JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

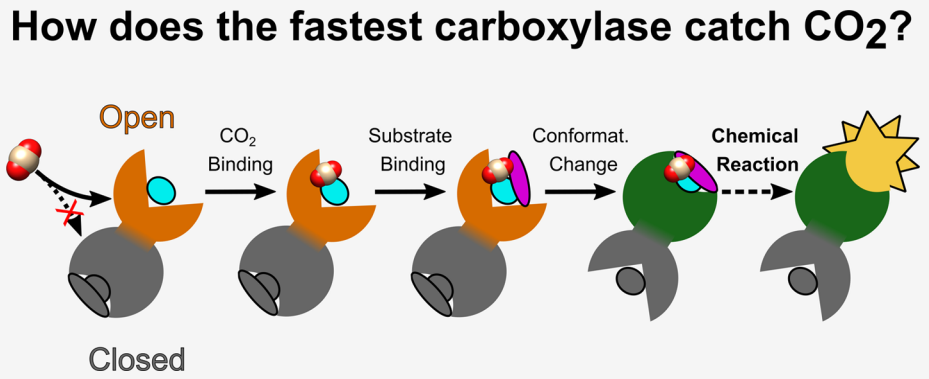
pubs.acs.org/jcim Article

Conformational Dynamics of the Most Efficient Carboxylase Contributes to Efficient CO_2 Fixation

Aharon Gomez, Tobias J. Erb, Helmut Grubmüller, and Esteban Vöhringer-Martinez*

Cite This: J. Chem. Inf. Model. 2023, 63, 7807−7815 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Crotonyl-CoA carboxylase/reductase (Ccr) is one of the fastest CO_2 fixing enzymes and has become part of efficient artificial CO_2-fixation pathways in vitro, paving the way for future applications. The underlying mechanism of its efficiency, however, is not yet completely understood. X-ray structures of different intermediates in the catalytic cycle reveal tetramers in a dimer of dimers configuration with two open and two closed active sites. Upon binding a substrate, this active site changes its conformation from the open state to the closed state. It is challenging to predict how these coupled conformational changes will alter the CO_2 binding affinity to the reaction's active site. To determine whether the open or closed conformations of Ccr affect binding of CO_2 to the active site, we performed all-atom molecular simulations of the various conformations of Ccr. The open conformation without a substrate showed the highest binding affinity. The CO_2 binding sites are located near the catalytic relevant Asn81 and His365 residues and in an optimal position for CO_2 fixation. Furthermore, they are unaffected by substrate binding, and CO_2 molecules stay in these binding sites for a longer time. Longer times at these reactive binding sites facilitate CO_2 fixation through the nucleophilic attack of the reactive enolate in the closed conformation. We previously demonstrated that the Asn81Leu variant cannot fix CO_2. Simulations of the Asn81Leu variant explain the loss of activity through the removal of the Asn81 and His365 binding sites. Overall, our findings show that the conformational dynamics of the enzyme controls CO_2 binding. Conformational changes in Ccr increase the level of CO_2 in the open subunit before the substrate is bound, the active site closes, and the reaction starts. The full catalytic Ccr cycle alternates among CO_2 addition, conformational change, and chemical reaction in the four subunits of the tetramer coordinated by communication between the two dimers.

INTRODUCTION

Fighting increasing CO_2 concentrations in the atmosphere demands a detailed knowledge of catalytic fixation mechanisms.[1,2] Nature fixes almost all atmospheric CO_2 (>95%) in the Calvin−Benson−Bassham cycle with the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), the most abundant protein on earth.[3−5] On average, RuBisCO has a CO_2 fixation rate of 1−10 CO_2 per second and reacts one in every five turnovers with oxygen.[6] In contrast, Crotonyl-CoA carboxylase/reductase (Ccr) from *Kitasatospora setae*[7] is up to 10 times faster and oxygen tolerant. Due to these features, Ccrs are used for efficient artificial CO_2-fixation pathways in vitro, paving the way for future applications.[8] The underlying mechanism of their efficiency, however, is not completely understood. Knowledge of how CO_2 fixation can be enhanced will provide protein engineers with new ways to design enzymes capable of transforming the greenhouse gas into valuable organic compounds.[8]

Recently, we studied the reaction mechanism used by Ccr to add CO_2 to its crotonyl-CoA substrate combining computational chemistry with biochemical characterization of the reaction kinetics and X-ray structures of the apo and

Received: September 11, 2023
Revised: November 21, 2023
Accepted: November 21, 2023
Published: December 4, 2023

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

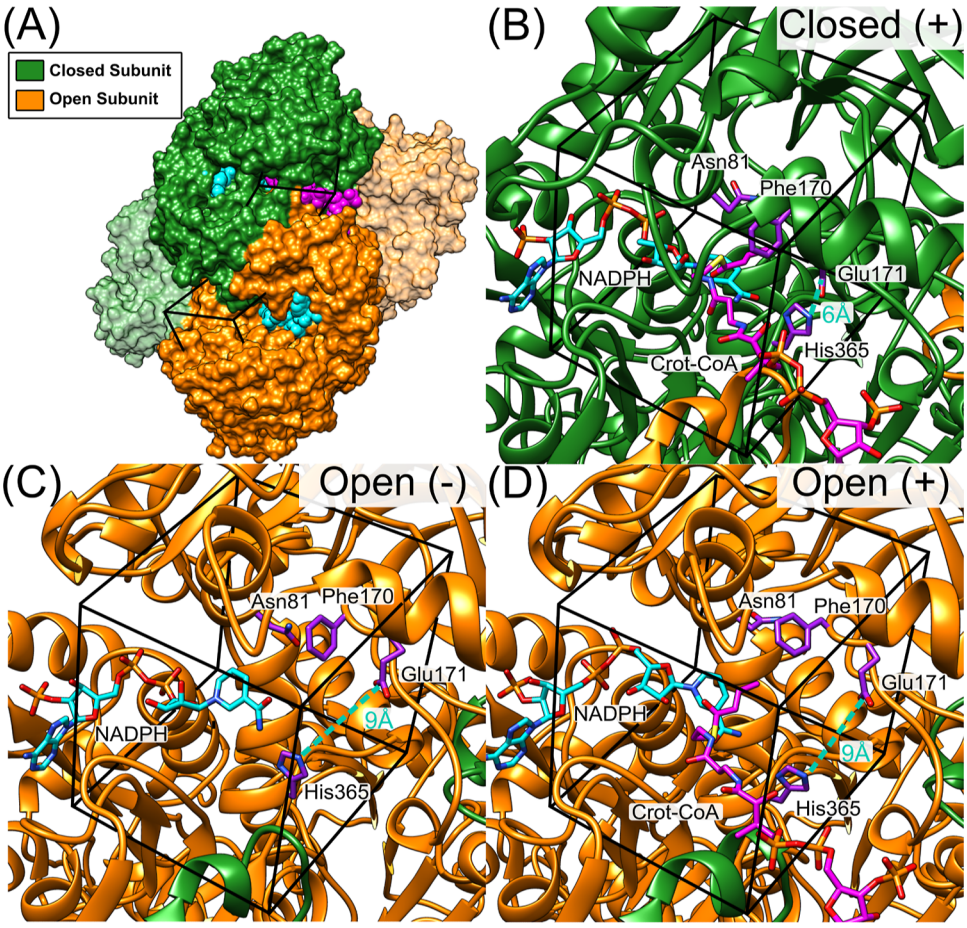


Figure 1. (A) Crystal structure of crotonyl-CoA carboxylase/reductase from *K. setae* (PDB ID: 6NA4) as a tetramer in a dimer of dimers configuration with closed (green) and open (orange) subunits. The black box defines an active site volume containing crotonyl-CoA (magenta) and NADPH (cyan) used to calculate the binding free energy. (B) Close-up of the closed(+) configuration. The subunit is closed, and the crotonyl-CoA molecule is in position for the reaction. The closed conformation is evidenced by the short distance between His365 and Glu171. (C) Close-up of the open(−) conformation with NADPH (cyan) and the four catalytic residues but no crotonyl-CoA. (D) Close up to the open(+) conformation modeled in silico. Crotonyl-CoA (magenta) was positioned as observed in the closed conformation close to NADPH (cyan). The open conformation is evidenced by the long distance between His365 and Glu171. Asn81 essential for the chemical reaction and Phe170 that shields water from the active site are shown as part of the four key catalytic residues (purple).

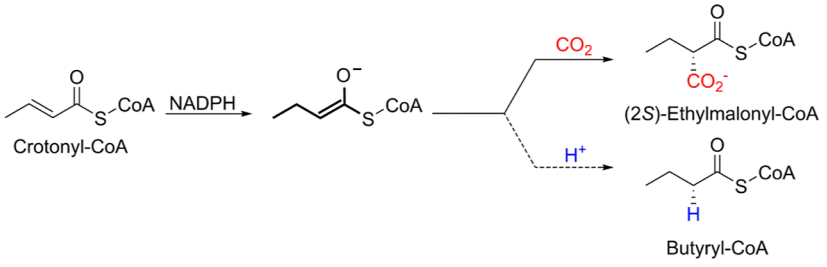


Figure 2. Reaction catalyzed by Ccr. The reactive enolate formed by NADPH reduction undergoes carboxylation to (2*S*)-ethylmalonyl-CoA in the presence of CO_2 and the reduction side reaction to butyryl-CoA in the absence of CO_2.

holoenzyme in the catalytic cycle.[7,9,10] The catalytic competent Ccr oligomer is a tetramer in a dimer of dimers configuration. Each dimer presents an open and closed subunit, as shown in Figure 1A. X-ray structures with either the product or a substrate analogue show the thioester tail only in the closed active sites [closed(+), Figure 1B], whereas the nicotinamide adenine dinucleotide phosphate (NADPH) coenzyme is present in all subunits in both crystal structures (PDB-ID: 6OWE, 6NA4).[7,9] The catalytically competent form is the closed conformation with substrates and NADPH in a favorable position for hydride transfer, as shown in Figure 1B. Although the substrate is not resolved in the active site of the open subunit [open(−), Figure 1C], electron density at the adenine end of the crotonyl-CoA suggests that it is bound to the enzyme but very flexible at the thioester end.[9] The flexibility of the substrate in the open subunit was confirmed further with computer simulations.[9] Hence, X-ray structures and computer simulations suggest that the enzyme undergoes conformational transitions between the open and closed states of its subunits and that these conformational dynamics are

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

essential for the function and efficiency of this enzyme. In particular, one would expect that the substrate first positions itself in the open subunit with NADPH before the subunit closes. Only the active site of the closed subunit maintains the substrate and the cofactor in a favorable conformation to trigger the chemical reaction.

In a previous study, we used high-level ab initio methods to study the reaction mechanism in the closed subunit assuming CO_2 already entered the catalytically active site. The reaction starts with the reduction of crotonyl-CoA at the β carbon atom of the thioester by NADPH, forming a reactive enolate (Figure 2) that represents the rate-limiting step. This enolate attacks the electrophilic CO_2 generating the carboxylated product (2*S*)-ethylmalonyl-CoA. Alternatively, it also forms a C2 adduct to avoid unfavorable side reactions with protons and store the reactive enolate as we have recently shown.[10]

Four conserved amino acids in the active site are essential for catalysis, as we have shown in site-directed mutational studies: Phe170, Asn81, Glu171, and His365(see Figure 1).[7] Phe170 shields the active site from water, and Asn81 is required for the reaction of CO_2 with the reactive enolate. A conserved water molecule acts as a bridge between Glu171 and His365 keeping both residues at a close distance (6 Å) in comparison to the open active site (9 Å) (see Figure 1B,C). The proximity of Glu171 alters the pK_a of His365 resulting in different protonation states at an experimental condition of pH 8.[10] Protonation changes of His365 are also induced by different conformations of the tetrameric enzyme.[9]

CO_2 fixation, therefore, involves binding of the substrate to the open subunit and conformational transitions to the closed state with protonation state changes of His365 before the chemical reaction can take place. This is followed by product release before a new cycle begins. We hypothesize that conformational changes and the presence of substrate would favor CO_2 binding to the active site promoting CO_2 fixation. In this case, conformational changes associated with the substrate and CO_2 binding are key steps in the catalytic cycle of the fastest carboxylase transforming this greenhouse gas into organic molecules.

Here, we used extensive molecular dynamics simulations of the Ccr tetramer to study its conformational dynamics and its effect upon CO_2 binding. We hypothesize that the conformational state of the active site and the presence of substrate modulate the CO_2 affinity. Prior to the reaction, the substrate positions itself in the open subunit, similar to the catalytically competent closed active site, before the closing of the subunit. To study this open active site with the substrate before closing of the subunit, we created an in silico model with an additional substrate molecule in the open subunit [open(+), Figure 1D]. We contrasted the likelihood of discovering a CO_2 molecule in binding sites close to the various active sites depicted in Figure 1: open(−), open(+), and closed(+). To determine whether conformation or the presence of the substrate increases the likelihood that the enzyme fixes CO_2 by reaction with the reactive enolate, we compare the average residence times of each CO_2 molecule in the various binding sites.

Our results show that CO_2 is preferentially bound to the open active site. Furthermore, substrate binding increases the time of the residence of CO_2 in a binding site close to His365. This binding site positions CO_2 favorably for the electrophilic attack on the enolate, once the active site adopts a closed conformation. Interestingly, our simulations of the Asn81Leu variant link the loss of the reactive binding sites to the observed inactivity of the CO_2 fixation.

Based on our conformational and substrate-dependent CO_2 binding, we propose a catalytic cycle for Ccr. This cycle starts with the addition of CO_2 to the open conformation followed by substrate binding. The presence of the substrate induces a conformational change to the closed conformation. CO_2 is well positioned in this conformation for the reaction with the enolate created by the substrate's NADPH reduction. After the chemical step, the active site opens to release the product, initiating a new catalytic cycle.

RESULTS

To establish the binding thermodynamics and kinetics of CO_2 to Ccr, we carried out all-atom molecular dynamics simulations. We considered different conformations of the enzyme and the presence or absence of the substrate. In addition, we calculated binding free energies from the relative probabilities to find a CO_2 molecule in the active site and the aqueous solution. We begin by comparing the thermodynamics and kinetics of CO_2 binding to the different active sites of the holoenzyme and discussing the effects of substrate binding in the closed conformation. Then, we study how the substrate in the unreactive open conformation alters CO_2 binding. Finally, we show that a single mutation (Asn81Leu) is capable of eliminating the CO_2 binding sites inactivating the carboxylase function of Ccr.

Substrate in the Closed Subunit Triggers CO_2 Binding in the Opposite Open Subunit of Each Dimer. To study the effect of substrate binding to crotonly-CoA carboxylase/reductase on the CO_2 concentration in the active sites, we started from crystal structures. In the binary complex without the substrate, the NADPH coenzyme is bound to the four subunits. In the catalytically active ternary complex a crotonyl-CoA substrate analogue occupies the two closed active sites. All-atom molecular dynamics simulations were carried out to obtain the probability to find CO_2 near the active sites compared with the bulk. The active site volume to monitor CO_2 molecules enclose the four essential residues for catalysis and the reactive parts of the coenzyme and the crotonyl-CoA substrate, as shown in Figure 1. Simulations were carried out at higher CO_2 concentrations than the one in the experiment (55 mM), to converge sampling of the CO_2 positions. Given the apolar nature of CO_2 and the low concentration, no CO_2 agglomeration in the bulk or at the active were observed. Binding free energies were calculated, as described in the Methods section and are shown in Figure 3A for the open and closed active sites of the binary and ternary complexes. The derived binding free energies are similar to the values reported for H_2 and O_2 molecules on the *Desulfovibrio fructosovorans* [NiFe]-hydrogenase, ranging from 0.5 to 1.0 k_BT units.[11] A comparison of the calculated binding free energies for the closed active sites in both complexes shows that the presence of a substrate, closed(+), disfavors CO_2 binding. Interestingly, the CO_2 binding affinity increases in the open active site without substrate, open(−), when the substrate is bound in the closed one.

We then addressed the effect of substrate binding in the open active site, creating an in silico model based on both crystal structures. Substrate binding to the open active site in the ternary complex decreases the CO_2 affinity slightly in open(+) compared to open(−). Our results demonstrate that the CO_2 concentration in the active site is modulated by the

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

conformation of the enzyme. The enzyme prepares the opposing open active site with CO_2 while the substrate reacts in the closed active site, according to the enzyme's most favorable binding affinity to open(−). After CO_2 binding, the active sites close triggering the chemical reaction.

Next, we attempted to identify the specific CO_2 binding sites in the closed and open states of the ternary complex shown in Figure 3B,C. To this aim, we calculated the CO_2 density on a grid as the ensemble average of all trajectories (see the Methods section). Binding sites in closed(+) are outside the active site (external), far from the reaction center located between the NADPH coenzyme (cyan) and substrate (magenta) (see Figure 3B). Both open(−) and open(+) active sites share three from four binding sites: substrate binding in open(+) eliminates the fourth one below NADPH in open(−). Interestingly, the two binding sites in the protein interior are near residues Asn81 and His365 (violet), which we identified to be crucial for carboxylation activity.[7] One binding site is solvent exposed at the entrance to the active site. In conclusion, the binding site analysis shows that CO_2 prefers to sit in the open active site near important residues His365 and Asn81 that have been experimentally linked to the carboxylation efficiency.

Substrate Binding in the Open Subunit Positions CO_2 for the Reaction. Our analysis of the binding thermodynamics reveals that CO_2 binds preferentially to the open active site and that substrate addition maintains binding sites close to those of His365 and Asn81. In the next step, we addressed the kinetics and asked if substrate binding alters the residence time of CO_2. For instance, a longer residence time would give the enzyme time to close the active site and initiate the reaction. We calculated the average time a CO_2 molecule stays in the binding site once it is bound (see the Methods section). Table 1 shows that the most solvent-exposed binding site, at the entrance of the active site, presents the smallest residence time because of its fast exchange possibilities to the solvent. The binding site near NADPH in open(−) has the largest residence time. Interestingly, substrate binding increases the residence time near the His365 residue, where CO_2 is positioned best to attack the enolate species once it is formed. The binding site near Asn81 presents residence times of several nanoseconds and is slightly reduced upon substrate binding. From the residence time analysis, we conclude that substrate binding increases the time that CO_2 is located in a reactive binding site to several tens of nanoseconds. This longer period would give the enzyme the time to close its active site and carry out the reaction.

Table 1. Residence Times τ in Nanoseconds Obtained for the Binding Sites Shown in Figure 3 of the Wild Type (WT) and the Asn81Leu CCr Variant (Error Estimation Obtained with Bootstrap Analysis)

Binding site	WT open(−)	WT open(+)	Asn81Leu open(+)
Asn81	8 ± 2	5 ± 2	
His365	11 ± 2	70 ± 20	
NADPH	60 ± 20		
entry	0.2 ± 0.1	0.2 ± 0.1	1.3 ± 0.1

To estimate the time required to close the active site, we considered the conformational dynamics of the ternary complex with the substrate in the two closed active sites (see the Supporting Information). The first eigenvector of the principal component analysis, which we previously identified, describes the conformational change to the open state.[9] The opening rate constant, assuming first-order kinetics, is 3.2×10^{-2} ns^{-1}. We calculated a half-life between 1 and 21.7 ns for closing the active site depending on the equilibrium constants between the closed and open states. Both values are considerably smaller than the residence time of CO_2 in binding site His365 from WT open(+), indicating that the conformational change occurs faster than CO_2 unbinding.

CO_2 Binding Site Elimination Explains Activity Loss of Asn81Leu Variant. Given that Asn81Leu variant is inactive for CO_2 fixation,[7] we asked how a single residue mutation alters the binding thermodynamics and kinetics. To this aim, we carried out the same simulations described above for the mutant and calculated binding affinity and binding sites, as shown in Figure 4. Unexpectedly, we did not observe a change in the binding affinity to the open and closed active sites of the ternary complexes as compared to the WT. Rather, the mutation eliminated binding sites near Asn81 and His365 leaving two solvent-exposed binding sites and one in the active site's entrance. One of them is shared between the wild-type and the Asn81Leu variant. The residence time shown in Table 1 of the mutant in this shared binding site is increased. This result implies that while the binding sites are exposed to the solvent maintaining affinities in the active site volumes, the critical affinities in the interior of the protein near the active site are reduced.

Taken together, these results suggest that the carboxylation activity loss of the Asn81Leu mutant is associated with the elimination of the CO_2 binding sites in the protein interior. When the Asn81Leu enzyme closes its active site for the reaction, then CO_2 is not likely to occupy a reactive position.

DISCUSSION AND CONCLUSIONS

From our results and our previous work on the reaction mechanism, we propose a new catalytic cycle for CO_2 fixation in Ccr that considers the four subunits, as shown in Figure 5A. Figure 5B describes the three steps taking place in different subunits: Prep: preparation of the subunit (blue) with bound CO_2 for the reaction: first products are released, and NADPH binds in the subunit of the other dimer (green). Accompanied by the product release a conformational change takes place where the subunit releasing the products (green) opens whereas the active site undergoing the reaction (blue) closes. The preparation finishes with substrate binding to the subunit as opposed to the closed one on the same dimer(gray). Rx: Once in the closed conformation (blue), the chemical reaction takes place following the reaction mechanism described by us previously.[10] $+CO_2$: after the reaction, this subunit (blue) opens partially. The open subunit of the other dimer without the substrate (green) accumulates CO_2 in the binding sites near His365 and Asn81 because of its highest binding affinity. In short, the full catalytic cycle consists of three steps taking place in each of the four subunits.

The proposed sequence of events composing the full cycle assumes communication between the dimers: product release and opening in one dimer triggers the closing of the active site of one subunit in the other dimer. We have previously shown that this communication is present in Ccr and can be altered by mutating residues at dimer interfaces.[9] Indeed, a triple mutant involving charged interface residues evidenced the decreased catalytic efficiency reflected in the catalytic rate constant and the Michaelis constant.

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

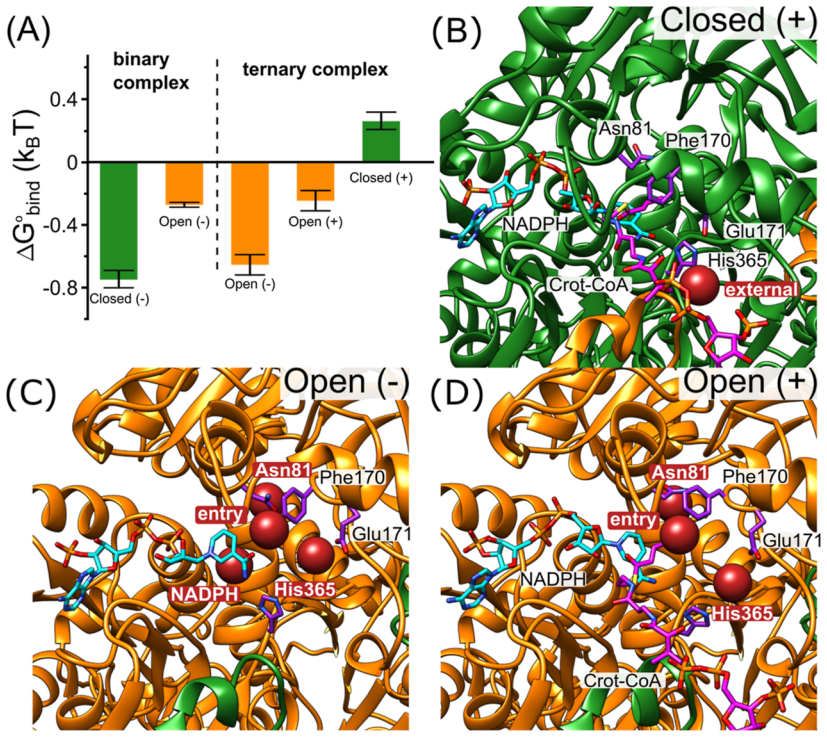


Figure 3. (A) CO_2 binding to the binary complex without the substrate and the catalytically active ternary complex. Binding free energies (k_BT) are shown for each subunit in either the open or closed conformation in the presence (+) or absence of substrate (−). Error bars are estimated with the bootstrap method. (B) Binding sites shown as red spheres in the closed(+) conformation are outside the active site located far from the four key catalytic residues (purple): Phe170, Asn81, His365, and Glu171. (C) Binding sites in the open conformation are in the protein interior near Asn81 and His365 residues and below NAPDH and one in the entry. (D) Substrate binding to the open configuration [open(+)] removes the binding site below the NADPH molecule maintaining the other two near Asn81 and His365 in the protein interior and the one in the entry.

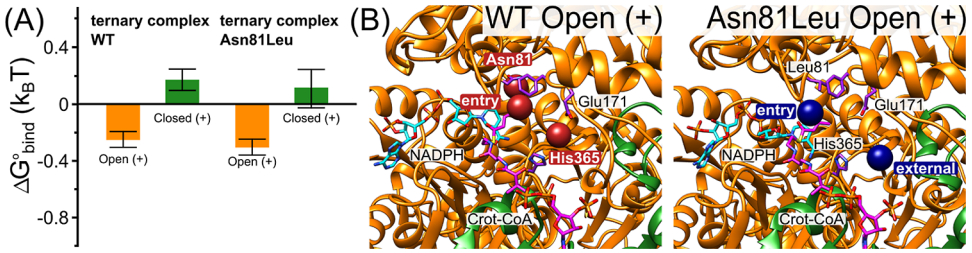


Figure 4. (A) CO_2 binding free energy (k_BT) comparison for the ternary complex with the substrate in the open(+) and closed(+) subunits for the wild type and Asn81Leu Ccr variants. (B) Binding site of the open(+) conformation for the WT and Asn81Leu variant. The buried binding site Asn81 and the reactive one near His365 disappear as a result of the mutation, leaving only the external binding site and the one at the entry of the active site.

Recently, we used Fourier transform infrared (FT-IR) spectroscopy to study CO_2 binding to Ccr.[12] We hypothesize that the communication between dimers and conformational changes plays an important role. Our computer simulations predicted that crotonyl-CoA substrate analogues eliminate CO_2 binding sites in the empty closed active site, inhibiting a possible anhydrase function observed experimentally. This prediction was confirmed by FT-IR spectroscopy. The agreement between computer simulations and spectroscopic monitoring of local changes within the active site strengthens the evidence that our methodology can provide insights into the thermodynamics and kinetics of CO_2 binding.

Conformational-dependent binding of small molecules is well established also in other proteins as hemoglobin, an efficient oxygen carrier with allosteric regulation. In the last 50 years, the allosteric addition of O_2 to the heme-group of the two identical $\alpha\beta$ dimers has been understood either as a conformational dependent binding constant to the T and R states or by sequential conformational changes of the subunits.[13] In earlier studies, various crystal structures of the unliganded (deoxy) and fully liganded hemoglobins were assigned either to the T or the R state. But more recently, structures with intermediate conformations have been solved.[13] Interestingly, these intermediate forms present a range of binding affinities. Hemoglobin, therefore, is a well-established example of the conformational-dependent binding affinity of small molecules to proteins. Its varying affinity is directly associated with its function as an oxygen carrier.

It is most likely that Ccr, similar to hemoglobin, also has intermediate states between the open and closed states of the crystal structure. In these conformations, the active site is partially closed as we propose in the CO_2 addition step. In contrast to hemoglobin, however, in Ccr the transitions between the states are not used for efficient and controlled gas

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

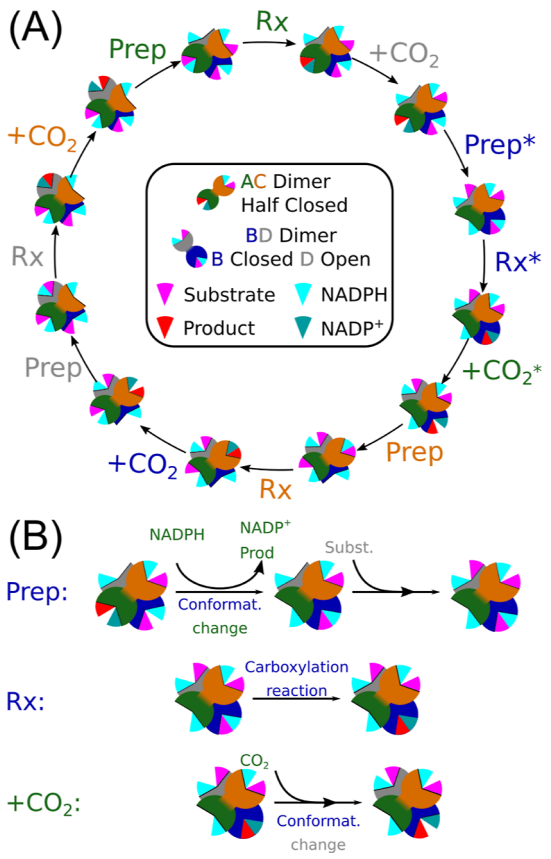


Figure 5. (A) Proposed catalytic cycle for CO_2 fixation in Ccr: The tetrameric structure is represented by two dimers with open or closed subunits (see colors in Figure 1) containing only NADPH(cyan triangle) or NADPH and substrate (pink triangle), which are transformed into NADP+(green triangle) and products(red triangle) after the chemical reaction in the closed subunits. Each step is characterized either by preparation for the reaction (Prep), the chemical reaction (Rx) or the addition of CO_2 to the open subunit(+CO_2) shown in (B). The color of the step denotes the subunit undergoing the processes.

molecule carriage triggered by environmental factors. Here, conformational changes rather concentrate and position CO_2 in the open active site followed by substrate binding, the closing of the active site and the chemical reaction. We conclude therefore that conformational dynamics of Ccr contribute to efficient CO_2 fixation.

The comparison to hemoglobin also demonstrates how the same conformational dynamics can optimize and modulate very different biochemical functions, indicating that the underlying mechanism of conformational regulation is probably a fairly universal factor that is important for a quantitative understanding of biomolecular function.

METHODS

System Preparation and Equilibration. The computational models were based on the crystal structure of the ternary complex of *K. setae* Ccr (KsCcr, PDB ID: 6NA4).[9] This corresponds to a tetrameric structure that adopts a pair of dimers geometry. Each dimer comprises a closed and an open subunit (subunits A/C and B/D), where the closed state is the catalytically competent active site conformation.[9] Each subunit contains a NADPH cofactor molecule, and the two closed subunits bear the substrate analogue butyryl-CoA (subunits A and B in Figure 1). The substrate analogue was modified to restore the original substrate crotonyl-CoA (i.e., one hydrogen atom was removed from the β-carbon and one from the α-carbon). To build the binary complex representative of the open(−) and closed(−) subunits, substrate molecules were removed from the closed subunits. The obtained coordinates are similar to the X-ray structure of the binary complex determined by some of us (PDB-ID: 6NA6).[9] To build the ternary in silico complex, representative for the open(+) and closed(+) subunits, substrate, and the complementary NADPH molecule were modeled in the open subunit fitting the oligomerization domains for the closed subunits in the open subunits of the enzyme.[9] The Asn81Leu variant was built using Chimera[14] modify structure tools.

Hydrogen addition and topology generation for the systems was done with GROMACS[15] 2019.3 software package. Protonation states of titratable residues except His365 (see below) were assigned using the Propka 3.1 program.[16,17] The interactions in the molecular system were described with the CHARMM36m force field[18] and the CHARMM TIP3P water model in combination with the CHARMM general Force Fields CGenFF[19−21] to describe CO_2, substrate, and cofactor. Force field parameters for CO_2 were validated with heptane/water partition coefficients and free energy calculations (see the Supporting Information). Specific parameters for NADPH and CoA fragments were taken from Pavelites et al.[22] and Aleksandrov and Field,[23] respectively. The protein system was embedded in a cubic box of water molecules leaving a 1 nm buffer region between the protein and box edges. Potassium and chloride ions were added to neutralize the system reaching a concentration of 125 mM. Energy minimization was performed using the steepest descent algorithm for 50,000 steps for each system followed by a 1 ns equilibration applying position restraints on the protein backbone and ligands in the *NVT* ensemble with a force constant of 1000 kJ nm^{-2} mol^{-1}, and subsequently for 5 ns in the *NPT* ensemble at 1 atm pressure and temperature of 298 K using an integration time step of 2 fs. The velocity rescaling thermostat[24] and Berendsen[25] pressure coupling were used with coupling coefficients of τ = 0.1 and 2 ps, respectively. All bond lengths of the protein and ligands were constrained using LINCS[26] with an expansion order of 4. Electrostatic interactions were calculated using particle mesh Ewald,[27] with a real space cutoff of 1.2 nm and a Fourier spacing of 0.16 nm. For the van der Waals interactions, a cutoff of 1.2 nm and a switching function starting at 1 nm was used.

As a consequence of His365 being exposed to different molecular environments depending on conformation and substrate binding [closed(+) vs open(+) vs open(−)], its protonation state was difficult to determine with the semiempirical method Propka.[17] His365 is rather solvent exposed in the open active site, whereas in the closed state, it interacts with negatively charged Glu171 via a conserved water molecule. To determine its protonation state, we performed alchemical free energy calculations for the four active sites: closed(−), closed(+), open(−), and open(+), as described in Figure 1. To assign the protonation state given the experimental pH = 8, we used the His365 pK_a shift compared to a reference histidine residue in a fully solvated peptide with methylated carboxy and amino terminals (pK_a = 6.5 for histidine in a fully solvated alanine pentapeptide). The pK_a shift of His365 in the open and closed active site results from

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

$$\Delta pK_a = \frac{\Delta\Delta G^o}{2.303RT} \quad (1)$$

To obtain the change in free energy, $\Delta\Delta G^o$, when a proton is transferred from the His365 residue in the open(−), open(+), closed(−), or closed(+) active site to the fully solvated histidine residue, we applied alchemical free energy calculations. The reference methylated histidine peptide was added at a distance at least 3.5 nm away from the His365 in the active site. This Nme-His-Ace peptide was added to the dodecahedric simulation box containing the protein, cofactor, substrate, and 1.5 nm of solvent buffer. Topologies for the protonated and deprotonated states of the two residues were created with the PMX toolset[28,29] from GROMACS 2019.3[15] using the CHARMM36m[30] force field and the CHARMM TIP3P water model[31] in combination with the parameters for substrate and cofactor described above. Tautomers and the most probable rotational conformer of His365 were assigned based on the local hydrogen bond interaction in the crystal structure. Systems preparation, minimization, and equilibration for subsequent alchemical free energy calculations followed the procedure described above. Equilibrium free energy calculations were performed for 5 ns per lambda value with 11 [0.00, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00] values for the Coulomb interactions, 9 [0.00, 0.13, 0.25, 0.38, 0.50, 0.63, 0.75, 0.88, 1.00] for the van der Waals interactions, and 7 [0.00, 0.15, 0.30, 0.50, 0.70, 0.85, 1.00] for the bonded interactions in a *NVT* ensemble with positions restraints over the protein peptide backbone and ligands at 298 K with stochastic dynamics. The free energy difference was estimated by bennet acceptance ratio[32] included in GROMACS as gmx bar.

The positive ΔpK_a shift of His365 in the substrate-bound open and closed active site, as shown in Table 2, demonstrate a

Table 2. Calculated ΔpK_a Shifts of His365 in the Four Different Active Sites

active site	ΔpK_a
closed(+)	+1.2 ± 0.1
open(+)	+0.9 ± 0.5
closed(−)	−0.6 ± 0.1
open(−)	−0.9 ± 0.1

stabilization of the positively charged histidine residue. The absence of a substrate seems to decrease His365 pK_a. Given neutral or slightly basic pH conditions, His365 would be neutral in the open(−) and closed(−) active sites (pK_a = 6.5 for histidine in a fully solvated alanine pentapeptide). For closed(+) and open(+), in contrast, the calculated pK_a shift suggests a doubly protonated His365. To validate this assumption, we performed 8 MD replicas of 100 ns using the less probable neutral form of His365 and the doubly protonated form in the closed(+) active site. But, during the simulations we observed that positively charged ions accumulate near Glu171 for the neutral form (see Figure S2). This positive charge accumulation demonstrates that the substrate together with the protein stabilizes the positive charge of His365 and confirms the ΔpK_a. Based on these results, His365 was used in the protonated state in the closed(+),open(+), and neutral in the substrate free open(−) and closed(−) active sites.

With these protonation states of His365, 100 CO_2 molecules were added, replacing water molecules. Given the number of water molecules the CO_2 concentration is ≃60 times larger than the experimental concentration of CO_2 (55 mM).[7] The increased concentration was required to sample all positions that CO_2 may adopt during the total simulation time. The minimization and equilibration protocol previously mentioned was repeated. After the pressure was equilibrated, production dynamics in the *NVT* ensemble were performed by keeping position restraints on the protein alpha carbons and NADPH heavy atoms to keep protein conformation fixed with a force constant of 1000 kJ nm^{-2} mol^{-1}. In the binary complex [open(−), closed(−)], the distance between Arg276 and NADPH's adenine moiety was restrained to their initial values. For the ternary complex [open(−), closed(+)] distances between residue Arg99 and crotonyl-CoA's phosphate groups, Arg303 and crotonyl-CoA's adenine moiety, the reactive crotonyl-CoA's C_β with NADPH's C_4 atoms, and His409 with NADPH's phosphate group for the closed subunits were restrained to their initial values. For the ternary complex [open(+), closed(+)] wild type and Asn81Leu variant, additional position restraints were added to the open(+) crotonyl-CoA molecules to explore the distribution of CO_2 in the defined state of the catalytic cycle of the enzyme. Multiple replicas were performed, generating a total of 20 μs of simulation time for the wild type systems and 10 μs for the Asn81Leu variant. The coordinates were saved every 100 ps resulting in up to 2×10^5 and 1×10^5 configurations available for analysis, respectively.

CO_2 Binding Free Energy. Binding free energies were obtained from extended MD simulations to achieve a sufficient sampling. Synthetic densities were calculated by convolution of the atomic positions with 3-D Gaussians to represent atomic densities. The widths of the Gaussians were defined by experimentally determined atomic scattering factors. The atomic density generation was performed with the gromaρs[33] analysis tools. This procedure was carried out for every configuration of the trajectory on a grid with a resolution of 0.1 nm and time averaged over trajectories. We selected 100 ns time windows to perform the analysis.

To compare the CO_2 concentration for different configurations of the active site inside the protein, we defined the solution concentration ($C_{solution}$) as a reference state. This concentration corresponds to the number of carbon dioxide molecules in the system inside the solvent accessible volume (SAV). The SAV accounts for the points in space that carbon dioxide molecules can access, and represents the volume of the simulation box without the volume occupied by the protein, NADPH and Crot-CoA. To account for the modification of the SAV due the movement of the side chains along the molecular dynamics simulation, we obtained the time averaged ⟨SAV⟩ to estimate the solution concentration

$$C_{solution} = \frac{N_{CO_2}}{\langle SAV \rangle} \quad (2)$$

To calculate ⟨SAV⟩ we selected a threshold isosurface value of 0.1 au, which recovered the overall shape of the side chains, cofactors, and substrates in the protein complex. The solvent accessible volume of each trajectory frame was defined by all grid points with a density below this threshold, corresponding to the volume not occupied by the protein, cofactor, and substrate atoms, thus accessible for the solvent and carbon

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815

dioxide molecules. ⟨SAV⟩ was obtained as the time average of 2×10^5 equally distant frames of a 20 μs trajectory.

To calculate the local CO_2 concentration in vicinity of the active site, we defined a box containing all key residues for the reaction, the nicotinamide ring and the crotonyl-CoA fragment up to the pyrophosphate (see Figure 1). We defined the local concentrations at the active sites by the number of $CO_{2,AS}$ molecules inside the box divided by the active site's accessible volume $\langle SAV \rangle_{AS}$. We obtained the active site $\langle SAV \rangle_{AS}$ following the same procedure described for the whole protein but considering only the volume of the active site box. For the number of carbon dioxide molecules inside active site $CO_{2,AS}$, we summed up all of the atomic densities inside the active site box and divided them by the density of one CO_2 molecule. This quantity gave us the number of CO_2 molecules inside the active site, which was then divided by the active site's $\langle SAV \rangle_{AS}$ to obtain the CO_2 concentration at the active site

$$C_{\text{Active site}} = \frac{N_{CO_{2,AS}}}{\langle SAV \rangle_{AS}} \tag{3}$$

This concentration was derived as the time average of the different configurations of CO_2 molecules inside the active site. This volume was corrected accounting only for the accessible volume of the solvent, which is common to the CO_2 molecules. From the ratio of local active site concentrations and the reference solution concentration, we calculated the binding free energy for the different configurations of the enzyme's active site via

$$\Delta G^{\circ}_{\text{bind}} = -k_B T \ln\left(\frac{C_{\text{Active site}}}{C_{\text{Solution}}}\right) \tag{4}$$

Binding Sites and Unbinding Kinetics. CO_2 binding sites were derived from the CO_2 density maps in the vicinity of the active sites. From the whole trajectory, we selected 100 ns time windows in which the average CO_2 concentration in the active site (box) is at least twice the bulk concentration (see the Supporting Information). The configurations of these time windows presented binding events sufficient to define binding sites and unbinding kinetics. Binding sites were defined using isosurfaces of the CO_2 concentration with isovalues 20 times larger than the bulk concentration (see Figure S4 for a comparison of different isovalues). The centers of these isosurface volumes define the center of the binding sites, as shown in Figure 3.

The residence time for CO_2 molecules was evaluated based on the average time span a CO_2 molecule remained bound to a particular binding site. To estimate the time span of the binding events, we counted the time length that the center of mass of each CO_2 molecule remains within a sphere with a radius two times a C═O bond distance of 2.4 Å (this corresponds approximately to half the distance between the closest binding sites). To obtain the time length, we built a binding state matrix $S_{N,c}$ for each of the active site conformations [open(+), open(−), and closed(+)]. This matrix had dimensions of the total number of CO_2 molecules times the total trajectory frames. The matrix elements $s_{n,c}$ took values of i = 1, 2, 3, ..., n when a specific CO_2 molecule n was bound to the binding site i or 0 otherwise. For each CO_2 molecule, we recorded the length of the binding events in which the molecule is bound to a specific binding site. The lengths of the binding event for each site were averaged to obtain its residence time. In some cases, a CO_2 molecule registered an unbinding event followed by new binding to the same binding site. This recrossings result from the geometric criteria used to define the bound state of a CO_2 molecule and lead to an underestimation of the residence time. To account for this effect, we performed an iterative protocol in which we used the previous residence time as threshold for recrossing events: if two binding events are separated by less than the residence time, they are appended and counted as only one event. The iterative cycles ends when there is no difference between residence time in two consecutive iterations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.3c01447.

Kinetic modeling of the conformational changes, validation of CO_2 force field parameters, ion distribution in active sites depending on the His365 protonation state, details on the definition of binding sites, and statistical analysis for convergence of binding affinity (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Esteban Vöhringer-Martinez − *Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepión 4030000, Chile;* orcid.org/0000-0003-1785-4558; Email: evohringer@udec.cl

Authors

Aharon Gomez − *Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepión 4030000, Chile;* orcid.org/0000-0003-4820-4697

Tobias J. Erb − *Department of Biochemistry and Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Marburg D-35043, Germany; LOEWE Center for Synthetic Microbiology (SYNMIKRO), Marburg 35032, Germany;* orcid.org/0000-0003-3685-0894

Helmut Grubmüller − *Department of Theoretical and Computational Biophysics, Max-Planck-Institute for Multidisciplinary Sciences, Göttingen D-37073, Germany;* orcid.org/0000-0002-3270-3144

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.3c01447

Notes

The authors declare no competing financial interest.

Initial structures for Ccr were taken from the PDB database with ID 6NA4 and 6NA6. All input parameters and topologies for the simulations with GROMACS are available under DOI:10.5281/zenodo.8336356. The gromaps analysis tool is freely available under Computational Biophysics Max Planck Tandem Group.

■ ACKNOWLEDGMENTS

The authors thank financial support provided by the Max-Planck Society through a Partner group and associated laboratory and the CONICYT PCI instituto Max Planck for terrestial microbiology Marburg MPG190003 collaboration project. T.J.E. received support from the Max Planck Society the European Research Council (ERC 637675 "SYBORG").

A.G. thanks ANID for the Doctorado Nacional grant no. 21190262.

REFERENCES

(1) Appel, A. M.; Bercaw, J. E.; Bocarsly, A. B.; Dobbek, H.; DuBois, D. L.; Dupuis, M.; Ferry, J. G.; Fujita, E.; Hille, R.; Kenis, P. J. A.; Kerfeld, C. A.; Morris, R. H.; Peden, C. H. F.; Portis, A. R.; Ragsdale, S. W.; Rauchfuss, T. B.; Reek, J. N. H.; Seefeldt, L. C.; Thauer, R. K.; Waldrop, G. L. Frontiers, Opportunities, and Challenges in Biochemical and Chemical Catalysis of CO_2 Fixation. *Chem. Rev.* **2013**, *113*, 6621−6658.

(2) Bierbaumer, S.; Nattermann, M.; Schulz, L.; Zschoche, R.; Erb, T. J.; Winkler, C. K.; Tinzl, M.; Glueck, S. M. Enzymatic Conversion of CO_2: From Natural to Artificial Utilization. *Chem. Rev.* **2023**, *123*, 5702−5754.

(3) Andersson, I. Catalysis and regulation in Rubisco. *J. Exp. Bot.* **2007**, *59*, 1555−1568.

(4) 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.

(5) Ellis, R. J. The most abundant protein in the world. *Trends Biochem. Sci.* **1979**, *4*, 241−244.

(6) 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.

(7) Stoffel, G. M. M.; Saez, D. A.; DeMirci, H.; Vögeli, B.; Rao, Y.; Zarzycki, J.; Yoshikuni, Y.; Wakatsuki, S.; Vöhringer-Martinez, E.; Erb, T. J. Four amino acids define the CO_2 binding pocket of enoyl-CoA carboxylases/reductases. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 13964−13969.

(8) Schwander, T.; Schada von Borzyskowski, L.; Burgener, S.; Cortina, N. S.; Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **2016**, *354*, 900−904.

(9) DeMirci, H.; Rao, Y.; Stoffel, G. M.; Vögeli, B.; Schell, K.; Gomez, A.; Batyuk, A.; Gati, C.; Sierra, R. G.; Hunter, M. S.; Dao, E. H.; Ciftci, H. I.; Hayes, B.; Poitevin, F.; Li, P.-N.; Kaur, M.; Tono, K.; Saez, D. A.; Deutsch, S.; Yoshikuni, Y.; Grubmüller, H.; Erb, T. J.; Vöhringer-Martinez, E.; Wakatsuki, S. Intersubunit Coupling Enables Fast CO_2-Fixation by Reductive Carboxylases. *ACS Cent. Sci.* **2022**, *8* (8), 1091−1101.

(10) Recabarren, R.; Tinzl, M.; Saez, D. A.; Gomez, A.; Erb, T. J.; Vöhringer-Martinez, E. Covalent Adduct Formation as a Strategy for Efficient CO_2 Fixation in Crotonyl-CoA Carboxylases/Reductases. *ACS Catal.* **2023**, *13*, 6230−6241.

(11) Wang, P.-h.; Best, R. B.; Blumberger, J. Multiscale Simulation Reveals Multiple Pathways for H_2 and O_2 Transport in a [NiFe]-Hydrogenase. *J. Am. Chem. Soc.* **2011**, *133*, 3548−3556.

(12) Gomez, A.; Tinzl, M.; Stoffel, G.; Grubmüller, H.; Erb, T. J.; Vöhringer-Martinez, E.; Stripp, S. Infrared Spectroscopy Reveals Metal independent Carbonic Anhydrase Activity in Crotonyl CoA Carboxylase/Reductase. *ChemRxiv* **2023**, chemrxiv:2022-00rfw-v2 preprint.

(13) Numoto, N.; Kawano, Y.; Okumura, H.; Baba, S.; Fukumori, Y.; Miki, K.; Ito, N. Coarse snapshots of oxygen-dissociation intermediates of a giant hemoglobin elucidated by determining the oxygen saturation in individual subunits in the crystalline state. *IUCrJ* **2021**, *8*, 954−962.

(14) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera−A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

(15) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.

(16) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.

(17) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK_a Values. *J. Chem. Theory Comput.* **2011**, *7* (7), 2284−2295.

(18) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.

(19) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671−690.

(20) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155−3168.

(21) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144−3154.

(22) Pavelites, J. J.; Gao, J.; Bash, P. A.; Mackerell, A. D., Jr. A molecular mechanics force field for NAD+ NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* **1997**, *18*, 221−239.

(23) Aleksandrov, A.; Field, M. Efficient solvent boundary potential for hybrid potential simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10503−10509.

(24) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(25) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(26) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(27) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(28) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan. *Angew. Chem., Int. Ed.* **2016**, *55*, 7364−7368.

(29) Gapsys, V.; de Groot, B. L. pmx Webserver: A User Friendly Interface for Alchemistry. *J. Chem. Inf. Model.* **2017**, *57*, 109−114.

(30) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.

(31) Boonstra, S.; Onck, P. R.; van der Giessen, E. CHARMM TIP3P Water Model Suppresses Peptide Folding by Solvating the Unfolded State. *J. Phys. Chem. B* **2016**, *120*, 3692−3698.

(32) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245−268.

(33) Briones, R.; Blau, C.; Kutzner, C.; de Groot, B. L.; Aponte-Santamaría, C. GROMaps: A GROMACS-Based Toolset to Analyze Density Maps Derived from Molecular Dynamics Simulations. *Biophys. J.* **2019**, *116*, 4−11.

https://doi.org/10.1021/acs.jcim.3c01447
J. Chem. Inf. Model. 2023, 63, 7807−7815