Structural and biochemical evidence of the glucose 6-phosphate-allosteric site of maize C₄-phosphoenolpyruvate carboxylase: Its importance in the overall enzyme kinetics

Rosario A. Muñoz-Clares¹*, Lilian González-Segura¹, Javier Andrés Juárez-Díaz², Carlos Mújica-Jiménez¹

¹Departamento de Bioquímica, Facultad de Química, ²Departamento de Biología Comparada, Facultad de Ciencias. Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México.

Running title: Glucose 6-phosphate-allosteric site of ZmPEPC-C₄

To whom correspondence should be addressed: Rosario A. Muñoz-Clares, Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México. Telephone number: +52 55 5622 3718; e-mail: clares@unam.mx
Abstract

Activation of phosphoenolpyruvate carboxylase (PEPC) enzymes by glucose 6-phosphate (G6P) and other phospho-sugars is of major physiological relevance. Previous kinetic, site-directed mutagenesis and crystallographic results are consistent with allosteric activation, but the existence of a G6P-allosteric site was questioned and competitive activation—in which G6P would bind to the active site eliciting the same positive homotropic effect as the substrate PEP—was proposed. Here we report the crystal structure of the PEPC-C4 isozyme from Zea mays with G6P well bound into the previously proposed allosteric site, unambiguously confirming its existence. To test its functionality, Asp239—which participates in a web of interactions of the protein with G6P—was changed to alanine. The D239A variant was not activated by G6P but, on the contrary, inhibited. Inhibition was also observed in the wild-type enzyme at concentrations of G6P higher than those producing activation, and probably arises from G6P binding to the active site in competition with PEP. The lower activity and cooperativity for the substrate PEP, lower activation by glycine and diminished response to malate of the D239A variant suggest that the heterotropic allosteric-activation effects of free-PEP are also abolished in this variant. Together, our findings are consistent with both the existence of the G6P-allosteric site and its essentiality for activation of PEPC enzymes by phosphorylated compounds. Furthermore, our findings suggest a central role of the G6P-allosteric site in the overall kinetics of these enzymes even in the absence of G6P or other phospho-sugars, because of its involvement in activation by free-PEP.

Keywords: phospho-sugars-allosteric activation; substrate-allosteric activation; protein crystal structure; site-directed mutagenesis; amino acid residues conservation
Introduction

Phosphoenolpyruvate carboxylases (PEPCs, EC 4.1.1.31) are mainly bacterial and plant enzymes that catalyze the formation of oxaloacetate and inorganic phosphate from phosphoenolpyruvate (PEP) and bicarbonate, in an essentially irreversible carboxylation reaction that requires Mg^{2+} ions. This reaction is the highly regulated first step in the assimilation pathway of atmospheric CO₂ in C₄ plants, such as maize. At physiological pH, the PEPC isozyme from leaves of these plants (PEPC-C₄) is allosterically activated by phospho-trioses or phospho-hexoses, such as D-glucose-6-phosphate (G6P), and inhibited by carboxylic acids such as L-malate or L-aspartate. Additionally, PEPC-C₄ isozymes from grasses are activated by glycine, serine or alanine. Activation by phospho-sugars constitutes a physiologically important kinetic property of PEPC enzymes, not only of those from plants, regardless of being of the C₃, C₄ or CAM types, but also from bacteria [reviewed in 1,2]. In addition to phospho-sugars, other phosphorylated compounds were also found to behave as activators of PEPC enzymes [3-5]. Among these, of particular kinetic and possibly physiological relevance is the free species, i.e. non-complexed with Mg^{2+} ions, of the substrate phosphoenolpyruvate (free-PEP) [6-8]. Thus, PEP itself behaves as both a substrate (as the complex PEP-Mg^{2+} and as free-PEP) and as an activator (free-PEP).

The allosteric nature of the activation of PEPC enzymes by G6P and other phosphorylated compounds has been widely accepted on the basis of a number of kinetic [4-12], crystallographic [13], docking [14] and site-directed mutagenesis studies [15,16], until Schlieper et al. [17] questioned the existence of a G6P-allosteric site proposing instead competitive activation, i.e. that G6P activation of PEPC arises from its binding to the active site where it simulates the positive homotropic allosteric effects of the substrate PEP, rather than from the positive heterotropic effects resulting from its binding to an allosteric site. Their proposal was supported by the crystal structure of the Flaveria trinervia C₄ isozyme (FtPEPC-C₄) (PDB code 4BXC) in which G6P is bound into the active site and not into the previously proposed G6P-allosteric site, where a sulfate anion from the
crystallization medium was found instead, leading them to propose this site as a sulfate-binding site of unknown physiological significance. Additionally, these authors reported a higher degree of activation by G6P at low PEP concentrations, which was interpreted as supporting evidence of the competitive activation suggested by them [17].

Knowing for certain the actual nature of the mechanism of PEPC activation by G6P and other phosphorylated compounds, i.e. allosteric activation or competitive activation, is of great importance because each of these two mechanisms would have different physiological effects. Indeed, it can be expected that competitive activation would be very inefficient and even could be detrimental for the plant under physiological conditions. Moreover, a model of competitive activation cannot explain the reported activation by free-PEP. Therefore, the nature of the activation by G6P and related phosphorylated compounds should be established without any doubt to fully understand the functioning of PEPC enzymes in vivo and to interpret the kinetic results obtained in vitro. Thus, with the aim of further investigating this important issue and solving the current controversy regarding the mechanism of activation of plant PEPC enzymes by G6P and other phosphorylated compounds, we obtained a crystal structure of ZmPEPC-C4 in which, for the first time, a G6P molecule was found to be bound in a crevice in the subunit/subunit interfaces of the dimeric units of the native tetramer, thus confirming the previously proposed location of the G6P-allosteric site. In addition, we tested the functionality of this allosteric-site by generating and kinetically characterizing the ZmPEPC-C4 D239A variant, which was not activated but inhibited by G6P. Our data also are consistent with the mutation causing the loss of activation by free-PEP, which has an important impact on the overall enzyme kinetics even in the absence of activators.

Materials and Methods

ZmPEPC-C4 expression, purification, site-directed mutagenesis, enzymatic assay and kinetic data analyses
BL21-CodonPlus(DE3)-RIL *E. coli* cells (Novagen) transformed with a pET32a(+) (Novagen) plasmid containing the full sequence of the gene coding for ZmPEPC-C₄ (accession number CAD60555) were used to overexpress the recombinant protein with an N-terminal His-tag, as reported [18]. The ZmPEPC-C₄ protein was purified using the Protino nickel-Tris(carboxymethyl)ethylene diamine resin (Macherey-Nagel, Germany) and then digested with enterokinase (New England Biolabs) following the procedure described in González-Segura *et al.* [18]. Protein concentration was determined spectrophotometrically by $A_{280}$, using the extinction coefficient of 111,730 M⁻¹ cm⁻¹ per subunit predicted on the basis of the amino acid sequence by ExPASy ProtParam [19].

To generate the ZmPEPC-C₄ D239A variant, the forward and reverse primers 5’ CCCCCGAGGcCGAAATGCCTATGGGATG and 5’ GCATTTCGgCCTCGGGGTGGGTTGTG 3’, respectively, were used. The above-mentioned plasmid was used as the template for the site-directed mutagenesis by PCR, using the QuickChange II site-directed mutagenesis kit (Agilent Technologies) and following manufacturer’s instructions. The DNA was sequenced to confirm that the desired mutation was present and that no other nucleotide changes occurred. Finally, the variant protein was overexpressed in BL21-CodonPlus(DE3)-RIL *E. coli* cells (Novagen). The overexpression and purification of both the recombinant wild-type and D239A ZmPEPC-C₄ proteins was achieved as reported [18].

ZmPEPC-C₄ activity was assayed spectrophotometrically at 30 °C in a coupled enzymatic assay with malate dehydrogenase by monitoring NADH oxidation at 340 nm ($ε_{340} = 6,220$ M⁻¹ cm⁻¹). The standard assay medium consisted of 100 mM HEPES-KOH buffer, pH 7.3, containing 0.2 mM NADH and five units of malate dehydrogenase. Unless otherwise stated, initial velocity assays were conducted at concentrations of substrates believed to be close to those prevailing *in vivo* under illumination conditions (0.1 mM bicarbonate, 0.4 mM free-Mg²⁺) as previously discussed by Tovar-Méndez *et al.* [20]. The amounts of total magnesium (as MgCl₂) and total PEP or total effector used...
to give the desired concentrations of the free and Mg$^{2+}$-complexed species were calculated as described in [7]. Enzyme activity in the presence of G6P, F6P or glycine was assayed at fixed 3 mM total-PEP; in the presence of malate at 2.5 mM total-PEP and 5 mM total-Mg$^{2+}$. Initial velocities are expressed in units (µmol of product formed per minute under our experimental conditions.). Each point shown in the figures is the average of duplicate or triplicate determinations. ORIGIN software (OriginLab Corporation) was used for kinetic data analysis by nonlinear regression and display. When activation was observed, saturation data were fitted to equation (1). When inhibition was observed, saturation data were fitted to equation (2). Substrate-saturation data were fitted to equation (3):

\[
v = v_0 \times I_{0.5}^h / (I_{0.5}^h + [I]^h) \quad (1)
\]

\[
v = v_0 + \{ (v_{a\ max} - v_0) \times [A]^h / (A_{0.5}^h + [A]^h) \} \quad (2)
\]

\[
v = V_{\ max} \times [S]^h / (S_{0.5}^h + [S]^h) \quad (3)
\]

where \(v\) is the experimentally determined initial velocity; \(v_0\) is the experimentally determined initial velocity in the absence of the allosteric effector; \(v_{a\ max}\) is the estimated maximum activity at saturating activator concentrations; \(V_{\ max}\) the estimated maximum velocity; \([S]\), \([A]\), and \([I]\) are the substrate, activator and inhibitor concentrations, respectively; \(A_{0.5}\) is the concentration of activator that gives half-maximum activation at fixed concentrations of substrate; \(I_{0.5}\) is the concentration of inhibitor that gives half-maximum inhibition at fixed concentrations of substrate; \(K_m\) is the concentration of substrate that gives half-maximum velocity; and \(h\) is the Hill number indicative of the degree of cooperativity in the binding of the ligand.

**Crystallization and X-ray data collection**

ZmPEPC-C$_4$ crystals were grown by hanging-drop vapor diffusion by mixing equal volumes of enzyme preparation and reservoir solution consisting of 0.17 M sodium acetate trihydrate, 0.085 M TRIS hydrochloride pH 8.5, 15% (w/v) polyethylene glycol (PEG) 4000, 0.1 M potassium/sodium
tartrate-4-hydrate, at 18 °C. Droplets of 2 μL of enzyme plus 2 μL of reservoir solution were prepared. The protein concentration was 10 mg ml⁻¹ in 0.010 M HEPES buffer, pH 7.5, 1 mM DTT and 10 mM MgCl₂. Prior to crystallization, 100 mM G6P, 100 mM glycine and 100 mM PGA were added to the enzyme preparation. The crystals were cryoprotected in reservoir solution plus 20% (v/v) glycerol and flash-cooled in N₂. X-ray diffraction data were collected at 100 K at the beamline 19-ID of the Advanced Photon Source of the Structural Biology Center at Argonne National Laboratory, Lemont, IL, USA.

**Structure determination and model refinement**

The data were integrated using XDS [21] and scaled and truncated with programs from the CCP4 suite [22]. Diffraction phases were determined by molecular replacement with the program Phaser [23] using the coordinates of the ZmPEPC-C₄ PDB code 5VYJ as starting model. Alternating cycles of automatic and manual refinement were carried out with the standard protocols of Phenix [24], monitoring the $R_{\text{work}}$ and $R_{\text{free}}$ split during the whole process. Phenix was also used for atomic positions and atomic displacement parameters. The simulated annealing omit map was calculated with Phenix following simulated annealing refinement at 3,000 K. The program Coot [25] was used to analyze the electron density maps ($2F_o-F_c$ and $F_o-F_c$). Structural figures and alignments were performed with Coot and PyMOL. The tetramer was generated through crystallographic symmetry operations in PyMOL.

**Docking simulations**

After removing the G6P bound into the G6P-allosteric site, flexible molecules of F6P, DHAP, or PEP were docked in the rigid ZmPEPC-C₄ 6MGI crystal structure. Docking simulations were performed using AutoDock Vina [26] and a box of 15 Å centered on Glu369. In each simulation experiment, the top then poses were visually inspected to select the one that fulfilled the criteria of a proper position.
of the phosphate group. The best solution was subjected to a 1000-step energy-minimization process using the Amber force field parameters in the UCSF Chimera program [27].

**Sequence analysis of PEPC enzymes**

We searched PEPC orthologs from plants and bacteria, performing a blastp search on the RefSeq collection of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) protein database as described [18]. The amino acid sequence of ZmPEPC-C$_4$ (GenBank accession CAD60555) was used as a query for PEPC-C$_4$; that of ZmPEPC-C$_4$ for PEPC-C$_3$ (GenBank accession PWZ12751), and that of PEPC from *Escherichia coli* (EcPEPC) (GenBank accession CAA29332) for bacterial sequences. All retrieved plant sequences belong to Viridiplantae. Plant PEPC-C$_4$ amino acid sequences were chosen based on the criteria of having serine at position equivalent to ZmPEPC-C$_4$ 780 [28] and belonging to a C$_4$ plant. All sequences that in addition to fulfill these criteria included the residues analyzed were considered in our study. Representative PEPC-C$_3$ sequences were retrieved from C$_3$ plants belonging to 32 different plant families and selected with the criterion of having an alanine at position equivalent to ZmPEPC-C$_4$ 780. Also, care was taken that their sequence identity varied between 83 and 74%. Only two complete PEPC sequences were retrieved from two CAM plants. Since it was not clear to us whether they were of the C$_3$-type or CAM-type, they were included in the analysis with the PEPC-C$_3$ sequences, although one of them has serine at position 780 and therefore it probably corresponds to a PEPC-CAM type isozyme. Given the great number of bacterial genomes, the analysis of bacterial PEPC sequences was performed using only sequences from ten representative organisms of each of the phyla Actinobacteria, Cyanobacteria, Firmicutes and Proteobacteria. To ensure the diversity of the sample analyzed, the degree of identity of the selected sequences with that of EcPEPC ranged from 79 to 31%. The conservation of residues of the G6P-allosteric site was investigated by multiple amino acid residues alignments using ClustalX2 [29] and
the alignments were refined by hand. Sequence logos were constructed using WebLogo 3 (http://weblogo.threeplusone.com) [30,31].

Results and Discussion

The ZmPEPC-C₄ 6MGI structure in complex with G6P, glycine and phosphoglycolate

The crystal structure of ZmPEPC-C₄ with the activators G6P and glycine bound into their respective allosteric sites and the PEP-analog phosphoglycolate (PGA) bound into the active site was solved in space group C222₁. This space group was the same as the previously seen for the ZmPEPC-C₄ structure with a sulfate anion bound in the putative G6P-allosteric site and no ligand bound in either the neutral amino acids-allosteric site or the active site (PDB code 1JQO) [13], but different from that of the ZmPEPC-C₄ structure with glycine bound into its allosteric site, no ligand in the G6P-allosteric site and acetate in the active site (PDB code 5VYJ) [18]. The new structure reported here was refined to a resolution of 2.98 Å. All data collection and refinement statistics are given in Table 1.

Overall, the fold of the protein polypeptide was the same as those described for the two others already reported ZmPEPC-C₄ structures mentioned above. The asymmetric unit is a dimer (Figure 1A) from which the tetramer was constructed by crystallographic symmetry (Figure 1B). The two subunits of the dimer have essentially the same conformation with an average root mean square deviation of 0.51 Å between their equivalent Cα atoms excluding the residues that do not show electron density, which are the same in the two subunits: the first 34 residues in the N-terminus region, residues 127-140, and residues 761-768. Residues 930-935 are not observed in subunit B but observed in subunit A with very high B-factors indicative of their high flexibility. The same flexible regions not showing electronic density can be observed in the subunits of the 1JQO and 5VYJ structures, with the exception of the residues 761-768, which have electronic density in 5VYJ but with very high B-factors values.
The structural superimposition of the Cα atoms of the subunit A of the 6MGI structure reported here with those of the previously reported 1JQO and 5VYJ (Supplementary Figure S1, A and B, respectively) shows that the subunits of these three ZmPEPC-C₄ crystal structures are practically identical regardless of having different ligands, or even no ligands, bound in their respective active and allosteric sites. The most important differences are with 5VYJ, in which—besides the different conformation of the 761-768 residues mentioned above—the conformation of the loop comprising residues 350-369, which contain four residues of the G6P-allosteric site (see below), is different from that observed in both 6MGI and 1JQO. Also, the dimeric arrangement observed in the three crystal structures is very similar (Supplementary Figure S1, C and D) and the most probable tetramer obtained by crystallographic symmetry of the observed 6MGI dimer is essentially identical to the one of 1JQO, also generated by crystallographic symmetry (Supplementary Figure S1E), and to that observed in the asymmetric unit of 5VYJ (Supplementary Figure S1F). The main polar interactions at the dimer/dimer interface in the 6MGI tetramer are two symmetric sets of ionic pairs between the sidechains of Glu493 and Arg498 of opposing subunits and two symmetric sets of hydrogen bonds between the carbonyl oxygen of Pro499 and the sidechain of Arg498, also of opposing subunits (not shown). The same interactions can be observed in the 5VYJ tetramer and can be predicted for the generated tetramer of 1JQO.

Besides the identification of the G6P-allosteric site, which will be described below, the 6MGI crystal structure permitted the observation of important properties of the active site, where an additional electron density consistent with the substrate analog PGA—added to the enzyme preparation previously to the crystallization procedure—was observed. This electron density was tentatively modeled with various components of the crystallization cocktail (G6P, glycine, ethylene glycol and glycerol), but a PGA molecule with occupancy of 0.7 gave the best fit. As shown in Figure 2A, the phosphate group of PGA interacts with the sidechains of Arg456, Glu566 and His177, which is the essential catalytic His residue. The carboxylate group is hydrogen bonded to the
sidechain of Ser602 and with the mainchain carbonyl group of Gly640. The interaction of His177 with the phosphate group of PGA is similar to that observed with one of the carboxylate groups of the inhibitor malonate bound into the active site of the archaeal-type PEPC of Clostridium perfringens [32]. This interaction probably takes place with PEP as well—as suggested by docking simulations using the 1JQO [13] and the 6MGI (not shown) structures—and probably is of great mechanistic relevance given the role of His177 in the PEPC reaction as the general base that stabilizes the carboxyphosphate intermediate and promotes carboxylation [33,34]. Thus, His177 appears to be correctly positioned inside the active site of 6MGI, a well as in those of 1JQO and 5VYJ, which show the same conformation of the active site (Supplementary Figure S2) despite of the fact that 1JQO has no ligand bound and that 5VYJ has acetate bound.

In the 6MGI structure, a glycine molecule was found bound into the neutral amino acids-allosteric site with occupancy of one (Figure 2B) The conformation of this allosteric site, even in the absence of ligand as in 1JQO, is conserved [18] and similar to that observed in 6MGI (Supplementary Figure S3A); the interactions that glycine makes with protein residues in 6MGI are identical to those observed in the 5VYJ structure previously reported [18] (Supplementary Figure S3B).

In contrast with the well formed conformations of the active site and of the allosteric sites for the two kinds of activators, which are the correct ones for carrying out the reaction or binding their respective activators, in 6MGI the inhibitory carboxylic acids-allosteric site is partially unformed and appears unable to bind the inhibitor malate, given the position of residues that participate in binding the inhibitor in other known PEPC structures [35, 36] (Figure 2C). The 1JQO and 5VYJ structures have a conformation of the carboxylic-acids allosteric-site essentially identical to that of 6MGI (Supplementary Figure S3, C and D). Interestingly, in spite of being present during crystallization, malate was not observed bound inside this allosteric-site in 1JQO [13], which is in accordance with not being properly formed.
The G6P-allosteric site of the ZmPEPC-C4 6MGI structure

For the first time in a PEPC crystal structure, a G6P molecule was observed bound in a crevice in the subunit/subunit interfaces of the dimeric units of the 6MGI structure, thus confirming the previously proposed location of the G6P-allosteric site [13]. The electron-density map 2Fo-Fc of G6P is well defined and was refined to an occupancy of one in the final structure. With this occupancy, the average B-factors of the modeled G6P molecule are very similar to those of its nearest neighbor protein atoms and no negative peaks of electronic density were observed. The simulated annealing omit map of G6P, countered at 3.0σ, clearly indicates the interactions made by G6P with protein residues (Figure 3A). The G6P-phosphate group interacts with the sidechains of Arg183, Ser185 and Arg231 as well as with the backbone NH group of Ser185 (ZmPEPC-C4 numbering), all of them of the same subunit. The residual positive charge of the N-terminus of α-helix to which Arg183 and Ser185 belong also contributes to the binding of the phosphate group. The glucose pyranose ring is hydrogen bonded to the Asp239 sidechain of one subunit as well as to the sidechains of Glu360, Trp362 and Lys363 and to the backbone NH of Phe361 of the opposing subunit. Additionally, the pyranose ring and the Phe361 sidechain form two CH-π bonds (Figure 3B), a kind of interaction which is known to be critical for the recognition and binding of carbohydrates by proteins [37]. The length of the CH-π bonds is within the range expected for these interactions [37]. The binding of G6P significantly contributes to the stabilization of the subunit/subunit interface of the dimeric units of the enzyme by bridging the two subunits through the hydrogen bonds that the activator makes with residues of opposing subunits. In addition, some of the residues of opposing subunits that form part of the G6P-allosteric site interact with other residues of the opposing subunit, thus not only participating in the allosteric regulation by G6P but also directly contributing to the stability of this interface.
To explore whether other reported activators of ZmPEPC-C₄—such as F6P, DHAP [4,38] and free-PEP [7,8]—could also be bound in the same allosteric site as G6P, as well as to learn the possible mode of their binding, we constructed in silico energy-minimized models with F6P, DHAP and free-PEP docked in this site. In the models, shown in Supplementary Figure S4, the three activators fit well into this site, without producing any steric clash, and make interactions with protein residues similar to those of the G6P molecule observed in the 6MGI structure reported here. The results of these simulations provide additional support to the notion that this region of the protein is indeed the allosteric site for the phosphorylated compounds known to be activators of the PEPC enzymes.

The important number of interactions between G6P and residues of the G6P-allosteric site observed in the 6MGI structure reported here—and proposed by the docking simulations of F6P, DHAP and PEP—makes clear that this allosteric site has evolved to be highly specific for its likely physiological ligands: G6P, other phospho-sugars and free-PEP, even though it can accommodate other phosphorylated compounds and even small anions analogs of the phosphate group of these compounds such as sulfate anions. The involvement of several of these residues, namely Arg183, Ser185, Arg231, and Glu360 (ZmPEPC-C₄ numbering), was predicted by docking simulations [14] using the ZmPEPC-C₄ 1JQO structure. Also, reported mutagenesis results indicate the importance of Arg183 and Arg231 [15,16] in the activation of ZmPEPC-C₄ by G6P, and chemical modification studies suggest the participation of a lysyl residue in the binding of G6P to its allosteric site in this enzyme [39]. Although the lysyl residue was not identified in the protein amino acid sequence, in view of the 6MGI structure, we believe that Lys363 might be this residue.

The conformation of the G6P-allosteric site with G6P bound observed in 6MGI and the one observed in 1JQO with sulfate bound are virtually identical (Figure 4A). The sulfate anion in the 1JQO structure makes the same interactions with the sidechains of residues Arg183, Ser185 and Arg231 and the backbone NH of Arg184 (Figure 4B), as does the G6P-phosphate group in the 6MGI
structure (Figure 3A). The residues 360-363, which are the ones interacting with the pyranose ring of G6P, also have the same conformations in both structures despite not making any interaction with the sulfate anion in 1JQO. But a structural superposition of the G6P-allosteric site of 6MGI and that of 5VYJ (Figure 4C) showed that although the conformation of the residues that interact with the phosphate group is similar in the latter structure, the conformation of the loop 350-369—which comprises the residues 360-363 that interact with the pyranose ring of G6P—is very different. In 5VYJ, part of this loop forms a short α-helix and the whole loop exhibits a large displacement that prevents the hydrogen bonds between Glu360 and Arg184 of the opposing subunit as well as those between Asp239 and Trp362 and Lys363, also of the opposing subunit (Figure 4D). The different conformation of the loop together with a minor but significant displacement of the α-helices of the other subunit facing this loop, results in a higher accessibility of the G6P-allosteric site entrance. This “open conformation” is observed in the four subunits of 5VYJ, regardless of the presence or not of an acetate molecule bound in the G6P-allosteric site. It appears that the 350-369 loop acts as a gating lid of the G6P-allosteric site. In 6MGI, as well as in 1JQO, this lid is in the “closed conformation”, while in 5VYJ is in the “open conformation” and exhibits an average B-factor value significantly higher than that of the whole structure, indicating its high flexibility, as it would be required for the binding and release of the activator. The local conformational change observed in 5VYJ is probably triggered by the binding of glycine to its allosteric site and suggests a role of the neutral amino acids in facilitating the access of the activator to the G6P-allosteric site, which is consistent with the known synergy between these two kinds of activators [5,20,38,40]. The crystallographic data suggest that G6P upon entering this site and making the appropriate interactions with protein residues through its phosphate group would trigger a conformational arrangement of the 350-369 loop, which would make possible the interactions with the activator pyranose ring, resulting in stabilization of its binding.
The finding by Schlieper et al. [17] of G6P bound into the active site and not in the previously proposed G6P-allosteric site in the crystal structure of the \( \text{FiPEPC-C}_4 \) isozyme, led these authors to question the existence of a G6P-allosteric site and to propose instead that this site is a sulfate-binding site of unknown physiological relevance. Indeed, a sulfate anion was observed bound in this site, as has been observed in the other PEPC structures obtained in the presence of ammonium sulfate: the \( \text{ZmPEPC-C}_4 \) 1JQO structure [13], three structures of \( \text{FiPEPC-C}_4 \) (PDB codes 4BXC, 4BXH, and 3ZGE) [17,36], and the PEPC-C\(_3\) structure of *Flaveria pringlei* (\( \text{FiPEPC-C}_3 \), PDB code 3ZGB) [36]. But while the sulfate anion in 1JQO was taken as an indication of the possible location of the G6P-allosteric site [13], in 4BXC the finding of sulfate instead of G6P was considered as a proof of the nonexistence of the G6P-allosteric site [17]. It has to be noted that, regardless of the important conformational differences between the G6P-allosteric sites of the \( \text{ZmPEPC-C}_4 \) and \( \text{FiPEPC-C}_4 \) reported structures (see below), a sulfate anion may bind to this site in any of these structures. Its binding to this site is consistent with the apparent desensitization of \( \text{ZmPEPC-C}_4 \) to G6P activation by increasing concentrations of ammonium sulfate [15], suggesting competition for the same activator-binding site. Nevertheless, it has also to be noted that sulfate binds differently in \( \text{ZmPEPC-C}_4 \) than in \( \text{FiPEPC-C}_4 \) structures, and that sulfate binding did not alter the different conformations of the G6P-allosteric site observed in these structures. The sulfate anion interacts with residues of only one subunit, as does the phosphate group of G6P, but its binding in the *Flaveria* enzymes is different from that of the sulfate anion or the G6P phosphate group in the \( \text{ZmPEPC-C}_4 \) structures, mainly due to movements in the \( \text{FiPEPC-C}_4 \) structures of the majority of the loops and secondary structure elements in this region of the subunit/subunit interface relative to its position in the 6MGI \( \text{ZmPEPC-C}_4 \) structure (Figure 5A and Supplementary Figure S5). In particular, the hydrogen bond of the sulfate anion with Arg366 of \( \text{FiPEPC-C}_4 \) (equivalent to Arg372 of \( \text{ZmPEPC-C}_4 \)) instead of with Arg226 (equivalent to Arg231 of \( \text{ZmPEPC-C}_4 \)) occurs because of the displacement of the N-terminus of the \( \alpha \)-helix that participates in binding the negatively charged phosphate group or sulfate anion.
(Figure 5A), and to the high flexibility of the sidechain of Arg226, which in the reported \textit{FtPEPC-C}_4 structures does not exhibit electronic density. Indeed, residues Arg227 and Thr228 of the Thr222-Thr231 loop (Thr227-Thr236 of \textit{ZmPEPC-C}_4) are not observed in the \textit{FtPEPC-C}_4 structures, and the residues of this loop that have electronic density exhibit very high \( B \)-factors, indicative of the high flexibility of the whole Thr222-Thr231 loop in these \textit{FtPEPC-C}_4 structures.

We think that the lack of binding of G6P to its allosteric site in the 4BXC \textit{FtPEPC-C}_4 structure plausibly is due to the loss of the interaction that its phosphate group would make with the sidechain of Arg226 and the different position of the important residues Arg183 and Ser185, so that G6P would bind in a different manner not able to trigger the proper conformation of the Ala345-Gln363 loop (equivalent to Ser350-Glu369 of \textit{ZmPEPC-C}_4) (Figure 5A), thus preventing any interaction of four (Glu354, Phe355, Trp356 and Lys357, equivalent to Glu360, Phe361, Tryp362 and Lys363 of \textit{ZmPEPC-C}_4) of the five residues that interacts with the G6P-pyranose ring in the 6MGI \textit{ZmPEPC-C}_4 structure (Figure 5B). The loss of all these important interactions, together with those between these residues and Asp234 (Asp239 of \textit{ZmPEPC-C}_4) that indirectly contributes to the stability of the complex G6P-enzyme, would result in a very weak binding, which could explain why G6P did not bind in its allosteric site in the 4BXC \textit{FtPEPC-C}_4 structure, in spite of the very high G6P concentration (200 mM) used for soaking the crystal.

\textit{Conservation of the G6P-allosteric site residues in plant and bacterial PEPC enzymes}

To further support our conclusion about the existence of the G6P-allosteric site in PEPC enzymes, we investigated the conservation in plant PEPC amino acid sequences of the eight residues that in 6MGI directly interact with the G6P molecule. The amino acid sequences used for this analysis were selected as described in the Materials and Methods section and are included in Supplementary Table S1 and S2. With the exception of Lys363 (\textit{ZmPEPC-C}_4 numbering)—which is conservatively changed to an arginine in a few monocot PEPC-\textit{C}_4 sequences—we found an absolute conservation of
these residues, regardless of the PEPC isozymes being of the PEPC-C₄, PEPC-C₃ or PEPC-CAM type (Figure 6), underlining the importance of this allosteric site in plant PEPCs. We also investigated the conservation of these residues in bacterial PEPC sequences from selected organisms (Supplementary Table S3). The two arginyl residues that the 6MGI structure shows interacting with the G6P-phosphate group (Arg183 and Arg231, ZmPEPC-C₄ numbering) are highly conserved, and the third residue (Ser185, ZmPEPC-C₄ numbering) is conservatively changed to a threonine in most of the bacterial sequences included in our analysis. Regarding the five residues that interact with the G6P pyranose ring, one of them (Asp239 ZmPEPC-C₄ numbering) is highly conserved in the bacterial PEPC sequences analyzed by us (in a few of them is conservatively changed to glutamate and in one to asparagine). The other four residues (Glu360, Phe361, Trp362 and Lys363) are not present in the bacterial enzymes because they belong to a loop that only exist in plant PEPCs. This high conservation suggests that important features of the G6P-allosteric site are also present in the bacterial PEPC enzymes, and supports the relevance of this site for the regulation of the activity of PEPC enzymes in general, with the exception of the archaeal PEPC-type where the G6P-allosteric site is absent [32].

**Biochemical evidence supporting the existence and functionality of the G6P-allosteric site in ZmPEPC-C₄**

In order to get biochemical evidence of the existence of the activating allosteric site for G6P and other phosphorylated compounds, we generated the variant ZmPEPC-C₄ enzyme D239A and tested its activation by G6P, F6P, and PEP, as well as its general kinetic properties. Of the eight residues that directly interact with G6P, we chose to change Asp239 because: (i) it was not changed in previous site-directed mutagenesis studies of the residues of the proposed G6P-allosteric site [15,16]; (ii) different from the previously changed residues, it interacts with the G6P-pyranose ring instead of the G6P-phosphate group and, as shown above, is the only one of the residues involved in binding
this ring that is absolutely conserved not only in plants but also in bacterial PEPC enzymes; (iii) the docking simulations performed by us suggested that it is also involved in the binding of others phospho-sugars and free-PEP; and (iv) it participates in a web of hydrogen bonds involving the pyranose ring and protein residues that would be lost in the variant enzyme. Thus, Asp239 is hydrogen bonded to the 2-OH group of the pyranose ring and to the sidechains of Trp362 and Lys363, which in turn are hydrogen bonded to the 2-OH and 3-OH groups, respectively, of this ring (Figure 3A). Indeed, through these hydrogen bonds, Asp239 is the residue most connected to G6P. Additionally, through its interactions with Trp362 and Lys363, Asp239 contributes to the correct position of Glu360 and Phe361, which are the two other protein residues hydrogen-bonded to the G6P-pyranose ring (Figure 3A). Note that these interactions of Asp239 could not be predicted in the reported molecular modeling experiments [14] because these studies were performed using the 1JQO structure where Asp239 was modeled as alanine.

Figure 7A shows that G6P and F6P not only did not activate but, on the contrary, inhibited the D239A variant enzyme, even at the lowest tested concentrations. G6P in the same concentration range exhibited a dual effect on the wild-type enzyme activity, being an activator at low and intermediate concentrations but an inhibitor at high concentrations (Figure 7B). These activatory and inhibitory effects of G6P on the wild-type enzyme suggest its binding to both the G6P-allosteric site, producing enzyme activation, and with much lower affinity to the active site, producing inhibition by competing with the substrate PEP, as was previously suggested [38]. Then, a plausible explanation of the G6P-saturation data of the D239A variant is that the activatory effect was abolished—which would be consistent with the existence of a non-functional G6P-allosteric site as a consequence of the mutation—while the inhibitory effect was preserved—which would be consistent with an active site still able to bind G6P. Activation of the D239A variant by F6P was also lost; instead of activation, a small inhibition by this compound was observed (Figure 7A), suggesting that F6P also binds to the
active site but with very low affinity, in agreement with our finding that F6P apparently does not inhibit the wild-type enzyme even at very high concentrations (Figure 7B).

The G6P- and F6P-saturation data of the D239A variant were best fitted to the equation that allows the estimation of the $I_{50}$ value of a total inhibitor (equation (1)), but the complex nature of the G6P-saturation data of the wild-type enzyme does not permit the fitting of a single saturation experiment to an equation simultaneously accounting for non-essential activation and total inhibition, so that to estimate apparent kinetic parameters of the wild-type enzyme for comparative purposes, we fitted the velocity data that do not exhibit apparent inhibition, i.e. those from zero to 15 mM G6P, to the equation corresponding to saturation by a non-essential activator (equation (2)). This equation was also used to fit the F6P-saturation data of the wild-type enzyme. The estimated kinetic parameters, given in Table 2, show that the affinity of the G6P-allosteric site for G6P is considerably higher than for F6P, consistent with the higher number of interactions made by G6P with residues of this site, as indicated by our crystallographic data and docking simulations (Figures 3 and Supplementary S4). This also seems to be the case for the site responsible for the inhibition by these compounds. In any case, the inhibition of the enzyme by ligands of the G6P-allosteric site is probably physiologically irrelevant given the high concentrations of the activators at which it takes place.

Assuming again that the inhibitory effects of the activators are due to their binding to the active site, we cannot at present explain the apparent positive cooperativity observed in the G6P-inhibition data of the variant D239A, given the apparently loss of this cooperativity in the saturation by PEP (see below). This finding deserves further research.

The total lack of activation of the D239A variant by G6P and F6P clearly indicates that the G6P-allosteric site is the one responsible for activation of the enzyme by phospho-sugars and related phosphorylated compounds. Although in the D239A variant the web of hydrogen bonds made by this residue are lost, the non-functionality of the G6P-allosteric site is not probably due to its inability to bind these compounds since the residues involved in binding the phosphate group probably have the
appropriate conformation for binding this group. It rather seems that the affinity of the G6P-allosteric site for the allosteric activators is very much reduced because the loss of the interactions of Asp239 with Trp362 and Lys363 would hamper the attainment of the proper “close conformation” of the loop 350-369 upon binding of the activator, and therefore would prevent the interactions of the residues in this loop with the G6P-pyranose ring. Also, the loss of the latter interactions could prevent the allosteric transition to occur. In other words, the activators probably bind to the variant enzyme very weakly and in a non-productive manner, i.e. in a manner that does not trigger the transmission of the allosteric signal. This hypothesis remains to be investigated.

To find out whether in the D239A variant free-PEP behaves as an allosteric activator that binds to the G6P-allosteric site, as proposed for the wild-type enzyme [7,8], we studied the kinetics of saturation of the wild-type and the D239A variant enzymes by total PEP at a fixed free-Mg$^{2+}$ concentration, as well as those by free-PEP at a constant Mg$^{2+}$-PEP concentration and of Mg$^{2+}$-PEP at a fixed free-PEP concentration (Figures 7C, D and E, respectively). In the saturation by total-PEP, the D239A variant showed a decrease in both $V_{\text{max}}$ and in the affinity for PEP. Also, while the wild-type enzyme exhibited a positive cooperativity in the binding of PEP, the variant D239A binds PEP non-cooperatively (Table 2). Qualitatively similar results were obtained when free-PEP was increased at constant 0.1 mM PEP-Mg$^{2+}$: the variant D239A showed lower $V_{\text{max}}$ and affinity for the substrate and the positive cooperativity observed in the wild-type enzyme was abolished (Figure 7D and Table 2). The lower activities observed for both enzymes in the free-PEP saturation experiments when compared with the saturation by total-PEP probably result from the limited availability of Mg$^{2+}$, which concentrations were increasingly lower as the free-PEP concentration rise in order to keep constant the concentration of Mg$^{2+}$-PEP at 0.1 mM. Together, these results suggest the loss in the variant enzyme of the heterotropic positive effects on $V_{\text{max}}$ and substrate affinity of free-PEP, which are to be expected, and in addition, suggest that the positive cooperativity observed in the wild-type enzyme arises from activation by free PEP. When saturating the enzyme with total-PEP at
a fixed free-$\text{Mg}^{2+}$ concentration, the concentrations of both $\text{Mg}^{2+}$-PEP and free-PEP are increased in a constant proportion, which would not result in a deviation of the Michaelis-Menten behavior if the enzyme binds both substrates in a non-cooperative manner, but cooperativity would result if the enzyme binds its allostERIC activator, free-PEP, cooperatively, as the other G6P-allosteric ligands do. The same effect would be responsible for the positive cooperativity observed when free-PEP was varied at constant $\text{Mg}^{2+}$-PEP. Thus, the apparent positive cooperativity in the binding of the substrate observed in the wild-type enzyme probably result from the increases in the affinity of the enzyme for its substrate upon binding of free-PEP to the G6P-allosteric site, and would be a reflection of its cooperative binding to this site. In the saturation by $\text{Mg}^{2+}$-PEP at 5 mM fixed free-PEP the variant did not show positive cooperativity in the binding of the varied substrate (Figure 7E and Table 2), as in the other two saturation experiments described above. This finding could be explained as before, i.e. binding of the substrate is non-cooperative if free-PEP does not have its allosteric heterotropic activatory effects as a consequence of a non-functional allosteric-G6P site. In the case of the wild-type enzyme the observed decreased degree of cooperativity is likely due to the fixed concentration of free-PEP being almost saturating for the G6P-allosteric site, thus producing an important proportion of the activated non-cooperative enzyme form.

Other kinetic features of ZmPEPC-C$_4$ were also significantly altered in the D239A variant: the marked decrease in both the affinity for glycine and in the degree of activation achieved by this activator (Figure 7F and Table 2), and a much lower affinity for the inhibitor malate (Figure 7G and Table 2). We believe that the diminished response to glycine observed in the D239A variant likely arise from the abolition in the variant enzyme of the allostERIC activation by free-PEP, which would result in the loss of the synergy previously observed between glycine and G6P, or between glycine and other activators that bind to the G6P-allosteric site [5,20,38,40]. Therefore, our results suggest that the activation by neutral amino acids in PEPC-C$_4$ from grasses importantly rely on the G6P-allosteric site, even in the absence of G6P, through the ability of this site to bind free-PEP, which
acting as an activator increases the affinity and activatory effects of neutral amino acids. The lower affinity for malate of the D239A variant was unexpected; it suggests a so far unknown complex relationship between the G6P- and the carboxylic acids-allosteric sites that seems to imply a kind of positive interaction between them. The poor ability of G6P to relieve malate inhibition in wild-type ZmPEPC-C_4 [20] would be consistent with dual effects of this activator, and probably of other ligands of the G6P-allosteric site including free-PEP, counteracting the carboxylic acids inhibition, by increasing the affinity of the active site for the substrate, and at the same time favoring their binding to the carboxylic acids-allosteric site. This possibility remains to be investigated.

Taken together, our kinetic results of the D239A variant and the wild-type ZmPEPC-C_4 enzymes provide additional support for the allosteric mechanism of activation by G6P. Also, these results are consistent with free-PEP acting as a substrate in both enzymes and as an allosteric activator that binds to the G6P-allosteric site in the wild-type enzyme, but not in the D239A variant.

Although Asp239 participates in interactions between two opposite subunits, the D239A variant altered kinetics couldn’t be attributed to changes in the protein association state, folding or stability. The recombinant D239A protein was obtained in the soluble fraction of E. coli extracts with a yield similar to that of the wild-type enzyme, indicating that is properly folded. Also, its native structure is tetrameric (Supplementary Figure S6A) and exhibits a thermal stability similar to the wild-type enzyme in thermal-shift experiments, with apparent transition temperatures ($T_m$) of 47.34 $^\circ$C for the wild-type and 48.14 $^\circ$C for the D239A variant (Supplementary Figure S6B).

**Final Remarks**

The widely accepted existence of an allosteric site for G6P and the allosteric nature of G6P activation were questioned on the basis of both structural and kinetic results [17]. In the present work, we provide new structural and kinetic data obtained with the ZmPEPC-C_4 isozyme that support that this site does exist and that it is responsible for the activation by G6P and other phosphorylated...
compounds, being free-PEP among them. Knowing for certain which of the two mechanisms proposed for the activation of PEPC enzymes by G6P—and by extension other physiologically relevant phospho-sugars and free-PEP—, i.e. allosteric activation or competitive activation, is of great importance because each mechanism would have different physiological effects. While in the competitive activation mechanism the activator is in fact an inhibitor and would produce inhibition if its concentration increases, particularly at low or moderate substrate concentrations [41], in the allosteric activation mechanism the activatory effect will be maintained after reaching its maximum at saturating concentrations of the activator, regardless of the substrate concentration. Because of this, the activator concentration range in which activation can be observed in the “competitive activation” mechanism is limited, and saturation of the active site by the “activator” would lead to zero enzyme activity. On the other hand, in the allosteric activation mechanism there is not such a restriction and the maximum activatory effect attained by saturation of the allosteric site by the activator persists even though the activator concentration is further increased. In the case of the PEPC enzymes, as observed in ZmPEPC-C₄, activation by G6P is followed by inhibition at higher activator concentrations, but, different from the competitive activation mechanism, the two opposite effects are a consequence of G6P binding to different sites, which greatly differ in their affinity for the activator, being that of the activatory site, the allosteric site, much greater than that of the inhibiting site, the active site. This allows the full expression of the activatory potential of the activator in an ample range of concentration before inhibition becomes significant. Moreover, the degree of activation of the enzyme elicited by two different activators that bind to the same allosteric site and are present in the enzyme environment at the same time—as is probably the case of the photosynthetic PEPC-C₄ isozymes that respond to several physiological phospho-sugars [4,9,38,42], which concentrations increase simultaneously under light conditions [4,43-46] and of those PEPC enzymes activated by free-PEP—will be different in the two mechanisms. While in the allosteric mechanism the activatory effects would be additive until saturation of the allosteric site by any of the activators is achieved, as
it has been experimentally found [38], in the case of competitive activation the degree of the activation elicited by any of them would decrease, even at relatively low concentrations, and inhibition of the enzyme activity would be observed at lower concentration of the activators, because more active sites would be occupied by any of them and less active sites would be available to the substrate. Therefore, our conclusion that activation of PEPC enzymes by phosphorylated compounds results from their binding to the G6P-allosteric site is physiologically sound, since under physiological conditions competitive activation would be very inefficient and even could be detrimental for the plant.

It has to be noted that the kinetic results that Schlieper et al. [17] reported as a proof of competitive activation—higher degree of activation of FtPEPC-C₄ by G6P at lower PEP concentrations—could equally support allosteric activation, because in both mechanisms, activation will be greater at low substrate concentrations, as has been repeatedly observed in the case of PEPC enzymes. This is because in the case of competitive activation, competition between the activator and the substrate increases as does the substrate concentration, or in the case of allosteric mechanism because the activator increases the affinity of the enzyme for the substrate, even in those enzymes in which the allosteric activator also increases the maximum velocity [47], as seems to be the case of the ligands of the G6P-allosteric site. Neither is the greater activatory effect of high concentrations of ethylene glycol on FtPEPC-C₄ at low PEP concentrations, reported by Schlieper et al. [17] in support for the competitive activation mechanism, conclusive evidence of the existence of this mechanism, since the activation by ethylene glycol probably arises from its action as a cosolute that preferentially stabilizes and shifts the equilibrium towards the active R-state, not as a ligand that specifically binds to the active or allosteric site. This is further indicated by the high concentrations required to observe activation. Indeed, only high concentrations of ethylene glycol protected the ZmPEPC-C₄ against dilution, activate the enzyme and overcome its inhibition by malate inhibition, and these effects were attributed to the changes that it produces in the solvent properties [48]. In a similar fashion was
interpreted the activation by another small cosolute, glycerol [10,38,49], which have additive effects with G6P, thus implying a different mechanism of activation [50].

In summary, the ZmPEPC-C4 6MGI crystal structure with G6P firmly bound into the previously proposed G6P-allosteric site, the conservation in PEPC enzymes of the residues of this site involved in binding the activator, and the altered general kinetics of the D239A variant, strongly indicate the existence of the G6P-allosteric site and the allosteric nature of activation of PEPC enzymes by G6P and other phosphorylated compounds, of which free-PEP likely is of most importance. Indeed, the results of the kinetic characterization of the D239A variant suggest a role of the G6P-allosteric site beyond its participation in the activation by G6P or other phospho-sugars due to the probably general ability of this site in PEPC enzymes in general to bind free-PEP.

**Accession number**

Coordinates and structure factors have been deposited in the Protein Data Bank with the accession number 6MGI.

**Author contributions**

RAMC conceived, designed, coordinated and supervised the study, analyzed the structural and kinetic data, and wrote the manuscript; LGS carried out crystallization, diffraction data collection, crystal structure determination and analysis, docking simulations, amino acid sequence analysis and figures preparation; CMJ performed protein expression, purification and kinetic experiments; JAJD performed site-directed mutagenesis. All authors revised and approved the final version of the manuscript.

**Acknowledgements**
Crystallographic results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are grateful to the staff of the beamline 19-ID for their support during data collection, particularly to Dr. Randy Alkire and Dr. Norma Duke.

**Funding**

This work was supported by DGAPA-UNAM (PAPIIT grant IN216911) and CONACYT (grant 252123) to R.A.M.-C.

**Competing interests**

The authors declare that they have no conflict of interest with the contents of this article.

**Abbreviations**

DHAP, dihydroxyacetone phosphate; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; PEP, phosphoenolpyruvate; PDB, Protein Data Bank; PEPC, phosphoenolpyruvate carboxylase; PGA, phosphoglycolate.
References

1. Chollet, R., Vidal, J. and O’Leary, M.H. (1996) Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 273–298

2. Andreo, C.S., González, D.H. and Iglesias, A.A. (1987) Higher plant phosphoenolpyruvate carboxylase: structure and regulation. FEBS Lett. 213, 1–8

3. Selinioti, E., Karabourniotis, G., Manetas, Y. and Gavalas, N.A. (1985) Modulation of phosphoenolpyruvate carboxylase by 3-phosphoglycerate: probable physiological significance for C₄-photosynthesis. J. Plant Physiol. 121, 353–360

4. Doncaster, D.H. and Leegood, R.C. (1987) Regulation of phosphoenolpyruvate carboxylase activity in maize leaves. Plant Physiol. 84, 82–87

5. Bandarian, V., Poehner, W.J., and Grover, S.D. (1992) Metabolite activation of crassulacean acid metabolism and C₄ phosphoenolpyruvate carboxylase. Plant Physiol. 100, 1411–1416

6. Rustin, P., Meyer, C.R. and Wedding, R.T. (1988) Identification of substrate and effector binding site of phosphoenolpyruvate carboxylase from Crassula argentea. A possible role of phosphoenolpyruvate as substrate and activator. J. Biol. Chem. 263, 17611–17614

7. Rodríguez-Sotres, R. and Muñoz-Clares, R.A. (1990) Kinetic evidence of the existence of a regulatory phosphoenolpyruvate binding site in maize leaf phosphoenolpyruvate carboxylase. Arch. Biochem. Biophys. 276, 180–190

8. Tovar-Méndez, A., Rodríguez-Sotres, R., López-Valentín, D.M. and Muñoz-Clares, R.A. (1998) Re-examination of the roles of PEP and Mg²⁺ in the reaction catalysed by the phosphorylated and non-phosphorylated forms of phosphoenolpyruvate carboxylase from leaves of Zea mays. Biochem. J. 332, 633–642

9. Wong, K.F. and Davies, D.D. (1973) Regulation of phosphoenolpyruvate carboxylase of Zea mays by metabolites. Biochem. J. 131, 451–458
10. Uedan, K. and Sugiyama, T. (1976) Purification and characterization of phosphoenolpyruvate carboxylase from maize leaves. *Plant Physiol.* **57**, 906–910

11. Mukerji, S.K. (1977) Corn leaf phosphoenolpyruvate carboxylases: inhibition of $^{14}$CO$_2$ fixation by SO$_3^{2-}$ and activation by glucose 6-phosphate. *Arch. Biochem. Biophys.* **182**, 360–365

12. Stiborová, M. and Leblová, S. (1985) Activation of maize (*Zea mays* L.) phosphoenolpyruvate carboxylase by glucose-6-phosphate and glycine: effects of pH and Mg$^{2+}$. *Photosynthetica* **19**, 177–182

13. Matsumura, H., Xie, Shirakata, Y.S., Inoue, T., Yoshinaga, T., Ueno, Y. et al. (2002) Crystal structures of C$_4$ from maize and quaternary complex of *E. coli* phosphoenolpyruvate carboxylases. *Structure* **10**, 1721–1730

14. Mancera, R.L. and Carrington, B.J. (2005) The molecular binding interactions of inhibitors and activators of phosphoenolpyruvate carboxylase. *J. Mol. Struct. Theochem.* **755**, 151–159

15. Takahashi-Terada, A., Kotera, M., Ohshima, K., Furumoto, T., Matsumura, H., Kai, Y. et al. (2005) Maize phosphoenolpyruvate carboxylase. Mutations at the putative binding site for glucose 6-phosphate caused desensitization and abolished responsiveness to regulatory phosphorylation. *J. Biol. Chem.* **280**, 11798–11806

16. Yuan, J., Sayegh, J., Mendez, J., Sward, L., Sanchez, N., Sanchez, S. Waldrop, G. et al. (2006) The regulatory role of residues 226-232 in phosphoenolpyruvate carboxylase from maize. *Photosynth. Res.* **88**, 73–81

17. Schlieper, D., Förster, K., Paulus, J.K. and Groth, G. (2014) Resolving the activation site of positive regulators in the plant phosphoenolpyruvate carboxylase, *Mol. Plant* **7**, 437–440

18. González-Segura, L., Mújica-Jiménez, C., Juárez-Díaz, J.A., Güemez-Toro, R., Martínez-Castilla, L.P. and Muñoz-Clares, R.A. (2018) Identification of the allosteric site for neutral amino acids in the maize C$_4$ isozyme of phosphoenolpyruvate carboxylase: the critical role of
Ser-100. *J. Biol. Chem.* **293**, 9945–9957

19. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I. and Appel, R.D. (2003) Bairoch, A.ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788

20. Tovar-Méndez, A., Mújica-Jiménez, C. and Muñoz-Clares, R.A. (2000) Physiological implications of the kinetics of maize leaf phosphoenolpyruvate carboxylase. *Plant Physiol.* **123**, 149–160

21. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132

22. Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763

23. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D. Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674

24. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N. et al. (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213–221

25. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132

26. Trott, O. and Olson, A. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461

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

28. Bläsing, O.E., Westhoff, P. and Svensson, P. (2000) Evolution of C4 phosphoenolpyruvate carboxylase in *Flaveria*, a conserved serine in the carboxyterminal part of the enzyme is a
major determinant for C4-specific characteristics. *J. Biol. Chem.* **275**, 27917–27923

29. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. et al. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics* **23**, 2947–2948

30. Crooks, G. E., Hon, G. Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190

31. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.* **18**, 6097–6100

32. Dharmarajan, L., Kraszewski, J.K., Mukhopadhyay, B. and Dunten, P.W. (2011) Structure of an archaeal-type phosphoenolpyruvate carboxylase sensitive to inhibition by aspartate. *Proteins* **79**, 1820–1829

33. Terada, K. and Izui, K. (1991) Site-directed mutagenesis of the conserved histidine residue of phosphoenolpyruvate carboxylase. *Eur. J. Biochem.* **202**, 797–803

34. Izui, K., Matsumura, H., Furumoto, T. and Kai, Y. (2004) Phosphoenolpyruvate carboxylase: a new era of structural biology. *Annu. Rev. Plant Biol.* **55**, 69–84

35. Kai, Y., Matsumura, H., Inoue, T., Terada, K., Nagara, Y., Yoshinaga, T. et al. (1999) Three-dimensional structure of phosphoenolpyruvate carboxylase: A proposed mechanism for allosteric inhibition. *Proc. Natl. Acad. Sci. USA* **96**, 823–828

36. Paulus, J.K., Schlieper, D. and Groth, G. (2013) Greater efficiency of photosynthetic carbon fixation due to single amino-acid substitution. *Nat. Commun.* **4**, 1518

37. Hudson, K.L. Barlett, G.J., Diehl, R.C., Agirre, J. Gallagher, T., Kiessling, L.L. et al. (2015) Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160

38. Mújica-Jiménez, C., Castellanos-Martínez, A. and Muñoz-Claires, R.A. (1998) Studies of the allosteric properties of maize leaf phosphoenolpyruvate carboxylase with the phosphoenolpyruvate analog phosphomycin as activator. *Biochim. Biophys. Acta* **1386**, 132–144.
39. Tovar-Méndez, A., Mújica-Jiménez, C. and Muñoz-Clares, R.A. (1997) Desensitization to glucose 6-phosphate of phosphoenolpyruvate carboxylase form maize leaves by pyridoxal 5'-phosphate. *Biochim. Biophys. Acta.* **1337**, 207–216

40. Gillinta, J. and Grover, S. (1995) Kinetic interactions of glycine with substrates and effectors of phosphoenolpyruvate carboxylase from maize leaves. *Photosynth. Res.* **45**, 121–126

41. Gerhart, H.C. and Pardee, A.H. (1963) The effect of the feedback inhibitor, CTP, on subunit interactions in aspartate transcarbamylase. Cold Spring Harb. Symp. **Quant. Biol.** **28**, 491–496

42. Coombs, J., Baldry, C.W. and Bucke, C. (1973) The C-4 pathway in *Pennisetum purpureum*. I. The allosteric nature of PEP carboxylase. *Planta* **110**, 95–107

43. Stitt, M. and Heldt, H.W. (1985) Generation and maintenance of concentration gradients between the mesophyll and bundle sheath in maize leaves. *Biochim. Biophys. Acta* **808**, 400–414

44. Stitt, M. and Heldt, H.W. (1985a) Control of photosynthetic sucrose synthesis by fructose-2,6-bisphosphate: intercellular metabolite distribution and properties of the cytosolic fructose bisphosphatase in leaves of *Zea mays* L. *Planta* **164**, 179–188

45. Stitt, M. and Zhu, X.G. (2014) The large pools of metabolites involved in intercellular metabolite shuttles in C₄ photosynthesis provide enormous flexibility and robustness in a fluctuating light environment. *Plant Cell Environ.* **37**, 1985–1988

46. Arrivault, S., Obata, T., Szecówka, M., Mengin, V., Guenther, M., Hoehne, M. et al. (2017) Metabolite pools and carbon flow during C₄ photosynthesis in maize: ¹³CO₂ labeling kinetics and cell type fractionation. *J. Exp. Bot.* **68**, 283–298

47. Traut, T.W. (2008) Allosteric regulatory enzymes pp. 201–207. Springer

48. Ogawa, N., Kai, T., Yabuta, N. and Izui, K. (1997) Phosphoenolpyruvate carboxylase of maize leaves: an improved method for purification and reduction of the inhibitory effect of malate by ethylene glycol and bicarbonate. *Plant Cell Physiol.* **38**, 76–80
49. Podesta, F.E. and Andreo, C.S. (1989) Maize leaf phosphoenolpyruvate carboxylase. Oligomeric state and activity in the presence of glycerol. Plant Physiol. 90, 427–433

50. Wedding, R.T., O’Brien, C.E. and Kline, K. (1994) Oligomerization and the affinity of maize phosphoenolpyruvate carboxylase for its substrate. Plant Physiol. 104, 613–616

FIGURE LEGENDS

Figure 1. The overall structure of the ZmPEPC-C_4 6MGI structure in complex with G6P, glycine and PGA. (A) Cartoon representation of the dimer observed as the asymmetric unit. The position of the two symmetrical G6P- and neutral amino acids-allosteric sites in the subunit/subunit interface is shown, as well as that of the two active sites. (B) Cartoon representation of the tetramer obtained by crystallographic symmetry. G6P, glycine (Gly) and PGA are depicted as spheres (carbon black, oxygen red and nitrogen blue).

Figure 2. Close-up of the active site and of the glycine- and malate-allosteric sites of the ZmPEPC-C_4 6MGI structure. (A) The active site of subunit A bound with PGA. (B) The neutral amino acids-allosteric site bound with glycine (Gly). The PGA and glycine electron density maps (2F_o-F_c) are contoured at 1σ (blue mesh) and the simulated annealing omit maps (F_o-F_c) are shown countered at 3.0 σ (green mesh). Amino acid sidechains, PGA and glycine are depicted as sticks with carbon atoms in green (subunit A), yellow (subunit B) or black (glycine and PGA), oxygen atoms in red and nitrogen atoms in blue. Hydrogen bonds are depicted as black dashed lines; cutoff is 3.5 Å. (C) The carboxylic acids-allosteric site of subunit A.

Figure 3. The allosteric-G6P site of the ZmPEPC-C_4 6MGI structure. (A) Stereoview of a close-up of the G6P-allosteric site bound with α-D-glucose 6-phosphate (G6P) (cartoon
representation). The G6P simulated annealing omit map ($F_o - F_c$) is shown countered at 3.0 σ (green mesh). Amino acid sidechains carbons are green (subunit A) or yellow (subunit B). Hydrogen bonds are depicted as black dashed lines; cutoff is 3.2 Å. The figure was made with PyMOL using the Wall-Eye Stereo Mode. (B) CH-π interactions of Phe361 side-chain with the C(3) and C(5) of the pyranose ring of G6P bound in the G6P-allosteric site of 6MGI. Bond distances are in Å.

**Figure 4. Structural comparison of the G6P-allosteric site of the ZmPEPC-C₄ 6MGI structure with those of ZmPEPC-C₄ 1JQO and ZmPEPC-C₄ 5VYJ structures.** (A) Cartoon representation of the superimposition of 6MGI bound with G6P on 1JQO bound with sulfate. (B) Close-up of the G6P-allosteric site of the 1JQO structure showing the interactions with sulfate [13]. (C) Cartoon representation of the superimposition of 6MGI bound with G6P on 5VYJ bound with acetate (ACT) [18]. Note the different conformation of the loop comprising residues 350-369, which makes the G6P-allosteric site of 5VYJ much more open and accessible than those of 6MGI and 1JQO. 6MGI carbons are shown yellow or green and 1JQO and 5VYJ carbons dark grey or white, depending on the subunit. (D) Close-up of the G6P-allosteric site of 5VYJ [18]. In panels (A) and (C), PGA bound in the active site of 6MGI and glycine (Gly) bound in the neutral amino acids-allosteric sites of 6MGI and 5VYJ structures are also shown. G6P, glycine and PGA are depicted with black carbons and ACT with cyan carbons. In panels (B) and (D) hydrogen bonds are depicted as black dashed lines; cutoff is 3.5 Å. PyMOL was used to generate the figures.

**Figure 5. Comparison of the G6P-allosteric site of the ZmPEPC-C₄ 6MGI structure with that of the FtPEPC-C₄ 4BXC structure.** (A) Superposition of the G6P-allosteric site of 6MGI (green and yellow carbons) with bound G6P and 4BXC (light and dark grey carbons) with bound sulfate. Glycine (Gly, black carbons) bound in 6MGI and sulfate (SO₄, yellow sulfur) bound in 4BXC are
also shown. Red dashed circles mark the regions without electronic density in 4BXC; loops with different conformation in 4BXC relative to 6MGI, which causes the displacement of critical residues in 4BXC and precludes the binding of G6P in the allosteric site of 4BXC. (B) Close up of the G6P-allosteric site of both structures showing the displacement of the sidechains involved in G6P-binding. 6MGI sidechains are depicted with green carbons and 4BXC sidechains with light grey carbons. The dimer of *Ft*PEPC-C$_4$ 4BXC was generated by crystallographic symmetry using monomer A.

**Figure 6. Conservation of residues directly involved in G6P binding to its allosteric site in the ZmPEPC-C$_4$ 6MGI structure.** Residues numbering corresponds to ZmPEPC-C$_4$ (CAD60555) for C$_4$ sequences, to ZmPEPC-C$_3$ (PWZ12751) for C$_3$ sequences and to EcPEPC (CAA29332) for bacterial sequences. The sequences analyzed are given in Supplementary Tables S1, S2 and S3. Two sequences from CAM plants were included in the C$_3$ logos; one of them probably corresponds to a PEPC-CAM type isozyme since instead of alanine has a serine at the position equivalent to Ser780 of ZmPEPC-C$_4$. The amino acids color scheme was according to the chemical properties of their sidechains: non-polar (P, A, V), black; polar uncharged (S, T, N, Q), green; aromatic (F, W), purple; cationic (K, R, H), blue; and anionic (D, E), red.

**Figure 7. Comparison of the kinetics of the D239A variant and wild-type ZmPEPC-C$_4$ enzymes.** (A and B) Saturation by G6P (closed symbols) or F6P (open symbols). (C, D and E) Saturation by total-PEP, free-PEP of Mg$^{2+}$-PEP, respectively. (F) Saturation by glycine. (G) Saturation by malate. In all panels initial velocity data are depicted as red symbols (D239A variant) or black symbols (wild-type). Assays were carried out at pH 7.4 and 30 °C, 0.4 mM free-Mg$^{2+}$, and 0.1 mM bicarbonate, except the saturation by malate experiments which were carried...
out at 2.5 mM total-PEP, 5 mM total-Mg\(^{2+}\) and 1 mM bicarbonate. The results are mean values from at least two independent experiments. The points in the figures are the experimentally determined values, whereas the curves are calculated from fits of these data to the appropriate equations, described in the Materials and Methods section.
Table 1. X-ray diffraction data collection and model-refinement statistics for ZmPEPC-C₄ in complex with G6P-Glycine-PGA

| Data collection$^a$ |  |
|---|---|
| Wavelength (Å) | 0.9795 |
| Space group | $C222_1$ |
| Unit cell dimensions |  |
| $a$, $b$, $c$ (Å) | 152.09, 170.68, 244.62 |
| Asymmetric unit | Two subunits |
| Resolution range (Å) | 29.62-2.98 |
| Unique reflections | 62,789 (8,596) |
| Total reflections | 337,329 (44,892) |
| Completeness (%) | 97.1 (88.3) |
| $R_{merge}$ (%) | 10.2 (50.6) |
| Mean $I/\sigma$ (I) | 12.1 (2.9) |
| Multiplicity | 5.4 (5.2) |
| $CC_{1/2}$ | 0.99 (0.85) |

| Refinement statistics |  |
|---|---|
| Resolution range (Å) | 29.62-2.98 |
| $R_{work}/R_{free}$ (%) | 20.83/24.35 |
| Number of atoms |  |
| Protein | 14,563 |
| G6P | 32 |
| PGA | 18 |
| Glycine | 10 |
| Ethylene glycol | 24 |
| Glycerol | 12 |
| Water | 21 |
| Average $B$-factors ($Å^2$) |  |
| Protein | 53.3 |
| G6P | 52.9 |
| PGA | 61.5 |
| Glycine | 62.8 |
| Ethylene glycol | 45.3 |
| Glycerol | 45.6 |
| Water | 40.6 |
| r.m.s. deviations from ideal stereochemistry$^b$ |  |
| Bond lengths (Å) | 0.003 |
| Bond angles ($^o$) | 0.570 |
| Ramachandran plot statistics |  |
| Residues in most favored regions (%) | 94.81 |
| Residues in allowed regions (%) | 4.08 |
| Residues in non-allowed regions (%) | 1.11 |

| Protein Data code | 6MGI |

$^a$Values in parentheses show the statistics for the highest resolution shell. $^b$r.m.s. indicates root mean square.
Table 2. Apparent kinetic parameters of wild-type and D239A variant ZmPEPC-C₄ enzymes

| Variable ligand | Parameter  | Wild-type          | D239A            |
|-----------------|------------|--------------------|------------------|
|                 | v₀         | 6.5                | 4.8              |
|                 | vₐ max     | 36.2 ± 1.0         | NA               |
|                 | A₀.₅       | 4.5 ± 0.2          | NA               |
|                 | I₀.₅       | NA                 | 15.4 ± 0.4       |
|                 | h          | 2.10 ± 0.17        | 1.56 ± 0.1       |
| G6P             |            |                    |                  |
| F6P             | v₀         | 5.7                | 6.6              |
|                 | vₐ max     | 36.2 ± 0.5         | NA               |
|                 | A₀.₅       | 14.2 ± 0.4         | NA               |
|                 | I₀.₅       | NA                 | 218.1 ± 15.1     |
|                 | h          | 2.00 ± 0.10        | 0.97 ± 0.14      |
| Total-PEP       | Vₐ max     | 27.5 ± 0.3         | 24.7 ± 0.6       |
|                 | S₀.₅ or Kₘ | 4.0 ± 0.1          | 8.4 ± 0.6        |
|                 | h          | 2.21 ± 0.16        | 1.07 ± 0.06      |
| Free-PEP        | Vₐ max     | 27.9 ± 1.3         | 19.6 ± 1.6       |
|                 | S₀.₅ or Kₘ | 4.9 ± 0.4          | 5.3 ± 1.0        |
|                 | h          | 1.75 ± 0.14        | 1.12 ± 0.12      |
| Mg²⁺-PEP        | Vₐ max     | 26.4 ± 0.1         | 16.8 ± 0.5       |
|                 | S₀.₅ or Kₘ | 0.20 ± 0.00        | 0.36 ± 0.02      |
|                 | h          | 1.24 ± 0.03        | 0.92 ± 0.09      |
| Gly             | v₀         | 5.6                | 5.2              |
|                 | vₐ max     | 42.3 ± 0.8         | 15.9 ± 0.5       |
|                 | A₀.₅       | 1.8 ± 0.1          | 2.4 ± 0.3        |
|                 | h          | 1.25 ± 0.11        | 1.04 ± 0.14      |
| Malate          | v₀         | 10.2               | 8.0              |
|                 | I₀.₅       | 0.38 ± 0.02        | 1.32 ± 0.01      |
|                 | h          | 1.21 ± 0.05        | 1.35 ± 0.03      |

aValues ± S.D. were estimated by the best fit to the appropriate equation, described in the Materials and Methods section. When fitting the data of the activation and inhibition
experiments, the value of $v_0$ was fixed at the experimentally determined values. Velocities ($v_0$, $V_{\text{max}}$, $v_{a \text{ max}}$ and $v_{i \text{ min}}$) are given as units/mg protein and $A_{0.5}$, $S_{0.5}$, $K_m$, and $I_{0.5}$ as mM. Assays were carried out at pH 7.3, 0.1 mM bicarbonate and 0.4 mM free-Mg$^{2+}$. Assays were conducted at fixed 3 mM total-PEP. $^b$Mg$^{2+}$-PEP was varied from 0.06 to 6.08 mM, free-PEP from 0.94 to 93.92 mM, and total-Mg$^{2+}$ from 0.47 to 6.49 mM. $^d$Assays were performed at fixed 0.1 mM Mg$^{2+}$-PEP; free-PEP varied from 1 to 80 mM and total-Mg$^{2+}$ from 0.33 to 0.107 mM. $^c$Assays were performed at fixed 5 mM free-PEP; Mg$^{2+}$-PEP varied from 0.05 to 5 mM, total-PEP from 5.05 to 10 mM, and total-Mg$^{2+}$ from 0.108 to 10.68 mM. $^f$Assays were performed at fixed 2.5 mM total-PEP, 5 mM total-Mg$^{2+}$ and 1mM bicarbonate. NA, not apply.
