Supporting Information

Stec 10.1073/pnas.1210754109

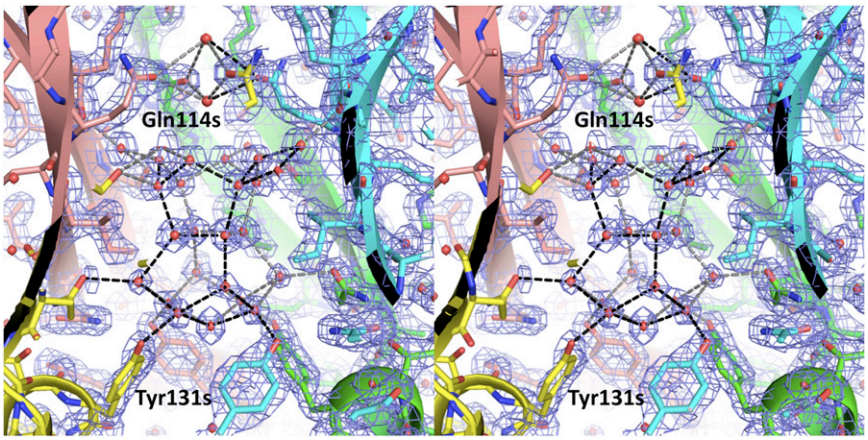

Fig. S1. The stereo representation of the eight-stranded β-barrel formed around the fourfold symmetry axis by the small subunits (marked in blue, green, beige, and yellow). In the center there is a water cluster forming an extensive clathrate around hydrophobic residues and anchored by the polar residues. One quarter of the network is connected by black dashed lines, and the rest of the network is in gray dashed lines. The barrel is reminiscent of an ion channel with two specificity filters formed by Tyr131 and Gln114 of the small subunit. The placement of the water molecules and the quality of the electron density provide support for the high quality of the data and the resulting models.

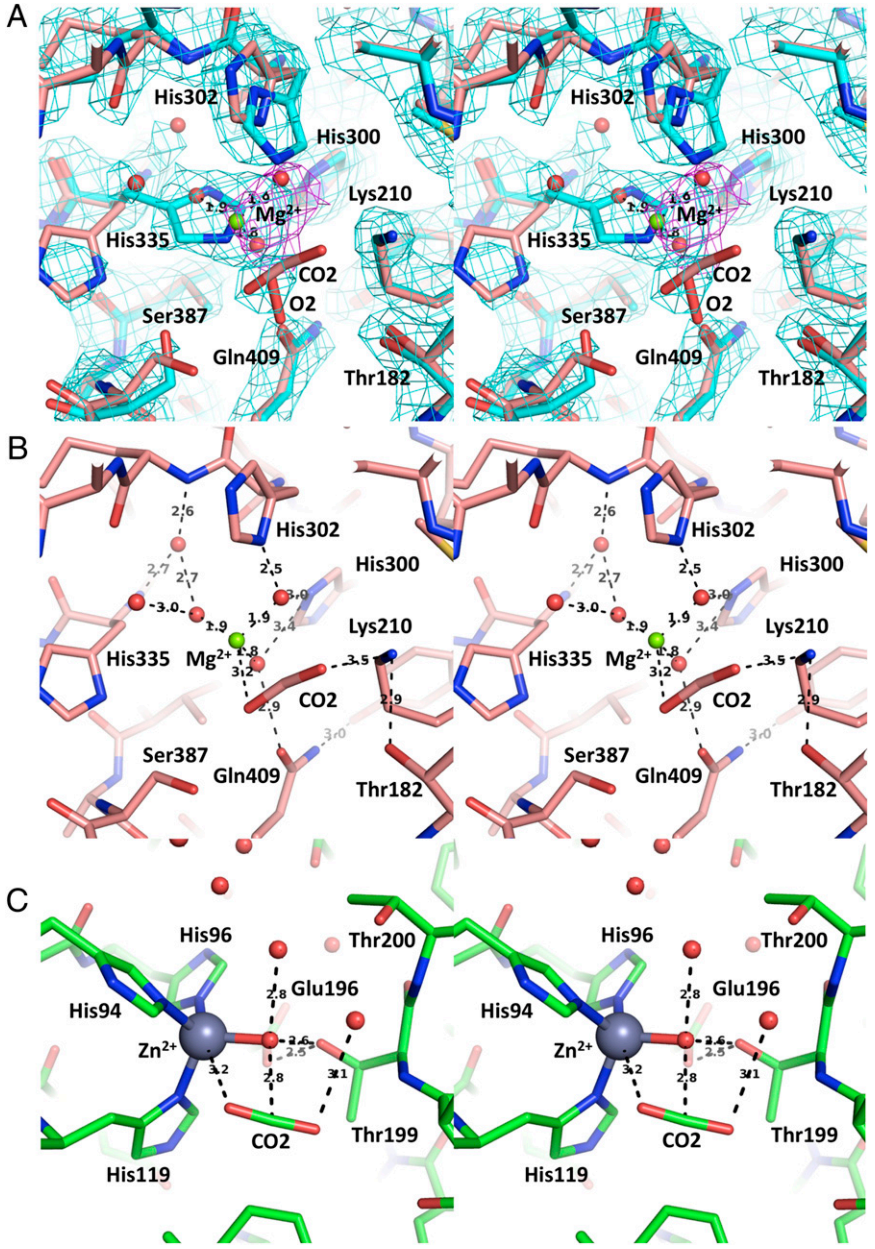

Fig. S2. (*A*) The bond-stick representation of the superposition of the refined final models of the apo-enzyme with O2 bound (in cyan) and activation complex with Mg^{2+} bound (in beige). Please note the conformational changes in the active site of the enzyme, especially the movement of His-335 and 302. The models are covered with the electron density computed with the data for the activation complex with Mg^{2+} and CO_2 bound but phased by the atomic positions of the apo-enzyme model. The difference electron density (in purple) cannot be explained by any addition to the apo model that is stereochemically correct. The apo-enzyme model does not describe the density well, and thus the explanation of the additional density requires conformational changes coupled to the Mg^{2+} ion binding. The resulting density is consistent with a metal ion and three water molecules as refined in the activation complex. The position of Mg^{2+} and the third water molecule is masked by the His335 side chain present in the apo-form. (*B*) The close-up view of the active site of the activation complex in bond-stick representation (carbon atoms in beige) showing the network of interactions of the metal ion with surrounding water molecules, the protein side chains, and the CO_2. (*C*) Detailed model of the active site of carbonic anhydrase (CA) (PDB ID code 2VVA) in bond-stick representation (carbon in green, Zn^{2+} ion in gray) with bound CO_2. The geometry of the bound CO_2 compares very favorably with the activation complex of RuBisCO with Mg^{2+} and CO_2 showing metal ion environments and surrounding water molecules. The distances of the metal ion to the bound CO_2 are highly similar in both complexes (3.2 Å) as are the planar angle between the metal and the CO_2 molecule (106° versus 109°). In both complexes (*B* and *C*), the attacking nucleophilic center (water hydroxide in CA and amino group of Lys210 in RuBisCO) is poised in a similar orientation for the initiation of the reaction. In opposition to a very stable Zn^{2+} binding in CA, the Mg^{2+} ion in RuBisCO is in a rare and unstable pseudotetrahedral environment surrounded by three water molecules (or OHs).

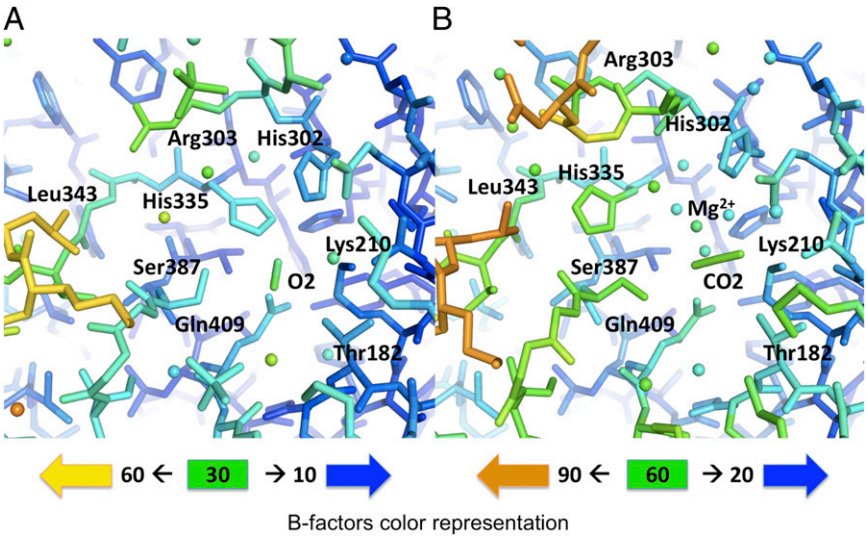

Fig. S3. The temperature factor representation of the active site. Blue represents low-temperature factors, and yellow represents high-temperature factors. (*A*) Model of the apo-enzyme structure (O_2 bound, B ~ 30 Å^2). (*B*) Model of the activation complex with Mg^{2+} and CO_2 bound, (B ~ 60 Å^2). Both models show large temperature-factor (mobility) gradients. The gaseous ligands show intermediate mobility fully comparable with the surrounding protein and solvent atoms. The model of the activation complex with Mg^{2+} and CO_2 bound shows higher temperature factors at the active site, which underscores the transient nature of the complex (despite the data having been collected on frozen crystals at 100 K).

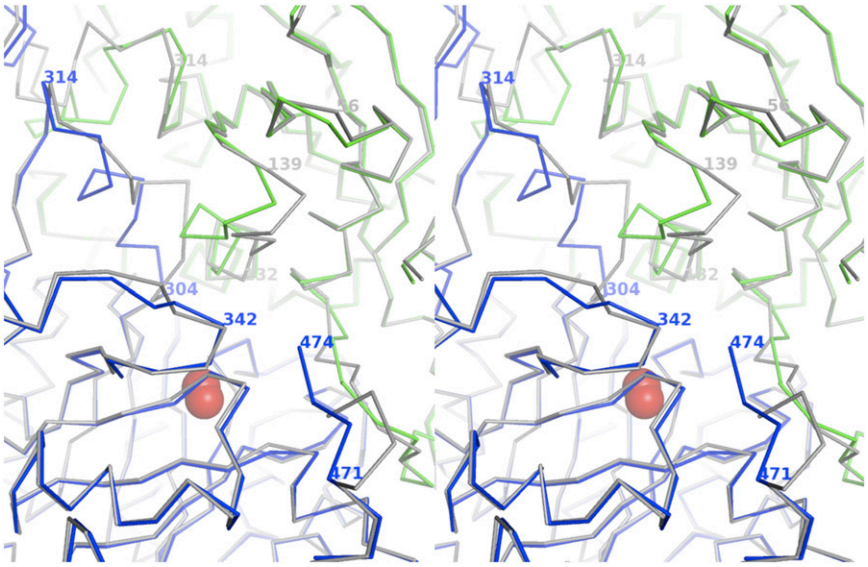

Fig. S4. The stereo representation of the α-carbon models of both superposed structures illustrates extensive conformational changes around both active sites. Blue and green represent the dimer of apo structure (O_2 bound), and gray represents both subunits of the activation complex with Mg^{2+} and CO_2 bound. The red spheres represent dioxygen bound a the active site.

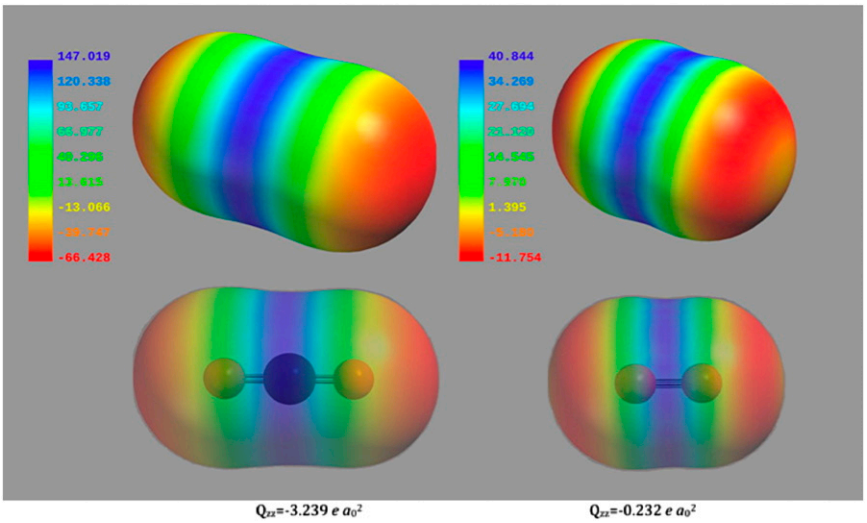

Fig. S5. The quadrupole moment of gaseous ligands is visualized by showing the molecular surface with mapped electrostatic potential. The color scheme follows commonly accepted conventions: blue, positive; red, negative. The properties of both small molecules were calculated and visualized in Spartan: (*Left*) CO_2. (*Right*) O_2. The values of the Qzz component of the quadrupole moment are listed with CO_2 having 15 times larger moment than O_2.

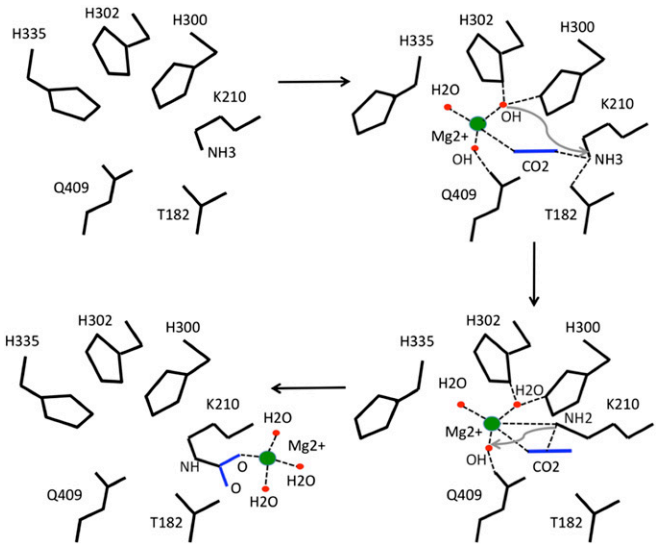

Fig. S6. Schematic of the activation reaction of RuBisCO. In step 1 (*Upper Left*), His-335 swings out to an alternative conformational state. In step 2 (*Upper Right*), a Mg^{2+} ion surrounded by two OH and a water molecule binds to the active site. In step 3 (*Lower Right*), after CO_2 binding to the metal ion, the Nε terminal atom of Lys210 swings in and positions close to the carbon and the metal ion. The transformation is aided by the proton transfer to one of the OH groups. In step 3, the second OH deprotonates the Nε atom that facilitates the nucleophilic attack of Nε on C of CO_2. In step 4 (*Lower Left*), Lys210 is carbamylated, and the Mg^{2+} is destabilized by a protonation of both OH groups to the H_2O, which leads to change of metal cation preference from tetrahedral to octahedral conformation and acquisition of a formal negative charge by the carbamyl group. Finally, the metal ion leaves an initial position driven by a conformational change in the carbamyl group of Lys210 to assume coordination with two additional carboxyl groups of Glu213 and Asp212, thus completing the activation step. Stage I has been based on our structure 4F0H and others (for example, 1IWA); stage II is based on our structure of the activation complex with Mg^{2+} and CO_2 (4F0K); stage III does not have any direct representation, but it was inferred from the carbonic anhydrase structure 2VVA where the amino group of the Lys210 is replaced by the activated water molecule bound to Mg^{2+}; stage IV, which represents the fully activated enzyme, has numerous analogs in other species, for example, 1AUS.

Table S1. Data collection and refinement statistics for the three models described

Parameter/data set	Apo + O2	Pre-activation + CO_2	Activation + H_2O
Space group	I422	I422	I422
Unit cell parameters			
a (Å)	136.44	136.14	136.32
b (Å)	136.44	136.14	136.32
c (Å)	121.67	121.29	121.52
Data collection			
Resolution (Å) (last shell)	50.0–1.96 (2.01–1.96)	50.0–2.05 (2.10–2.05)	50.0–2.25 (2.30–2.25)
I/s(I) (last shell)	25.4 (4.6)	17.4 (1.6)	12.1 (1.3)
Completeness (%) (last shell)	95.2 (98.5)	97.9 (80.8)	98.9 (91.8)
Rsym (last shell)	0.046 (0.33)	0.114 (0.611)	0.101 (0.751)
Refinement			
Resolution (Å)	40.0–1.96	40.0–2.05	40.0–2.25
No. of reflections	37,497	33,570	25,853
R-factor	0.166	0.181	0.163
Rfree	0.222	0.241	0.234
Protein B ($Å^2$)	20.4	34.0	35.3
Water molecules	334	407	353
Water molecules B ($Å^2$)	23.4	43.1	37.8
Ligands	O_2	Mg, CO_2	Mg, H_2O
Ligands B ($Å^2$)	32.2	57.8	45.8
Other ligand	PO4	Glycerol, Cl	Glycerol, Cl
Other ligand B ($Å^2$)	34.8	52.7	50.5
rms deviation from ideal			
Bond lengths (Å)	0.021	0.021	0.019
Bond angles (°)	1.906	1.849	1.774
Chiral angles (°)	0.133	0.115	0.114
PDB ID code	4F0H	4F0K	4F0M