Published in final edited form as:

Biochemistry. 2018 December 04; 57(48): 6688–6700. doi:10.1021/acs.biochem.8b00963.

Structural Control of Nonnative Ligand Binding in Engineered Mutants of Phosphoenolpyruvate Carboxykinase

Henry Y. H. Tang^{†,‡,⊗}, David S. Shin[†], Greg L. Hura^{†,§}, Yue Yang^{||}, Xiaoyu Hu[⊥], Felice C. Lightstone^{||}, Matthew D. McGee[#], Hal S. Padgett[#], Steven M. Yannone[†], and John A. Tainer^{*,†,▽}

[†]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[‡]Department of Chemistry, University of California, Berkeley, California 94720, United States

[§]Department of Biochemistry and Chemistry, University of California, Santa Cruz, California 95064, United States

^{||}Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, California 94550, United States

[⊥]Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

[#]Novici Biotech LLC, Vacaville, California 95688, United States

[▽]Department of Molecular and Cellular Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, United States

Abstract

Protein engineering to alter recognition underlying ligand binding and activity has enormous potential. Here, ligand binding for *Escherichia coli* phosphoenolpyruvate carboxykinase (PEPCK), which converts oxaloacetate into CO₂ and phosphoenolpyruvate as the first committed step in gluconeogenesis, was engineered to accommodate alternative ligands as an exemplary system with structural information. From our identification of bicarbonate binding in the PEPCK active site at the supposed CO₂ binding site, we probed binding of nonnative ligands with three oxygen atoms arranged to resemble the bicarbonate geometry. Crystal structures of PEPCK and point mutants with bound nonnative ligands thiosulfate and methanesulfonate along with strained ATP and reoriented oxaloacetate intermediates and unexpected bicarbonate were determined and analyzed. The mutations successfully altered the bound ligand position and orientation and its specificity: mutated PEPCKs bound either thiosulfate or methanesulfonate but never both. Computational calculations predicted a methanesulfonate binding mutant and revealed that release of the active site ordered solvent exerts a strong influence on ligand binding. Besides nonnative ligand binding, one mutant altered the Mn²⁺ coordination sphere: instead of the canonical octahedral ligand arrangement, the mutant in question had an only five-coordinate arrangement. From this work,

*Corresponding Author: Department of Molecular and Cellular Oncology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030. Telephone: 713-563-7725. Fax: 713-794-3270. jtainer@mdanderson.org.
⊗H.Y.H.T.: School of Biological Sciences, The University of Auckland, 3A Symonds St., Auckland 1010, New Zealand.
The authors declare no competing financial interest.

critical features of ligand binding, position, and metal ion cofactor geometry required for all downstream events can be engineered with small numbers of mutations to provide insights into fundamental underpinnings of protein–ligand recognition. Through structural and computational knowledge, the combination of designed and random mutations aids in the robust design of predetermined changes to ligand binding and activity to engineer protein function.

Graphical Abstract

Escherichia coli phosphoenolpyruvate carboxykinase (PEPCK, UniProtKB entry P22259) is a metabolic enzyme that catalyzes the first committed step of gluconeogenesis. PEPCK decarboxylates oxaloacetate (OAA) and phosphorylates it with ATP, generating phosphoenolpyruvate (PEP) and ADP.¹ The reverse reaction, in which PEP, ADP, and CO₂ react to generate OAA, happens readily *in vitro*.² The overall reaction is as follows:

Divalent metal cations are required for catalysis: Mg²⁺ with either Mn²⁺ or Ca²⁺ for optimal catalytic rates.^{3,4} The structure of PEPCK is characterized by X-ray crystallography as a bilobed enzyme with the active site located between the two lobes.⁵ As seen in crystal structures, the active site is modular: nucleotide binding is in one region, while binding of CO₂ and PEP occurs near the bottom of the cleft. Because of this modularity of the active site, modification of one region of the active site should not directly affect the others (see Figure 1a). As such, the CO₂ binding site of PEPCK is an exemplary target for mutagenesis to alter ligand specificity and understand molecular recognition.

In this work, we first analyzed an existing PEPCK structure with CO₂ bound⁶ and present a corrected interpretation of the electron density at the supposed CO₂ binding site. With the updated model of bicarbonate binding at that position, we explored the binding of other ligands containing three oxygen atoms arranged in a similar geometry. We identified a nonnative ligand, thiosulfate, which binds in the active site of wild-type (WT) PEPCK. Exploiting this observation and further building on existing crystal structures, we generated rationally designed mutants that either altered the thiosulfate orientation in the binding pocket or lost thiosulfate binding but gained methanesulfonate binding, thereby altering the ligand specificity and orientation in the active site. We combined six novel X-ray structures with biophysical and biochemical measurements to examine ligand interactions. The results

provide and test insight into the binding of these nonnative ligands in both WT and mutant PEPCK.

MATERIALS AND METHODS

Cloning and Expression.

The gene encoding PEPCK was amplified via polymerase chain reaction from genomic *E. coli* DNA. A C-terminal six-His tag was added through the reverse primer. The amplified fragment was TOPO-cloned into a pET-21b vector and verified by restriction digests and sequencing.

PEPCK point mutants were generated using gene splicing by overlap extension (SOEing) and cloned using sequence- and ligation-independent cloning (SLIC).⁷ G209S K212C and K212I F216V double mutants were isolated from shuffled DNA libraries generated using GRAMMR technology for high-resolution genetic reassortment (Novici Biotech LLC). Briefly, the GRAMMR library was generated by shuffling a set of sequence-verified plasmids containing specific point mutations into various combinations. Each clone contained unique combinations of approximately two mutations at the specified positions.

Plasmids encoding wild-type and mutant PEPCK were transformed into Rosetta 2 competent cells (Novagen) for expression and purification.

Purification.

Overnight cultures of *E. coli* Rosetta 2 containing PEPCK in a pET-21b vector were grown at 37 °C while being shaken at 250 rpm in Luria-Bertani broth supplemented with 100 µg/mL ampicillin. Overnight cultures were transferred into larger-volume cultures, grown until the OD₆₀₀ reached 0.500, induced with 1 mM isopropyl β-D-1-thiogalactopyranoside, and grown for an additional 16 h at 16 °C. Cells were then harvested by centrifugation at 8000g for 20 min at 4 °C and frozen at −20 °C until use.

Frozen pellets were resuspended in buffer A [20 mM Tris (pH 8.0) and 200 mM NaCl], and cells were lysed by sonication. Cellular debris was cleared from the lysate by centrifugation at 40000g for 30 min at 4 °C.

PEPCK was batch-bound to Ni-NTA resin (5 PRIME) by being rocked at 4 °C for 2 h (2 mL of 50:50 resin slurry per liter of culture). The resin was then washed with 10 column volumes (CV) of buffer A and then 5 CV of buffer A with 25 mM imidazole. PEPCK was eluted with 5 CV of buffer A with 250 mM imidazole.

The Ni-NTA column eluate was loaded onto an S-200 16/60 gel filtration column (GE Healthcare) pre-equilibrated with buffer A. Fractions containing the main peak were pooled and concentrated in a Vivaspin 20 concentrator with a 10K molecular weight cutoff (GE Healthcare); 50 µL aliquots were plunged into liquid nitrogen and frozen at −80 °C until they were ready for use.

Enzyme Biochemical Assays.

Biochemical activity measurements of the purified PEPCK protein variants were taken through a coupled enzymatic assay. To probe the activity in the reverse direction, the conversion of PEP to OAA, malate dehydrogenase was used as a readout enzyme. A standard curve of reduced nicotinamide adenine dinucleotide (NADH) was created at 340 nm. Next, a 100 μ L reaction mixture consisting of 100 mM NaHCO₃, 10 mM ADP, 2.5 mM MgCl₂, 2.5 mM MnCl₂, 100 mM Tris (pH 7.5), 1 unit of malate dehydrogenase, and 1 mM NADH was created; 0.5 μ g of purified PEPCK enzyme was added, except in the no protein negative control. The reactions were then started by the addition of PEP to a final concentration of 5 mM, except for for the no substrate negative control. Spectrophotometric readings at 340 nm were collected on a BMG Labtech PolarSTAR Omega UV/visible plate reader every 20 s for 5 min. Experiments were performed in triplicate. The resulting absorption versus time plots showed a linear decrease over time until a minimum was reached, usually around 3 min. The linear portion was fit with a linear regression, and the slope of the line was converted to a specific activity (micromoles per minute per milligram of enzyme).

Crystallization and Data Collection.

PEPCK crystals (unless specifically mentioned below) were grown by hanging-drop vapor diffusion. Solutions of 10–15 mg/mL protein, 20 mM Tris (pH 8.0), 200 mM NaCl, 5 mM MgCl₂, 5 mM MnCl₂, 5 mM ADP, 5 mM PEP, and 100 mM sodium thiosulfate, sodium methanesulfonate, sodium ethanesulfonate, or sodium 2-mercaptoethanesulfonate were combined with an equal amount of a precipitant containing 24% polyethylene glycol (PEG) 3350, 100 mM Bis-Tris (pH 5.5), and 400 mM NaCl and placed in a 20 °C incubator. Rodlike crystals formed in approximately 3 days.

Before data collection on the G209S K212C double mutants, crystals were derivatized with xenon to identify hydrophobic pockets. Crystals were pressurized to 200 psi in a xenon derivatizer (Hampton Research) for 2 min and flash-frozen.

Crystals of the K212I F216V mutant were initially grown under the same conditions that are described above, but only fine needles too small for diffraction experiments were observed. Crystals of the mutant were grown by hanging-drop vapor diffusion with a new well solution consisting of 0.16 M ammonium sulfate, 0.08 M sodium acetate (pH 4.8), 20% PEG 4000, and 15% (v/v) glycerol. The resulting crystals were in large plates that diffracted with acceptable statistics.

Crystallographic data for the Y207F, G209N, G209S, G209S K212C, and K212I F216V mutants were collected at the SIBYLS beamline at the Advanced Light Source in Berkeley, CA,⁸ while data for the WT–thiosulfate complex were collected at beamline 8.3.1 of the Advanced Light Source. Before data collection, crystals were cryoprotected by being allowed to sit for 1 min in the precipitant solution with 20% glycerol. The crystals diffracted to 1.79–1.12 Å, and images were recorded on an ADSC Q-315r detector (SIBYLS beamline) or Dectris Pilatus 6M detector (beamline 8.3.1). Two hundred degrees of frames were collected and used for processing for the majority of crystals. Images were processed

using XDS. The structures were determined by molecular replacement using Phaser⁹ and refined using phenix.refine, both in the PHENIX software suite.¹⁰ Fitting was performed in Coot.¹¹

The WT PEPCK enzyme was found to contain two copies per asymmetric unit. There is electron density indicative of nonnative ligand in both copies, but density in chain A was of better quality and thus was used for analysis.

Computational Prediction of Mutant–Methanesulfonate Binding and Solvent Analysis.

Thirty mutants were prepared using the MOE¹² suite of programs based on G209S–methanesulfonate crystal structures. Methanesulfonate was docked into mutants followed by single-point MM-GBSA rescoring calculations, all of which were performed using MOE. Ten top-ranked methanesulfonate binding poses for each mutant were saved as the initial structures for MD simulations, resulting in a total of 300 mutant–methanesulfonate systems. The AMBER ff99SB force field¹³ was used to model the protein system, while the parameters for modeling methanesulfonate were obtained from the generic AMBER force field (GAFF) for small molecules. The charge for methanesulfonate was derived following the restrained electro-static potential (RESP) procedure^{14,15} using Gaussian 03.¹⁶ Also, to improve the modeling of the Mn²⁺ complex, MCPB.py was used to generate the force constant parameters.¹⁷ For each mutant–methanesulfonate system, a geometry minimization was conducted prior to slowly heating the system to 300 K and relaxing it at 1 atm. After that, a 25 ns *NPT* production simulation was performed. Subsequently, for each system, MMGBSA.py was implemented to calculate the methanesulfonate binding free energy against the entire MD trajectory. For each mutant, the average binding free energy of 10 systems was calculated and ranked in order to make predictions. In addition, because the small size of the ligand presents a challenge to the free energy evaluation using docking and/or MM-GB/SA, for top-ranked mutant–methanesulfonate complexes, a conformational search was conducted using the LowModeMD function of MOE, and the interaction energy between the ligand and receptor was calculated for each obtained conformation. The interaction energy and the average binding free energy calculated using MMGBSA.py were used to cross-screen mutants to improve prediction accuracy.

To understand the impact of the ordered solvent on the ligand preference between WT PEPCK and the G209S mutant, we also implemented the solvent analysis using MOE. Potential water sites within 10 Å of the ligand site were calculated for each receptor (assuming no ligand binding yet).

RESULTS

Reinterpretation of the CO₂-Bound PEPCK Structure.

To better understand CO₂ binding in the native reaction of PEPCK, we retrieved and visually inspected the electron density maps associated with CO₂-bound PEPCK⁶ [Protein Data Bank (PDB) entry 2OLQ] (Figure 2). The deposited structure had a water molecule located adjacent to the CO₂ molecule that was hydrogen bonded to residues R65 and Y207. The distances from the modeled water to the oxygen atoms of CO₂ were 3.0 and 3.1 Å, which

are reasonable distances for hydrogen bonding. However, inspection of the electron density map revealed that the water molecule was placed outside the density and furthermore showed a positive difference peak adjacent to the placed water molecule (Figure 2a). Fitting the water molecule into that positive peak placed it 2.4 and 2.6 Å from the CO₂ oxygen atoms. The distances along with the three oxygen atoms and the single carbon atom led us to model bicarbonate into that density. Refinement using the deposited structure factors showed bicarbonate binding with 82% occupancy, with no appreciable difference density. In addition to hydrogen bonding with the R65 and Y207 residues responsible for CO₂ hydrogen bonding (2.98 and 3.01 Å, respectively), the third oxygen atom of bicarbonate hydrogen bonds to a water molecule in the Mn²⁺ coordination sphere at a distance of 3.0 Å (Figure 2b). The crystallization condition contained acetate that would also fit into the density but with a loss of one hydrogen bond.

Sulfonate Binding in PEPCK.

From our identification of a carboxylate-containing molecule binding in the active site of WT PEPCK, we reasoned that we could test the implied bicarbonate binding and possible control of nonnative ligand recognition by examining ligands resembling bicarbonate. We therefore decided to explore binding of the sulfonate-containing (–SO₃[–]) compounds thiosulfate and methanesulfonate because of their chemical stability and the similarity of this moiety with bicarbonate for both geometry and electrostatics.

An overall view of PEPCK and its active site with the structures and mutations examined in this work (Figure 1b) shows the agreement of the backbones across the various mutants indicating they retain the enzyme fold and tertiary structure.

WT–Thiosulfate Crystals.

WT PEPCK was crystallized with the nonnative ligand thiosulfate, yielding crystals (PDB entry 6AT4) whose data collection and model statistics are listed in Table 1. Importantly, this crystal form allows structural definition of the carboxylate binding pocket of PEPCK and its mutants, including nonnative bound ligands (Figure 3).

Electron density indicative of thiosulfate was identified in the carboxylate binding pocket in chain A. The $2mF_o - DF_c$ electron density from a composite omit map contoured to 2σ reveals the thiosulfate ligand (Figure 3a). The sulfonate portion of the molecule lacked symmetric density, and density surrounding one oxygen atom was larger than the others, suggesting binding flexibility or that thiosulfate in WT PEPCK binds in the carboxylate binding pocket with more than one orientation. However, upon refinement with the additional conformations, the resulting occupancies in the subsequent orientations were lower than 20%. The occupancy of the thiosulfate ligand in the carboxylate binding pocket is 68%. Hydrogen bonds between the oxygen atoms of the sulfonate (–SO₃) moiety on thiosulfate and side chains in the active site mediate the protein–ligand interaction (Figures 3a and 4a). The two residues previously reported to play a role in hydrogen bonding, R65 and Y207,⁶ also make bonds with thiosulfate, interacting with NE of R65 (2.9 Å) and OH of Y207 (2.6 Å). This can be attributed to the similar O–O distances between CO₂ and the O–S–O bond on thiosulfate: The ideal O=C=O bond length is 2.3 Å, while the O–S–O distance

for the atoms interacting with R65 and Y207 in this thiosulfate molecule is 2.5 Å. The thiosulfate oxygen atoms also hydrogen bond with the amino group of K212 mediated through a water, the carbonyl oxygen of T394 mediated through a bound water as well, the amino group of K213 (3.1 Å), and NH1 of R333 (2.9 Å) (Figure 4a). These interactions led to thiosulfate in the WT PEPCK carboxylate pocket being oriented with the sulfur atom pointing to the center of the N-terminal lobe of PEPCK.

Overall, the enzyme adopts a closed conformation that overlays well with WT PEPCK with ATP bound. This extends and questions previous studies that stated a nucleotide is required to reach the closed conformation, as no ATP was present in the crystallization condition: thiosulfate was present in the ATP binding site and appears to be sufficient for closing the system.

Crystallizing WT PEPCK with both thiosulfate and ATP along with Mn^{2+} and Mg^{2+} required for ATP binding yielded crystals that contained only thiosulfate. Thiosulfate was present at a concentration (100 mM) 20-fold greater than that of ATP (5 mM), yet it was still interesting to see thiosulfate preferred over ATP in its native binding pocket. Aligning the two models yielded a root-mean-square deviation (RMSD) of only 1.1 Å, yet there was one key noticeable difference: the cap (residues 386–406) over the active site was in the closed position when thiosulfate was bound but in the open position in the ATP/ Mg^{2+} / Mn^{2+} complex.

Collectively, these experiments provide novel data for better characterizing the carboxylate binding pocket in PEPCK. We reasoned that the ability of this pocket to accommodate a much larger tetrahedral ligand coupled with a library of mutants with slightly altered pockets also provides a starting point for engineering nonnative ligand binding and catalysis into this enzyme. We therefore set out to characterize several different point mutants bound to thiosulfate as well as to methanesulfonate to test a second nonnative ligand.

Identifying Regions To Mutate.

From the hydrogen bonding observed with thiosulfate binding in PEPCK along with carboxylate binding, we hypothesized that altering the hydrogen bonding network would impart different ligand specificity to the active site of PEPCK. We therefore mutated residues responsible for hydrogen bonding in the carboxylate binding pocket of the native enzyme: R65 and Y207. Point mutants R65H, R65Q, Y207F, Y207L, Y207I, and Y207V were made. However, crystal structures lacked thiosulfate in the carboxylate pocket (results not shown).

Next, instead of the direct alteration of the hydrogen bonding network through the existing residues, added hydrogen bonding capability was conferred through mutation of residues adjacent to R65 and Y207: a loop spanning T63–K70, which includes hydrogen bonding residue R65, and an α -helix immediately following Y207 from the G209–F216 sequence were mutated with 5–10 alternative hydrophobic amino acids in multiple positions in these regions. Several of these mutations were assayed and, through X-ray crystallography, were found to bind to nonnative ligands as described in subsequent sections.

Activity Assays of Select PEPCK Mutants.

Mutant enzymes containing a point mutation targeting the aforementioned residues were generated and characterized for native activity in the reverse direction (OAA formation). A biochemical assay of all four G209 mutants showed a loss of activity, while the impact of the Y207 mutations depended on the identity of the mutation as described below and shown in Figure 5.

From the assay results, it is evident that both the position and the identity of the mutations play roles in influencing and disrupting native activity. There is a range of effects from this small collection of point mutations, with the Y207L mutation showing minimal to no impact on native reverse activity to 3-fold reduced activity in the Y207F mutant to slight activity above the background for the remaining Y207 mutants. Finally, none of the G209 mutants showed activity above the background.

Overall, the biochemical activity data for the point mutants provide information that complements the crystal structures. Flexibility combined with hydrogen bonding and electrostatics all act in controlling ligand binding and orientation. The change in the native reaction rate due to mutation of certain residues highlights catalytically crucial regions and also suggests those that are of interest for engineering efforts.

The Y207F Mutant Shows Oxaloacetate and Thiosulfate Binding in Its Carboxylate Binding Pocket.

Crystals of Y207F PEPCK (PDB entry 6AT3) contained two copies per asymmetric unit (ASU) with statistics listed in Table 1. Density resembling oxaloacetate occurs near the carboxylate binding pocket of chain A of the PEPCK Y207F mutant (Figure 6). This structure uncovers unique information about oxaloacetate binding. Previously, there has only been one structure in the Protein Data Bank of PEPCK with OAA bound (2PXZ), which is not associated with an article. In that model, OAA is located away from the CO₂ binding site and the ATP binding site, toward the base of the Ω -loop (residues 386–406). This is inconsistent with existing knowledge because the carboxylate group being removed from OAA should be positioned in the carboxylate binding pocket. The density map was also unconvincing for this ligand. In contrast, our structure provides convincing electron density data for oxaloacetate binding in a location consistent with the biochemistry.

OAA binding was also reported in the rat cytosolic PEPCK.¹⁸ In those structures, the OAA was coordinated to the Mn²⁺ ion. In 2QF1, OAA was found in two positions around the manganese ion, while in 2QF2, the density was modeled with both OAA and pyruvate. The OAA orientation placed the portion of OAA originating from CO₂ pointed toward the Arg and Tyr that are involved in CO₂ hydrogen bonding.

In our Y207F mutant, OAA forms hydrogen bond interactions primarily with a water molecule bound to N331 and T394 (Figure 6). O4 of OAA hydrogen bonds to the bound water (3.2 Å), and O1 to NE2 of H232 (3.2 Å) and NZ of K213 (3.0 Å). The OAA is oriented so that the carboxylate moiety being removed pointed away from R65 and Y207. The OAA no longer interacts with the 207 side chain, because of the Y207F point mutation. OAA binding also involves H232, which is involved in Mn²⁺ coordination in WT PEPCK.

In this mutant, there is no evidence to suggest metal binding, even though the mutation is not directly involved in the coordination sphere and Mg^{2+} and Mn^{2+} are present in the crystallization conditions.

There is density opposite O3, which could indicate an alternate orientation of OAA where O1 and O2 would be positioned where O4 and O5 currently are located and with O3 pointing toward R65. However, the proposed alternate conformation is a small contribution: the methylene moiety of OAA is expected to have density weaker than that of the oxygen atoms, and that is reflected with the current positioning. The minor conformation of OAA places the CO_2 leaving group (O1 and O2) closer to the residues that are implicated in WT PEPCK hydrogen bonding to CO_2 : residues 65 and 207. Thus, the disruption of the hydrogen bond donor abilities from the Y207F mutation changes the orientation of the OAA substrate from that of the native enzyme.

The thiosulfate in the carboxylate binding site of Y207F chain B is not in the same position as in the WT enzyme (Figure 3). While the ligand is near R65, it is no longer interacting with the Phe side chain now at position 207, and hydrogen bonding between the ligand and the protein is through H232 and S250 as well as through a bound water molecule (Figure 4b). Thus, compensatory hydrogen bonds are made with other side chains, showing that loss of the hydrogen bond donor at residue 207 alters the position and orientation of the thiosulfate ligand.

G209N Binding to Thiosulfate.

PEPCK point mutant G209N was crystallized in the presence of thiosulfate, ADP, PEP, Mg^{2+} , and Mn^{2+} . The resulting data collection statistics are listed in Table 1 (PDB entry 6AT2).

Electron density for thiosulfate was identified in the carboxylate binding site (Figure 3c). The thiosulfate molecule in the CO_2 binding site is in a different position and orientation compared to those of thiosulfate in the WT enzyme because of the bulkier Asn side chain now at position 209: the larger side chain pushes thiosulfate so that its O2 atom interacts with ND2 of N331 (3.2 Å) (Figure 4c). The sulfur atom on thiosulfate in this mutant points toward the active site cleft and the C-terminal lobe.

ATP was found to be the nucleotide bound in the active site, even though ADP and PEP were added. This is because the transfer of phosphate from PEP to ADP is energetically favored and spontaneous over the crystallization time scale. The conformation of ATP in this mutant resembles those previously characterized as being bound to PEPCK.

The G209S Point Mutation Alters the Specificity for Methanesulfonate.

The G209S mutant was crystallized under conditions identical to those of G209N, giving rise to crystals (PDB entry 6ASI) as described in Table 1.

Density for the methanesulfonate ligand is seen in the carboxylate pocket of this model (Figure 3d). Postrefinement, the occupancy of the ligand increases to 87%. The absence of the electron-rich sulfur atom compared to thiosulfate is unambiguous. Hydrogen bonding of

the ligand once again involves R65 and the serine mutation (Figure 4d). In this mutant, the methanesulfonate methyl group points along the cleft running between the two lobes.

ATP binding was also observed in this mutant, with the molecule and residues involved in Mg^{2+}/Mn^{2+} coordination overlaying well with the WT enzyme structure.

The G209S mutant that crystallized in the presence of thiosulfate yielded crystals with an empty carboxylate binding pocket (data not shown), indicative of the ligand selectivity exhibited by this enzyme. Efforts to crystallize WT PEPCK and G209N structures with methanesulfonate bound also yielded crystals with empty CO_2 binding pockets.

The G209S K212C Mutation Reverts Selectivity to Thiosulfate.

Crystals of the G209S K212C double mutant were obtained under the same conditions that were used for the previous mutants, leading to crystals (PDB entry 6ASM) as described in Table 1. This double mutant crystallized in the presence of thiosulfate and methanesulfonate in two separate crystal trials, but only the thiosulfate condition yielded electron density in the CO_2 binding pocket (Figure 3e).

ATP binding in this mutant had one change. Surprisingly, the Mn^{2+} coordination sphere was no longer octahedral as previously observed; K213 is now pointing away from Mn^{2+} and instead hydrogen bonds to the nonnative ligand thiosulfate. The resulting manganese coordination sphere now has five members in a square pyramidal geometry.

The addition of the K212C mutation to the G209S active site reverted the ligand selectivity back to thiosulfate, even though the G209S single mutant had methanesulfonate bound. The thio S of THJ points toward ATP in the active site.

Alteration of the Mn^{2+} Coordination Sphere in the G209S K212C Mutant.

The G209S K212C mutant influenced the ATP binding region. In the G209S (Figure 7a) crystal structure and the WT PEPCK structure, Mn^{2+} has an octahedral coordination sphere consisting of an oxygen from the ATP γ -phosphate, Lys213, His232, Asp269, and two waters. In the mutant (Figure 7b), Lys213 is no longer coordinating. Instead, Lys213 is now hydrogen bonding to a bound water molecule positioned where the amino group on the Lys would be. Instead of the previously seen octahedral coordination sphere, a coordination sphere with a square pyramidal configuration was observed. Although it seems like the other oxygen on the side chain of Asp269 could complete the octahedral sphere, the 2.81 Å distance between that oxygen and the Mn^{2+} does not support coordination between those two atoms.

Computational Prediction of Binding in Other Mutants.

All mutants that bound to the nonnative ligands thiosulfate and methanesulfonate thus far have involved mutations at the G209 position. The involvement of other residues could also lead to nonnative binding. However, expanding the mutational search space to span all mutations in the three regions of interest would not be feasible experimentally. To increase throughput allowing more comprehensive consideration of mutational space, mutants were

screened *in silico* using molecular docking, molecular dynamics (MD), and molecular mechanics-generalized Born surface area (MM-GB/SA) methods.

Because of the small size of thiosulfate and methanesulfonate, using only traditional docking and/or single-point MM-GB/SA calculations cannot yield reliable predictions. To improve prediction accuracy, we implemented MD-based MM-GB/SA rescoring, which calculates the average binding free energy of each mutant–ligand complex against the entire MD simulation trajectory, thus taking into account receptor flexibility, as seen to be important for peptide ligands to proteins.¹⁹ In addition, we also implemented ligand–receptor interaction energy analysis to overcome the difficulty associated with small ligand size. Using such a cross-screening approach, we ranked mutants by their binding scores and interaction energies with the nonnative ligand methanesulfonate (Table 2). One of the top computational hits, the K212I F216V double mutant, was identified and showed crystallo-graphic evidence of methanesulfonate binding as discussed in the following section. In this case, the binding score alone would not have ranked double mutant K212I F216V high enough to suggest experimental testing. Its calculated average binding free energy is slightly less favorable than that of G209S, but its interaction energy is slightly more favorable than that of G209S. Therefore, considering both scores together would have ranked this double mutant as one of the top predictions.

Evidence for Methanesulfonate Binding in K212I F216V.

The K212I F216V double mutant predicted to bind methanesulfonate was crystallized under a condition different from that used for the other mutants; use of the Bis-Tris condition yielded fine needles unsuitable for diffraction studies. This new condition yielded plates rather than rodlike crystals resulting from the other condition. There are two molecules in the ASU of these crystals (PDB entry 6ASN) as in the WT enzyme, but the space group for this mutant is $P2_1$ (Table 1).

Although this mutant was crystallized in the presence of ADP, PEP, Mg^{2+} , and Mn^{2+} , none of these ligands or cofactors were present in the electron density. From the difference map, SO_4^{2-} was found to bind to Thr255 in the ATP binding region. This is explained in part by the change in crystallization to one that contains ammonium sulfate as a precipitant. The protein maintained an overall open position that was consistent with the WT enzyme's apo form.⁵

Because of the crystallization conditions, two possibilities for the ligand can be built into the electron density shown in Figure 3f: (1) methanesulfonate, the predicted ligand, and (2) sulfate, from the crystallization condition. By examining the electron density, B -factors of the ligand atoms, and the residual difference peaks after refinement with both possible ligands, we assign methanesulfonate as the bound ligand despite the much higher concentration of sulfate present in the condition.

Methanesulfonate bound in the K212I F216V mutant with an occupancy of 77%, slightly lower than that seen for some other mutants. Hydrogen bonding is observed for two of the three oxygen atoms on the sulfonate group, O4 to NE of R65 with a distance of 2.81 Å and O5 to OH of Y207 (2.8 Å) (Figure 4f). Such a hydrogen bonding pattern is also observed in

the computationally predicted model, which predicted distances of 2.9 Å between O4 and NE of R65 and 2.8 Å between O5 and OH of Y207. While ligand binding in all of the other aforementioned cases involved all three oxygen atoms on the sulfonate moiety, only two-thirds of the oxygen atoms are involved here. This may explain why the density is less clear: the interaction between the ligand and protein is not as strong.

DISCUSSION

CO₂ Binding in the PEPCK Active Site.

Because the first step of the PEPCK mechanism is the decarboxylation of oxaloacetate¹ and PEPCK is capable of catalyzing the reverse reaction,^{2,20} it must contain a CO₂ binding region in the active site and utilize CO₂ rather than bicarbonate in its reaction unless the decarboxylation is directly coupled to bicarbonate formation. A crystal structure with CO₂ bound was reported by Cotelesage et al.⁶ However, as we have shown, the density is indicative of bicarbonate binding rather than CO₂. These crystallographic results do not necessarily invalidate the previous knowledge of the PEPCK reaction. The equilibrium between CO₂ (aq) and H₂CO₃ (aq) occurs on a short time scale (seconds)²¹ relative to crystal growth (days), making it challenging to select for one form of carbon dioxide over the other to be bound in the crystal. Thus, another method for probing the binding of CO₂ in the PEPCK active site, such as computational modeling, is required to fully understand the mechanism of this protein. Although not definitive, our data are consistent with the enzyme active site binding to bicarbonate rather than CO₂.

We selected two nonnative ligands, thisulfate and methanesulfonate, to probe the PEPCK CO₂ binding pocket. These were chosen because of the similarity of the arrangement of oxygen atoms in the sulfonate moiety to that of bicarbonate. Through the binding of these ligands and rationally engineered mutations, we can gain a better understanding of the ligand selectivity and orientation control in this pocket.

Altering Ligand Binding through Direct and Neighboring Mutations.

From this work, we showed that design of even single-residue mutations can have drastic effects on ligand recognition in proteins. Although point mutations are generally thought to have a minimal effect on the protein structure, this work shows that they can lead to nonnative ligand recognition and binding in the mutated region and furthermore also have long-range effects on nucleotide binding in PEPCK. Specificity can thus depend upon single-residue mutations in the active site cleft. This observation has implications for evolution, medical health, and protein design for synthetic biology.

The first three mutant PEPCK structures presented in this article (Y207F, G209N, and G209S) are direct mutations, as they interact with the nonnative ligand through the mutated side chain. These were rationally designed mutants, with Y207F altering the native hydrogen bonding capability conferred through the Tyr207 side chain and G209N and G209S conferring new hydrogen bonding capability by the addition of a new side chain at a position where one did not exist previously.

One observation arising from these high-resolution data sets is that the mutations need not necessarily be directly involved in ligand binding to have a critical impact. We refer to these sites as “neighboring mutations”. The α -helix forming part of the CO₂ pocket seems to be especially prone to impact from such neighboring mutations. The first position is K212. While the G209S point mutant was determined by X-ray crystallography to bind only to methanesulfonate, the G209S K212C double mutant bound exclusively to thiosulfate. The residue 212 side chain is not involved directly in hydrogen bonding interactions with the ligand in either of these cases; position 212 is recessed from the CO₂ binding pocket. In the case of this double mutant, mutation of K212 changes the behavior of K213, which is in the CO₂ pocket: the coordination sphere of the Mn²⁺ cation containing K213 is altered, likely changing the local electrostatics of this pocket.

The second instance in which neighboring mutations affect ligand selectivity is the K212I F216V mutant. Neither of these positions has side chains that protrude into the CO₂ binding pocket, yet this mutant binds methanesulfonate. The effects of this mutation would have been hard to predict directly from the structure alone, suggesting that enzyme engineering efforts based solely on a rational design approach may not effectively cover the mutation space to determine effects of mutations; pairing such rational design efforts with a somewhat random element, such as libraries of variants, may therefore be more productive. The following sections delve into the factors behind the protein–ligand interaction in PEPCK and its mutants.

Role of Hydrogen Bond Donor Species in Ligand Binding.

Thiosulfate was found to bind in the WT PEPCK active site with at least two conformations as determined by the shape of the electron density around the ligand. Addition of a hydrogen bond donor at the G209 position restricted the possible orientations of thiosulfate to one pose. As has been seen before,^{22,23} enzyme engineering is capable of increasing ligand affinity.

Investigating the hydrogen bonding networks in these mutants further, we looked at not only the number of hydrogen bonds formed but also the disorder of the side chains involved in hydrogen bonding. In WT PEPCK, besides hydrogen bonding to the R65 and Y207 residues involved in native ligand binding, the two added residues of K213 and R333 are involved. In the case of methanesulfonate binding in the G209S mutant, the only side chains involved are R65 and the newly mutated G209S, along with water molecules in the Mn²⁺ coordination sphere. The side chains on K213 and R333 are among the largest and most flexible among the 20 naturally occurring amino acids, with 24 and 33 rotamers, respectively, so affinity could be decreased due to the disorder of those side chains. It turns out that R333 is more ordered (*B*-factor of 10–15 Å²) than the rest of the protein, but K213 is more disordered (*B*-factor of 27 Å² at the terminal amine), consistent with the greater restraints on Arg movements from the large planar guanidinium group in Arg. Ser in the G209S mutant is a much shorter side chain, and R65 is involved in hydrogen bonding of two-thirds of the oxygens of the sulfonate moiety in methanesulfonate, resulting in a very well-ordered side chain for hydrogen bonding. Along with the tightly bound waters in the coordination sphere, the positions of the substituents involved in hydrogen bonding of methanesulfonate seem to

be much more fixed. The relationship between the *B*-factors of the hydrogen bonding residues and ligand order provides a possibility of engineering increased affinity by reducing the level of side chain disorder in proteins, perhaps by mutagenesis toward shorter side chains. For example, shortening Glu or Gln side chains involved in hydrogen bonding to Asp or Asn, respectively, by removing a methylene ($-\text{CH}_2-$) group could increase ligand affinity. However, that could lead to a shift in ligand position and possibly also alter the orientation. We expect that employing *B*-factors for design considerations may be generally useful, but in such efforts, it may be important to consider the impact of crystal contacts and correct for their influence on ordering otherwise flexible regions.²⁴ In some cases, the types of side chain rigidity and flexibility seen here can be rationally exploited to increase ligand pocket size and interaction specificity²⁵ or to sculpt substrate orientation to block or promote activities.²⁶

Ligand Selectivity and Orientation.

The effects of ligand selectivity upon the change of one residue in the CO_2 binding pocket are intriguing. Experimental evidence has shown that the PEPCK mutants in this study are surprisingly selective in their ligand binding, capable of binding to either thiosulfate or methanesulfonate but never both. This experimentally validated observation provides a foundation for further investigations of this phenomenon to illuminate more general cases of specificity both for understanding recognition specificity and for protein engineering and design.

Enzymes have evolved to have high ligand selectivity; the primary reason is that binding of the wrong substrate could lead to catalysis of an undesired reaction causing a loss of energy or toxicity. In the context of the cell, that could have catastrophic consequences. With this in mind, there are many mechanisms for obtaining this selectivity. First, geometric fit (sterics) plays a key role in ligand identification. Addition of something as small as a methyl group can have profound effects on ligand affinity. However, that is not the case here. Although there is a size difference between thiosulfate and methanesulfonate, it is not a substantial difference in binding specificity, especially because we know that we are able to accommodate carboxylate and thiosulfate in the active site.

As we can largely rule out sterics in this case, a likely contributor to ligand recognition for thiosulfate versus methanesulfonate is the effect on the adjacent ordered solvent. Upon ligand binding, entropy gained via the release of weakly bound water molecules, along with any new favorable protein–ligand enthalpic interactions, may outweigh any enthalpic penalty of removing bound water molecules. Notably, clear electron density may simply show a site is more favorable than others nearby for a localized water molecule.²⁷ Correctly modeling this protein–water interface remains a challenge and source of error for both X-ray crystallography²⁸ and X-ray scattering²⁹ structural models but is important for better understanding ligand binding as suggested here.

In WT PEPCK, thiosulfate binding displaced four loosely bound water molecules, and the space occupied by one of its sulfur atom possibly squeezed the space for another loosely bound water molecule (WAT_A in Figure 8). The orientation of thiosulfate allows a potential hydrogen bond between one of its oxygen atoms and the hydroxyl group of Tyr207. In the

G209S mutant, however, the Tyr207 hydroxyl group points in an opposite direction, leaving a loosely bound water molecule (WAT_B in Figure 8) in the pocket. In addition, the Ser209 hydroxyl group not only inhibits binding of water molecule WAT_A present in the WT active site but also favors the placement of a potential hydrogen bond acceptor. As a result of bound solvent changes upon methanesulfonate binding, the sulfonate group moves closer to Ser209, displacing two loosely bound water molecules and leaving a potential hydrogen bond acceptor, while the hydrophobic methyl moiety displaced WAT_B along with another loosely bound water molecule. The oxygen atoms of methanesulfonate displace two other loosely bound water molecules and form hydrogen bonds with the hydroxyl group of Tyr207 and two water molecules identified in the crystal structure.

In addition to altering ligand selectivity, the mutations introduced in these experiments also impact ligand orientation and position. In some cases, it is apparent why the mutation is leading to these changes. For example, via comparison of ligand binding in the G209N and G209S mutants, the thiosulfate ligand in the G209N structure is hydrogen bonding to the Asn side chain, while methanesulfonate in the G209S structure interacts with the Ser side chain. Asn is bulkier than Ser, so the side chain is already occupying the space that was occupied by the ligand as seen in WT PEPCK and the G209S mutant. This in turn pushes the ligand out further to minimize steric repulsion. At this new position, thiosulfate is capable of hydrogen bonding to N331, a residue that was not involved in ligand binding in WT PEPCK. We explored this further by generating the G209S N331Q mutation to see if pushing out the carboxamide functionality of residue 331 through the introduction of a methylene spacer would act as a compensatory mutation of sorts to bring about ligand binding interactions that more closely resemble the binding mode seen in the G209N mutant. Unfortunately, crystals of the G209S N331Q double mutant did not show binding of thiosulfate or methanesulfonate (data not shown), illustrating that this complex system can behave in nonintuitive ways.

Analyzing the orientation of thiosulfate across the various mutants showed that the thio S atom of thiosulfate (S-SO_3^{2-}) had a tendency to point toward positively charged groups. This effect likely reflects the electrostatic attraction between the weak negative charge of the aforementioned sulfur atom and the positive charge of various side chains. The partial negative nature of the sulfur can be justified through resonance structures of the thiosulfate anion as there are resonance structures that place the negative charge on the sulfur. This trend can be seen in the crystal structures of the mutants containing thiosulfate: the anions tend to point toward arginine most often and sometimes lysine. This phenomenon is not observed with methanesulfonate as the methyl group of that anion lacks negative charge. In such electrostatic interactions, individual charges can have a 10-fold effect with consequences for both interaction kinetics and specificity.^{26,30}

Finally, ligand selectivity in the K212I F216V double mutant is surprising because neither of these mutations was expected to be able to hydrogen bond to the ligand, and indeed, that is the case in the crystal structure. From this mutation, we can infer that electrostatic effects neighboring the binding site have a strong role to play in determining ligand selectivity in this system and likely many others, as well. These data thus suggest that further investigation into the roles of electrostatics in protein–ligand interaction will prove to be useful for future

protein engineering experiments. This is furthermore one area that could greatly benefit from added computational input.

Altered Mn²⁺ Coordination Sphere in the G209S K212C Mutant.

Mn²⁺ ions in enzymatic systems often are in an octahedral coordination sphere around the metal ion.^{31,32} The octahedral coordination sphere is also what has been observed in the PEPCK system with one exception, the G209S K212C double mutant. In this mutant, a lysine that is normally in the octahedral coordination sphere points away from the metal, resulting in a coordination sphere that is one ligand short.

Energetics of ATP Binding versus Thiosulfate at the ATP Position.

Nucleotide binding in PEPCK is a key event leading to the closure of the enzyme's active site. As such, one would presume that it is a relatively robust feature in the system. Thus, the ATP affinity would be substantially higher than those of smaller ligands, because more interactions usually result in a higher binding affinity, yet upon cocrystallization of the WT enzyme in the presence of thiosulfate and ATP along with the divalent cations required for nucleotide binding, we found thiosulfate bound in the ATP binding site. When thiosulfate is bound, the divalent cations in turn are not found in the active site. Normally, several oxygen atoms on the β - and γ -phosphates of ATP complete the octahedral coordination spheres of the Mg²⁺ and Mn²⁺ cations. As crystallography is known to capture stable, low-energy states, this result indicates that the thiosulfate-bound structure with free ATP with metal ions in solution must be a state with an energy lower than that of the ATP-bound structure with thiosulfate in solution. This can be attributed to the minimization of the Gibbs free energy of the system: although the enthalpy from the ATP-protein interactions is lost, the entropy gain from the free nucleotide and metals must be greater than the bond energies. This is most likely attributed to the additional degrees of rotation that can be freely attained by free ATP in solution. However, it is furthermore consistent with the possible strain we observe in the ATP-bound enzyme structure.

As seen by the results presented here, the high-resolution ATP-bound PEPCK crystal structure shows that ATP binds in the *syn* conformation. As this conformation with a nucleoside oriented over the ribose sugar is less stable and higher in energy, it negates some of the enthalpic gain from ATP binding. Upon ATP binding, π -stacking is also observed between the aromatic system of adenine and the side chain of R449; the absence of this binding means this favorable interaction is also lost. Unexpectedly, ATP binding in the presence of thiosulfate was restored for mutants. In all of the mutants described in the Results except for the K212I F216V mutant, ATP/Mg²⁺/Mn²⁺ binding was observed, yet thiosulfate was present under the crystallization conditions at the same concentration as under the WT enzyme crystallization condition. The mutations made were not directly in the ATP binding site, because the purpose was to engineer novel binding in the CO₂ pocket. These observations underscore the relevance of the neighboring mutation phenomenon observed in the CO₂ pocket for the accurate prediction of mutational outcomes.

CONCLUSION

Overall, this work examined the effects of engineered mutations in the active site of *E. coli* phosphoenolpyruvate carboxykinase and showed structurally how these mutations led to nonnative ligand binding. Notably, targeted point mutations altered the species, position, and orientation of the ligands. Our results uncover the key role of hydrogen bonding networks in PEPCK ligand recognition and positioning and the substantial impact neighboring mutations can have on these phenomena. Altogether, our results establish the power of re-engineering ligand specificity by targeted mutations given an initial starting structure.

ACKNOWLEDGMENTS

The authors thank S. G. Remesh for assistance with enzymatic assays, S. Classen for support and assistance with crystallographic data collection and refinement expertise, K. Burnett for aiding protein expression, and S. Classen and T. Ogorzalek for feedback on the manuscript. Y.Y. and X.H. thank Dr. Xiaohua Zhang and Dr. Sergio Wong for helpful suggestions. Y.Y., X.H., and F.C.L. thank Livermore Computing for computer time.

Funding

This work was conducted at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy (DOE), Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by the DOE Office of Biological and Environmental Research. Added support came from KAUST and High-End Instrumentation Grant S10OD018483. J.A.T. is supported by National Institutes of Health Grant R35CA22043, a Robert A. Welch Chemistry Chair, and the Cancer Prevention and Research Institute of Texas. This work was supported by a Department of Energy ARPA-E REMOTE grant. Part of the computational work was performed under the auspices of the DOE by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Release number LLNL-JRNL-735122.

REFERENCES

- (1). Matte A, Tari LW, Goldie H, and Delbaere LTJ (1997) Structure and Mechanism of Phosphoenolpyruvate Carboxykinase. *J. Biol. Chem* 272, 8105–8108. [PubMed: 9139042]
- (2). Utter MF, and Kurahashi K (1954) Purification of Oxalacetic Carboxylase from Chicken Liver. *J. Biol. Chem* 207, 787–802. [PubMed: 13163066]
- (3). Sudom A, Walters R, Pastushok L, Goldie D, Prasad L, Delbaere LTJ, and Goldie H (2003) Mechanisms of activation of phosphoenolpyruvate carboxykinase from *Escherichia coli* by Ca²⁺ and of desensitization by trypsin. *J. Bacteriol* 185, 4233–42. [PubMed: 12837799]
- (4). Delbaere LTJ, Sudom AM, Prasad L, Leduc Y, and Goldie H (2004) Structure/function studies of phosphoryl transfer by phosphoenolpyruvate carboxykinase. *Biochim. Biophys. Acta, Proteins Proteomics* 1697, 271–8.
- (5). Matte A, Goldie H, Sweet RM, and Delbaere LT (1996) Crystal structure of *Escherichia coli* phosphoenolpyruvate carboxykinase: a new structural family with the P-loop nucleoside triphosphate hydrolase fold. *J. Mol. Biol* 256, 126–143. [PubMed: 8609605]
- (6). Cotelesage JJH, Puttick J, Goldie H, Rajabi B, Novakovski B, and Delbaere LTJ (2007) How does an enzyme recognize CO₂? *Int. J. Biochem. Cell Biol.* 39, 1204–10. [PubMed: 17475535]
- (7). Jeong JY, Yim HS, Ryu JY, Lee HS, Lee JH, Seen DS, and Kang SG (2012) One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies. *Appl. Environ. Microbiol* 78, 5440–5443. [PubMed: 22610439]
- (8). Classen S, Hura GL, Holton JM, Rambo RP, Rodic I, McGuire PJ, Dyer K, Hammel M, Meigs G, Frankel KA, and Tainer JA (2013) Implementation and performance of SIBYLS: A dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. *J. Appl. Crystallogr* 46, 1–13. [PubMed: 23396808]
- (9). McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ (2007) Phaser crystallographic software. *J. Appl. Crystallogr* 40, 658–674. [PubMed: 19461840]

- (10). Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 66, 213–221. [PubMed: 20124702]
- (11). Emsley P, Lohkamp B, Scott WG, and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 66, 486–501. [PubMed: 20383002]
- (12). Molecular Operating Environment (MOE), version 2013.08 (2013) Chemical Computing Group Inc., Montreal.
- (13). Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, and Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* 65, 712–725. [PubMed: 16981200]
- (14). Bayly CI, Cieplak P, Cornell W, and Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem* 97, 10269–10280.
- (15). Cieplak P, Cornell WD, Bayly C, and Kollman PA (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem* 16, 1357–1377.
- (16). Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr., Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, and Pople JA (2004) Gaussian 03, revision C.02, Gaussian, Inc., Wallingford, CT.
- (17). Li P, and Merz KM (2016) MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model* 56, 599–604. [PubMed: 26913476]
- (18). Sullivan SM, and Holyoak T (2007) Structures of rat cytosolic PEPCK: Insight into the mechanism of phosphorylation and decarboxylation of oxaloacetic acid. *Biochemistry* 46, 10078–10088. [PubMed: 17685635]
- (19). Schnecke V, Swanson CA, Getzoff ED, Tainer JA, and Kuhn LA (1998) Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. *Proteins: Struct., Funct., Genet.* 33, 74–87. [PubMed: 9741846]
- (20). Tortora P, Hanozet GM, and Gueritore A (1985) Purification of Phosphoenolpyruvate Carboxykinase cerevisiae and Its Use for Bicarbonate from Saccharomyces Assay. *Anal. Biochem* 144, 179–185. [PubMed: 3885783]
- (21). Khalifah RG (1971) The Carbon Dioxide Hydration Activity of Carbonic Anhydrase. *J. Biol. Chem* 246, 2561–2573. [PubMed: 4994926]
- (22). Posner BA, Li L, Bethell R, Tsuji T, and Benkovic SJ (1996) Engineering specificity for folate into dihydrofolate reductase from Escherichia coli. *Biochemistry* 35, 1653–1663. [PubMed: 8634297]
- (23). Röthlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff E, Zanghellini a, Dym A, Albeck S, Houk KN, Tawfik DS, and Baker D (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453, 190–195. [PubMed: 18354394]
- (24). Tainer JA, Getzoff ED, Alexander H, Houghten RA, Olson AJ, Lerner RA, and Hendrickson WA (1984) The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature* 312, 127–134. [PubMed: 6209578]
- (25). Garcin ED, Arvai AS, Rosenfeld RJ, Kroeger MD, Crane BR, Andersson G, Andrews G, Hamley PJ, Mallinder PR, Nicholls DJ, St-Gallay SA, Tinker AC, Gensmantel NP, Mete A, Cheshire DR,

Connolly S, Stuehr DJ, Åberg A, Wallace AV, Tainer JA, and Getzoff ED (2008) Anchored plasticity opens doors for selective inhibitor design in nitric oxide synthase. *Nat. Chem. Biol* 4, 700–707. [PubMed: 18849972]

- (26). Tsutakawa SE, Thompson MJ, Arvai AS, Neil AJ, Shaw SJ, Algasaier SI, Kim JC, Finger LD, Jardine E, Gotham VJB, Sarker AH, Her MZ, Rashid F, Hamdan SM, Mirkin SM, Grasby JA, and Tainer JA (2017) Phosphate steering by Flap Endonuclease 1 promotes 5'-flap specificity and incision to prevent genome instability. *Nat. Commun* 8, 15855. [PubMed: 28653660]
- (27). Kuhn LA, Siani MA, Pique ME, Fisher CL, Getzoff ED, and Tainer JA (1992) The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. *J. Mol. Biol* 228, 13–22. [PubMed: 1447777]
- (28). Holton JM, Classen S, Frankel KA, and Tainer JA (2014) The R-factor gap in macromolecular crystallography: an untapped potential for insights on accurate structures. *FEBS J.* 281, 4046–4060. [PubMed: 25040949]
- (29). Rambo RP, and Tainer JA (2013) Super-Resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology. *Annu. Rev. Biophys* 42, 415–441. [PubMed: 23495971]
- (30). Getzoff ED, Tainer JA, Weiner PK, Kollman PA, Richardson JS, and Richardson DC (1983) Electrostatic recognition between superoxide and copper, zinc superoxide dismutase. *Nature* 306, 287–290. [PubMed: 6646211]
- (31). Castagnetto JM, Hennessy SW, Roberts VA, Getzoff ED, Tainer JA, and Pique ME (2002) MDB: the Metalloprotein Database and Browser at The Scripps Research Institute. *Nucleic Acids Res.* 30, 379–382. [PubMed: 11752342]
- (32). Cvetkovic A, Menon AL, Thorgersen MP, Scott JW, Poole FL II, Jenney FE Jr., Lancaster WA, Praissman JL, Shanmukh S, Vaccaro BJ, Trauger SA, Kalisiak E, Apon JV, Siuzdak G, Yannone SM, Tainer JA, and Adams MWW (2010) Microbial metalloproteomes are largely uncharacterized. *Nature* 466, 779–782. [PubMed: 20639861]
- (33). Laskowski RA, and Swindells MB (2011) LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model* 51, 2778–2786. [PubMed: 21919503]

Figure 1.

Overview of the PEPCK structure highlighting mutation sites. (a) View of the PEPCK active site cleft relative to the entire protein. The cyan surface represents the ATP binding site, while the magenta surface outlines the CO₂ binding site. ATP (green sticks), Mg²⁺ (green sphere), and Mn²⁺ (purple-gray sphere), along with the nonnative ligands thiosulfate and methanesulfonate from the mutants in this work, are overlaid. (b) View of the PEPCK active site cleft with the mutations discussed in this work explicitly shown. There is good

agreement of the backbones across the various mutants, showing that the mutations have not disrupted the overall tertiary structure. ATP is shown for reference.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

Electron density in the CO₂-bound structure of PEPCK that reflects bicarbonate rather than CO₂. The $2mF_o - DF_c$ maps around the ligand are shown contoured at 1.0σ (blue), while the $mF_o - DF_c$ difference maps are contoured at 3.0σ , with positive peaks colored green and negative peaks colored red. (a) Analysis of the maps calculated for the previously deposited model and structure factors for PDB entry 2OLQ show positive density adjacent to the CO₂ molecule and a water molecule lying outside of the density. (b) Refinement of the structure with bicarbonate modeled into the density shows a better fit with no extraneous difference density peaks and a better explanation of the experimental data.

Figure 3.

View of the CO₂ binding pocket of PEPCK and its mutants with a nonnative ligand bound.

The $2mF_o - DF_c$ composite omit maps, colored blue and contoured at 2.0σ (thiosulfate, panels a–c and f) or 1.2σ (panels d and e), surround the nonnative ligand in the pocket.

Yellow dashed lines represent hydrogen bonds. Red spheres indicate ordered water molecules. (a) CO₂ binding pocket of WT PEPCK with thiosulfate bound. Hydrogen bonding interactions involve the two side chains involved in native CO₂ binding, R65 and Y207, as well as R333 and K212 mediated through a bound water. (b) The PEPCK Y207F mutant binds thiosulfate in a position similar to that of WT PEPCK but interacts with different side chains. Hydrogen bonding is accomplished through R65, H232, and S250, possibly because of the lack of hydrogen bonding capability on F207. (c) Thiosulfate binding in the PEPCK G209N mutant. N209 and N331 are also involved in hydrogen bonding this time. The orientation and position of the ligand have changed relative to those of WT PEPCK. (d) Methanesulfonate binding in the G209S mutant, viewed from a slightly rotated angle for the sake of clarity. This position is similar to that of the WT enzyme, but the orientation has changed. R65 is involved in hydrogen bonding while Y207 is not in this

case. (e) Thiosulfate binding in the G209S K212C double mutant. The addition of the mutation at K212 changed the ligand specificity in the pocket from methanesulfonate to thiosulfate. Hydrogen bonding of thiosulfate to S209 and K213 alters the orientation of the ligand relative to WT PEPCK and also G209S. (f) Methanesulfonate binding in the K212I F216V double mutant. The two mutations are to hydrophobic residues away from the immediate binding site. Only two-thirds of the oxygens on methanesulfonate are hydrogen bonded to the protein.

Figure 4.

Hydrogen bonding networks in the PEPCK CO₂ binding pockets with a nonnative ligand bound shown in two-dimensional representations. Green dashed lines represent hydrogen bonding, while blue spheres represent bound waters. Distances are reported in angstroms. (a) WT PEPCK with thiosulfate bound. (b) Y207F mutant with thiosulfate bound. (c) G209N mutant with thiosulfate bound. There is direct interaction between the ligand and the newly added side chain N209. (d) G209S mutant with methanesulfonate bound. (e) G209S K212C mutant that binds to thiosulfate. Mutation at position 212 reverted ligand specificity to thiosulfate. (f) K212I F216V double mutant with methanesulfonate bound. Figures were generated through LigPlot+. ³³

Figure 5.

Select PEPCK point mutants show a range of activities when assayed in the direction of OAA formation. The activities of selected mutants at positions Y207 and G209 are compared to WT activity. All four G209 mutants assayed show no discernible activity above background, while the effects of the Y207 mutations depend on the identity of the mutation. The Y207L mutant showed activity comparable to that of the WT enzyme; Y207F had one-third the activity of the WT enzyme, and Y207I, Y207L, and Y207N mutants showed slight but detectable activity above background.

Figure 6.

Oxaloacetate binding in the Y207F mutant. Interactions between the bound ligand and side chains are illustrated with yellow dashes. The $2mF_o - DF_c$ composite omit map, contoured at 1.5σ , is colored blue. Red spheres indicate ordered water molecules. Oxaloacetate is positioned so the carboxyl group that is decarboxylated is positioned between the two residues involved in hydrogen bonding, R65 and Y207. In this mutation, there is no interaction between the side chain at position 207 and oxaloacetate. Labels on oxaloacetate show the atom names for the various oxygen atoms on the molecule.

Figure 7.

Manganese(II) coordination sphere in PEPCK that is altered by mutations. Coordination spheres for manganese(II) in PEPCK are shown along with electron density for the ligands (blue, contoured at 2.0σ , $2mF_o - DF_c$ map) and the manganese ion (orange, contoured at 12.0σ , $2mF_o - DF_c$ map). (a) G209S Mn²⁺ octahedral coordination sphere, identical to previously determined structures of WT PEPCK. The ligands include three side chains, Lys213, His232, and Asp269, along with the γ -phosphate of ATP, and two bound waters. (b) Mn²⁺ coordination sphere in the G209S K212C double mutant that shows one major difference from WT enzyme behavior. K213 is no longer coordinated to Mn²⁺, leaving a five-coordinate sphere.

Figure 8. Computational analysis of bound water molecules in (A) thiosulfate–WT and (B) methanesulfonate–G209S complexes. Residues Gly209 and Ser209 are colored red in panels a and b, respectively, while residue Tyr207 is colored blue in each panel. The less thermodynamically stable predicted water sites are shown as red spheres, and the number on the sphere represents the potential affinity gained upon displacement of the labeled water, expected to result in an entropic gain from water release and equal or improved enthalpic ligand energy. The higher the number, the more red the sphere and the more energy gained by releasing the water. The transparent surface surrounding thiosulfate represents the space filling of the ligand, and one of the sulfur atoms is squeezing WAT_A (8.702). The oxygen atom close to Tyr207 displaced another marginally stable water molecule (4.765) and forms a hydrogen bond with the hydroxyl group of Tyr207. Upon introduction of the G209S mutation, the increased size of the side chain of residue 209 eliminated WAT_A, and the hydroxyl group of residue 207 points in an opposite direction, resulting in a less favorable energy for WAT_B (10.100). The methanesulfonate methyl group displaces WAT_B.

Table 1.

X-ray Diffraction Data Collection and Refinement Statistics^a

	PEPCK WT-thiosulfate	Y207F-thiosulfate	G209N-thiosulfate	G209S-methanesulfonate	G209S K212C-thiosulfate	K2121 F216V-methanesulfonate
resolution range (Å)	78.19–1.33 (1.38–1.33)	47.33–1.46 (1.507–1.455)	46.27–1.44 (1.50–1.44)	46.66–1.79 (1.85–1.79)	47.04–1.55 (1.61–1.55)	47.30–1.55 (1.61–1.55)
space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>C</i> 12 ₁	<i>C</i> 12 ₁	<i>C</i> 12 ₁	<i>P</i> 12 ₁ 1
unit cell dimensions	94.72 Å, 103.61 Å, 119.16 Å, 90°, 90°, 90°	94.67 Å, 105.86 Å, 120.49 Å, 90°, 90°, 90°	125.17 Å, 94.36 Å, 46.56 Å, 90°, 96.40°, 90°	125.31 Å, 93.31 Å, 46.53 Å, 90°, 96.23°, 90°	125.71 Å, 94.08 Å, 46.60 Å, 90°, 96.40°, 90°	60.84 Å, 75.61 Å, 71.53 Å, 90°, 94.70°, 90°
total no. of reflections	1645290 (56741)	1607635 (146925)	694878 (65420)	203400 (19707)	294030 (15661)	632783 (32215)
no. of unique reflections	243952 (14564)	208965 (20200)	92651 (8788)	48917 (4789)	76017 (6235)	85603 (5170)
multiplicity	6.7 (3.9)	7.7 (7.3)	7.5 (7.4)	4.2 (4.1)	3.9 (2.5)	7.4 (6.2)
completeness	0.92 (0.55)	0.99 (0.95)	0.96 (0.92)	0.98 (0.96)	0.98 (0.80)	0.91 (0.55)
mean <i>I</i> (σ)	19.15 (1.18)	11.88 (0.90)	10.55 (1.30)	25.19 (4.68)	13.12 (0.83)	12.75 (1.04)
Wilson <i>B</i> -factor (Å ²)	16	21	16	17	17	18
<i>R</i> _{merge}	0.05 (0.99)	0.08 (1.72)	0.126 (1.64)	0.044 (0.304)	0.084 (1.17)	0.113 (1.67)
<i>R</i> _{meas}	0.06 (1.14)	0.091 (1.849)	0.136 (1.762)	0.051 (0.349)	0.097 (1.477)	0.121 (1.815)
CC _{1/2}	0.999 (0.462)	0.999 (0.457)	0.998 (0.644)	0.999 (0.937)	0.998 (0.319)	0.998 (0.333)
CC*	1 (0.795)	1 (0.792)	0.999 (0.885)	1 (0.984)	1 (0.695)	1 (0.707)
no. of reflections used in refinement	243938 (14562)	208588 (19857)	92630 (8778)	48915 (4788)	75999 (6223)	85582 (5154)
no. of reflections used for <i>R</i> _{free}	1999 (119)	1985 (186)	1999 (190)	1949 (191)	1995 (163)	2007 (121)
<i>R</i> _{work} (%)	13.4 (30.5)	19.9 (41.1)	17.1 (34.8)	16.7 (22.8)	17.0 (35.7)	17.4 (34.9)
<i>R</i> _{free} (%)	15.0 (32.7)	21.9 (40.2)	19.1 (33.0)	18.6 (25.5)	18.8 (36.9)	19.3 (33.5)
CC(work)	0.966 (0.755)	0.964 (0.686)	0.970 (0.818)	0.960 (0.918)	0.968 (0.621)	0.966 (0.636)
CC(free)	0.966 (0.661)	0.962 (0.685)	0.966 (0.773)	0.956 (0.908)	0.968 (0.524)	0.952 (0.545)
no. of non-hydrogen atoms	9221	9085	4470	4291	4566	4373
macromolecules	8270	8162	4082	4041	4056	4089
ligands	25	29	46	38	53	20
protein residues	1069	1069	535	529	527	527
RMSD for bonds (Å)	0.005	0.007	0.008	0.005	0.010	0.009
RMSD for angles (deg)	0.80	0.91	1.03	0.82	1.08	0.97

	PEPCK WT-thiosulfate	Y207F-thiosulfate	G209N-thiosulfate	G209S-methanesulfonate	G209S K212C-thiosulfate	K212I F216Y-methanesulfonate
Ramachandran favored (%)	97.75	97.46	97.93	97.90	98.5	98.09
Ramachandran allowed (%)	2.2	2.25	1.88	1.91	1.34	1.72
Ramachandran outliers (%)	0.09	0.28	0.19	0.19	0.19	0.19
rotamer outliers (%)	0.46	0.59	0.72	0.49	0.24	0.7
Clashscore	1.96	0.87	1.24	1.64	2.12	1.49
average <i>B</i> -factor (Å ²)	21	29	20	21	23	25
macromolecules	20	28	19	21	22	25
ligands	21	33	17	18	21	35
solvent	33	36	25	22	32	28

^a Statistics for the highest-resolution shell are shown in parentheses.

Table 2.Top-Ranked Mutants in Terms of Methanesulfonate Binding Energies^a

system	average binding free energy (kcal/mol)	interaction energy (kcal/mol)
G209S	−13.3	−147.6
G209S K212C	−14.9	−133.2
H67Q	−14.1	−132.4
K212V F216S	−13.6	−133.3
K212I F216V	−12.9	−154.4
H67E G209S	−12.2	−143.1

^aG209S is listed as a reference. The other systems were sorted on the basis of their average binding free energies. K212I F216V is one of three mutants that has both average binding free energy and interaction energy predicted to be comparable with those of G209S and has crystallographically validated binding.