Carbon dioxide fixation in RuBisCO is protonation state dependent and irreversible

Oscar A. Douglas-Gallardo,[†,‡] Juliana A. Murillo-López,[†] Javier Oller,[†,¶] Adrian J. Mulholland,[§] and Esteban Vöhringer-Martinez[*,†]

†*Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Chile.*

‡*Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom.*

¶*Departamento de Química Orgánica y Fisicoquímica, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile.*

§*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom.*

E-mail: evohringer@udec.cl

Abstract

Most CO_2 from the atmosphere is assimilated in photosynthetic organisms by the ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO) enzyme as part of the Calvin cycle. Despite its relevance and many efforts in the last few decades, the mechanistic picture of the catalytic CO_2 fixation reaction is still under debate. Here, we combine QM/MM molecular dynamics simulations with high-level electronic structure methods and the projector-embedding approach to provide reference values for the activation and reaction free energies of the catalytic CO_2 fixation reaction. Our results show that carboxylation is protonation state dependent and irreversible, making the

back reaction (decarboxylation reaction) highly unfavourable. The carbamylated lysine residue, Kcx201, coordinated to the magnesium (II) cation in the active site plays a central role shuffling protons from and to the substrate creating the proper reactive enolate species that adds CO_2. The emerging microscopic picture that involves several protonation equilibria and the free energy profile of the CO_2 fixation reaction provides insights that may be used in the future to improve enzymatic efficiency in RuBisCO.

Keywords

CO_2 fixation reaction, RuBisCO, QM/MM molecular dynamics, projector-embedding approach and post Hartree-Fock electronic structure methods.

Introduction

Anthropogenic emission of carbon dioxide (CO_2) has increased significantly since the industrial era contributing to the rise of the average global temperature through the greenhouse effect.[1–3] The main sink of atmospheric CO_2 is the photosynthesis via the Calvin-Benson-Bassham cycle that represents the primary organic carbon source in the biosphere.[4–8] RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) is the enzyme catalyzing the first cycle step of CO_2 fixation and is considered the most abundant protein in plants.[5,6,8–10] Because of its key role in biology one might expect it to be highly efficient and fast. RuBisCO, however, presents catalytic rate constants 80 times slower than most metabolic enzymes (k_{cat}=1-10 s^{-1}).[11] The elementary chemical steps associated with this comparable slow conversion of CO_2 are still not completely understood giving rise to several theoretical and experimental studies focused on its molecular mechanisms.[7,8,12–17]

RuBisCO is composed of eight large and eight small sub-units (L_8S_8) in its form I that is mainly found in plants, algae and the majority of photosynthetic organisms.[7,8,18,19] The large sub-units are catalytic active and form a toroid around a four-fold symmetry axis

(Figure 1A).[7,8,18,19] The small sub-units are on top and bottom of the toroid around the symmetry axis and stabilize the quaternary structure of the enzyme.[7,8,18,19] They do not affect the catalytic activity of the enzyme and recently they have been postulated as a CO_2 reservoir.[20] Two large sub-units are arranged anti-parallel, creating the two catalytic sites for CO_2 fixation at their interface (see Figure 1B). The active sites of these dimers consist mainly of the first coordination sphere of magnesium (II) cation, see Figure 1C. This metal is coordinated with one oxygen atom from carbamylated lysine, two oxygen atoms from the ribulose-1,5-bisphosphate (RuBP) substrate, one oxygen atom from glutamate and one from aspartate residue (Lys201, Glu204 and Asp203 according to numbering in RuBisCO from spinach, PDB-ID:8RUC).[7,18] The sixth coordination site of the magnesium (II) cation is occupied with the carboxylate group from 2-carboxy-*D*-arabinitol-1,5-biphosphate (2CABP) in the X-Ray structure of RuBisCO from spinach (PDB-ID:8RUC).[7,18]

The CO_2 fixation reaction starts with the carbamylated Lys201 residue abstracting the proton from the C3 carbon atom according to deuterium kinetic isotope effects of the substrate and solvent reported by Tchekerz and collaborators[12,16,21] (Figure 1, see current accepted mechanism). After proton rearrangements the reactive nucleophilic enolate adds CO_2 at the C2 carbon atom and forms 2-carboxy-3-keto-*D*-arabinitol-1,5-bisphosphate (CKABP) which differs by the keto functional group at the C3 atom from the hydroxy group in 2CABP observed in the crystal structure. The two 3-phospho-*D*-glycerate products are formed after hydration of the CKABP intermediate and a carbon-carbon bond scission with a stereoselective protonation.[7,13,14]

The sequence of elementary steps in the reaction mechanism were studied computationally with Density Functional Theory (DFT) calculations using different fragment models and compared to measured rate constants and deuterium or carbon isotope effects.[7,13,15,17,22] After Clealand and collaborators first proposed the mechanism,[23] Cummins, Gready and coworkers have provided a potential energy profile of the elementary steps using DFT calculations and optimized structures of fragments of the active site with increasing number of

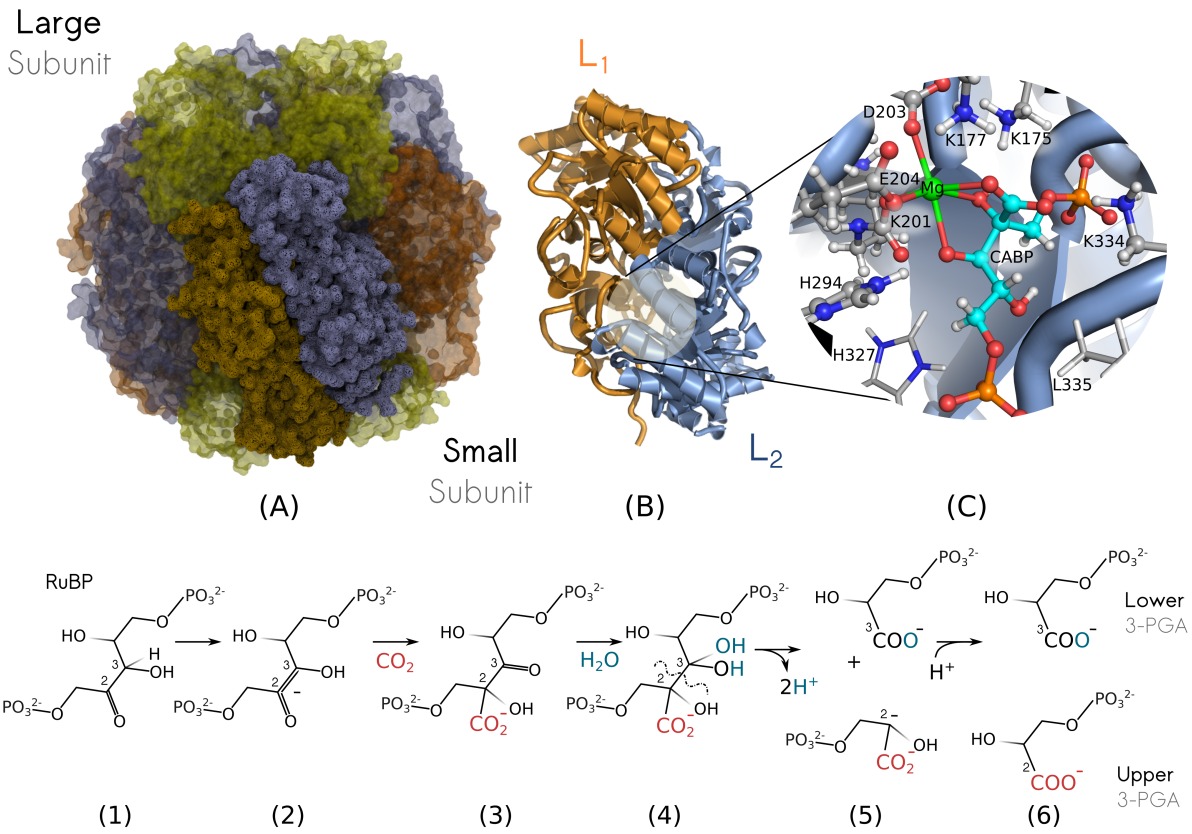

Figure 1: Quaternary structure of RuBisCO (form I) hexa-decamer complex with the catalytic active dimer composed of two large sub-units containing the active site associated with CO_2 fixation through the elementary steps shown at the bottom. (A) Schematic representation of the whole enzymatic complex highlighting the eight large sub-units in orange-blue and the eight small sub-units in yellow below and above of the toroid. (B) Catalytic active dimer of two large sub-units forming the catalytic site at their interface. Only one of the two catalytic site was considered at QM/MM level. (C) RuBisCO active site highlighting the first and second magnesium coordination sphere in a ball and stick representation. The magnesium cation is shown as a green colored sphere.

atoms (up to 343).[13,15,22] Gready and her group have also performed QM/MM calculations, including the rest of the enzyme and the solvent by MM methodology using semi-empirical electronic structure method: these calculations showed significant differences in the potential energy for some steps.[22] Their most recent study suggest that proton wires play a mayor role and starts with an active site bearing a protonated glutamate residue (Glu60).[13] The protonation of this glutamate residue, however, seems unlikely considering its acidic nature and a positively charged lysine residue (Lys335) at a distance of 3.3 Å. In summary, these studies

suggest a wide range of potential energy barriers (5-16 kcal mol^{-1}) for the first CO_2 fixation reaction, with values dependent on the size of the active site model, the DFT functional and basis set, and the inclusion of additional water molecules in the active site.

Based on measured D_2O solvent isotope effects and generic kinetic relationships, Tchekerz and collaborators derived the activation free energy for the carboxylation reaction of the enediolate at 16.4 kcal mol^{-1}.[16] The authors state, however, that this value does not result from their observation but is rather an estimation based on previous data introduced in their kinetic model. The only observable rate constant involves the whole sequence of elementary steps and corresponds to a 16.8 kcal mol^{-1} activation barrier.[16] This activation barrier might, however, correspond to subsequent elementary steps such as the hydration and bond scission dominating the observed rate constant (see Figure 1).

From the above mentioned studies, it is evident that an accurate calculation of the free energy change of the first elementary step involving CO_2 is lacking. This becomes even more important based on a recent revision of the reaction mechanism reported by Flamholz *et al.*[11] Analyzing measured values of catalytic rate and Michaelis constants of various organisms, the authors conclude that the free energy barrier for the CO_2 fixation step is not coupled to the following elementary steps. Furthermore, this step could be the target of bio-engineering efforts to increase carboxylation over the oxygenation side reaction in RuBisCO. The authors propose an equilibrium between the substrate and the enolate species, which might vary between different organisms, giving rise to different carboxylation efficiencies. It remains, however, unknown which protonation states participate in this equilibrium and which one reacts with CO_2. After the formation of the enolate, protons have to be exchanged between the substrate and nearby histidine residue (His294) or the carbamylated lysine (Kcx201) residue to form the observed product.[12,24–26]The relevant protonation states and the activation barrier leading to CO_2 fixation are required to fully characterize the first elementary steps in RuBisCO.

Here, we provide calculations of the activation barrier and reaction free energy of the CO_2

fixation reaction from different protonation states in the active site resulting from the initial enolate formation. Proton exchange between substrate and enzyme residues precedes the fixation of CO_2 and creates the reactive enolate species. Assuming a fast exchange between the protonation states, we identify one of them to react with CO_2 in an irreversible reaction with an activation free energy barrier of 4.7 kcal mol^{-1}. The reaction is exergonic (-12.2 kcal mol^{-1}), which limits the decarboxylation reaction in accordance with previous experimental findings. Our results provide a detailed microscopic picture of the steps associated with CO_2 fixation. From this picture, new ideas for future protein engineering efforts to increase RuBisCO carboxylation may result that could contribute to reduce atmospheric CO_2 levels and increase simultaneously the crop yields.

Computational Methods

Initial preparation of RuBisCO model

The form I of RuBisCO is found in plant, algae and the majority of photosynthetic organisms.[7,8,18,24] This enzymatic variety is a hexadecamer (L_8S_8, Mr= 550,000) built up of eight large (L, 475 residues) and eight small (S, 123 residues) sub-units. These sub-units are spatially organized as four dimers (L_2) of L sub-units capped on the tips with two tetramers of S sub-units, see Figure 1A.[7,8,18] Each pair of L sub-units contains two active sites localized in the interface between them. All L and S sub-units have identical three-dimensional arrangements, and RuBisCO's general geometric configuration can be summarized by the formula $4(L_2S_2)$. The S sub-units do not seem to have an important effect on the chemical reactions, and are not included in our RuBisCO model.[24] Both active sites are within a TIM barrel at opposite sides and each of them is formed of residues stemming from both L sub-units. Here, only a dimer of L sub-units (L_2) was considered to explore the carboxylase activity of RuBisCO enzyme. A schematic representation of the simulated RuBisCO model is shown in the Figure 1B where the secondary structure in both L sub-units is highlighted.

The initial coordinates used in our enzymatic model were taken from the high-resolution X-ray (1.6Å) crystal structure of activated spinach (*spinacia oleracea*) RuBisCO with 2-carboxy-D-arabinitol bisphosphate (2-CABP) bound (PDB ID: 8RUC).[18] The PROTOSS software was used to predict the protonation states of all titratable amino-acid residues within the enzyme.[27] Protonation states of His294 and Kcx201 within the active site were modified to explore different reaction paths, maintaining the remaining protonation states of all other residues as predicted by PROTOSS.[27]

All molecular dynamics (MD) simulations were carried out using the AMBER16 software package.[28] The all-atom AMBER force field ff14sb[29] was used for both MM and QM/MM MD simulations. The Generalized AMBER Force Field (GAFF)[30] was employed for the parameters corresponding to the substrate (modified 2-CABP, CKABP) and the non-standard carbamylated lysine (Kcx201) residue (both located within the active site) with AM1-BCC atomic charges. The carboxylated analogue (2-CABP) present in the initial crystal structure was modified into the carboxylated product (structure 3, Figure 1, CKABP) that results from CO_2 addition to activated RuBP anionic species. Each active site present in the crystallized structure contains a magnesium (II) cation hexa-coordinated with six oxygen atoms from aspartate (Aps203), glutamate (Glu204) and carbamylated lysine (Kcx201) amino acid residues, each of which interact by means of a single oxygen atom (monodentate coordination). The remaining coordination positions are occupied by three monodentate oxygen atoms from the substrate: one of them corresponds to the newly formed carboxylate group (2'-carboxyl oxygen) and the other two oxygen atoms to the hydroxyl group at C2 and C3 (later named O2 and O3, respectively), see active site in Figure 1C. The carbamylated lysine (Kcx201) is located under the five-member ring formed between the magnesium cation and carbon (C2, C3) and oxygen (O2,O3) atoms belonging to the substrate. This location is opposite to the incoming CO_2 molecule and at the same side of the hydrogen atom at the C3 carbon atom. The geometric disposition turns out to be relevant when a general bases for the enolization process is assessed (see text below).

We also tested a different simulation setup in which one water molecule was included within the first magnesium coordination sphere, replacing the position occupied by one oxygen atom of the carboxylate group as suggested by Gready.[7] The initial coordinates of the oxygen atom that directly interacts with magnesium (II) cation were used to locate the new water molecule. The carboxylate group was rotated in order to avoid overlap with the water molecule. 8 water molecules were kept within the active site based on the initial crystal structure 8RUC.[7,18]

Molecular Mechanics (MM) Molecular Dynamics Simulations

A set of energy minimizations and molecular mechanics (MM) simulations were performed to equilibrate (300 K, 500 ps and 1 bar, 5 ns) the system using a Langevin thermostat and Monte Carlo barostat, respectively. The TIP3P water model[31] was used for the water molecules, applying periodic boundary conditions. For long-range interactions, a cutoff of 8 Å was employed in combination with the PME method. A cubic simulation box, with a minimum distance of 11 Å between the enzyme and the box edge in each direction, was employed. To investigate conformational states of the protein, several production MM MD simulations (100–300ns) were carried out. A harmonic restraint (force constant 2.0 kcal mol^{-1} $Å^{-2}$) was applied for both non-standard carbamylated lysine residues and substrates in the MM equilibration stages. These restraints, however, were released in the production MM and QM/MM molecular dynamics simulations.

Minimum Free Energy Path QM/MM simulations.

The activation and reaction free energies were calculated from QM/MM molecular dynamics simulations. In all cases, the DFTB3 semiempirical hamiltonian with the 3ob parameter set[32] was used to describe QM atoms in the enzyme's active site. The selected QM zone contains around 115 atoms, including the magnesium (II) cation and its first and second coordination

spheres (see Figure 1C as an example of a reduced version of the QM zone). The starting point in all QM/MM MD simulations was the caboxylated product of the activated RuBP intermediate which was previously equilibrated at QM/MM and MM levels. To obtain the enediolate intermediate steered MD simulations were performed elongating the bond between the carboxylate (or CO_2) carbon atom and the substrate C2 carbon atom from 1.6 to 3.2 Å. This final distance turned out to be large enough to obtain an almost linear CO_2 molecule which is evidence of a weak interaction with the substrate. A representative structure of the enediolate reactive state was obtained performing a new QM/MM equilibration over the final structure from the steered MD simulations. From equilibrated conformations of the reactant and product, the adaptive finite temperature string method[33,34] was employed to find the minimum free-energy path (MFEP) that connects these states. A set of collective variables (CV) was chosen to describe the reaction coordinate (s) associated with carboxylation reaction. In all the cases, the bond between the CO_2 carbon atom and the C2 carbon atom of the substrate was taken into account, together with the hybridization change of the substrate (C_2). In those systems without a water molecule interacting with the magnesium cation, an additional CV was included to describe the interaction between the CO_2 oxygen atom and the magnesium (II) cation.

Adiabatic mapping calculations

The energetic changes associated with the carboxylation reaction catalyzed by RuBisCO enzyme have also been explored by means of adiabatic mapping calculations. Within this approach, a reduced number of atoms within an active zone are optimized at QM/MM level leaving the majority of the remaining atoms frozen to reduce the computational cost by including them as charge points. An active zone of 16 Å around the magnesium (II) cation was selected and optimized at the B3LYP/6-31G(d)//AMBER-ff14sb level. Only a thin layer of water molecules 8 Å from the enzyme's surface was considered. The selected QM zone includes the magnesium (II) cation and its first and second coordination sphere. The

minimum potential energy path was obtained along the reaction coordinate defined as the distance between the CO_2 carbon atom and the substrate C2 carbon atom. To assess the performance of several DFT xc-functionals, a set of single point calculations over optimized structures was carried out. For each optimization window, the reaction coordinate was restraint with a force constant of 800 hartree a_0^{-2}. All adiabatic mapping calculations were performed using the ChemShell package,[35,36] combining ORCA[37] for the electronic structure calculations and AMBER[28] for the MM parameters used to run with DL-poly molecular dynamics.[35,36]

Projector-embedding approach with post-HF electronic structure calculations

Activation barriers and reaction energies were also calculated using several high level ab-initio electronic structure methods embedded into a DFT environment by a projector-embedding approach.[17,38–41] With this technique, it is possible to obtain activation and reaction energies at a high-level of theory with accuracy and at a low computational cost. This approach reduces errors from DFT xc-functional and allows a post-HF treatment of the reactive core of the catalyzed reaction describing the surrounding atoms with DFT xc-functional. This approach can remove the uncertainty associated with the dependence of the barriers on the choice of DFT xc-functional. A technique for basis set truncation was also used to reduce the computational cost. To obtain these values for activation and reaction energies, different single point calculations on the previously optimized structure at B3LYP/6-31G(d)//AMBER-ff14sb level were carried out with the Molpro package.[42,43] The enzymatic environment was taken into account by including the MM point charges from the AMBER ff14sb force field.[29] The QM region was subdivided into subsystems A and B, representing the high-level QM methods and DFT subsystems, respectively. To select the size of subsystem A in the embedding calculation, we tested four different regions that increasingly include some residues that are important for WF correlation in the reacting moiety. Figure S2 on Supporting In-

formation shows the different model fragments and the activation energy variation of system I for different SCS-MP2-in-DFT embedding calculations. Recently, a similar analysis on a RuBisCO fragment model was reported by us employing the projector embedding approach on this relevant metallo enzyme.[17]

Boltzmann Weighted Reactivity Descriptors from conceptual DFT

The chemical reactivity of both enolate species (*Enolate 2 and 3*) shown in Figure 2 was studied within conceptual Density Functional Theory using local reactivity descriptors. The Fukui function was calculated under the linear energy model approach for the electrophilic attack $f^-_{N_0}(\mathbf{r})$ on the RuBP substrate as the derivative of electron density with respect to the number of electrons, N, at constant external potential $v(\mathbf{r})$

$$\left(\frac{\partial \rho_N(\mathbf{r})}{\partial N}\right)_{v(\mathbf{r})} |_{N=N_0^-} = \rho_{N_0}(\mathbf{r}) - \rho_{N_0-1}(\mathbf{r}) = f^-_{N_0}(\mathbf{r}) \tag{1}$$

$\rho_N(\mathbf{r})$ is the electronic density of the molecule with N electrons. $\mathbf{r}$ describes a position in the molecular coordinate frame defined by nuclei's position.

To calculate the Fukui function we used the frontier molecular orbital (FMO) approach, neglecting the orbital relaxation due to the change in the total number of electrons, where the Fukui function $f^-_{N_0}(\mathbf{r})$ is approximated to the electronic density of the highest occupied molecular orbital:

$$f^-_{N_0}(\mathbf{r}) = \rho_{N_0}(\mathbf{r}) - \rho_{N_0-1}(\mathbf{r}) = \rho^{HOMO}(\mathbf{r}) \tag{2}$$

The Fukui function was condensed to atoms with the Hirshfeld-I partitioning method using the fragment of molecular response approach to assign the reactivity of specific atoms in the RuBP substrate.

The effect of conformational changes in the active site and the enzymatic environment on the reactivity of the substrate was assessed by our recently introduced Boltzmann weighted

atom condensed Fukui function[44] defined as:

$$\langle f_A^- \rangle = \frac{1}{Z(N,V,T)} \int_{D(v)} d\mathbf{R}_1 ... d\mathbf{R}_N f_A^-(\mathbf{R}_1, ..., \mathbf{R}_N) e^{-\beta U(\mathbf{R}_1,...,\mathbf{R}_N)} \quad (3)$$

where $\beta = \frac{1}{k_B T}$, T is the temperature and k_B the Boltzmann constant. $f_A^-(\mathbf{R}_1, ..., \mathbf{R}_N)$ is the atom condensed Fukui function obtained at one specific configuration of nuclei given by $\mathbf{R}_i$. The term $U(\mathbf{R}_1, ..., \mathbf{R}_N)$ is the potential energy of the system describing the interactions of the atoms at position $\mathbf{R}_i$ and $Z(N,V,T)$ the configurational partition function. The Boltzmann weighted atom condensed Fukui function was calculated as an ensemble average of atom condensed Fukui functions considering representative uncorrelated configurations of the molecular system. A canonical ensemble of the reactive molecule and its environment is considered for which the different configurations or microstates are found using QM/MM molecular dynamics simulations with the parameters defined above (DFTB3/amberff14sb). 50 different representative structures or configurations for protonation state I and II were used in the QM/MM energy calculations. The studied system were divided in a reactive zone described by QM level of theory, applying the link atom scheme and the rest is described by MM through atomic charges. QM/MM calculations were carried out using the interface between AMBER and Terachem.[45] The QM region was described at the ωB97X / def2-svp level of theory and for the MM region with the amberff14sb force field. From the polarized electron density of each structure, we calculated the atom condensed Fukui function f^- to evaluate the Boltzmann weighted atom condensed Fukui function $\langle f^- \rangle$.

For comparison we also consider the molecule in a vacuum optimizing the nuclei position at the ωB97X / def2-svp level of theory using the Orca package.[37] The optimized geometries were used to calculate the condensed Fukui function f^- with Chemtools software.[46]

Results

To characterize the elementary steps associated with CO_2 fixation in RuBisCO, we first use reactivity descriptors to analyze possible proton exchange reactions between the enolate intermediate (structure 2, Figure 1) and enzyme residues that would lead to the observed product (CO_2 addition on C2 RuBP, structure 3, Figure 1). Then, for the two most probable protonation states, we calculate the free energy along the reaction path with QM/MM molecular dynamics simulations using a dimer formed of two large subunits as molecular model of RuBisCO enzyme. In the final step, we benchmark the obtained activation barrier and reaction free energy with high level ab-initio electronic structure methods using a projector-embedding approach.[17,38–41]

Enzyme assisted proton exchange between hydroxyl oxygen atoms in the substrate creates reactive species for the carboxylation reaction

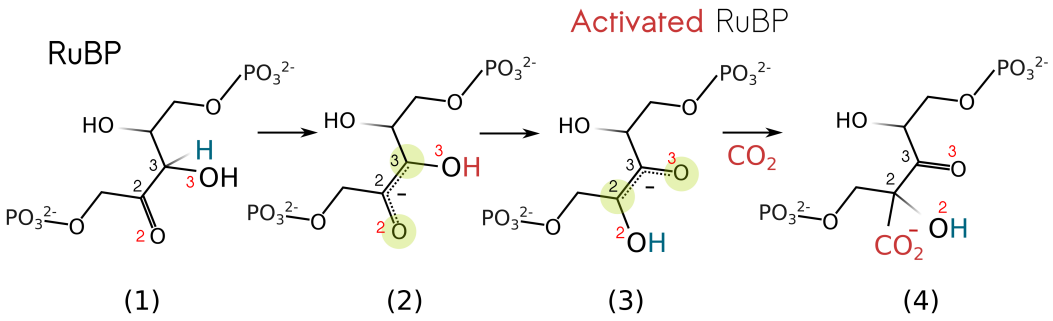

Figure 2: Schematic representation of possible structures associated with intermediate enolates that connect ribulose-1,5-bisphosphate (RuBP) substrate(1) with the caboxylated product (4): (2) initial cis-2,3-enediolate and the activated cis-2,3-enediolate (3) leading to the observed product. The green circle highlight different nucleophilic sites in the two enolates and the proton changing from O2 to O3 is highlighted with in blue and red, respectively.

To activate RuBisCO for catalysis, first Lys201 is carbamylated (Kcx201) and the magnesium (II) cation and substrate bind to the active site accompanied with the closure of

loop 6 (activated enzymatic form).[7,47] The CO_2 fixation reaction starts with the proton abstraction from the C3 carbon atom creating a reactive enolate (blue proton in Figure 2), named here as *Enolate 2*. CO_2 forms a new carbon-carbon bond with the C2 carbon atom of the anionic substrate as evidenced from analysis of the products. To study if the enolate species formed by initial proton abstraction from C3 (*Enolate 2*) presents the correct reactivity on the C2 carbon atom to form the observed product, we carried out QM/MM molecular dynamics simulations combined with reactivity descriptors based on conceptual density functional theory. The most reactive site of the enolate was identified based on atom condensed Fukui functions, which are Boltzmann averaged over all configurations of the enzyme-substrate Michaelis complex.[44] These reactivity descriptors include the polarizing effect of the enzyme on the substrate's electron density and identify the most nucleophilic site in the substrate as the atom with the maximal value in $\langle f^- \rangle$. In the initially formed enolate species *Enolate 2* in Figure 2, the C3 atom is the most nucleophilic site followed by the oxygen atom O2 (see Table 1). Based on this result, the reaction with CO_2 would not lead to the observed 2-carboxyketone because CO_2 would react with C3 and not the C2 carbon atom. The nucleophilicity on either C3 or C2 atom, however, will also depend on the protonation state of both oxygen atoms O2 and O3. To test this dependence, we created an alternative enolate species named *Enolate 3* in the active site where a proton was transferred from the O3 oxygen atom to O2 (molecule (3) Figure 2). Interestingly, the most reactive atom in this enolate species is C2 which would result in the correct observed product after CO_2 addition (see Table 1). Our reactivity descriptors, therefore, suggests that after proton abstraction on C3 a proton exchange between the two oxygen atoms O2 and O3 has to take place before CO_2 addition in order to create the correct reactive specie for the carboxylation reaction.

After establishing the protonation state of the enolate fixing CO_2, we asked what role the residues and the magnesium (II) cation in the active site play in modulating the reactivity of the initially formed (*Enolate 2*) and the reactive enolate intermediate (*Enolate 3*). As

Table 1: Atoms most prone to an electrophilic attack in enolates 2 and 3 (see Figure 3) represented by Boltzmann weighted atom condensed Fukui function $\langle f^- \rangle$ at 300 K (ωB97X / def2-svp) in RuBisCO and the atom-condensed f^- values for the minimum energy structure of the same enolates in a vacuum. (Standard deviations of the weighted values in RuBisCO are shown in parentheses).

	Enolate 2		Enolate 3	
Atom	f^- (vac.)	$\langle f^- \rangle$ (enzyme)	f^- (vac.)	$\langle f^- \rangle$(enzyme)
C_2	0.12	0.14 (0.01)	0.29	0.32 (0.02)
O_2	0.24	0.26 (0.03)	0.15	0.12 (0.01)
C_3	0.33	0.38 (0.02)	0.15	0.17 (0.01)
O_3	0.17	0.09 (0.01)	0.28	0.26 (0.02)

comparison, we evaluated the reactivity descriptors without the polarizing effect of the enzyme environment in a vacuum. For both enolate species the obtained values on the carbon atoms are smaller indicating that the electrostatic field in the catalytic cavity created by charged residues and the Mg^{2+} cation increases the reactivity of the substrate favoring CO_2 addition.

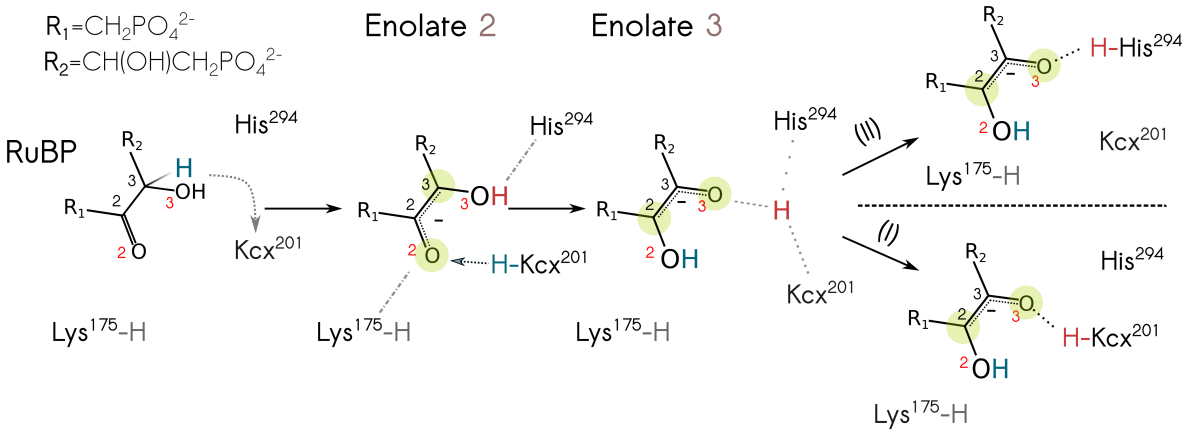

Figure 3: Proton exchange reactions in the active site between the RuBP substrate and residues Lys175, His294 and carbamylated Lys201. After proton abstraction on the C3 atom and reprotonation of oxygen atom O2, the proton of hydroxyl oxygen atom O3 is transferred to His294 leading to state II or to carbamylated Lys201 resulting in state I.

Besides increasing the reactivity, charged residues in the active site with acidic or basic side chains may also participate in proton exchange reactions between the two oxygen atoms O2 and O3 to form the reactive *Enolate 3*. Experiments with deuterated water and substrate did not present an increased isotope effect for the carboxylation and enolate formation

reaction, which rules out solvent contribution for this step of the reaction mechanism.[12,16] The residues close to the substrate capable of donating a proton to the O2 atom are the protonated carbamylated Lys201 (Kcx201) holding the initial abstracted proton associated with the first reaction step or the positively charged Lys175 (see Figure 3). Lys175, however, was discarded due to its large distance of 3.2 Å to the O2 atom.[18] This residue contributes to catalysis, stabilizing the negative charge on O2 when the initial *Enolate2* is produced.[7] The protonated carbamylated Lys201 (Kcx201), however, might become close to the O2 atom after it has abstracted the proton from C3 and could donate its proton to the negatively charged O2 atom to form a gemdiolate.

After carbamylated Lys201 (Kcx201) transfers its proton to oxygen atom O2 the proton from oxygen atom O3 has to be transferred to a basic residue to form the reactive enolate species (*Enolate 3*). One possibility is His294 leading to positively charged His294 residue and a deprotonated carbamylated Lys201 (state II in Figure 3). The other possibility is that Kcx201 accepts the proton from O3 after having transferred its proton to O2 (state I in Figure 3). Which of the two reaction paths takes place depends on the distance to both residues in a specific configuration of the residues and substrate in the active site. It is probable that an equilibrium between these two protonation states exists in which both interchange. To address this interchange by computer simulations is beyond the scope of this work and would require accurate electronic structure methods with low computational cost and exhaustive sampling of all configurations in the active site. Here, we assume that both protonation states can be present in the active site and that the proton exchange between them is a fast process. Under this quasi-equilibrium approximation, CO_2 fixation will be the rate-limiting step and its activation barriers for both protonation states will dictate the kinetics of the system.

Activation free energy barrier of the carboxylation reaction depends on the protonation states of His294 and carbamylated Lys201

The formation of the reactive enolate (Enolate 3) species leading to the observed product is compatible with two different protonation states of the residues in the active site (see states I and II in Figure 3). Assuming a fast exchange between the two protonation states, CO_2 fixation will mainly occur from the protonation state with the lowest activation barrier. To establish the reaction with the lowest free energy barrier, we determined the minimum free energy path between the reactive enolate (*Enolate 3*) and the carboxylated product in the two protonation states (see Figure 4). We used the adaptive string method[33,34] with a semi-empirical DFTB3 Hamiltonian describing 86 atoms of the active site in QM/MM molecular dynamics simulations. The free energy barrier for state I with protonated carbamylated Lys201 (Kcx201) is 3.1 kcal mol^{-1} lower than state II leading to a more stable carboxylated product compared to the reactive enolate intermediate. CO_2 fixation occurs when the carbamylated Lys201 is protonated and His294 deprotonated. A small activation barrier and a negative reaction free energy imply a mostly irreversible formation of the carboxylated product. The irreversible CO_2 addition step is in agreement with the very slow decarboxylation of the carboxylated product.[48]

To rationalize the difference in the activation barrier between both protonation states, we analyzed the electronic structure of the two reactants. We hypothesized that in state I the neutral carbamylated Lys201 (Kcx201) would result in a less positive magnesium cation abstracting less electron density from the O2 atom. A reduced negative charge on O2 favors electron density accumulation on the C2 atom and the nucleophilic attack on CO_2. In state II the negatively charged Kcx201 subtracts electron density of the magnesium cation resulting in a more negative O2 atom. Contrary to state I, a more negative O2 atom reduces the nucleophilicity of C2 and therefore reactivity towards CO_2. To confirm this hypothesis, we calculated the average atomic charges of the O2 atom obtained from the QM/MM molecular dynamics simulations in the protonation states I and II. Indeed, the Mulliken atomic charge

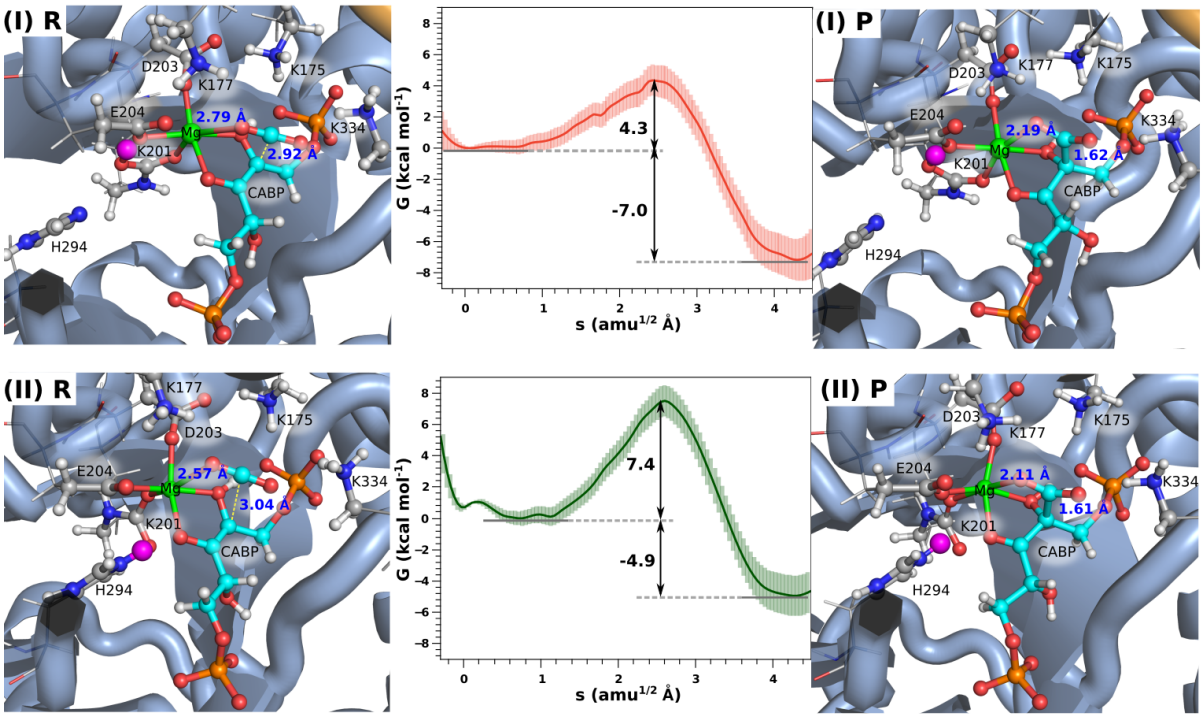

Figure 4: Carboxylation reaction of the activated ribulose-1,5-bisphosphate enolate obtained with different protonation states of His294 and carbamylated Lys201 (see Figure 3). Representative snapshots of reactant and product are shown for state I and II highlighting the involved proton in magenta. The enolate is shown in cyan and magnesium (II) cation in green with distances of CO_2 to the C2 carbon atom of the substrate. The Gibbs free energy along the Minimum free-energy path (MFEP) connecting the enolate and the carboxylated product are shown in the middle with the activation and reaction free energy for each protonation state I and II.

of the O2 atom in protonation state II is more negative than in state I (see Figure S3 in the Supporting Information). The different distribution of the substrate's electron density in the two protonation states, therefore, contributes to the smaller activation barrier observed in protonation state I.

Based on these results, carbamylated Lys201 (Kcx201) initially abstracts the proton from the C3 atom of the substrate, participates in the proton exchange between O2 and O3 to create the reactive enolate and it also reduces the activation barrier for CO_2 fixation. After formation of the carboxylated product it is probable that it looses its proton to His294 assisting the hydration reaction as proposed recently by Gready et al.[13,22]

In the crystal structure of the carboxylated product analogue the sixth coordination site

of the magnesium (II) cation is occupied by an oxygen atom of the carboxyl group. It has been postulated, that before CO_2 addition the empty coordination site would be occupied by a water molecule.[7] To test how this alternative coordination of the magnesium (II) cation alters the carboxylation reaction, we added a water molecule at the sixth coordination site of magnesium in the reactant with the activated enolate anion. Then, we calculated the free energy change along the minimum free energy path to the carboxylated product for both protonation states I and II. The activation barrier for the protonation state I with the protonated Lys201 residue was more than doubled with the additional water molecule coordinating magnesium (see Figure S1 in Supporting Information). For the state II with protonated His294 and negatively charged Lys201, the barrier decreases by 1.4 kcal mol^{-1}. These results suggest that a water molecule hinders CO_2 fixation in state I where a strong interaction between the entering CO_2 molecule and the magnesium (II) cation is expected. The resulting free energy barrier is still larger than the lowest barrier observed for state I without coordinating water molecule, excluding this path for the CO_2 fixation step.

Our QM/MM molecular dynamics simulations support a mechanism in which the carbamylated Lys201 (Kcx201) first abstracts the proton from the C3 atom and donates it back to the O2 atom. After protonating the O2 atom it is probable that an equilibrium between two protonation states exists. In the first state, the negatively charged carbamylated Lys201 becomes protonated and in the second one the proton is transferred to His294. From these two possible protonation states only the one with protonated Kcx201 residue will react with CO_2 because of its significantly smaller barrier. The reaction is irreversible due to the negative reaction free energy leading to a stable carboxylated product.

Projector-based embedding post-HF electronic structure calculations provide accurate activation barriers for the CO_2 fixing reaction

After having established the reaction mechanism, we used high-level electronic structure methods to reduce possible errors from the semi-empirical DFTB3 method and establish reference values for the activation barrier and reaction free energy of the carboxylation reaction in state I with a protonated (Kcx201) residue. First, we calculated the minimum potential energy path in the enzyme with the ChemShell[35,36] package along the reaction coordinate defined as the distance between C2 of the substrate and the carbon atom in CO_2 at the B3LYP/6-31G(d) level of theory. We performed electronic structure calculations of the optimized structures with several DFT functionals, Hartree-Fock and post-HF methods combined with various basis sets for reactant, transition state and products. The error of the activation barrier ($\Delta\Delta E^{\ddagger}_{\mathrm{QM/MM}}$) and reaction free energy ($\Delta\Delta E^{\circ}_{\mathrm{QM/MM}}$) was approximated as the difference between the potential energy barrier and reaction energy obtained with the DFTB3 semi-empirical Hamiltonian and the DFT, HF or post-HF method.

For most DFT xc-functionals the correction of the activation barrier is below 2 kcal mol^{-1}, with exception of PBE which would result in a barrier-less reaction with the smallest basis set after applying the correction to the free energy barrier (see Figure 5). The correction from the Hartree-Fock method would more than double the barrier. These two methods, however, present limiting cases. DFT xc-functionals lacking an explicit HF exchange contribution tend to underestimate activation barriers and the Hartree-Fock method tends to overestimate them due to the lack of electron correlation. The other hybrid DFT xc-functionals incorporating different amount of HF exchange are found between these two limits and their values are partially modulated by the amount of HF exchange incorporated. Accounting for 86 atoms in the active site, our best estimate of the barrier correction at the post-HF level is 2.3 kcal mol^{-1} obtained with the SCS-MP2 method and an augmented

double zeta basis set shown as dotted lines in Figure 5. We have already shown in the past that this level of theory is accurate enough to characterise the carboxylation reaction on a RuBisCO fragment model.[17] This reference value is not matched by any DFT xc-functional; but it might still possess an error contribution from the small basis set. The large system size, however, prevented the use of triple zeta basis set at the MP2 or SCS-MP2 level to improve the accuracy of the correction.

Together with the activation barrier, we also evaluated the correction of the reaction free energy with DFT and post-HF methods. The corrections are larger in magnitude than the ones for the barrier as shown on the left of Figure 5. In all cases reaction free energies are more negative than the ones obtained with the DFTB3 method. We observed large variations in the corrections for different basis sets of up to 12 kcal mol^{-1}. For a given basis set, different DFT xc-functionals provide a wide range of values without matching the SCS-MP2 reference value with the augmented double zeta basis set for the correction of the reaction free energy of -5.1 kcal mol^{-1}. One possible explanation of the large variations between functionals and basis sets is the large contribution of electron correlation energy difference between the CO_2 molecule and carboxylated products established recently by us.[49] These difference is hardly reproduced by any DFT xc-functional. Extended basis sets in combination with post-HF methods are necessary to reach the chemical accuracy of 1 kcal mol^{-1}.[49]

The large variations of the correction predicted from the DFT xc-functionals make a reliable correction of the free energy difficult and for the post-HF method a triple zeta basis set would be required to provide accurate reference values. The system size, however, prevented us from using larger basis sets with our computational resources.

To provide reference values of activation barrier and reaction energy at post-HF level with triple zeta basis set, we used the projector-embedding method.[17,38,39] As we have previously mentioned, we already used this methodology[17] to explore the carboxylation reaction in RuBisCO for a small fragment similar to Gready et al. study.[7] This method reduces the computational cost using post-HF methods for a smaller fraction of the atoms in the QM re-

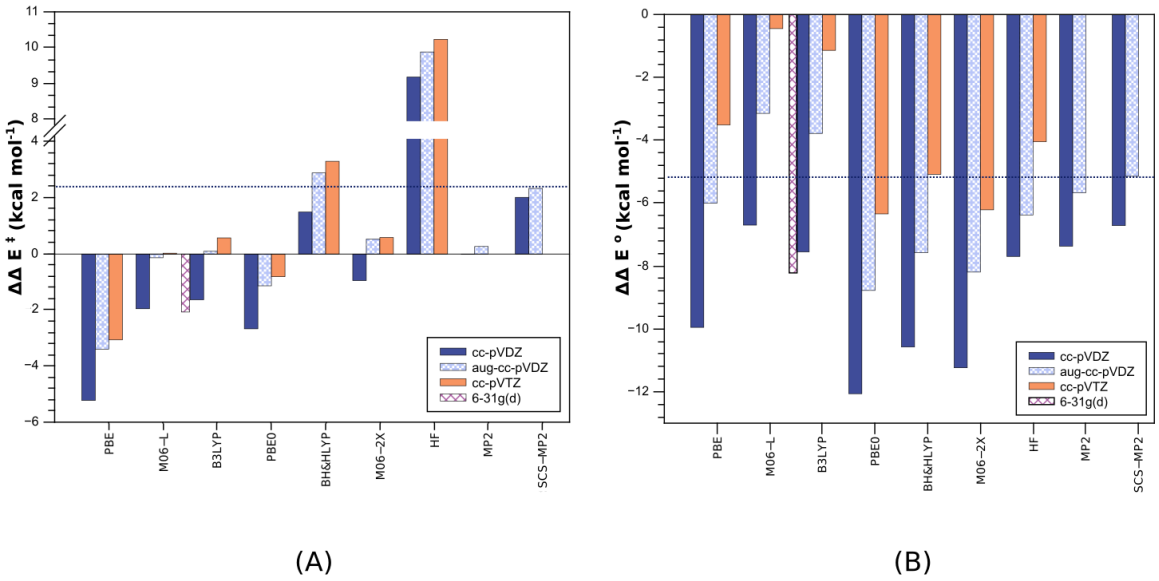

Figure 5: Correction to the potential energy barriers and reaction energies for the carboxylation reaction with protonation state I, using different DFT, HF, and post-HF methods and basis sets. The best estimate at the SCS-MP2/aug-cc-pVDZ level is indicated by the dotted line.

gion and the rest of the atoms in the region are described by a DFT xc-functional.[17,38,39] The electrostatic polarizing effect of the rest of the enzyme is still included within the QM/MM methodology. The atoms included in the high-level region that are required to capture the electronic changes of the reaction by the post-HF method were varied accounting for key residues in the active site until the difference to the full SCS-MP2 calculation was below 1 kcal mol^{-1} for the decarboxylation reaction (see Figure S2 in Supporting Information). We observed a large basis set dependence for the correction of the activation barrier and the reaction energy increasing the basis set to triple zeta (see Table 2). With the larger basis set the results obtained with the computationally less demanding PBE functional within the projector-embedding method were within chemical accuracy of the B3LYP results with the larger basis set. Based on these results, the activation free energy barrier for the carboxylation reaction at the DFTB3 level is accurate within 1 kcal mol^{-1} and the reaction free energy is more negative by up to 5 kcal mol^{-1} .

It has been observed experimentally, that metals with d-orbitals slow down the carboxy-

Table 2: Correction to the activation and reaction free energies obtained from projector-embedding calculations with different functionals and basis sets.

Method	$\Delta\Delta E^{\ddagger}_{\mathrm{QM/MM}}$		$\Delta\Delta E^{\circ}_{\mathrm{QM/MM}}$	
	cc-pVDZ	cc-pVTZ	cc-pVDZ	cc-pVTZ
B3LYP	-1.6	0.6	-7.6	-1.1
PBE	-5.2	-3.1	-10.0	-3.5
SCS-MP2 - in - B3LYP	-0.2	0.4	-8.7	-5.2
SCS-MP2 - in - PBE	-0.7	0.2	-9.8	-5.8

lation reaction.[12,50] As a final step, we tested whether changing the magnesium (II) cation by a zinc (II) cation with d-orbitals changes the activation barrier for the carboxylation reaction. All d-orbitals are filled with electrons in the zinc (II) cation matching the closed valence shell of the magnesium (II) ion. Other metals studied in the experiments have only partially filled d-orbitals and would require more advanced electronic structure calculations. We first optimized the coordination sphere of the zinc (II) cation at the B3LYP/6-31G(d,p) level of theory for the reactant and transition state. The obtained activation barrier in projector-based embedding calculations (at the SCS-MP2-in-B3LYP level with the cc-pVTZ basis set) is 3 kcal mol^{-1} larger than for the magnesium (II) cation. These results suggest that a closed-shell cation with filled d-orbitals would slow down the carboxylation reaction. Measured rate constants of RuBisCO with other metals as Mn^{2+} and Co^{2+} confirm that the carboxylation reaction is slower compared to the reaction with magnesium.[50] Even when the effect with Zn (II) cation can be less favorable for carboxylation reaction, we can not rule out its potential effect on the oxygenation reaction. An alternative way to increase the efficiency of RuBisCO enzyme may be to discourage the secondary oxygenation reaction.

Combining QM/MM molecular dynamics simulations with high-level ab-initio electronic structure calculations we provide accurate free energy barriers and reaction free energies for the carboxylation reaction. These values describe the changes in the electronic energy of the carboxylation reaction at the SCS-MP2/cc-pVTZ level using the projector-based embedding method of an active site composed by 86 atoms including the polarizing effect of the enzyme and solvent. Additionally, our reference values account for the multiple conformations of

the enzyme residues and water molecules and possible entropy changes in the reaction. The resulting activation free energy barrier and reaction free energy are 4.7 and -12.2 kcal mol^{-1}, respectively. The carboxylation reaction, therefore, has a small barrier compared to the experimentally observed catalytic rate constant and is irreversible leading to a stable carboxylated product. Before CO_2 is fixed different proton exchanges between the substrate, carbamylated Lys201 and His294 take place from which only one protonation state leads to the carboxylation reaction. Our results provide a microscopic picture for the proposed mechanism by Flamholtz and collaborators, in which the difference in the carboxylation reaction of different RuBisCO forms is attributed mostly to different populations of reactive enolates in the active site because of mutations or quaternary structure.

Discussion

Our activation and reaction free energy match the potential energies reported recently by Gready and collaborators using similar protonation states of the active site residues.[14] Yet, their values depend critically on the size of the fragment model used in the calculation and do not include the dynamics of the enzyme and the entropic contribution to the free energy. It is likely that error cancellation stemming from the electronic structure calculation and the neglect of the enzyme and solvent entropy lead to similar numerical values although they describe different physico-chemical properties.

We observe a significant difference of more than 10 kcal mol^{-1} to the most recent study including water wires by the same authors.[13] We suggest that additional water molecules placed in the initial preparation of the structure and the protonated Glu60 residue are the main factors explaining the discrepancy. Protonation of Glu60 in the active site would require a considerable free energy considering that its deprotonated form is stabilized by the positively charged Lys334 residue at 3.3 Å distance.

Although no measured value for the activation free energy of the first elementary step ex-

ists, our results are in agreement with various experimental observations. We assign residue His294 a participation in the proton exchange equilibrium after initial enolate formation and we suggest that it possibly also assist subsequent elementary steps involving hydration and bond scission. One would expect that mutation of this residue would alter the enolization equilibria and the carboxylation reaction. Indeed, Harpel and collaborators have shown that H294Q and H294N mutant in *Rodospirillum Rubrum* are greatly impaired in carboxylation and enolization exchange activity relative to wild type RuBisCO.[51] Additionally, the predicted increase in the activation barrier by the zinc (II) cation filled with d-orbitals also agrees with the measured slower carboxylation reaction by Tchekerz,[12] although the nature of the transition metals used in the experiment (Mn^{2+}) differ.

The participation of the H3 proton in the carboxylation fixation reaction is confirmed by recent experiments using deuterated solvent and RuBP substrate.[12] Analyzing the deuterium isotope effects on the enolization reaction and the carboxylation reaction the authors conclude that H3 of RuBP participates in the geometry of the transition state associated with CO_2 addition.[12]

Conclusions

Starting from the crystal structure of RuBisCO with the carboxylated product analogue (2CABP), we created a model of the catalytic active dimer of the two large sub-units in aqueous solution conserving expected protonation states of the residues in the active site. Using QM/MM molecular dynamics simulation we calculated the minimum free energy path connecting the enolate species with the carboxylated product to obtain the activation and reaction free energy of two possible protonation states I and II (see Figure 3 and 4). Our results show that only the protonation state in which the carbamylated Lys201 residue is protonated (state I) contributes to CO_2 fixation because of the smaller free energy barrier. Based on high-level ab-initio calculations the activation free energy barrier is 4.7 kcal mol^{-1}

and the reaction is exergonic (-12.2 kcal mol^{-1}) leading to an irreversible carboxylation reaction in agreement with previous experimental observations.

The measured rate constant (k_{cat}) for the whole sequence of elementary steps associated with RuBisCO results in a barrier of 16.8 kcal mol^{-1} that overcomes our value. The CO_2 addition elementary step is therefore not rate limiting. It is probable that the hydration or C-C bond breaking elementary steps constitute the rate limiting step determining the value of k_{cat}.

Our results support the recently proposed mechanism by Flamholtz *et al.*[11] in which the active site fluctuates between nonreactive and reactive states prior to the addition of either CO_2 or O_2. Based on our results, protonation state I in Figure 3 would correspond to the reactive and state II to the nonreactive state. More interesting, Flamholtz *et al.* proposed that RuBisCOs from different organisms present varying occupancy of both states resulting from evolutionary tuning. Our assignment of reactive and nonreactive states to specific microscopic protonation states represents a starting point to address how different RuBisCO forms might alter the protonation equilibrium and at the same time provides new insights for enzyme engineering efforts that could stabilize the reactive state.

Acknowledgement

The authors want to thank Simon Bennie for his contributions in the calculations with the projector-embedding method. O.A.D-G., J.O and E.V.M acknowledge financial support by Comisión Nacional de Investigaciones Científicas y Tecnológicas (CONICYT-Chile). O.A.D.-G. thanks Fondo Nacional de Desarrollo Cientifico y Tecnologico (Fondecyt) for his postdoctoral fellowship No. 3170029 and European Molecular Biology Organization (EMBO) for his EMBO short-term fellowship STF 8022 to visit Prof A. J. Mulholland's Lab. J.O thanks Fondecyt for his postdoctoral fellowship No. 3210287. E. V.-M. and J.M thanks for his Fondecyt grant No. 1200369, and the financial support provided by CONICYT PCI insti-

tuto Max Planck for terrestial microbiology Marburg MPG190003 collaboration project and Max-Planck Society through a Partner Group. A.J.M. thanks the Engineering and Physical Sciences Research Council (EPSRC) UK for support (grant number EP/M022609/1).

Supporting Information Available

Figure S1 shows the free energy profiles of the carboxylation reaction with a water molecule occupying the sixth coordination site of the magnesium (II) cation, Figure S2 displays the convergence of the activation energy correction with size of the high-level region in the projector-based embedding electronic structure calculations and Figure S3 provides representative snapshots and average atomic charges of the carbon and oxygen atom of the RuBP substrate dependent on the protonation state of Kcx201.

References

(1) Vitillo, J. G. Magnesium-based systems for carbon dioxide capture, storage and recycling: from leaves to synthetic nanostructured materials. RSC Adv. **2015**, 5, 36192–36239.

(2) Mikkelsen, M.; Jorgensen, M.; Krebs, F. C. The teraton challenge. A review of fixation and transformation of carbon dioxide. Energy Environ. Sci. **2010**, 3, 43–81.

(3) Yang, H.; Xu, Z.; Fan, M.; Gupta, R.; Slimane, R. B.; Bland, A. E.; Wright, I. Progress in carbon dioxide separation and capture: A review. J. Envioron. Sci. **2008**, 20, 14 – 27.

(4) Calvin, M. The Path of Carbon in Photosynthesis. Angew. Chem. Int. Ed. **1962**, 1, 65–75.

(5) Erb, T. J. Carboxylases in Natural and Synthetic Microbial Pathways. Appl. Environ. Microbiol. **2011**, 77, 8466–8477.

(6) Portis, A. R.; Parry, M. A. J. Discoveries in Rubisco (Ribulose 1,5-bisphosphate carboxylase/oxygenase): a historical perspective. Photosynth. Res. **2007**, 94, 121–143.

(7) Kannappan, B.; Gready, J. E. Redefinition of Rubisco Carboxylase Reaction Reveals Origin of Water for Hydration and New Roles for Active-Site Residues. J. Am. Chem. Soc. **2008**, 130, 15063–15080.

(8) Tommasi, I. C. The Mechanism of Rubisco Catalyzed Carboxylation Reaction: Chemical Aspects Involving Acid-Base Chemistry and Functioning of the Molecular Machine. Catalysts **2021**, 11, 813.

(9) Bar-On, Y. M.; Milo, R. The global mass and average rate of rubisco. Proc. Natl. Acad. Sci. USA **2019**, 116, 4738–4743.

(10) Erb, T. J.; Zarzycki, J. A short history of RubisCO: the rise and fall of Nature's predominant CO2 fixing enzyme. Curr. Opin. Biotechnol. **2018**, 49, 100 – 107.

(11) Flamholz, A. I.; Prywes, N.; Moran, U.; Davidi, D.; Bar-On, Y. M.; Oltrogge, L. M.; Alves, R.; Savage, D.; Milo, R. Revisiting Trade-offs between Rubisco Kinetic Parameters. Biochemistry **2019**, 58, 3365–3376.

(12) Bathellier, C.; Yu, L.-J.; Farquhar, G. D.; Coote, M. L.; Lorimer, G. H.; Tcherkez, G. Ribulose 1,5-bisphosphate carboxylase/oxygenase activates O2 by electron transfer. Proc. Natl. Acad. Sci. USA **2020**, 117, 24234–24242.

(13) Cummins, P. L.; Gready, J. E. Kohn–Sham Density Functional Calculations Reveal Proton Wires in the Enolization and Carboxylase Reactions Catalyzed by Rubisco. J. Phys. Chem. B **2020**, 124, 3015–3026.

(14) Kannappan, B.; Cummins, P. L.; Gready, J. E. Mechanism of Oxygenase-Pathway Reactions Catalyzed by Rubisco from Large-Scale Kohn–Sham Density Functional Calculations. J. Phys. Chem. B **2019**, 123, 2833–2843.

(15) Cummins, P. L.; Kannappan, B.; Gready, J. E. Ab Initio Molecular Dynamics Simulation and Energetics of the Ribulose-1,5-biphosphate Carboxylation Reaction Catalyzed by Rubisco: Toward Elucidating the Stereospecific Protonation Mechanism. J. Phys. Chem. B **2019**, 123, 2679–2686.

(16) Tcherkez, G. G. B.; Bathellier, C.; Stuart-Williams, H.; Whitney, S.; Gout, E.; Bligny, R.; Badger, M.; Farquhar, G. D. D2O Solvent Isotope Effects Suggest Uniform Energy Barriers in Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Catalysis. Biochemistry **2013**, 52, 869–877.

(17) Douglas-Gallardo, O. A.; Shepherd, I.; Bennie, S. J.; Ranaghan, K. E.; Mulholland, A. J.; Vöhringer-Martinez, E. Electronic structure benchmark calculations of CO2 fixing elementary chemical steps in RuBisCO using the projector-based embedding approach. J. Comput. Chem. **2020**, 41, 2151–2157.

(18) Andersson, I. Large Structures at High Resolution: The 1.6 Å Crystal Structure of Spinach Ribulose-1,5- Bisphosphate Carboxylase/Oxygenase Complexed with 2-Carboxyarabinitol Bisphosphate. J. Mol. Biol. **1996**, 259, 160–174.

(19) Andersson, I.; Taylor, T. C. Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. Arch. Biochem. Biophys. **2003**, 414, 130–140.

(20) van Lun, M.; Hub, J. S.; van der Spoel, D.; Andersson, I. CO2 and O2 Distribution in Rubisco Suggests the Small Subunit Functions as a CO2 Reservoir. J. Am. Chem. Soc. **2014**, 136, 3165–3171.

(21) Tcherkez, G. G. B.; Farquhar, G. D.; Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proc. Natl. Acad. Sci. USA **2006**, 103, 7246–7251.

(22) Cummins, P. L.; Kannappan, B.; Gready, J. E. Revised mechanism of carboxylation of ribulose-1,5-biphosphate by rubisco from large scale quantum chemical calculations. J. Comput. Chem. **2018**, 39, 1656–1665.

(23) Cleland, W. W.; Andrews, T. J.; Gutteridge, S.; Hartman, F. C.; Lorimer, G. H. Mechanism of Rubisco: The Carbamate as General Base. Chem. Rev. **1998**, 98, 549–562.

(24) Andersson, I. Catalysis and regulation in Rubisco. J. Exp. Bot. **2008**, 59, 1555–1568.

(25) King, W. A.; Gready, J. E.; Andrews, T. J. Quantum Chemical Analysis of the Enolization of Ribulose Bisphosphate: The First Hurdle in the Fixation of CO2 by Rubisco. Biochemistry **1998**, 37, 15414–15422.

(26) Mauser, H.; King, W. A.; Gready, J. E.; Andrews, T. J. CO2 Fixation by Rubisco: Computational Dissection of the Key Steps of Carboxylation, Hydration, and C-C Bond Cleavage. J. Am. Chem. Soc. **2001**, 123, 10821–10829.

(27) Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M. Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. J. Chem. Inf. Model. **2014**, 6, 12.

(28) Case, D. A.; Cerutti, D.; Cheatham, T. AMBER 2016. **2016**,

(29) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput. **2015**, 11, 3696–3713.

(30) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. J. Comput. Chem. **2004**, 25, 1157–1174.

(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. **1983**, 79, 926–935.

(32) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. J. Phys. Chem. B **2015**, 119, 1062–1082.

(33) Zinovjev, K.; Tuñón, I. Adaptive Finite Temperature String Method in Collective Variables. J. Phys. Chem. A **2017**, 121, 9764–9772.

(34) Zinovjev, K.; Ruiz-Pernía, J. J.; Tunon, I. Toward an Automatic Determination of Enzymatic Reaction Mechanisms and Their Activation Free Energies. J. Chem. Theory Comput. **2013**, 9, 3740–3749.

(35) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell—a modular software package for QM/MM simulations. Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2014**, 4, 101–110.

(36) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations. J. Phys. Chem. A **2009**, 113, 11856–11865.

(37) Neese, F. Software update: the ORCA program system, version 4.0. Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2018**, 8, e1327.

(38) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. A Projector-Embedding Approach for Multiscale Coupled-Cluster

Calculations Applied to Citrate Synthase. J. Chem. Theory Comput. **2016**, 12, 2689–2697.

(39) Ranaghan, K. E.; Shchepanovska, D.; Bennie, S. J.; Lawan, N.; Macrae, S. J.; Zurek, J.; Manby, F. R.; Mulholland, A. J. Projector-Based Embedding Eliminates Density Functional Dependence for QM/MM Calculations of Reactions in Enzymes and Solution. J. Chem. Inf. Model. **2019**, 59, 2063–2078.

(40) Ranaghan, K. E.; Morris, W. G.; Masgrau, L.; Senthilkumar, K.; Johannissen, L. O.; Scrutton, N. S.; Harvey, J. N.; Manby, F. R.; Mulholland, A. J. Ab Initio QM/MM Modeling of the Rate-Limiting Proton Transfer Step in the Deamination of Tryptamine by Aromatic Amine Dehydrogenase. J. Phys. Chem. B **2017**, 121, 9785–9798.

(41) Zhang Xinglong,; Bennie Simon J.,; van der Kamp Marc W.,; Glowacki David R.,; Manby Frederick R.,; Mulholland Adrian J., Multiscale analysis of enantioselectivity in enzyme-catalysed 'lethal synthesis' using projector-based embedding. R. Soc. Open Sci. 5, 171390.

(42) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2012**, 2, 242–253.

(43) Werner, H.-J. et al. The Molpro quantum chemistry package. J. Chem. Phys. **2020**, 152, 144107.

(44) Oller, J.; Saez, D. A.; Vöhringer-Martinez, E. Atom-Condensed Fukui Function in Condensed Phases and Biological Systems and Its Application to Enzymatic Fixation of Carbon Dioxide. J. Phys. Chem. A **2020**, 124, 849–857.

(45) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ufimtsev, I. S.; Martínez, T. J.

TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. J. Chem. Phys. **2020**, 152, 224110.

(46) Heidar-Zadeh, F.; Richer, M.; Fias, S.; Miranda-Quintana, R. A.; Chan, M.; Franco-Pérez, M.; González-Espinoza, C. E.; Kim, T. D.; Lanssens, C.; Patel, A. H.; Yang, X. D.; Vöhringer-Martinez, E.; Cárdenas, C.; Verstraelen, T.; Ayers, P. W. An explicit approach to conceptual density functional theory descriptors of arbitrary order. Chem. Phys. Lett. **2016**, 660, 307–312.

(47) Stec, B. Structural mechanism of RuBisCO activation by carbamylation of the active site lysine. Proceedings of the National Academy of Sciences **2012**, 109, 18785–18790.

(48) Pierce, J.; Andrews, T. J.; Lorimer, G. H. Reaction intermediate partitioning by ribulose-bisphosphate carboxylases with differing substrate specificities. J .Biol. Chem. **1986**, 261, 10248–10256.

(49) Douglas-Gallardo, O. A.; Saez, D. A.; Vogt-Geisse, S.; Vöhringer-Martinez, E. Electronic structure benchmark calculations of inorganic and biochemical carboxylation reactions. J. Comput. Chem. **2019**, 40, 1401–1413.

(50) Christeller, J. T. The effects of bivalent cations on ribulose bisphosphate carboxylase/oxygenase. Biochem. J. **1981**, 193, 839–844.

(51) Harpel, M. R.; Larimer, F. W.; Hartman, F. C. Multiple catalytic roles of His 287 of Rhodospirillum rubrum ribulose 1,5-bisphosphate carboxylase/oxygenase: Active-site histidine of Rubisco. Protein Science **1998**, 7, 730–738.

Graphical TOC Entry