Binding Sites of Bicarbonate in PEP Carboxylase

*Nicolas Chéron**

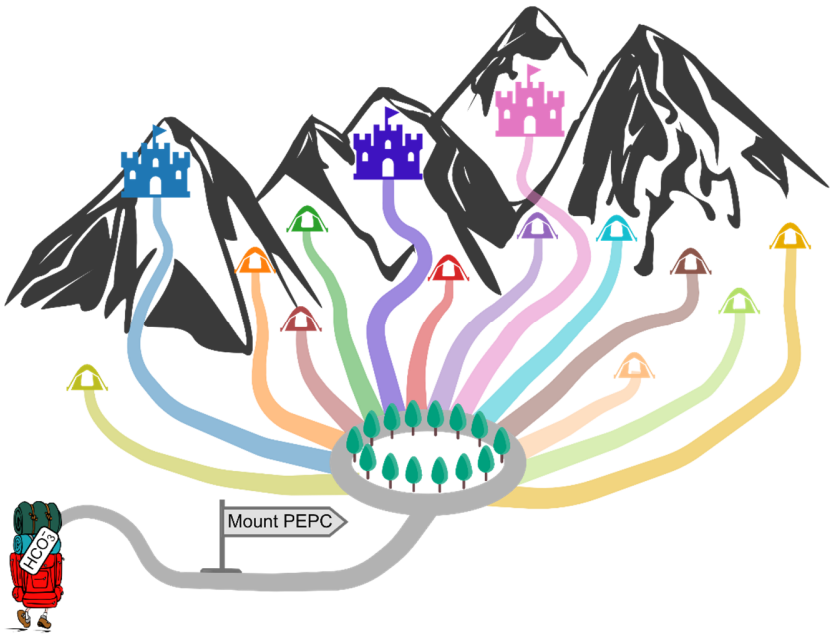
PASTEUR, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France

Corresponding Author:

*Email: nicolas.cheron@ens.psl.eu

ABSTRACT (141/150 words): PhosphoEnolPyruvate Carboxylase is used in plant metabolism for fruit maturation or seed development, as well as in the C4 and CAM mechanisms in photosynthesis where it is used for the capture of hydrated CO_2 (bicarbonate). To find the yet unknown binding site of bicarbonate in this enzyme, we have first identified putative binding sites with non-equilibrium molecular dynamics simulations, and then ranked these sites with alchemical free energy calculations with corrections of computational artefacts. 14 pockets where bicarbonate could bind were identified, with three having realistic binding free energies with differences with the experimental value below 1 kcal/mol. One of these pockets is found far from the active site at 14 Å, whereas in the two others bicarbonate is in direct interaction with the magnesium ion. The study of mutant K606N allowed to discriminate between these two pockets and to identify the binding site of bicarbonate in PhosphoEnolPyruvate Carboxylase.

TOC GRAPHICS



KEYWORDS: Molecular dynamics, Binding free energy, Enzymes, Photosynthesis, Carbon dioxide.

Introduction

Every carbon atom in materials or living organisms that surrounds us was part of a CO_2 molecule in the past, and has been captured by some autotrophic organism during the second step of photosynthesis (Calvin-Benson cycle[1]). This complicated process transforms carbon from an inorganic state (CO_2) to an organic state by forming a C-C bond, and is thus directly or indirectly the prime source of growth of all living organisms. Three different mechanisms of carbon-capture are used during photosynthesis: C3, C4 and CAM (Crassulacean Acid Metabolism). 97% of plant varieties belongs to the C3 family and simultaneously captures and integrates the atmospheric CO_2 in the organism *via* the enzyme Rubisco[2]. In the C4 and CAM families, first the enzyme PEP Carboxylase (PEPC) catalyses the reaction between hydrated CO_2 (HCO_3^-) and phosphoenolpyruvate (PEP) to form oxaloacetate (see Figure 1). This metabolite is then converted and transported to another cell where CO_2 is released and incorporated in the organism by Rubisco during the Calvin-Benson cycle[2]. Separating the capture of CO_2 by PEPC from its integration in the organism by Rubisco in two different cells prevents a costly side reaction also catalysed by Rubisco when in contact with O_2 (photorespiration), and C4 plants are then more efficient than C3 plants for carbon capture, especially in hot environments[3,4]. Even if CO_2-capture with the C4 or CAM mechanisms is found in only 3% of plant varieties, C4 plants represent 23% of terrestrial carbon fixation[2], including plants of huge importance such as *sugarcane*, *maize* (which is the most grown cereal in the world[5], used to feed humans and livestock), *millet* (which is one of the only edible plant that can grow in arid regions) or *switchgrass* (which is a serious candidate for developing the production of biocarburant[6]). The reaction catalysed by PEPC is thus at a central corner for both food and resource productions.

In the context of an ever increasing global population (which is expected to grow from 7 billion in 2011 to 9 billion in 2050[7]), there will be a need for improved crop yields to provide food, cloth, or sources of clean energy. Rubisco has long been considered the rate-limiting step of photosynthesis, however this is probably true only in conditions of high light, high temperature, and for C3 organisms[8,9]. Indeed, when other enzymes of the Calvin-Benson cycle such as sedoheptulose-1,7-bisphosphatase (SBPase) are overexpressed in different organisms, the photosynthetic rates are increased which lead to more biomass[10–12], suggesting that Rubisco is not the bottleneck. In contrast, the overexpression of Rubisco in rice can lead to a decrease of the CO_2 assimilation rate at 25°C (and an increase at 40°C)[13]. Moreover, the estimated rate of transformations of metabolites from ribulose-5-phosphate-3-epimerase (RPE) is estimated to be slightly lower to the one from Rubisco[14]. These results are in line with the statement that despite a slow kinetic rate, "Rubisco is not really so bad"[15,16]. Thus, the reality is probably more complex than the textbook description of a single enzyme which would be the bottleneck of photosynthesis and of plant growth. In ambient conditions for C4 organisms, all enzymes from the Calvin-Benson cycle are involved in controlling the flux of carbon[8,17] and without a fully efficient cycle the production of reactants can be limited. Hence, the paradigm for plant engineering has changed from being mainly focused towards Rubisco (with limited success), to an increasing interest towards the C4-mechanism (e.g., the C4 Rice Project[18]). However, to engineer plants, one needs to fully understand their molecular components (proteins) and knowledge on some enzymes of the photosynthesis cycle is still not complete. We propose here to focus on PEPC which is at the beginning of the C4-mechanism and may be involved in the overall control of carbon flux[9,19]. In addition to photosynthesis in C4 plants, we underline that PEPC is used in all plants metabolism for functions such as fruit maturation, seed development or amino-acid synthesis and is thus a central enzyme for the plant kingdom[20].

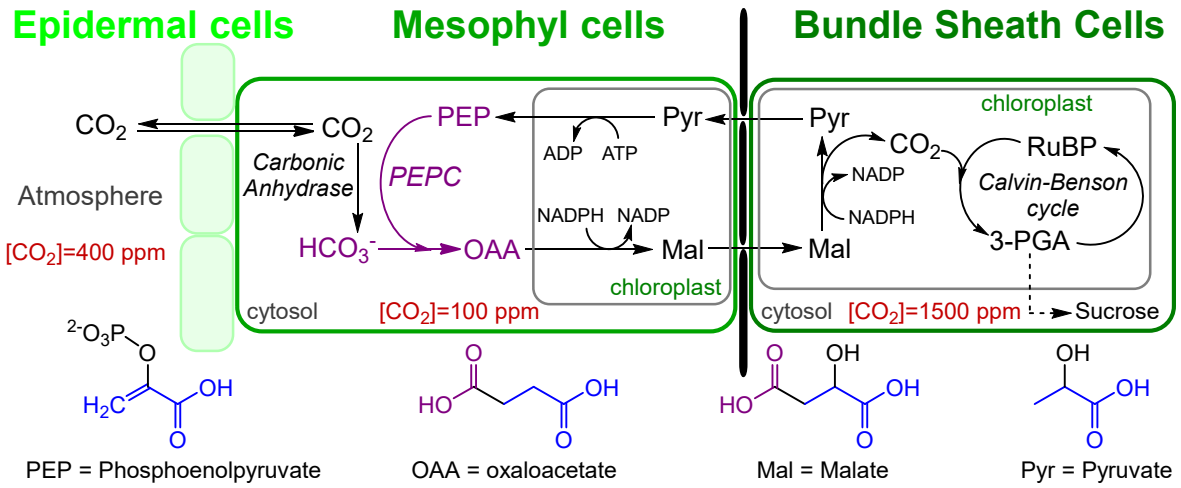


Figure 1. Simplified mechanism of carbon capture and fixation in C4 plants[2].

Despite its important role for Nature, the mechanism of carbon-capture by PEPC in C4 plants is still not fully elucidated: (1) global motions of the enzymes are still not understood, such as the movements of loop II (residues 761-768 in maize numbering) during the catalytic cycle. This loop is supposed to serve as a lid to protect the active site and to exist in two conformations[21], however it was found fully complete in only four of the 17 published crystallographic structures of PEPC, and always in the same conformation. Thus, no one knows its behaviour along the cycle. (2) The understanding of regulation in PEPC is still not complete, for example the role of phosphorylation at a molecular level and its consequences on the enzyme's structure[21]. (3) The chemical mechanism of the enzyme is supposed to be a three-step mechanism[21–23], however extracting microscopic information specific to each step from the macroscopic measurements has proven to be difficult and the free energy profile of the reaction is still not known. Thus, the concerted or sequential characters of the different steps, as well as the identity of the rate-determining step, are unknown. Since experimental approaches have not yet been able to answer these questions, one could rely on computational techniques. However, crucial details on PEPC are still lacking and currently

preclude the use of simulations to answer these questions, such as the localisation of one of the two substrates. To date, crystallographic studies couldn't provide enough data to overcome this hurdle, and no structural NMR studies on PEPC were published. Moreover, available mutagenesis assays also don't provide enough data to firmly conclude on the localisation of bicarbonate. We propose to address and unlock this question which will then allow to computationally study PEPC and to get more insights on this important enzyme.

Some small featureless ligands, such as CO_2, are sometimes said to be "slippery" because they can't bind to an enzyme and don't form a Michaelis complex[24]. However, Michaelis constants can be measured for bicarbonate in PEPC, which is consistent with the fact that the ligand is negatively charged and that there are several positively charged residues in the active site of PEPC. Thus, we envisioned that it should be possible to identify an effective binding site for bicarbonate in PEPC and in the current article, which is the first of a series on PEPC, we seek to answer to the question: "Where does bicarbonate bind in PEP Carboxylase?". Since to date no experimental techniques were able to answer this question, we relied on computational tools and in particular on molecular dynamics (MD) simulations. Some computational tools were designed to identify binding sites, such as FTsite, DeepSite, FPocket or CASTp to name a few that are still available[25–28]. However, these tools couldn't be used here because bicarbonate is too small to provide significant results.

Presentation of PEPC

The PEP Carboxylase (PEPC, E.C. 4.1.1.31) has been known for decades and has already been reviewed[21,29–32]. Its structure was first resolved in 1999 for bacteria (*E. Coli*[33]) and in 2002 for plants (*Zea Mays*[23]). It is a large enzyme (870 to 1150 residues depending on the organism[21]) whose functional form is a tetramer. The contact surface areas between one monomer and its two

neighbors are 450 and 3000 Å^2, and the structure is thus better described as a dimer of dimers[21] (see structures in SI). In the current article, we focused on a dimer of the enzyme from *Zea Mays* to save computational time. In 2002, the structure 1JQN was resolved with Mn^{2+} as an analog to Mg^{2+} and the ligand 3,3-dichloro-2-phosphonomethyl-acrylic acid (DCDP) as an analog to PEP, which allowed to accurately locate the position of these ligands. In their vicinity, one can find three important loops: (i) the loop II (residues 761-768 in maize numbering) which is supposed to be a mobile loop that is closed during the catalytic cycle[21]; (ii) the loop I (residues 640-649) which can change of conformation upon aspartate binding to displace Arg647 away from the active site[21]; (iii) the loop 174-184 which can change of conformation upon glucose-6-phosphate (G6P) binding to place His177 within the active site[21]. Aspartate and G6P serve respectively as allosteric inhibitor and activator, and in addition neutral aminoacids such as glycine can also serve as allosteric activators. Thus, PEPC is an enzyme that is highly regulated, which is usually a sign of importance for the organism. Structure of the active site and of binding site of allosteric effectors are shown in SI. The proposed chemical mechanism and the phosphorylation mechanism are not discussed here and the interested reader is invited to consult reviews on PEPC[21,29–32].

Building the system

This article is the first computational study performed on full PEPC (with the exception of some docking studies performed previously[34–36]) and is the first of a series, we will thus describe in some details the building of the system. We provide in SI a table with an analysis of the 17 available crystallographic structures. We started with the structure 5VYJ[35] (from *Zea Mays*) resolved in 2018 at 3.3Å of resolution, since it is the first structure with the complete loop 761-768. However, loop 126-140 is missing in 5VYJ, we thus used the UCSF Chimera's interface to Modeller to build

missing parts of the enzyme[37,38]: one adjacent residue was allowed to move, five models were generated, and the loop modeling protocol was set to DOPE[39]. To decide which model to choose, we aligned the structure 6U2T with complete models of 5VYJ and picked the model the closest to the loop 126-140 from 6U2T (from visual inspection). Residues 932/934/935 are also missing in 5VYJ: they were modelled in a similar way, and the model with the best DOPE score was retained. We then aligned the structure 1JQN[23] (from *E. Coli*) with 5VYJ to obtain the position of the two analogs Mn^{2+} and DCDP, which were then manually modified to Mg^{2+} and PEP. From 5VYJ, acetate ions were removed and glycines (an allosteric activator) that are bound in their actual binding sites were kept (i.e. one glycine molecule was removed out of the five that were resolved). The build structure did not contain glucose-6-phosphate (G6P, an allosteric activator) since glycines were present, nor aspartate (an allosteric inhibitor). Finally, we kept chains A and B for further investigations.

All structures were solvated in rhombic dodecahedron boxes, with a distance of at least 8 Å between two replicas of the system. This corresponds to edges of a triclinic box of roughly 155, 155 and 110 Å, which are filled by ~72000 water molecules (for a total of ~250000 atoms). Under physiological conditions, the enzyme is surrounded by a media with concentrations of 0.1 mM of bicarbonate, 3 mM of PEP, 0.4 mM of free Mg^{2+}, and 20 mM of L-malate[40]. These concentrations correspond to respectively 0.1 molecule of bicarbonate in the box, 3.9 molecules of PEP, 0.5 magnesium(II) and 25.9 molecules of malate: we have thus included in the simulation box 0 free bicarbonate, 4 free PEP, 0 free magnesium(II) and 26 free malate. In addition to these molecules, we have neutralized the charge with sodium ions and added a concentration of 0.1 M of sodium chloride (leading to respectively ~250 sodium and ~150 chloride). Computational details are described at length in SI.

Identification of binding sites

To identify the binding site of bicarbonate, we started by simulating a monomer of PEPC surrounded with 200 molecules of bicarbonate ([HCO_3^-]~0.42 M, i.e. 4200 times higher than the physiological concentration): no significant zone of high density near the active site could be observed with this approach (see SI for detailed results). Thus, we relied on another strategy where bicarbonate was pushed inside the protein with steered MD simulations. Starting conformations were manually created by placing bicarbonate at 50 different localisations around PEPC: 35 with the axis between bicarbonate and PEP aligned with the loop II, and 15 with this axis aligned with the β-barrel of PEPC (see Figure 2 and Figure SI-6). For each starting position, we performed two steered MD simulations along different reaction coordinates: (i) the distance between C of bicarbonate and P from PEP, (ii) the distance between C of bicarbonate and C2 from PEP (which is the sp^2 carbon bound to the carboxyl and the phosphate). During these simulations, the distance between bicarbonate and PEP decreased at a rate of 1.0 Å/ns and was restrained with a force constant of k=10000 kJ/mol/nm^2. The initial distances ranged from 22 to 56 Å and the simulation length was chosen so that the final distance reached ~4 Å; thus, the steered MD simulations lasted from 15 to 58 ns. To get insights into the penetration process, free energy profiles along the reaction coordinate were computed and results are discussed in SI. At the end of these 100 simulations, this procedure allowed to obtain bicarbonate at different localisations in the active site, but bicarbonate was still not in actual binding sites.

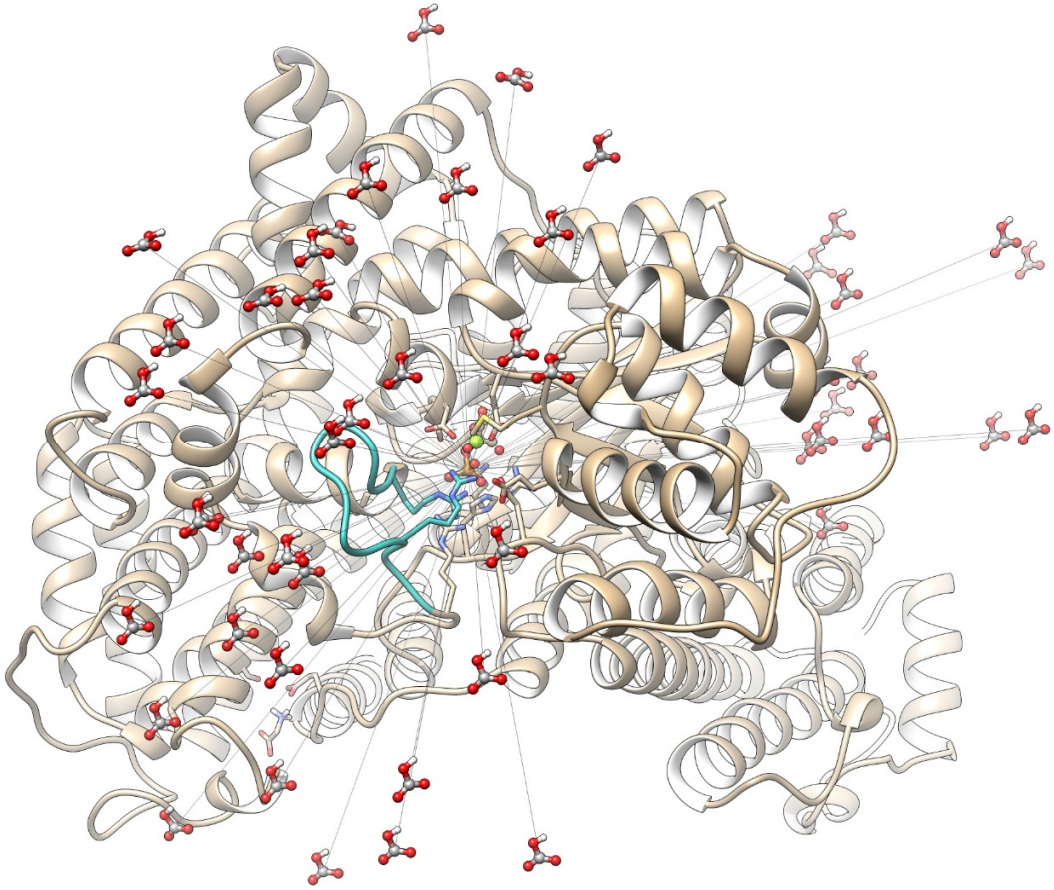


Figure 2. Super-imposition of starting positions of bicarbonate before steered MD simulations (in each simulation, only one bicarbonate was present). The lines joining C of bicarbonate and P of PEP are shown in grey. Loop II is in cyan. Additional images are available in SI (Figure SI-6).

Following the steered MD simulations, 100 ns-long free MD simulations without any constraint were performed to let bicarbonate sample the active site region: we observed that during these simulations the ligand ended up in 14 different pockets, which we labelled P1 to P14 and which are shown in Figure 3 as well as in Figure SI-10. A localisation in the active site is considered as a pocket when bicarbonate stayed there at least 50 ns (a more thorough discussion on the pockets is proposed in SI, where we describe how we identified and validated them). Eight of the free simulations were extended to 200 ns because a change of bicarbonate position occurred during the

second half of the simulation (after 50 ns), thus bicarbonate had not reached a pocket according to our definition. Two pockets were populated each in a single simulation, whereas one of the pockets (P1) was populated 31 times (see Table 1). During 10 simulations, bicarbonate unbound from the protein and diffused in the solvent. We observed that in some of these 10 simulations, bicarbonate was first in pocket P1 before unbinding: this provided contradictory results, since on one hand pocket P1 is the most populated pocket which seemed to indicate it is the most favourable pocket with the lowest binding free energy, and on the other hand bicarbonate could escape from pocket P1 meaning that the interactions were not strong enough to retain bicarbonate. During these free MD simulations, we observed barely no exchange between pockets which means that thermodynamic equilibrium had not been reached and that the absolute binding free energy in each pocket had to be computed to compare and rank all of them. We first tried to extract the binding free energy from a physical path by pulling bicarbonate outside the active site and computing the free energy profile along that path with umbrella sampling simulations. This approach was not conclusive (see details in SI), and we have thus decided to perform alchemical transformations coupled with free energy perturbations to accurately compute the absolute binding free energies in each pocket.

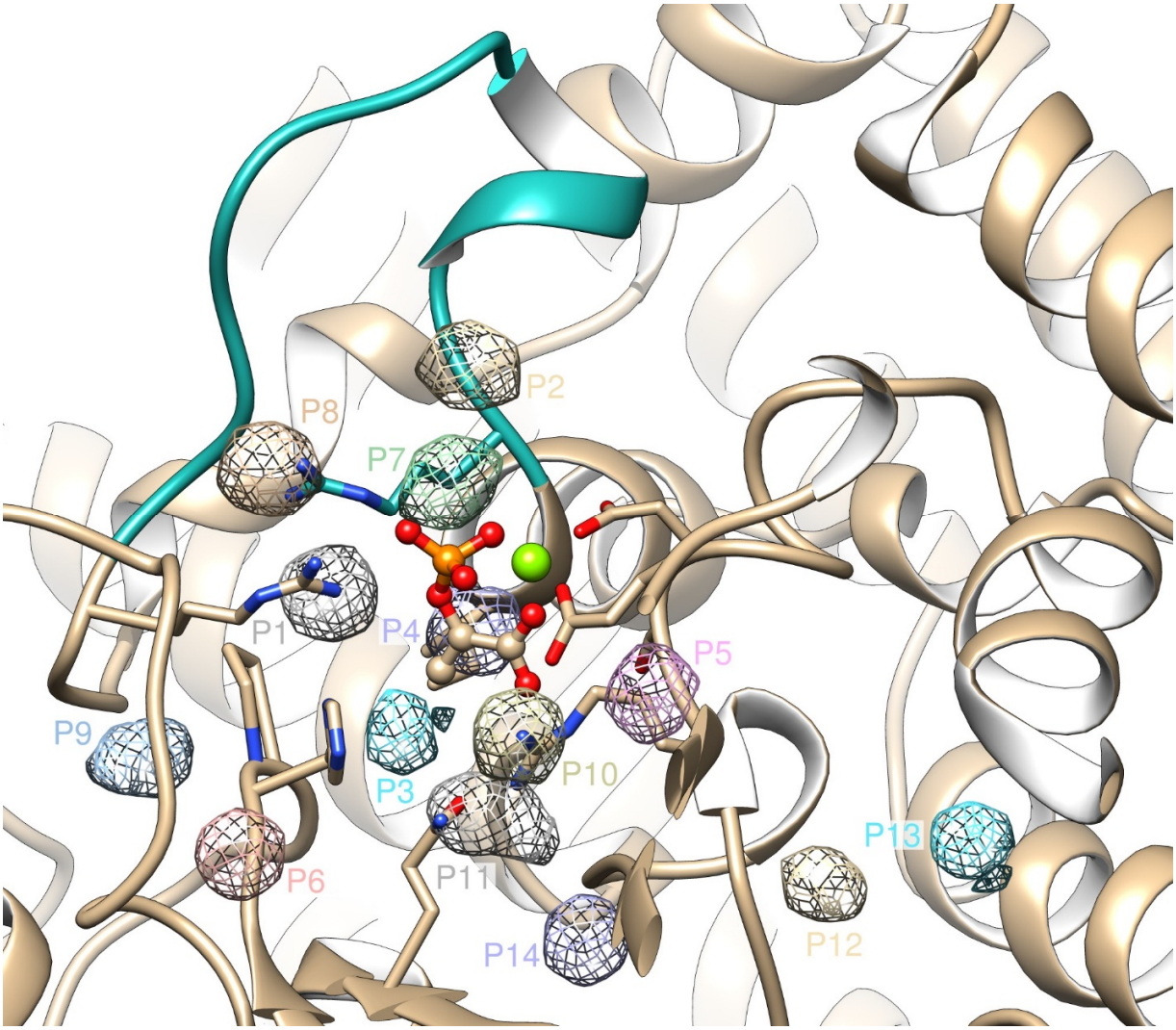


Figure 3. Identified binding sites of bicarbonate. The densities of bicarbonate in simulations representative of each pocket were analysed with GROMaps[41] and are here represented as isosurfaces with all the same threshold. PEP is in ball and stick, Mg^{2+} is in green, loop II is in cyan. Additional images are available in SI (Figure SI-10).

Alchemical transformations with free energy perturbations

Since the free energy is a state function, the followed path is not relevant to compute the difference of free energy between two states. Thus, one can change the chemical nature of compounds and

transfer them through a non-chemical way from bulk solvent to the binding site of an enzyme to compute the binding free energy: this kind of approach is called *alchemical transformations*. It is nowadays widely used, especially for drug design, and it has been numerously reviewed[42–48]. Using this approach, we have decomposed the binding free energy with a thermodynamic cycle where four terms must be computed independently: (1) the desolvation of the ligand, (2) the cost of adding restraints, (3) the binding free energy with restraints, (4) the gain in removing the restraints. We will briefly discuss these terms here; they are discussed at length with more computational details in SI. The cost of adding restraints can be obtained with an analytical function since it is only linked to the *ratio* of occupied volume with and without the restraints. The gain in removing the restraints is usually null for good ligands in good pockets and is often omitted. However, since we are here investigating pockets that can be not favourable, the gain in removing the restraints was also computed explicitly by progressively changing the amount of restraints. The desolvation of the ligand and the binding free energy with restraints were obtained through alchemical transformations of the ligand with free energy perturbations (FEP): in this kind of calculations, a series of simulations in which a coupling parameter λ varies are performed (for example, λ goes from 0 to 1 by steps of 0.1). If the variations of λ are small enough, one can compute the difference of free energies between two neighbour states, and by cumulating the differences one can get the difference of free energies between the two end states. This was used to make the ligand disappear from bulk solvent (desolvation free energy) and then appear in the enzyme (binding free energy). We have here used 21 alchemical windows to switch off the intermolecular interactions between bicarbonate and its environment (solvent or PEPC). To actually compute the differences of free energies, two approaches can be used efficiently: Thermodynamic Integration (TI) or the Bennett Acceptance Ratio (BAR). We have compared the

two, and for the later we have used the more modern multistate BAR (MBAR)[49]. The python packages alchemlyb and pymbar[49] were used to process the simulations and perform the calculations.

Assessing the convergence of FEP calculations is crucial to provide significant results, and is again discussed at length in SI. The convergence analysis was performed with MBAR, before comparing the results with those from TI. For each set of simulations (i.e. for each starting point used for the alchemical transformations), we have first performed a forward/backward analysis to assess the needed length of the simulations to reach convergence. The goal was that the first half and the second half of the used data provide comparable results, meaning that this set of simulations is converged with that amount of data. The data that were used for this analysis were extracted from the end of the simulations with an increasing size of the time bloc (i.e. last 10 ns, last 11 ns, last 12 ns, and so on), and we aimed at the longest time bloc that provides a difference between the two halves below 1 k_BT[47]. If the objectives could not be reached, the simulations were extended by 5 ns. At the end, all simulations for the binding free energy with restraints lasted between 20 and 40 ns, used at least 10 ns of data, and the computed errors from MBAR were on average 0.03 kcal/mol, ranging from 0.02 to 0.04 kcal/mol. We have then verified that the phase space overlaps between each neighbouring windows were high enough, and found that they are 23% on average on the first off-diagonal matrix, with a lowest value of 7% (whereas the recommended lower acceptable value is 3%[47]).

When alchemical transformations are performed on charged species, they must be corrected from computational artefacts (called finite-size effects) prior comparisons with experimental data. Indeed, since the total charge of the system changes between the two end states, the background potential in the simulation boxes changes as well, which influences the potential energies, hence

the computed free energies. Different approaches exist to correct these artefacts[50,51], and we relied on the one proposed by Rocklin *et al.*[50] where four effects are corrected: (1) the electrostatic interactions between the reference box, the periodic replicas and the neutralizing background, (2) the undersolvation due to the solutes in the periodic replicas, (3) the solvent-excluded volume of the solutes and the fact they are not point charges, (4) the difference of reference for the electrostatic potential between simulations and bulk experiments. The corrections must be performed for both the desolvation of the ligand and the binding process, and they are applied *a posteriori* through equations that are recalled in SI. One set of calculations per pocket was performed, on 50 different snapshots. Standard deviations between snapshots of a pocket are always low (below 0.07 kcal/mol, see Table SI-10), which illustrates that a single snapshot could have been used. Moreover, we observed that the correction values between the 14 pockets are all comparable (range of 0.2 kcal/mol).

To compare the binding free energies in each pocket, several independent replicas were used for some pockets. In such cases, the starting points for the replicas were extracted every 10 ns from the end of the free MD simulations (e.g. at 100 ns, 90 ns, 80 ns, and so on). Since FEP calculations are intensive in terms of computational resources, we have followed the following procedure: (i) we first performed the full calculations with three replicas for the first six identified pockets (P1 to P6). The lowest binding free energy was $\Delta G = -6.7$ kcal/mol with standard deviations between 0.5 and 2.7 kcal/mol. We thus envisioned that meaningful pockets to analyse should have a binding free energy lower than –4.0 kcal/mol (= –6.7+2.7). (ii) We then performed calculations for one replica in each remaining pocket. In these eight pockets, two had a binding free energy lower than –4.0 kcal/mol in the first calculations (P7 and P13), and we thus performed four additional replicas for them. (iii) Two additional set of calculations were performed for pocket P5 to lower the

standard error. The full set of data for each replica with the MBAR and the TI analysis are presented in Tables SI-4 and SI-5.

From the calculations of binding free energies of different replicas for each pocket, we computed the average binding free energy in each pocket. In pocket P7, one data seemed to be an outlier (see Table SI-4), and we have thus performed a Dixon's Q-test statistical analysis to unambiguously identify putative outliers in the data. In such a test, the value Q=gap/range is defined as the ratio between the gap between one value and its closest one, and the range of the values in the series. Q is computed for the highest and the lowest values of a series, and if Q is higher than a tabulated number that depends on the number of data points, then the corresponding value can be considered as an outlier. For example, if we set the confidence at 99% and with five values in the series, if Q > 0.821 then the value must be discarded. From all pockets, only the first value from pocket P7 was found to be an outlier (Q=0.845 with MBAR and Q=0.881 with TI), and the molecular explanation of it is described in SI. Thus, for pocket P7, we have used four data.

Overall the data from MBAR and TI were found to be highly comparable (see Figure SI-21). The average binding free energies from the two analysis methods are reported in Table 1 with the standard error, the number of replicas that were used, and the probability to find bicarbonate in each pocket. We also report in Table 1 the number of free MD simulations where bicarbonate ended in each pocket. Three pockets out of the 14 were found to be meaningful: P5, P7 and P13. The computed binding free energies are -5.4 ± 1.1, -5.7 ± 0.1 and -6.1 ±0.9 kcal/mol with MBAR data, and -5.2 ± 0.9, -5.4 ± 0.1, -6.1 ± 0.9 kcal/mol with TI data. Experimentally, only the Michaelis constant K_M of bicarbonate is available with values of 0.10 mM[52] and 0.12 mM[53] reported. These values can be approximated to the dissociation constant K_d and converted to binding free energies of -5.5 and -5.4 kcal/mol (see SI). Thus, the computed binding free energies of pockets P5 and P7

are within 0.3 kcal/mol to the experimental value with either MBAR or TI, and are within 0.7 kcal/mol for P13. With data from MBAR, the pockets are found to be populated at respectively 15, 24 and 54%, whereas with data from TI they are populated at 15, 19 and 59%. Pocket P3 is populated at 6% with both methods and all the other ones are populated at most 0.4%. From these computational data, we thus conclude that pockets P5, P7 and P13 should be considered as putative binding sites for bicarbonate, with probabilities that are comparable. P13 has the lowest binding free energies, but P5 and P7 have values very close to the experimental one.

Pocket	Free MD	MBAR				TI			
		$\langle \Delta G \rangle$	$\sigma/\sqrt{N}$	N	%	$\langle \Delta G \rangle$	$\sigma/\sqrt{N}$	N	%
P1	31	0.6	0.6	3	0.0	0.7	0.6	3	0.0
P2	4	-1.7	0.6	3	0.0	-1.5	0.5	3	0.0
P3	3	-4.8	0.8	3	5.8	-4.7	0.8	3	6.5
P4	18	-1.2	1.6	3	0.0	-0.9	1.5	3	0.0
P5	**2**	**-5.4**	**1.1**	**5**	**15.2**	**-5.2**	**0.9**	**5**	**14.8**
P6	3	2.5	0.3	3	0.0	2.7	0.3	3	0.0
P7	**11**	**-5.7**	**0.1**	**4**	**24.0**	**-5.4**	**0.1**	**4**	**19.4**
P8	3	-3.2	–	1	0.4	-3.1	–	1	0.4
P9	4	-2.5	–	1	0.1	-2.3	–	1	0.1
P10	2	4.7	–	1	0.0	4.9	–	1	0.0
P11	4	1.2	–	1	0.0	1.0	–	1	0.0
P12	1	0.8	–	1	0.0	1.1	–	1	0.0
P13	**3**	**-6.1**	**0.9**	**5**	**54.3**	**-6.1**	**0.9**	**5**	**58.7**
P14	1	-1.7	–	1	0.0	-1.9	–	1	0.1

Table 1. Number of free MD simulations where bicarbonate ended in each pocket and average binding free energies in each pocket with MBAR and TI data. N is the number of replicas used to compute the average binding free energy after Dixon Q-test. The column "%" gives the probability to find bicarbonate in each pocket.

As can be seen in Figure 3 and Figure SI-10, pocket P13 is buried inside the protein with a distance between magnesium and bicarbonate of 14.4 Å. Interactions between the negatively charged bicarbonate and the positively charged Arg610 and Arg641 explain the low binding free energy for this pocket (see Figure 4). Bicarbonate may serve as an anchor between these two residues to

restrict the relative motion between α-helix 26 (made of residues 610-629) and loop I (made of residues 640-649) and activate the enzyme, without being consumed during the catalytic cycle. A similar behavior occurs with Rubisco, where a lysine must first be activated to a carbamate through a reaction with a CO_2 molecule, and then the catalytic cycle occurs with reaction with the CO_2 molecules as substrate. The hypothesis of bicarbonate serving as an allosteric activator of PEPC could be tested by mutating the positions 610 and/or 641 and measuring the Michaelis constants. However, even if bicarbonate can bind in pocket P13, in the following we will not investigate it further because (i) either our hypothesis is true and bicarbonate is itself an allosteric activator in pocket P13: this means that the reactive bicarbonate substrate will bind to P5 or P7 and we can focus on these pockets, (ii) or our hypothesis is wrong and bicarbonate could bind to pocket P13 before a reactive event since it has a lower binding free energy: in such a case, prior the chemical reaction, bicarbonate will first transfer to pockets P5 or P7 and we can focus on these pockets.

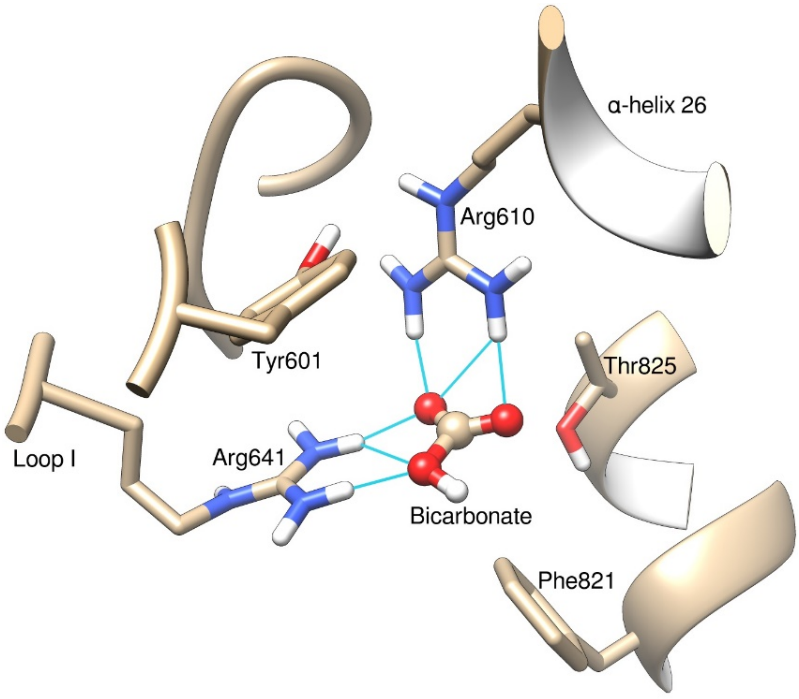


Figure 4. Zoom on the position of bicarbonate in pocket P13.

Pockets P5 and P7 are close to PEP and magnesium, with distances between magnesium and bicarbonate of 2.8 and 3.1 Å respectively (and 3.8 Å between the two pockets). Analysis of the free MD simulations revealed that the orientation of bicarbonate in these two pockets did not change along the trajectories and that the ligand kept its interactions with the protein in each simulation. A cluster analysis was performed on the last 50 ns of the free MD simulations and representative structures are displayed Figure 5. We observed that the orientation of PEP within the active site is clearly different in the two pockets, the two planes formed by PEP in each pocket forming an angle of ~60°. For each pocket, two of the free MD simulations where bicarbonate ended in these pockets were extended by 1 µs, and we didn't observe any exchange between the pockets during the four simulations of 1 µs. Moreover, the previous observations related to the stability of the interactions during the simulations were confirmed.

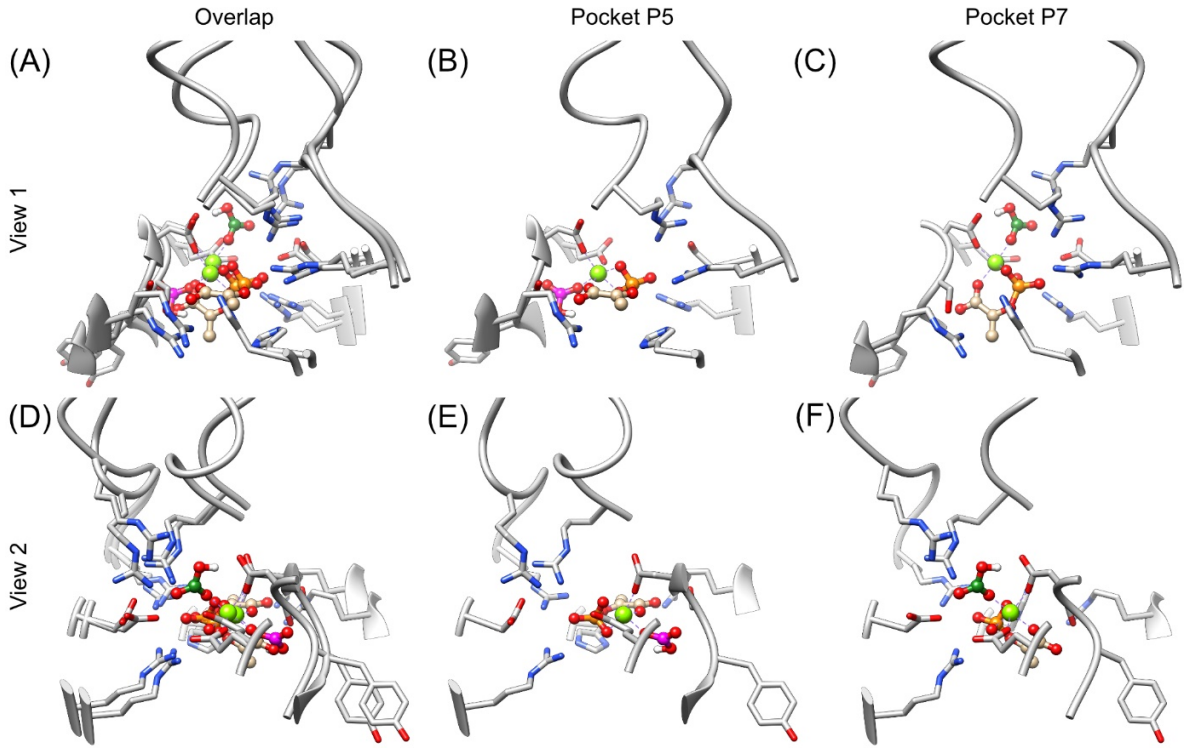


Figure 5. Overlap and zoom on the pockets P5 and P7 after alignment of the proteins, and seen from two angles. Ligands are in spheres, carbons from PEP are in light brown, carbon from bicarbonate in pocket P5 is in pink, carbon from bicarbonate in pocket P7 is in green. The protein is in grey.

In each snapshot presented Figure 5, the magnesium ion can be considered in an octahedral environment, with three of the coordination sites occupied by PEP atoms (two oxygen from the phosphate moiety and one from the carboxylate). In Pocket P5, an oxygen atom from bicarbonate is in *anti* position from an oxygen of the phosphate from PEP, whereas in pocket P7 the same oxygen atom from bicarbonate is in *anti* position from the carboxylate of PEP. The two remaining sites of coordination for magnesium are occupied by oxygen atoms from Glu566 and Asp603. Thus, in pocket P7, the bicarbonate is close to the phosphorus atom of PEP which can be beneficial for the reactivity according to the consensus mechanism which involves the formation of a carboxyphosphate intermediate[21,29,31,32]. However, since this mechanism has not been firmly demonstrated, this information cannot be used to decipher the preference of bicarbonate between pockets P5 and P7, and we cannot conclude on the most probable site of binding between these two pockets from these data.

To compare pockets P5 and P7, we turned to mutants of PEPC. Upon mutations of residues Lys762, Arg763 and Arg764 in *E. Coli* and maize (which are located on loop II), a side-reaction involving water molecules from solvent occurs, which suggest that loop II cannot protect anymore the active site[21]. Thus, the conformation of loop II is impacted by mutations, and we didn't use mutations made on this loop. Instead, we used data from Dong *et al.* who mutated Lys606[52]; the strongest effect was observed when this residue was mutated to Asn with $K_M(HCO_3^-)$ going from 0.10mM to 2.5mM, i.e. binding free energy increasing from -5.5 kcal/mol to -3.6 kcal/mol (ΔΔG

= +1.9 kcal/mol). Starting from representative snapshots of pockets P5 and P7, the mutation was performed with UCSF Chimera[37], the structures were build again (solvation, addition of ions, etc…) and were equilibrated during 500 ns. The same procedure as already described for FEP calculations was then performed (see SI for details). Six snapshots for each pocket were extracted at 500, 490, 480, 400, 390 and 380 ns, and the average between the first three and the last three were compared to ensure that comparable values are obtained, meaning that the mutants' structures are well equilibrated. With data from MBAR, we have calculated binding free energies of -4.4 ± 0.7 kcal/mol for pocket P5 and -6.5 ± 0.7 kcal/mol for pocket P7, i.e. ΔΔG of +1.0 and -0.8 kcal/mol respectively. With data from TI, we have calculated binding free energies of -4.2 ± 0.7 kcal/mol for pocket P5 and -6.4 ± 0.7 kcal/mol for pocket P7, i.e. ΔΔG of +1.0 and -1.0 kcal/mol respectively. Thus, only data for pocket P5 are compatible with the experimental data which found a positive ΔΔG of +1.9 kcal/mol, and we can conclude that P5 is the best candidate for the binding site of bicarbonate in PEP Carboxylase. Structures of the system build from 5VYJ with the bicarbonate in pockets P5 or P7 are provided in SI as PDB files.

In the current article, we sought to find the binding site of bicarbonate in PEP Carboxylase in order to study in the future questions such as loop motions, allosteric regulation or chemical reactivity. Using non-equilibrium MD simulations (steered-MD), we have identified 14 putative binding sites. The binding free energy of bicarbonate in each site was then accurately computed with alchemical free energy calculations, and we found that three sites are meaningful with computed binding free energies within 1 kcal/mol to experimental values. One site was found to be a candidate for allosteric activation by restricting the motion between an α-helix and a loop. To discriminate the two remaining pockets, we studied the mutant K606N and found that only the pocket P5 is compatible with experimental data. However, since bicarbonate is a small ligand with

not a lot of binding features, we could also envision that it can diffuse and fluctuate between different positions within the active site and we cannot firmly rule out that bicarbonate could bind in both pockets P5 and P7; only experiments will provide a definitive answer. Our next studies on PEPC will focus on the prediction of the closed conformation of loop II, the allosteric regulation and the phosphorylation of PEPC. We end this article by pointing out that, as already mentioned, loop II is supposed to be in a closed conformation during the catalytic cycle (acting as a lid) since upon mutations on this loop the activity of a side reaction increased by tenfold[21], and k_{cat} and $K_M(HCO_3^-)$ respectively decreased and increased. Thus, the experimental binding free energies that we have used here as a comparison were obtained with the loop probably in a closed conformation (since they were derived from K_M) whereas our simulations were performed with the open conformation. Once the closed conformation of loop II will be identified, we will have to replicate the calculations of the binding free energy of bicarbonate in some pockets to see if a difference arises.

ASSOCIATED CONTENT

Supporting Information. The following files are available free of charge: presentation of PEP Carboxylase, computational details, details on simulations with a high concentration of bicarbonate, discussions on (i) pushing the bicarbonate inside the protein, (ii) the sampling of the active site, (iii) pulling bicarbonate outside the protein, (iv) free energy perturbations with details on each term of the thermodynamic cycle, the convergence, the outlier, and the corrections due to the charged ligands (PDF). Structure of the enzyme with bicarbonate in pockets P5 and P7 are provided as PDB files. Scripts to perform the corrections of computational artefacts are provided as bash and python scripts.

AUTHOR INFORMATION

Nicolas Chéron (corresponding author): ORCID: 0000-0002-4090-5897

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

The author thanks Julien Henri for fruitful discussions on Calvin-Benson enzymes (Sorbonne Université, Paris), Elise Duboué-Dijon and Jérôme Hénin (IBPC, Paris) for fruitful discussions regarding constraints in free energy calculations and analysis, Daniel Borgis (ENS-PSL, Paris) for fruitful discussions regarding electrostatics, and Damien Laage (ENS-PSL, Paris) for various fruitful scientific discussions. This work was granted access to the HPC resources of CINES and IDRIS under the allocation 2020-077156 made by GENCI.

REFERENCES

(1) Bassham, J. A.; Benson, A. A.; Calvin, Melvin. The Path of Carbon in Photosynthesis. *J. Biol. Chem.* **1950**, *185* (2), 781–787. https://doi.org/10.1016/S0021-9258(18)56368-7.
(2) Sage, R. F.; Sage, T. L.; Kocacinar, F. Photorespiration and the Evolution of C4 Photosynthesis. *Annu. Rev. Plant Biol.* **2012**, *63* (1), 19–47. https://doi.org/10.1146/annurev-arplant-042811-105511.
(3) Endo, T.; Mihara, Y.; Furumoto, T.; Matsumura, H.; Kai, Y.; Izui, K. Maize C4-Form Phosphoenolpyruvate Carboxylase Engineered to Be Functional in C3 Plants: Mutations for Diminished Sensitivity to Feedback Inhibitors and for Increased Substrate Affinity. *J. Exp. Bot.* **2008**, *59* (7), 1811–1818. https://doi.org/10.1093/jxb/ern018.
(4) Kellogg, E. A. C4 Photosynthesis. *Curr. Biol.* **2013**, *23* (14), R594–R599. https://doi.org/10.1016/j.cub.2013.04.066.
(5) Ranum, P.; Peña-Rosas, J. P.; Garcia-Casal, M. N. Global Maize Production, Utilization, and Consumption. *Ann. N. Y. Acad. Sci.* **2014**, *1312* (1), 105–112. https://doi.org/10.1111/nyas.12396.
(6) Schmer, M. R.; Vogel, K. P.; Mitchell, R. B.; Perrin, R. K. Net Energy of Cellulosic Ethanol from Switchgrass. *Proc. Natl. Acad. Sci.* **2008**, *105* (2), 464–469. https://doi.org/10.1073/pnas.0704767105.
(7) Bloom, D. E. 7 Billion and Counting. *Science* **2011**, *333* (6042), 562–569. https://doi.org/10.1126/science.1209290.

(8) Raines, C. A.; Harrison, E. P.; Ölçer, H.; Lloyd, J. C. Investigating the Role of the Thiol-Regulated Enzyme Sedoheptulose-1,7-Bisphosphatase in the Control of Photosynthesis. *Physiol. Plant.* **2000**, *110* (3), 303–308. https://doi.org/10.1111/j.1399-3054.2000.1100303.x.
(9) Furbank, R. T.; Hatch, M. D.; Jenkins, C. L. D. C4 Photosynthesis: Mechanism and Regulation. In *Photosynthesis: Physiology and Metabolism*; 2000; pp 435–457.
(10) Lefebvre, S.; Lawson, T.; Fryer, M.; Zakhleniuk, O. V.; Lloyd, J. C.; Raines, C. A. Increased Sedoheptulose-1,7-Bisphosphatase Activity in Transgenic Tobacco Plants Stimulates Photosynthesis and Growth from an Early Stage in Development. *Plant Physiol.* **2005**, *138* (1), 451–460. https://doi.org/10.1104/pp.104.055046.
(11) Simkin, A. J.; Lopez-Calcagno, P. E.; Davey, P. A.; Headland, L. R.; Lawson, T.; Timm, S.; Bauwe, H.; Raines, C. A. Simultaneous Stimulation of Sedoheptulose 1,7-Bisphosphatase, Fructose 1,6-Bisphophate Aldolase and the Photorespiratory Glycine Decarboxylase-H Protein Increases CO2 Assimilation, Vegetative Biomass and Seed Yield in Arabidopsis. *Plant Biotechnol. J.* **2017**, *15* (7), 805–816. https://doi.org/10.1111/pbi.12676.
(12) Hammel, A.; Sommer, F.; Zimmer, D.; Stitt, M.; Mühlhaus, T.; Schroda, M. Overexpression of Sedoheptulose-1,7-Bisphosphatase Enhances Photosynthesis in Chlamydomonas Reinhardtii and Has No Effect on the Abundance of Other Calvin-Benson Cycle Enzymes. *Front. Plant Sci.* **2020**, *11*.
(13) Qu, Y.; Sakoda, K.; Fukayama, H.; Kondo, E.; Suzuki, Y.; Makino, A.; Terashima, I.; Yamori, W. Overexpression of Both Rubisco and Rubisco Activase Rescues Rice Photosynthesis and Biomass under Heat Stress. *Plant Cell Environ.* **2021**, *44* (7), 2308–2320. https://doi.org/10.1111/pce.14051.
(14) Meloni, M.; Fanti, S.; Tedesco, D.; Gurrieri, L.; Trost, P.; Fermani, S.; Lemaire, S. D.; Zaffagnini, M.; Henri, J. Structural and Functional Characterization of Chloroplast Ribulose-5-Phosphate-3-Epimerase from the Model Green Microalga Chlamydomonas Reinhardtii. bioRxiv September 30, 2022, p 2022.09.29.510120. https://doi.org/10.1101/2022.09.29.510120.
(15) 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.* **2006**, *103* (19), 7246–7251. https://doi.org/10.1073/pnas.0600605103.
(16) Bathellier, C.; Tcherkez, G.; Lorimer, G. H.; Farquhar, G. D. Rubisco Is Not Really so Bad: Rubisco Efficiency. *Plant Cell Environ.* **2018**, *41* (4), 705–716. https://doi.org/10.1111/pce.13149.
(17) Ding, F.; Wang, M.; Zhang, S.; Ai, X. Changes in SBPase Activity Influence Photosynthetic Capacity, Growth, and Tolerance to Chilling Stress in Transgenic Tomato Plants. *Sci. Rep.* **2016**, *6* (1), 32741–32754. https://doi.org/10.1038/srep32741.
(18) Furbank, R. T.; von Caemmerer, S.; Sheehy, J.; Edwards, G. C4 Rice: A Challenge for Plant Phenomics. *Funct. Plant Biol.* **2009**, *36* (11), 845. https://doi.org/10.1071/FP09185.
(19) von Caemmerer, S.; Furbank, R. T. Modeling C4 Photosynthesis. In *C4 Plant Biology*; Elsevier, 1999; pp 173–211. https://doi.org/10.1016/B978-012614440-6/50007-0.
(20) O'Leary, B.; Park, J.; Plaxton, W. C. The Remarkable Diversity of Plant PEPC (Phosphoenolpyruvate Carboxylase): Recent Insights into the Physiological Functions and Post-Translational Controls of Non-Photosynthetic PEPCs. *Biochem. J.* **2011**, *436* (1), 15–34. https://doi.org/10.1042/BJ20110078.

(21) Izui, K.; Matsumura, H.; Furumoto, T.; Kai, Y. Phosphoenolpyruvate Carboxylase: A New Era of Structural Biology. *Annu. Rev. Plant Biol.* **2004**, *55* (1), 69–84. https://doi.org/10.1146/annurev.arplant.55.031903.141619.
(22) Janc, J. W.; Urbauer, J. L.; O'Leary, M. H.; Cleland, W. W. Mechanistic Studies of Phosphoenolpyruvate Carboxylase from Zea Mays with (Z)- and (E)-3-Fluorophosphoenolpyruvate as Substrates. *Biochemistry* **1992**, *31* (28), 6432–6440. https://doi.org/10.1021/bi00143a011.
(23) Matsumura, H.; Xie, Y.; Shirakata, S.; Inoue, T.; Yoshinaga, T.; Ueno, Y.; Izui, K.; Kai, Y. Crystal Structures of C4 Form Maize and Quaternary Complex of E. Coli Phosphoenolpyruvate Carboxylases. *Structure* **2002**, *10* (12), 1721–1730. https://doi.org/10.1016/S0969-2126(02)00913-9.
(24) Tcherkez, G. Modelling the Reaction Mechanism of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase and Consequences for Kinetic Parameters. *Plant Cell Environ.* **2013**, *36* (9), 1586–1596. https://doi.org/10.1111/pce.12066.
(25) Ngan, C.-H.; Hall, D. R.; Zerbe, B.; Grove, L. E.; Kozakov, D.; Vajda, S. FTSite: High Accuracy Detection of Ligand Binding Sites on Unbound Protein Structures. *Bioinformatics* **2012**, *28* (2), 286–287. https://doi.org/10.1093/bioinformatics/btr651.
(26) Jiménez, J.; Doerr, S.; Martínez-Rosell, G.; Rose, A. S.; De Fabritiis, G. DeepSite: Protein-Binding Site Predictor Using 3D-Convolutional Neural Networks. *Bioinformatics* **2017**, *33* (19), 3036–3042. https://doi.org/10.1093/bioinformatics/btx350.
(27) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10* (1), 168. https://doi.org/10.1186/1471-2105-10-168.
(28) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res.* **2018**, *46* (W1), W363–W367. https://doi.org/10.1093/nar/gky473.
(29) Chollet, R.; Vidal, J.; O'Leary, M. H. Phosphoenolpyruvate Carboxylase: A Ubiquitous, Highly Regulated Enzyme in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47* (1), 273–298. https://doi.org/10.1146/annurev.arplant.47.1.273.
(30) Vidal, J.; Chollet, R. Regulatory Phosphorylation of C4 PEP Carboxylase. *Trends Plant Sci.* **1997**, *2* (6), 230–237.
(31) Kai, Y.; Matsumura, H.; Izui, K. Phosphoenolpyruvate Carboxylase: Three-Dimensional Structure and Molecular Mechanisms. *Arch. Biochem. Biophys.* **2003**, *414* (2), 170–179. https://doi.org/10.1016/S0003-9861(03)00170-X.
(32) Lepiniec, L.; Thomas, M.; Vidal, J. From Enzyme Activity to Plant Biotechnology: 30 Years of Research on Phosphoenolpyruvate Carboxylase. *Plant Physiol. Biochem.* **2003**, *41* (6–7), 533–539. https://doi.org/10.1016/S0981-9428(03)00069-X.
(33) Kai, Y.; Matsumura, H.; Inoue, T.; Terada, K.; Nagara, Y.; Yoshinaga, T.; Kihara, A.; Tsumura, K.; Izui, K. Three-Dimensional Structure of Phosphoenolpyruvate Carboxylase: A Proposed Mechanism for Allosteric Inhibition. *Proc. Natl. Acad. Sci.* **1999**, *96* (3), 823–828. https://doi.org/10.1073/pnas.96.3.823.
(34) Mancera, R. L.; Carrington, B. J. The Molecular Binding Interactions of Inhibitors and Activators of Phosphoenolpyruvate Carboxylase. *J. Mol. Struct. THEOCHEM* **2005**, *755* (1–3), 151–159. https://doi.org/10.1016/j.theochem.2005.08.014.
(35) González-Segura, L.; Mújica-Jiménez, C.; Juárez-Díaz, J. A.; Güémez-Toro, R.; Martinez-Castilla, L. P.; Muñoz-Clares, R. A. Identification of the Allosteric Site for Neutral Amino

Acids in the Maize C4 Isozyme of Phosphoenolpyruvate Carboxylase: The Critical Role of Ser-100. *J. Biol. Chem.* **2018**, *293* (26), 9945–9957. https://doi.org/10.1074/jbc.RA118.002884.
(36) Muñoz-Clares, R. A.; González-Segura, L.; Juárez-Díaz, J. A.; Mújica-Jiménez, C. Structural and Biochemical Evidence of the Glucose 6-Phosphate-Allosteric Site of Maize C4-Phosphoenolpyruvate Carboxylase: Its Importance in the Overall Enzyme Kinetics. *Biochem. J.* **2020**, *477* (11), 2095–2114. https://doi.org/10.1042/BCJ20200304.
(37) 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* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.
(38) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *54* (1), 5.6.1-5.6.37. https://doi.org/10.1002/cpbi.3.
(39) Shen, M.; Sali, A. Statistical Potential for Assessment and Prediction of Protein Structures. *Protein Sci.* **2006**, *15* (11), 2507–2524. https://doi.org/10.1110/ps.062416606.
(40) Tovar-Méndez, A.; Mújica-Jiménez, C.; Muñoz-Clares, R. A. Physiological Implications of the Kinetics of Maize Leaf Phosphoenolpyruvate Carboxylase. *Plant Physiol.* **2000**, *123* (1), 149–160. https://doi.org/10.1104/pp.123.1.149.
(41) 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* (1), 4–11. https://doi.org/10.1016/j.bpj.2018.11.3126.
(42) Jorgensen, W. L.; Thomas, L. L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4* (6), 869–876. https://doi.org/10.1021/ct800011m.
(43) Mobley, D. L.; Klimovich, P. V. Perspective: Alchemical Free Energy Calculations for Drug Discovery. *J. Chem. Phys.* **2012**, *137* (23), 230901. https://doi.org/10.1063/1.4769292.
(44) Chipot, C. Frontiers in Free-Energy Calculations of Biological Systems. *WIREs Comput. Mol. Sci.* **2014**, *4* (1), 71–89. https://doi.org/10.1002/wcms.1157.
(45) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57* (12), 2911–2937. https://doi.org/10.1021/acs.jcim.7b00564.
(46) Mobley, D. L.; Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annu. Rev. Biophys.* **2017**, *46* (1), 531–558. https://doi.org/10.1146/annurev-biophys-070816-033654.
(47) Mey, A. S. J. S.; Allen, B. K.; Macdonald, H. E. B.; Chodera, J. D.; Hahn, D. F.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A.; Scheen, J.; Shirts, M. R.; Tresadern, G.; Xu, H. Best Practices for Alchemical Free Energy Calculations. *Living J. Comput. Mol. Sci.* **2020**, *2* (1), 18378. https://doi.org/10.33011/livecoms.2.1.18378.
(48) Song, L. F.; Merz, K. M. Jr. Evolution of Alchemical Free Energy Methods in Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60* (11), 5308–5318. https://doi.org/10.1021/acs.jcim.0c00547.
(49) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129* (12), 124105. https://doi.org/10.1063/1.2978177.
(50) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. Calculating the Binding Free Energies of Charged Species Based on Explicit-Solvent Simulations Employing Lattice-Sum

Methods: An Accurate Correction Scheme for Electrostatic Finite-Size Effects. *J. Chem. Phys.* **2013**, *139* (18), 184103. https://doi.org/10.1063/1.4826261.
(51) Lin, Y.-L.; Aleksandrov, A.; Simonson, T.; Roux, B. An Overview of Electrostatic Free Energy Computations for Solutions and Proteins. *J. Chem. Theory Comput.* **2014**, *10* (7), 2690–2709. https://doi.org/10.1021/ct500195p.
(52) Dong, L.-Y.; Ueno, Y.; Hata, S.; Izui, K. Effects of Site-Directed Mutagenesis of Conserved Lys606 Residue on Catalytic and Regulatory Functions of Maize C4-Form Phosphoenolpyruvate Carboxylase. *Plant Cell Physiol.* **1997**, *38* (12), 1340–1345. https://doi.org/10.1093/oxfordjournals.pcp.a029127.
(53) Muramatsu, M.; Suzuki, R.; Yamazaki, T.; Miyao, M. Comparison of Plant-Type Phospho Enol Pyruvate Carboxylases from Rice: Identification of Two Plant-Specific Regulatory Regions of the Allosteric Enzyme. *Plant Cell Physiol.* **2015**, *56* (3), 468–480. https://doi.org/10.1093/pcp/pcu189.