioned and unprotonated following transfer of a proton to O2 during enolization.[25] Therefore, it is ideally suited to accept a proton from O2, opening the path for transfer of the O_w proton to O2. We were not able to find any transition structure where the nascent carboxyl of the lower 3-P-glycerate was protonated. The negatively charged carboxylate appears to be the most suitable leaving group for the C−C cleavage (see **8a** and **8b** in Schemes 3 and 4). Our proposed pathway that avoids hypervalent O2 is supported by a comparison of the activation energies (ΔE_a^{CC}) necessary for the bond breaking. Previous (nonenzymatic) studies predicted 56 and 71 kcal mol^{-1} (refs 15 and 16, respectively) for the cleavage with proton transfer from O_w to O2. Zhan et al.[12] calculated ΔE_a^{CC} including solvent interactions for the same arrangement to be 41 kcal mol^{-1}. We found $\Delta E_a^{CC} = 30$ (28) kcal mol^{-1} for pathway I (Table 1).

On the other hand, protonation of the upper carboxylate is essential for a low-energy C2−C3 bond cleavage (Figure 4). In pathway I, the proton is supplied by the enzyme environment. The hydrogen-bonded Lys-334[27,29] seems ideally positioned for this task. The protonation leads to electron deficiency of the central CS atom. The neighboring C2 atom becomes polarized and induces a charge transfer from the lower part of the RuBP to form a double C2−CS bond (**7a, 8a** in Figures 4 and 5). This assists the heterolytic C2−C3 bond cleavage and leads to a significant negative charge on C2, which facilitates the final protonation of C2 in the next step.

The alternative transition structure **8b** of pathway II arises from form II of the *gem*-diol where both of the O atoms attached to C3 are coordinated by the metal. In this case, the concerted proton transfer from O_w is destined to the uncoordinated OS atom (Scheme 4). This transition structure is somewhat analogous to that of alternative 2 suggested by Safont et al.,[13] although the inclusion of the Mg ion and the carbamate model in our calculation leads to differences in the identities of the protons involved. As discussed above, the proton originally bound to O3 was abstracted before the carboxylation and therefore is not present in this transition state. Neither is the second proton derived from substrate water; it is transferred to the carbamate before the C2−C3 cleavage, increasing the electron density on the newly formed lower carboxyl group and simultaneously weakening the bond between O_w and its proton. Despite these structural differences and the different calculation methods, the C2−C3 bond distance estimated for the transition structure of Safont et al.[13] (2.0 Å) is the same as in our model for **8b** (Table 1). The activation energies computed by the two approaches are also in astonishing agreement: between 30 and 44 kcal mol^{-1} for the model of Safont et al.[13] compared with 37 (32) kcal mol^{-1} for pathway II in our model (Table 1).

The choice between pathway I and pathway II will be influenced by the arrangement of acidic and basic groups within the active site (see Figure 1). Pathway I would be favored by the presence of an acidic group capable of protonating the noncoordinated oxygen of the carboxylate group (e.g., protonated Lys-334) and a proton donor close to O_w (e.g., protonated His-327) to facilitate the proton transfer to O2. In contrast, a repulsive interaction of the environment near the same carboxylate group (e.g., unprotonated Lys-334) might force the rotations about the CS/C2 and C2/C3 axes necessary to bring O_w into coordination with the metal and to stabilize the carboxylate group by forming a hydrogen bond with O_w-H_w' (**6b, 7b**, Scheme 4). If protonated, Lys-334 seems ideally suited for protonating the carboxylate group derived from the CO_2 substrate in pathway I. Simultaneously, neutral Lys-175 could be protonated by the O2 proton, reversing the movement of this proton that occurred during enolization[25] and preparing the way for eventual transfer of this proton to C2. These coupled proton movements convert **6a** to **7a**. A favorable environment

(38) Pierce, J.; Tolbert, N. E.; Barker, R. *Biochemistry* **1980**, *19*, 934−942.

(39) Duff, A. P.; Andrews, T. J.; Curmi, P. M. *J. Mol. Biol.* **2000**, *298*, 903−916.

(40) Tapia, O.; Andrés, J.; Safont, V. S. *J. Phys. Chem.* **1996**, *100*, 8543−8550.

for pathway II, however, could result from a direct proton transfer from protonated Lys-334 to neutral Lys-175.

Stereospecific Protonation of C2. In either case, the newly protonated Lys-175 can then act as a proton donor to protonate C2 stereospecifically, as proposed on the basis of structural analysis[31] and mutagenesis data.[30,41] Therefore, both transition structures meet the structural requirements of the enzyme. In both pathways, the rearrangement of the charge distribution of the intermediate and its surroundings supports the protonation of C2 after passing the transition structures **8a** or **8b**. As described above, C2 becomes negatively polarized after the bond to C3 is broken. The shift of the charge is likely to induce the approach of Lys-175 to C2, facilitating the stereospecific protonation. As the *Re*-face of C2 is still sterically protected by close contact to C3, the protonation can only occur on the *Si*-face, leading to D-3-PGA.

Conclusion

We investigated Rubisco-catalyzed carboxylation computationally using an abbreviated model of the substrate and active site and derived a model for the catalytic sequence that is consistent with experimental observations. Inclusion in the model of two features characteristic of the active site—a magnesium ion and a carbamylated Lys side chain—substantially decreased activation energies associated with the calculated transition structures. CO_2 is added directly to the almost planar enediolate of RuBP without forming a Michaelis complex. We could not find any evidence that carboxylation was concerted with hydration; transition states for the two steps were discrete and separated by a β ketoacid intermediate. The carbamylated Lys-201 plays an important role in the whole catalytic cycle. In addition to its critical role in enolization,[25] it is important for stabilizing the addition of the substrate water and in facilitating the cleavage of the C2−C3 bond of the substrate. The enzyme might realize one of two alternatives for C−C cleavage, starting from two conformations of the hydrated *gem*-diol that differ in the way they coordinate with the metal (**5a** and **5b)**. In either case, the activation energy for the bond breaking is remarkably low (28−37 kcal/mol^{-1}) but still sufficient to limit the overall catalytic rate at high CO_2 concentration.

To perform the first-principle calculations, our model was limited to a relatively small number of atoms, and questions about the roles of other active site residues in the reaction sequence remain. We will address these in a forthcoming QM/MM study.

Acknowledgment. This work was supported by the German Academic Exchange Service (DAAD) within the Hochschulsonderprogramm III. We thank Dr. Alistair Rendell for his helpful comments on computational aspects. We gratefully acknowledge the Australian National University Supercomputing Facility for generous grants of computer time.

JA011362P

(41) Lorimer, G.; Hartman, F. C. *J. Biol. Chem.* **1988**, *263*, 6468−6471.