UCLA
UCLA Previously Published Works

Title

A PII-Like Protein Regulated by Bicarbonate: Structural and Biochemical Studies of the Carboxysome-Associated CPII Protein.

Permalink

https://escholarship.org/uc/item/4t4379qb

Journal

Journal of Molecular Biology, 428(20)

Authors

Wheatley, Nicole

Eden, Kevin

Ngo, Joanna

et al.

Publication Date

2016-10-09

DOI

10.1016/j.jmb.2016.07.015

Peer reviewed

A PII-like protein regulated by bicarbonate: structural and biochemical studies of the carboxysome-associated CPII protein

Nicole M. Wheatley[1], Kevin D. Eden[2], Joanna Ngo[2], Justin S. Rosinski[2], Michael R. Sawaya[1], Duilio Cascio[1], Michael Collazo[1], Hamidreza Hoveida[2], Wayne L. Hubbell[2,3], and Todd O. Yeates[1,2]

[1]UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095

[2]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095

[3]Jules Stein Eye Institute, University of California, Los Angeles, CA 90095

Abstract

Autotrophic bacteria rely on various mechanisms to increase intracellular concentrations of inorganic forms of carbon (i.e. bicarbonate and CO_2) in order to improve the efficiency with which they can be converted to organic forms. Transmembrane bicarbonate transporters and carboxysomes play key roles in accumulating bicarbonate and carbon dioxide, but other regulatory elements of carbon concentration mechanisms in bacteria are less understood. In this study, after analyzing the genomic regions around α-type carboxysome operons, we characterize a protein that is conserved across these operons but which has not been previously studied. On the basis of a series of apo- and ligand-bound crystal structures and supporting biochemical data, we show that this protein, which we refer to as the carboxysome-associated PII protein (CPII), represents a new and distinct subfamily within the broad superfamily of previously studied PII regulatory proteins, which are generally involved in regulating nitrogen metabolism in bacteria. CPII undergoes dramatic conformational changes in response to ADP binding, and the affinity for nucleotide binding is strongly enhanced by the presence of bicarbonate. CPII therefore appears to be a unique type of PII protein that senses bicarbonate availability, consistent with its apparent genomic association with the carboxysome and its constituents.

Graphical abstract
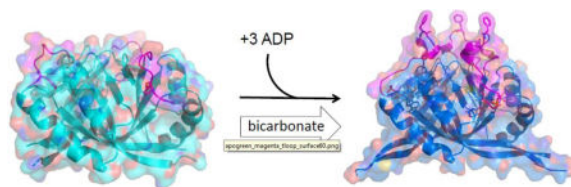
Corresponding author: Todd O. Yeates, 611 Charles Young Dr. East, Los Angeles, CA 90095-1569, (tel) 310-206-4866, yeates@mbi.ucla.edu.

Keywords

nitrogen regulatory PII proteins; bicarbonate; allostery; carbon concentrating mechanism; carboxysome; Rubisco; nucleotide-binding; CPII; DUF2309; NADH oxidoreductase subunit 5

INTRODUCTION

The enzyme RuBisCO performs a key reaction in global carbon fixation, attaching CO_2 to the five-carbon ribulose bisphosphate to give two molecules of the three-carbon 3-phosphoglycerate.[1–5] Owing to the limited catalytic rate and selectivity of RuBisCO for CO_2,[6–9] microorganisms that use RuBisCO to fix CO_2 have evolved sophisticated carbon concentrating mechanisms (CCMs) to improve the enzyme's catalytic efficiency.[8,10–15]

Diverse bacterial CCMs include elements that operate in two stages. In the first stage, bicarbonate is accumulated at high concentration in the cytosol.[15–17] This can occur by inward pumping of bicarbonate[15] (bicarbonate transport in symport with Na^+ ions[16,18,19] or by coupling to ATP hydrolysis[20,21]) or by electron-transport-coupled[22–24] hydration of CO_2 from outside the cell to bicarbonate inside the cell. The negative charge carried by bicarbonate prevents its diffusion back across the membrane bilayer. The second stage, wherein inorganic carbon is fixed by chemical reactions into organic form, occurs within the carboxysome. Carboxysomes, graphically depicted in Figure 1A, are roughly icosahedral, intracellular microcompartments that co-encapsulate the two enzymes RuBisCO and carbonic anhydrase.[25–28] The proteinaceous shell of the carboxysome is made up of a few thousand copies of a few homologous but distinct proteins from the BMC (bacterial microcompartment) family of proteins.[29–34] BMC shell proteins assemble into hexameric units, each of which bears a central pore for molecular transport[30,33–35], and these hexamers further assemble edge-to-edge to make a complete carboxysome shell. A high concentration of cytoplasmic bicarbonate drives its net diffusion through the shell protein pores and into the carboxysome, where it is dehydrated to carbon dioxide and fixed into organic form by RuBisCO before it can escape. Over the last several years, many of the components that operate directly in the context of bacterial CCMs have been characterized.[36–41] These include the structural and functional characterization of most of the constituents of the carboxysome. On the other hand, a higher-level understanding of how the carboxysome and other elements of the CCM are regulated and integrated into the cellular metabolism of bacteria is less clear.

Hints about more nuanced aspects of carboxysome and CCM function have come from analyzing the genomes of carboxysome-containing bacteria and the genomic regions around their carboxysome operons.[36,42,43] Interspersed among genes known to be more or less directly associated with carboxysome function, other conserved genes appear, which are surmised to be functionally linked to the carboxysome or CCMs more generally. Two types of carboxysomes (α and β) exist in nature; their distribution falls largely along phylogenetic lines, with the α-type occurring in chemoautotrophs and some (α-type) cyanobacteria and the β-type occurring exclusively in β-type cyanobacteria[44,45]. Across the known α-carboxysome operons and surrounding genomic regions, a few genes coding for proteins of

unknown function have been noted.[36,42] One of these, CbbQ with ATPase activity, was characterized in a recent study[46]. Another uncharacterized protein is the focus of the present work. Proteins from this family (about 110 amino acids in length) could be classified by sequence comparison algorithms as being remote members of the PII superfamily, a group of diverse regulatory proteins, among which a number are known to regulate nitrogen metabolism.

Nitrogen regulatory PII proteins (which we refer to here as nrPIIs) are a ubiquitous class of nucleotide-binding signal transduction molecules capable of inhibiting or activating target protein activities, often through the allosteric influence of α-ketoglutarate. Alpha-ketoglutarate is a central metabolite whose inverse relationship with the concentration of primary nitrogen metabolites (e.g. ammonia) makes it a reliable signal for nitrogen limiting conditions.[47–50] Numerous X-ray crystallographic studies have illuminated the structural mechanisms by which differential ATP- and ADP-binding promotes binding and non-binding conformational states in nrPII proteins.[51–53] NrPIIs and many PII-like proteins bind to ATP, ADP and α-ketoglutarate[54,55], though some exceptions have been reported; a PII-like protein from *Haloferax* binds additionally to AMP[56], and a *Bacillus* PII-like protein binds cyclic-di-AMP and not ATP or ADP[57]. Alpha-ketoglutarate exerts its allosteric regulatory effects via interaction with Mg^{2+}-nucleotide bound nrPII, thereby influencing its conformation and the interactions it makes with other proteins. NrPII proteins are known to modulate the activities of a variety of primary nitrogen metabolic transcription factors, enzymes and transporters. Examples of established nrPII-regulated activities include glutamine synthesis[58,59], ammonium transport[60] and nitrogen fixation[61,62]. A study in *Synechocystis* PCC 6803 also suggested a connection between a nrPII protein and nitrate transport; phosphorylation of nrPII was affected by growth on nitrate[63]. NrPII proteins are implicated in the regulation of primary carbon metabolism, as well. In land plants and bacteria, a central carbon metabolic enzyme acetyl-CoA carboxylase is directly regulated by nrPIIs[64,65]. In cyanobacteria, nrPIIs are thought to regulate nitrogen assimilation in a carbon-dependent manner[63,66]. In an α-cyanobacteria species lacking nitrate/nitrite uptake activities (thus relying on ammonia as a nitrogen source), a role for nrPII in regulating bicarbonate uptake has been proposed but not experimentally demonstrated[67]. Additionally, the distantly related PII-type protein SbtB from β-cyanobacteria binds directly to and inhibits the activity of sodium-bicarbonate transporters when those two proteins are expressed together in a recombinant *E. coli* system.[68]

In the work presented here, we provide the first structural and biochemical characterization of a unique type of PII protein whose gene occurs in a conserved fashion across α-type carboxysome operons in chemoautotrophic bacteria. We show that its three-dimensional structure and nucleotide binding properties are affected dramatically by bicarbonate binding, thereby tying its activities to carboxysome function and the regulation of one-carbon metabolism in autotrophic bacteria.

RESULTS

Conservation of a PII-like Gene in α-Carboxysome Operons

An analysis of genomic sequence data shows that a protein with distant similarity to nrPII proteins is associated with α-type carboxysome operons in various lithoautotrophic bacteria as a member of a novel four-member gene cluster (Fig. 1B). Two of the four members are predicted to be integral transmembrane proteins. The first is often annotated as 'NADH-ubiquinone oxidoreductase chain 5 homolog' (NDH-5), referring to the outer-most proton-pumping subunit of respiratory Complex I.[69,70] Interestingly, homologs of this membrane protein are also encoded near cyanobacterial β-carboxysome genes and have been shown to be involved in CO_2 concentrating mechanisms in those organisms.[22–24] The second, a ~500 amino-acid low-sequence-complexity transmembrane protein, does not contain recognizable domains or sequence motifs. The third member, predicted to be soluble, belongs to the DUF 2309/PF10070 family. The DUF2309 family comprises large proteins of unknown function with no homologous protein structures available in the Protein Data Bank. The NDH-5/PII four member cluster described here is one of several gene clusters that contain NDH-5 and DUF2309. The NDH-5 and DUF2309 proteins are found genetically coupled as an operon throughout the domain Bacteria and within the class Halobacteria of the Euryarchaeota. According to the ELMB database, there are >5000 occurrences of the NDH-5/DUF2309 pair, most of which are distributed among 11 bacterial phyla, especially concentrated among the Firmicutes (Gram positive) and Proteobacteria (Gram negative). Furthermore, to our knowledge DUF2309 is always accompanied by an NDH-5 homolog. One of the most common NDH-5/DUF2309-based clusters contains a single third gene annotated as a 'hydrogen-peroxide inducible genes activator' or more generically as a 'LysR transcriptional regulator'. Based on sequence comparison methods (see Methods), the fourth member of the novel four-gene cluster reveals remote but statistically detectable homology to the ubiquitous PII signal-transducer proteins. An analysis of these carboxysome-associated PII-like proteins reveals divergent PII-like motifs (B-loop and C-loop) thought to be critical for nitrogen regulatory PII activity[71–74], suggesting a distinct or specialized function. Estimates of evolutionary divergence place the carboxysome-associated PII-like proteins in a distinct clade (Fig. 1C). We therefore introduce the name CPII for these specialized carboxysome-associated PII proteins, which have not been studied previously.

Crystal Structure of Apo-CPII

In order to characterize a CPII protein in detail, we solved a series of crystal structures of CPII from *Thiomonas intermedia* strain K12 (Genbank accession ADG29527, locus tag Tint_0114). The overall three-dimensional structure of CPII (Fig. 2) is similar to structures reported for well-studied nrPII proteins and to a structure deposited in the PDB (2CZ4) for a PII-like protein from *Thermus thermophilus*.[74–78] PII proteins are homotrimeric with nucleotide binding sites at three identical subunit interfaces. The CPII secondary structure consists of two consecutive but interdigitated β-α-β motifs connected by a flexible loop that comprises the so-called "T-loop". The two consecutive β-α-β motifs are organized so that β1 and β4, α1 and α2, and β3 and β2 align in antiparallel orientation (Fig. 2). In contrast to nrPII, the β2 strand in CPII is kinked in the middle by an additional asparagine residue (Asn-41) forming a β-bulge. For the purposes of comparison to nrPIIs, the notations β2.1

and β2.2 are chosen to specify the first half of β2 and the second half of β2, respectively. The tertiary structure of CPII is completed as β1:β4 and β2:β3 come together through the β1:β3 antiparallel β-sheet. The resulting four-stranded β-sheet (β2:β3:β1:β4) has one hydrophobic face and one more polar face. The hydrophobic face minimizes solvent exposure by abutting against α1 and α2. The other side of the four-strand β-sheet is partly exposed, facing a very small internal cavity in the trimer. The interfaces between monomers are stabilized by antiparallel β-sheet interactions involving β4-β4 and β2.1–β2.2. The structure of an apo CPII trimer is shown in two orientations in Figure 2. This description holds true for most PII and PII-like protein structures. Between CPII and canonical nrPII proteins, however, considerable deviation occurs in the coils that create the adenosine-nucleotide binding pocket, namely the B-loop, C-loop, and T-loop, as discussed below.

Our crystal structure of CPII with no ligands bound (apo-CPII, PDB ID 5D4L) reveals a flexible T-loop from Gly-46 to Phe-57 (Fig. 2). A consensus sequence of the 16 amino acid T-loop of CPII can be described (Fig. 3A) as G[**S**RA][**H**Q]G[FIM][HY] [**E**G]GH[LV]MFNED[**D**EA], where residues in bold font represent the most prevalent amino acid type among alternative possibilities. Conserved residues within the T-loop may dictate specificity for target effectors; T-loops often mediate direct protein-protein interactions that PII proteins make with their partners, though an exception has been noted for one particular interaction between nrPII (GlnZ) and DraG in *Azospirillum brasilense*.[79] In our apo-CPII structures, the T-loop adopts different conformations in different crystal-packing environments. If stabilized through crystal contacts, the T-loop usually lies as a coil along the length of the empty nucleotide binding pocket, with Phe-57 filling part of the adenine binding site (see PDB 5D4L chain A, 5D4O chain C). When T-loops are not stabilized by crystal contacts, electron density is weak or absent from His-48 to Met-56, indicating flexibility or multiple alternative conformations (see PDB 5D4L chain B, 5D4N chain C,). Phenylalanine-57, however, is still present in the adenine binding pocket in structures with disordered T-loops.

Other key loop regions are likewise dissimilar between CPII and well-studied nrPIIs. The sequence-conserved CPII B-loop (the coil connecting α2 and β4) is significantly shorter than the nrPII B-loop (Fig. 3B), suggesting altered nucleotide specificity.[53,80] CPII does not contain any Walker-A-like motif spanning its B-loop. In place of the nrPII B-loop motif [**T**S][**G**S]xxG[E**D**]GKhh (h being any hydrophobic residue)[71,81,82], the CPII B-loop contains a short but conserved sequence H[T**S**]G spanning residues 89–91, with the Ser-90 hydroxyl group pointing into the nucleotide binding cleft (Fig. 2, 3B). In all PII and PII-like sequences, the B-loop makes close, direct contacts with bound nucleotides.

A C-loop motif analogous to the one in nrPII proteins is present only in ligand-bound forms of CPII, but not in apo-CPII. The absence of a canonical C-loop in apo-CPII (Fig. 3C) suggests again that CPII behaves differently from nrPIIs with respect to nucleotide binding.[72] The nrPII C-loop is a small C-terminal hairpin with two conserved arginine residues that serve to interact with the phosphate groups of bound ATP. In place of the C-loop motif, the elongated N-terminus of apo-CPII extends along the length of the C-terminus as an anti-parallel β-sheet. Also, instead of the conserved RxR sequence motif of the nrPII C-loop, CPII contains a conserved RxxKF sequence beginning at Arg-102. In apo-crystal

structures of CPII, electron density for residues 1, 2 and 104–108 is weak or absent. Figure 3D compares nrPII β2 strand with that of CPII, highlighting the β-bulge induced by Asn-41. Figure 3E compares the three-dimensional trimeric assemblies of nrPII and CPII proteins.

Ligand-Induced Conformational Changes in CPII

X-ray crystallography was used to visualize the effects of ligands on the structure of CPII. A series of crystal structures were determined with diffraction resolutions between 1.7 – 2.3 Å from crystals bound by combinations of ADP, AMP, $MgCl_2$, bicarbonate and/or its structural analog acetate. (See Table 1 and Supplemental Table 1 for details regarding crystal conditions and ligands.) Several of these structures are shown in Figure 4, highlighting the effects of ligand binding. ADP- and bicarbonate-bound structures reveal drastic conformational changes in both the C-terminus and the T-loop compared to apo-structures.

The CPII T-loop Phe-57 disassociates from the interior of a nucleotide binding cleft upon occupation by ADP, and is re-positioned to the apex of the raised T-loop with its hydrophobic phenyl ring solvent-exposed (Fig 4A,5). All three T-loops (Fig 4A1–A3) extend out and perpendicular to the triangular plane of the CPII trimer in the structure identified by PDB code 5D4P. This rearrangement corresponds to a movement in the T-loop of almost 20 Å. Similarly elevated T-loops are observed in various ADP-bound structures (Fig. 4B1 and 4C1). AMP also displaces T-loops in CPII, although they are more often disordered than perpendicularly elevated (Fig 4 B2 and C2). In addition to displacing the T-loop from the nucleotide binding cleft, nucleotide binding stabilizes the C-terminus as a short α-helix. The changes that occur in the T-loop and C-loop upon nucleotide binding are highlighted in Figure 5. The amino group of Lys-105 is positioned near the α-phosphate of the nucleotide and the oxygen atom of the Gly-91 (conserved B-loop residue). Across the distinct crystal forms studied, distances between Lys-105 and the α-phosphate vary between 2.7 and 3.0 Å, while distances to the carbonyl group (of Gly-91) vary more widely from 2.8 Å (PDB 5D4O) to as high as 4.7 Å (PDB 5D4P).

Bicarbonate molecules bind in nooks formed by Gly-101 and Arg-102, which immediately precede the C-loop α-helix (Fig. 4,6). The presence of bicarbonate promotes the geometry of the C-loop helix through interaction with Arg-102. Bicarbonate binding appears to jointly promote the formation of the C-loop helix. Figure 6 compares C-loops with and without bicarbonate, revealing variation in the helical geometry and re-orientation of Arg-102. Although the C-loop Lys-105 directly contacts the ADP α-phosphate with or without bicarbonate or acetate, the helical geometry of the C-terminal segment is promoted by interaction of bicarbonate with Arg-102. Without bicarbonate or acetate bound, the side chain of Arg-102 extends away from the protein core and into the solvent.

We analyzed the dramatic ADP-dependent conformational changes in CPII in the context of conformational changes that have been observed before in well-studied nrPII proteins. In some cases, crystal structures of ADP-bound nrPIIs in complex with target effectors show T-loop conformational changes as large as those found between apo-CPII to ADP-bound CPII crystal structures in this study[51,52]. In the case of CPII, the changes are peculiar both in terms of the articulated surface shape of the T-loops and in the hydrophobicity created by exposed T-loop residues Leu-55, Met-56 and Phe-57. The ADP-bound form thus appears to

be poised for interactions with other molecules, though more detailed conclusions cannot be drawn at present. The capacity for intermolecular interactions at this site is manifested by the tendency of these extended phenylalanine residues (at ADP-bound clefts) to participate in fortuitous intermolecular contacts with other trimers in some crystal forms. The C-termini of typical nrPII proteins are known to interact with nucleotide phosphates of ATP and ADP, but they are not known to interact with bicarbonate. However, we note that two acetate molecules were modeled in the structure of an nrPII protein GlnK2 from A. *fulgus* (PDB ID 3NCQ), though the biological significance of this observation (if any) is unknown.

Measurement of Nucleotide Binding by Electron Paramagnetic Resonance (EPR)

Initial attempts to measure nucleotide binding by CPII using isothermal titration calorimetry were complicated by a tendency of the protein to precipitate in the instrument, apparently as a consequence of stirring. We therefore devised an SDSL-EPR-based (Site Directed Spin Labeling- Electron Paramagnetic Resonance) method for measuring nucleotide binding.[83,84] Supplemental Figure 1 provides a detailed example of our analytical strategy. Our crystal structures (discussed below) suggested that nucleotide binding stabilizes the C-terminus of the protein in an α-helical conformation. We therefore monitored the EPR spectrum of a nitroxide side chain (known as R1[83,84]) introduced via an engineered C-terminal cysteine, Cys-109, with the expectation that nucleotide binding would be reflected by a reduced mobility of the R1 spin-label with corresponding changes in the EPR spectrum. The EPR spectra of CPII-C109-R1 in the presence of various concentrations of ATP, ADP or AMP were recorded. ADP and AMP produced changes reflecting a concentration-dependent decrease in R1 mobility with saturation behavior, while ATP did not appreciably change the EPR spectrum of C109R1 (Fig. 7A and Supplemental Figure 2). From the resulting binding curves (Fig. 7B) the K_D values for ADP and AMP were measured to be 1.6 and 1.8 mM respectively. Those values were not substantially affected by the presence of $MgCl_2$ nor acetate (Supplemental Figure 2).

We also investigated the possibility of identifying other effector molecules that might influence nucleotide binding by CPII. Because α-ketoglutarate is an effector for many nrPII proteins, we tested the effects of that metabolite on the binding of nucleotides by CPII but found little to no effect (Supplemental Figure 2). However, we discovered instead that nucleotide binding was strongly and consistently affected by bicarbonate in a concentration-dependent manner. Figure 7B shows the binding curves of ADP to CPII-R1 in a range of concentrations of bicarbonate (10 μM, 5 mM, 10 mM, 50 mM HCO_3^-). At high concentrations of bicarbonate, the affinity of CPII for ADP is increased by a factor of about 15, to 230 uM. This result suggested that, at physiological concentrations of ADP, CPII might fluctuate between partially- and fully-bound forms depending on intracellular bicarbonate concentrations. To further probe this possibility, we monitored ADP binding in a solution of constant concentrations of protein and ADP (10 μM CPII-R1 and 2 mM ADP) as a function of increasing bicarbonate (Fig. 7B and 7C). As expected, the binding affinity of CPII for ADP is strongly influenced by bicarbonate concentrations in the range of 5 to 10 mM bicarbonate. Likewise, at a fixed physiologically relevant concentration of 2 mM ADP, nucleotide binding is strongly promoted by increasing bicarbonate concentration; e.g. from about 50% saturated in the absence of bicarbonate to about 90% saturated in 10 mM

bicarbonate. Similar binding results were obtained between CPII, bicarbonate and AMP (Supplemental Figure 3). The binding data demonstrate a thermodynamic coupling between nucleotide binding and bicarbonate binding, consistent with an allosteric relationship inferred from the ligand-bound crystal structures described above. Both binding events appear to stabilize similar conformational changes in the protein, especially involving the C-loop helix. Nucleotide binding stabilizes the C-terminal helix through nucleotide α-phosphate interaction with Lys-105, while bicarbonate binding stabilizes the helix through interactions with Arg-102. An additional EPR experiment, in the absence of nucleotide, confirmed that high bicarbonate concentrations appear to stabilize the same conformation of the C-terminal helix as nucleotide binding (Supplemental Figure 4).

In order to more firmly establish that the conformational changes seen by EPR upon nucleotide addition are due to nucleotide binding, we performed equilibrium dialysis experiments. ADP binding experiments were performed in triplicate in the absence and presence of bicarbonate, allowing K_D values to be calculated. The measured K_D values were 14 (+/− 4) mM in the absence of bicarbonate and 200 (+/− 100) uM in the presence of 50 mM bicarbonate. The equilibrium dialysis results confirm the strong dependence of ADP affinity on the presence of bicarbonate. The numerical K_D values are similar from EPR and equilibrium dialysis in the presence of bicarbonate, though there is an important deviation between K_D values for the two experimental methods in the absence of bicarbonate. The nucleotide binding measurements by dialysis were made on the basis of A_{260} values. This was made possible by the minimal influence from the protein at that wavelength as CPII lacks tryptophan residues, though we note that the precision of the K_D measurements were not as high by this approach in comparison to the EPR-based measurements described above. Beyond the difference in precision of the two methods as we applied them here, this may reflect that the two experiments are reporting on distinct phenomena – nucleotide binding and ordering of the C-loop – which are coupled but not identical.

AMP Modeled into Various Crystal Structures

We observed that CPII can bind both AMP and ADP using crystallographic and biochemical methods. Interestingly, several crystallization trials that aimed to co-crystallize ADP and MgCl$_2$ with CPII (5D4N, 5DRK and 5D4O, see Table 1 and Supplemental Table 1) produced crystals with electron density that could only be modeled by AMP at one of the three nucleotide binding sites, even though AMP was not added to the crystallization conditions. We are unable to determine whether or not ADP was catalytically hydrolyzed to AMP by CPII, or if some ADP non-enzymatically hydrolyzed in the crystallization conditions after several days of incubation at room temperature. We considered that the electron density modeled as AMP could instead be an ADP with a structurally disordered β-phosphate. However, in view of the high resolution of the diffraction data, the electron dense nature of the phosphate group, and the relative flatness of a difference electron density map in that spot, we judged this to be unlikely. Difference electron density maps (Fig. 8) revealed no density protruding from the α-phosphate group of the modeled AMP. In a test map generated by refinement of the structure without modeled AMP, positive difference density corresponding to the α-phosphate of AMP appears at a height of 19σ (standard deviations above the mean), while no positive density higher than 1.5σ appears in any position that

might correspond to a β-phosphate. This eliminates the possibility that ADP occupies that nucleotide-binding cleft. Further studies are required to explain the unexpected presence of AMP in the CPII binding sites in these crystal forms.

DISCUSSION

In this study we describe a novel protein sub-family, CPII, which has diverged from a ubiquitous superfamily of regulatory PII proteins known from previous work to be involved in various facets of nitrogen metabolism. Interestingly, CPII's genomic associations indicated a connection to CO_2 concentration and carbon fixation. We used structural and biochemical methods to reveal unique features of CPII compared to other PII proteins, with the binding of CPII to bicarbonate being particularly noteworthy. Those findings shed light on how CPII might operate in the cell and how its activities relate to inorganic carbon concentration in bacteria.

A series of crystal structures indicate that CPII binds to and signals through nucleotide-dependent conformational changes. We discovered that ADP binding causes major conformational changes in CPII as large as 20 Å, with the most dramatic changes occurring in the T-loop region (Fig. 4 and Fig. 5). The ADP-bound conformation (PDB ID 5D4P, bound to three ADP molecules) is notable for the peculiar nature of the surface it creates on one side of the trimer, where The T-loops expose phenylalanine, along with other hydrophobic T-loop residues, near the three-fold axis of symmetry. This creates an articulated hydrophobic surface, which we surmise corresponds to a site of interaction with a downstream signaling partner. Among well-studied nitrogen regulatory PII proteins, signaling occurs through interactions of protruding T-loop amino-acid side chains with protein partners such as the ammonium transporter AmtB[51], acetyl-glutamate kinase[53] and PipX[52]. The cognate partner or partners of CPII will need to be defined in subsequent studies. If CPII is encoded near its target effector, as is often the case for nrPII proteins, then the NADH dehydrogenase subunit 5-like and the DUF2309 protein are plausible candidates as target effectors for CPII. Unfortunately, all attempts at heterologous production of the NDH/DUF2309/low-complexity proteins in *E. coli* failed, whether expressed separately or in combinations.

The ligand binding properties of CPII were explored in order to corroborate the structural data on nucleotide binding and to look for effector molecules that might show a thermodynamic coupling to nucleotide binding events, as would be expected for allosteric regulation. We found that EPR spectroscopy was a facile approach for extracting binding affinities based on the effects that ligand binding had on the mobility of a spin label attached near the presumptive binding site cleft.[84] Binding affinities in the low millimolar range were found for ADP and also AMP, but no binding was detected for ATP. Most importantly, the nucleotide binding affinities increased by a factor of about 15 to give K_D values of 230uM in the presence of bicarbonate, with nearly full effect exhibited in the 10 mM concentration range for bicarbonate. This result establishes an allosteric signaling mechanism for CPII involving bicarbonate, which is particularly significant given that a key motivation of the present work was to uncover possible connections to carbon metabolism for this gene that is conserved across carboxysome and carbon concentrating operons. Estimates of intracellular

concentrations of ADP and AMP range roughly between 0.2 mM – 1.5 mM and 0.04 - 0.3 mM, respectively,[85] and bicarbonate concentrations may exceed 50 mM during active carbon concentration. Thus, under physiological conditions, increasing bicarbonate concentrations could shift the equilibrium position for CPII binding to ADP to a more fully bound state, with concomitant protein conformational changes driving the association of CPII with downstream effectors.

These studies on CPII provide early clues about function and mechanism, while opening up several new avenues for further exploration. If comparable to nrPIIs, the relationship between nucleotide binding and CPII function is likely defined by the nucleotide's effect on the positioning of the T-loop. Our structural findings indicate that ADP binding displaces the T-loop bearing Phe-57 and also promotes conversion of the C-terminal protein segment to a helical form by way of an interaction with Lys-105. The concomitant effect on the position of Arg-102 favors binding of bicarbonate, and vice-versa. These mechanistic features tie nucleotide and bicarbonate concentrations to an altered conformation of CPII that we presume interacts with target effectors. Based on the apparent genomic association of CPII with the transmembrane NDH-5-like protein, DUF2309, and a low-complexity transmembrane protein of unknown function, together with the known regulation of ammonium transport by the interaction of well-studied nrPII proteins with the membrane transporter, AmtB, we hypothesize that CPII could exert its regulatory function through interactions with its genetically associated transmembrane proteins, which may be involved in molecular transport events that are not yet understood. The use of nucleotide binding as a regulatory feature in CPII would parallel other well-studied systems, while the bicarbonate-induced conformational changes in CPII relate to its dedicated role in regulating carbon concentration and carbon metabolism. Attempts were made to analyze the effects that binding partners might have on CPII, but as noted above, heterologous production of candidate CPII target effectors (the NDH-5-like protein, DUF2309, and the low-complexity transmembrane protein) were unsuccessful. A clearer picture of which other proteins directly interact with CPII, most likely in the fully-ADP-bound state promoted by bicarbonate binding, will be important in clarifying its cellular activities and its downstream signaling pathway.

METHODS

Analysis of CPII Genes and their Genomic Context

A comparative genomics analysis of CPII was performed using the following online programs: SEED[86], STRING[87], Clustal Omega[88], NCBI protein domain search and NCBI protein BLAST tool. Sequences of canonical nrPII proteins used for comparisons to CPIIs were extracted from phylogenetically distinct proteins with structures deposited in the Protein Data Bank. CPII proteins (NCBI protein ID from *T. intermedia*: ADG29527) were selected from operons that included both a DUF2309 protein (NCBI protein ID from *T. intermedia*: WP_013121928) and NADH:ubiquinone oxidoreductase like protein (NCBI protein ID from *T. intermedia*: WP_013121926) which together were near α-carboxysome operons. The radial phylogenetic tree was generated using the online program Phylogeny.fr[89] (http://phylogeny.lirmm.fr/phylo_cgi/index.cg) in August 2015. The radial

phylogenetic tree of chosen nrPII and CPII sequences was generated with the phylogenetic analysis One Click method using default parameters.

Cloning and Protein Expression of CPII

A gBlock (purchased from Integrated DNA Technologies, Inc.) encoding the *T. intermedia* CPII amino acid sequence with a C-terminal 6-Histidine tag was assembled by the Gibson method[90] into the plasmid pET28a between the NcoI and HindIII restriction enzyme sites. A sequence-verified clone was transformed into *E. coli* strain BL21(DE3) cells. Cells were grown overnight in 100 mL LB broth, 15 μM kanamycin, with shaking at 200 rpm at 37°C. A 15 mL aliquot of the overnight culture was used to inoculate LB media in 1L baffled flasks. These inoculates were grown with shaking at 37°C until the optical density, measured by absorbance 600 nm, reached ~0.8. At this point, protein expression was induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM, and shaken at 23°C for 24 hours. Afterwards, bacterial cell pellets were frozen and stored at −20°C.

The 6His-SUMO-CPII construct was cloned so that a tagless version of CPII could be purified. CPII-encoding gBlocks were assembled into a pET-vector that had previously been engineered to contain an N-terminal hexahistidine SUMO[91]-encoding gene. The N-terminal CPII methionine directly follows the two glycine residues recognized by ULP1 protease. Cleavage occurs C-terminal to the second glycine. 6His-SUMO-CPII was expressed similarly to 6His-CPII.

Purification of CPII

One liter of cells from 6His-CPII were resuspended at 4°C in 50 mL the buffer 50 mM Tris-HCl pH 7.6, 300 mM NaCl, one pill of Roche complete EDTA-free protease inhibitor cocktail, and lysed by sonication. Cell debris was pelleted by centrifugation at 16,500 rpm in an SS-34 rotor for 30 minutes at 4°C. The resulting supernatant was filtered through a 0.22 μm Whatman membrane filter, washed with 30 mL wash-buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 100 mM imidazole), and eluted in one step with elution buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 500 mM imidazole). Purified CPII was dialyzed into 10 mM Tris-HCl pH 7.6, 50 mM NaCl. The first purification step of 6His-SUMO-CPII was performed as above. After elution of 6-His-SUMO-CPII, the fraction was diluted 6-fold with 50 mM Tris-HCl pH 7.6, 300 mM NaCl and incubated overnight at room temperature with approximately 1–3 mg of 6His-ULP1 (6His-ubiquitin like protease 1). The solution was then run through a nickel column equilibrated with 50 mM Tris pH 7.6, 300 mM NaCl. SDS-PAGE confirmed that the flow-through contained pure CPII. Samples were dialyzed into appropriate buffers, concentrated and stored at 4°C or −20°C until further use.

Protein Crystallization

Hanging-drop crystallization screens were performed using the nanoliter liquid handling Mosquito from TTP LabTech at the High Throughput Macromolecular Crystallization Facility at UCLA. Crystallization conditions were screened using the JCSG, Structure Screen I and II and Wizard crystallization screens. For all trials, the concentrations of 6His-CPII and CPII were between 25–30 mg/mL (2.0–2.5 mM CPII monomer) as determined by

absorbance at 280 nm using the theoretical extinction coefficient 0.3480 M^{-1} cm^{-1}. When other ligands were included in co-crystallization conditions, their concentrations were the following: 10 mM MgCl$_2$, 12 mM ADP, 12 mM AMP, 12 mM ATP, 20 mM NaHCO$_3^-$ pH 7.0, 10–100 mM 2-oxogluatarate pH 7.0. In order to control pH and prevent the possibility of pH-dependent hydrolysis of nucleotides, solutions were buffered by 100 mM Tris pH 7.2 and 50 mM NaCl. When nucleotides were absent, buffers contained 10 mM Tris pH 7.2 and 50 mM NaCl. Sodium bicarbonate and sodium α-ketoglutarate solutions were prepared in 50 mM Tris and titrated to a final pH of 7.2 with NaOH. Sodium adenosine diphosphate stock solutions were prepared by resuspension of powdered sodium ADP (Sigma Aldrich). Final concentrations of ADP solutions were determined using a Nano-drop spectrophotometer at 260 nm. The theoretical extinction coefficient we used for ADP is 15400 M^{-1}cm^{-1}. The final pH of ADP solutions were estimated using Millipore pH strips.

Optimizations of initial crystal hits were performed using an oil immersion sitting drop technique in 96 well plates. The technique employed was similar to batch crystallization, but with features allowing for variation in protein concentration and evaporation rates. Eight diluted solutions of CPII were prepared by mixing the original CPII protein stock with buffer to give final concentrations that were 94%, 88%, 82%, 76%, 70%, 64%, 58%, and 52% of the original value. These eight CPII diluted solutions were prepared along the A–H axis in a single column of a 96-well plate. Using the TTP Laptech Mosquito Crystal machine, 400 nl of the diluted range of CPII solutions were transferred from the reservoir column to each well of a Greiner microbatch plate. All 96 drops were then overlaid with 200 nl of the original crystallization condition, which was obtained directly from the well of the original crystal screen condition. A gradient of paraffin-silicon oil mixtures, ranging from 100% paraffin to 9.1% paraffin in silicon oil, were prepared along the 12 columns of the 96-well microplate. A 12-channel pipette was used to overlay all 96 sitting-drops with 15 μL of oil. The recessed wells of the microbatch plate enabled the oil to cover the drops without mixing between wells, so that the rate at which the drops lost water would vary along each row of the microplate.

Phasing and Atomic Refinement

Diffraction data were collected both in-house and at the Advanced Photon Source Argonne National Laboratory synchrotron. Our initial structure of apo-CPII was determined by molecular replacement using the SbtB structure from *T. thermophilus* H8, PDB ID 2CZ4, as a search model. The subsequent crystal structures were determined by molecular replacement using apo-CPII monomers as search models. Molecular replacement, model building and atomic refinement procedures were performed with PHASER[92], COOT[93], and BUSTER[94] or Phenix[95], respectively. Data quality and atomic refinement statistics are shown in Table I. A summary of deposited crystal structures and corresponding crystallization conditions is provided in Table S1.

Site-Directed Spin Labeling, EPR Spectroscopy

Purified protein containing an engineered cysteine mutation at site 109 was cloned using Gibson assembly with gBlock from ITD, and purified by nickel column as previously described. It was solvent exchanged with sample buffer (25mM Tris-HCl pH 7.2, 200 mM

NaCl, and 10% glycerol) via a pre-equilibrated PD-10 desalting column from GE Lifesciences. The protein concentration was then determined by UV absorbance at 280nm and an equimolar amount of spin label S-(1-oxy-2,2,5,5,-tetramethylpyrroline-3-methyl)-methanethiosulfonate, more commonly referred to as MTSSL, was added and incubated at room temperature for 10 minutes to generate the nitroxide side chain designated R1[84]. After incubation, the sample was desalted with a freshly equilibrated PD-10 column and was then concentrated to approximately 800 μM using a 10 KDa concentrator and stored at −20°C for subsequent preparations. Nucleotides were prepared at stock concentration of 100 mM with a final pH of 7.5 as determined by Millipore pH strips.

All EPR Spectra were collected at room temperature using a Varian E-109 Model EPR spectrometer equipped with a 2-loop-1-gap Resonator at X-band microwave frequency.[96] Continuous wave (CW) spectra were typically the average of 8 to 16 scans. Samples were typically 12 to 15 μL in volume and consisted of 10 μM spin labeled protein, sample buffer, and varying concentrations of nucleotide. To determine the fraction of nucleotide-bound protein in each sample, the spectra were fit using custom LabView software to weighted linear combinations of nucleotide free and nucleotide fully-bound lineshapes. (Supplemental Figure 1) A plot of the fraction bound versus nucleotide concentration was then fit to a two-state binding model to calculate the binding constant under each condition.

Equilibrium Dialysis

Purified CPII and ADP concentrations were measured immediately prior to experiment setup. Protein was placed in either 50ul or 100ul dialysis buttons (Hampton Research) in triplicate with or without 50mM sodium bicarbonate, respectively. Buttons were sealed with dialysis membranes (Spectrum Labs) possessing 6–8kD MW cutoffs and incubated in 10ml volumes of nucleotide solution. The system was equilibrated at room temperature under conditions of gentle mixing by orbital rotation. After approximately sixteen hours, the solutions both inside and outside the dialysis button were assessed for nucleotide concentration using A_{260} readings with a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific), being corrected for CPII A_{260} contributions inside the button by subtracting values measured from a control dialysis equilibrium button prepared without nucleotide but containing identical protein concentrations. ADP concentration determined in the outer volume constitutes $[ADP]_{free}$ while the ADP concentration contained within the button represents $[ADP]_{total}$. Concentration of bound complex was calculated as [CPII-ADP] = $[ADP]_{total} − [ADP]_{free}$. $[CPII]_{total}$ is user-defined at the onset of the experiment and confirmed by A_{280} measurements taken from the control dialysis equilibrium button. Free protein concentration was calculated as $[CPII]_{free} = [CPII]_{total} − [CPII-ADP]$. The binding constant was calculated from a one-site binding model as $K_D = ([CPII]_{free} * [ADP]_{free})/[CPII-ADP]$. Measurement uncertainty was determined by propagating fractional errors (estimated from triplicate measurement variations) throughout the calculations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations and Symbols

BMC	bacterial microcompartment (hexameric) shell protein
RuBisCO	Ribulose-1,5-bisphosphate carboxylase/oxygenase
CPII	carboxysome-associated PII protein
nrPII	nitrogen regulatory PII protein
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SDSL-EPR	Site Directed Spin Labeling- Electron Paramagnetic Resonance
ULP1	Ubiquitin-like-specific protease 1

References

1. Andersson I. Catalysis and regulation in RuBisCO. J Exp Bot. 2007; 59:1555–1568. [PubMed: 18417482]

2. Andersson I, Backlund A. Structure and function of RuBisCO. Plant Physiol Biochem PPB Société Fr Physiol Végétale. 2008; 46:275–291.

3. Martin W, Schnarrenberger C. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr Genet. 1997; 32:1–18. [PubMed: 9309164]

4. Sims PA. 'Big-Picture' worksheets to help students learn and understand the Pentose Phosphate Pathway and the Calvin Cycle. J Chem Educ. 2014; 91:541–545.

5. Spreitzer RJ, Salvucci ME. Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. Plant Biol. 2002; 53

6. Bauwe H, Hagemann M, Kern R, Timm S. Photorespiration has a dual origin and manifold links to central metabolism. Curr Opin Plant Biol. 2012; 15:269–275. [PubMed: 22284850]

7. Kozaki A, Takeba G. Photorespiration protects C3 plants from photooxidation. Nature. 1996; 384:557–560.

8. Wingler A, Lea PJ, Quick WP, Leegood RC. Photorespiration: metabolic pathways and their role in stress protection. Philos Trans R Soc Lond B Biol Sci. 2000; 355:1517–1529. [PubMed: 11128005]

9. Raven JA, Cockell CS, Rocha CLDL. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. Philos Trans R Soc Lond B Biol Sci. 2008; 363:2641–2650. [PubMed: 18487130]

10. Smith EC, Griffiths H. A pyrenoid-based carbon-concentrating mechanism is present in terrestrial bryophytes of the class *Anthocerotae*. Planta. 1996; 200:203–212.

11. Keeley JE, Rundel PW. Evolution of CAM and C4 carbon-concentrating mechanisms. Int J Plant Sci. 2003; 164:S55–S77.

12. Espie GS, Kimber MS. Carboxysomes: cyanobacterial RuBisCO comes in small packages. Photosynth Res. 2011; 109:7–20. [PubMed: 21556873]

13. Giordano M, Beardall J, Raven JA. CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. Annu Rev Plant Biol. 2005; 56:99–131. [PubMed: 15862091]

14. Badger MR, Price GD. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot. 2003; 54:609–622. [PubMed: 12554704]

15. Ogawa T, Kaplan A. Inorganic carbon acquisition systems in cyanobacteria. Photosynth Res. 2003; 77:105–115. [PubMed: 16228369]

16. Omata T, Takahashi Y, Yamaguchi O, Nishimura T. Structure, function and regulation of the cyanobacterial high-affinity bicarbonate transporter, BCT1. Funct Plant Biol. 2002; 29:151–159.

17. Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L. Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl Acad Sci. 2004; 101:18228–18233. [PubMed: 15596724]

18. Felce J, Saier MH Jr. Carbonic anhydrases fused to anion transporters of the SulP family: evidence for a novel type of bicarbonate transporter. J Mol Microbiol Biotechnol. 2004; 8:169–176. [PubMed: 16088218]

19. Shibata M, et al. Genes essential to sodium-dependent bicarbonate transport in cyanobacteria function and phylogenetic analysis. J Biol Chem. 2002; 277:18658–18664. [PubMed: 11904298]

20. López-Igual R, Picossi S, López-Garrido J, Flores E, Herrero A. N and C control of ABC-type bicarbonate transporter Cmp and its LysR-type transcriptional regulator CmpR in a heterocyst-forming cyanobacterium, *Anabaena* sp. Environ Microbiol. 2012; 14:1035–1048. [PubMed: 22221957]

21. Omata T, et al. Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. Proc Natl Acad Sci. 1999; 96:13571–13576. [PubMed: 10557362]

22. Battchikova N, Zhang P, Rudd S, Ogawa T, Aro EM. Identification of NdhL and Ssl1690 (NdhO) in NDH-1L and NDH-1M Complexes of *Synechocystis* sp. PCC 6803. J Biol Chem. 2005; 280:2587–2595. [PubMed: 15548534]

23. Battchikova N, Eisenhut M, Aro EM. Cyanobacterial NDH-1 complexes: Novel insights and remaining puzzles. Biochim Biophys Acta BBA - Bioenerg. 2011; 1807:935–944.

24. Klughammer B, Sültemeyer D, Badger MR, Price GD. The involvement of NAD(P)H dehydrogenase subunits, NdhD3 and NdhF3, in high-affinity CO_2 uptake in *Synechococcus* sp. PCC7002 gives evidence for multiple NDH-1 complexes with specific roles in cyanobacteria. Mol Microbiol. 1999; 32:1305–1315. [PubMed: 10383770]

25. Rae BD, et al. Cyanobacterial carboxysomes: microcompartments that facilitate CO_2 fixation. J Mol Microbiol Biotechnol. 2013; 23:300–307. [PubMed: 23920493]

26. Shively JM, English RS. The carboxysome, a prokaryotic organelle: a mini-review. Can J Bot. 1991; 69:957–962.

27. Cannon GC, et al. Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl Environ Microbiol. 2001; 67:5351–5361. [PubMed: 11722879]

28. Yeates TO, Kerfeld CA, Heinhorst S, Cannon GC, Shively JM. Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat Rev Microbiol. 2008; 6:681–691. [PubMed: 18679172]

29. Cai F, et al. The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. PLoS ONE. 2009; 4:e7521. [PubMed: 19844578]

30. Yeates TO, Thompson MC, Bobik TA. The protein shells of bacterial microcompartment organelles. Curr Opin Struct Biol. 2011; 21:223–231. [PubMed: 21315581]

31. Wheatley NM, Gidaniyan SD, Liu Y, Cascio D, Yeates TO. Bacterial microcompartment shells of diverse functional types possess pentameric vertex proteins. Protein Sci. 2013; 22:660–665. [PubMed: 23456886]

32. Tanaka S, Sawaya MR, Phillips M, Yeates TO. Insights from multiple structures of the shell proteins from the β-carboxysome. Protein Sci. 2009; 18:108–120. [PubMed: 19177356]

33. Kerfeld CA, et al. Protein structures forming the shell of primitive bacterial organelles. Science. 2005; 309:936–938. [PubMed: 16081736]

34. Tanaka S, et al. Atomic-level models of the bacterial carboxysome shell. Science. 2008; 319:1083–1086. [PubMed: 18292340]

35. Tsai Y, et al. Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. PLoS Biol. 2007; 5:e144. [PubMed: 17518518]

36. Roberts EW, Cai F, Kerfeld CA, Cannon GC, Heinhorst S. Isolation and characterization of the *Prochlorococcus* carboxysome reveal the presence of the novel shell protein CsoS1D. J Bacteriol. 2012; 194:787–795. [PubMed: 22155772]

37. Wheatley NM, Sundberg CD, Gidaniyan SD, Cascio D, Yeates TO. Structure and identification of a pterin dehydratase-like protein as a RuBisCO assembly factor in the alpha-carboxysome. J Biol Chem. 2014; jbc.M113.531236. doi: 10.1074/jbc.M113.531236

38. Onizuka T, et al. The rbcX gene product promotes the production and assembly of ribulose-1,5-bisphosphate carboxylase/oxygenase of *Synechococcus* sp. PCC7002 in *Escherichia coli*. Plant Cell Physiol. 2004; 45:1390–1395. [PubMed: 15564522]

39. Baker, et al. The correlation of the gene *csoS2* of the carboxysome operon with two polypeptides of the carboxysome in *Thiobacillus neapolitanus*. Arch Microbiol. 1999; 172:233–239. [PubMed: 10525740]

40. Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS. Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. Proc Natl Acad Sci U S A. 2010; 107:2455–2460. [PubMed: 20133749]

41. Long BM, Badger MR, Whitney SM, Price GD. Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem. 2007; 282:29323–29335. [PubMed: 17675289]

42. Jorda J, Lopez D, Wheatley NM, Yeates TO. Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria. Protein Sci. 2012; 22:179–95. [PubMed: 23188745]

43. Gaudana SB, Zarzycki J, Moparthi VK, Kerfeld CA. Bioinformatic analysis of the distribution of inorganic carbon transporters and prospective targets for bioengineering to increase Ci uptake by cyanobacteria. Photosynth Res. 2014; 126:99–109. [PubMed: 25399051]

44. Rae BD, Long BM, Badger MR, Price GD. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev MMBR. 2013; 77:357–379. [PubMed: 24006469]

45. Whitehead L, Long BM, Price GD, Badger MR. Comparing the in vivo function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. Plant Physiol. 2014; 165:398–411. [PubMed: 24642960]

46. Sutter M, et al. Structural Characterization of a Newly Identified Component of α-Carboxysomes: The AAA+ Domain Protein CsoCbbQ. Sci Rep. 2015; 5

47. Muro-Pastor MI, Reyes JC, Florencio FJ. Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. J Biol Chem. 2001; 276:38320–38328. [PubMed: 11479309]

48. Vázquez-Bermúdez MF, Herrero A, Flores E. 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. FEBS Lett. 2002; 512:71–74. [PubMed: 11852054]

49. Doucette CD, Schwab DJ, Wingreen NS, Rabinowitz JD. α-ketoglutarate coordinates carbon and nitrogen utilization via enzyme I inhibition. Nat Chem Biol. 2011; 7:894–901. [PubMed: 22002719]

50. Hart Y, et al. Robust control of nitrogen assimilation by a bifunctional enzyme in *E. coli*. Mol Cell. 2011; 41:117–127. [PubMed: 21211727]

51. Conroy MJ, et al. The crystal structure of the *Escherichia coli* AmtB–GlnK complex reveals how GlnK regulates the ammonia channel. Proc Natl Acad Sci. 2007; 104:1213–1218. [PubMed: 17220269]

52. Llácer JL, et al. Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. Proc Natl Acad Sci. 2010; 107:15397–15402. [PubMed: 20716687]

53. Llácer JL, et al. The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. Proc Natl Acad Sci. 2007; 104:17644–17649. [PubMed: 17959776]

54. Ninfa AJ, Jiang P. PII signal transduction proteins: sensors of α-ketoglutarate that regulate nitrogen metabolism. Curr Opin Microbiol. 2005; 8:168–173. [PubMed: 15802248]

55. Forchhammer K, Lüddecke J. Sensory properties of the PII signalling protein family. FEBS J. 2016; 283:425–437. [PubMed: 26527104]

56. Palanca C, et al. The structure of a PII signaling protein from a halophilic archaeon reveals novel traits and high-salt adaptations. FEBS J. 2014; 281:3299–3314. [PubMed: 24946894]

57. Gundlach J, et al. Identification, characterization, and structure analysis of the cyclic di-AMP-binding PII-like signal transduction protein DarA. J Biol Chem. 2015; 290:3069–3080. [PubMed: 25433025]

58. Adler SP, Purich D, Stadtman ER. Cascade control of *Escherichia coli* glutamine synthetase. Properties of the PII regulatory protein and the uridylyltransferase-uridylyl-removing enzyme. J Biol Chem. 1975; 250:6264–6272. [PubMed: 239942]

59. van Heeswijk WC, et al. An alternative PII protein in the regulation of glutamine synthetase in *Escherichia coli*. Mol Microbiol. 1996; 21:133–146. [PubMed: 8843440]

60. Coutts G. Membrane sequestration of the signal transduction protein GlnK by the ammonium transporter AmtB. EMBO J. 2002; 21:536–545. [PubMed: 11847102]

61. Dodsworth JA, Leigh JA. NifI inhibits nitrogenase by competing with Fe protein for binding to the MoFe protein. Biochem Biophys Res Commun. 2007; 364:378–382. [PubMed: 17950693]

62. Drepper T. Role of GlnB and GlnK in ammonium control of both nitrogenase systems in the phototrophic bacterium *Rhodobacter capsulatus*. Microbiology. 2003; 149:2203–2212. [PubMed: 12904560]

63. Hisbergues M, Jeanjean R, Joset F, Tandeau de Marsac N, Bédu S. Protein PII regulates both inorganic carbon and nitrate uptake and is modified by a redox signal in *Synechocystis* PCC 6803. FEBS Lett. 1999; 463:216–220. [PubMed: 10606724]

64. Feria Bourrellier AB, et al. Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. Proc Natl Acad Sci U S A. 2010; 107:502–507. [PubMed: 20018655]

65. Gerhardt ECM, et al. The bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. Mol Microbiol. 2015; 95:1025–1035. [PubMed: 25557370]

66. Lee HM, Flores E, Forchhammer K, Herrero A, Tandeau De Marsac N. Phosphorylation of the signal transducer PII protein and an additional effector are required for the PII-mediated regulation of nitrate and nitrite uptake in the Cyanobacterium *Synechococcus* sp. PCC 7942. Eur J Biochem FEBS. 2000; 267:591–600.

67. Palinska KA, et al. The signal transducer PII and bicarbonate acquisition in *Prochlorococcus marinus* PCC 9511, a marine cyanobacterium naturally deficient in nitrate and nitrite assimilation. Microbiology. 2002; 148:2405–2412. [PubMed: 12177334]

68. Du J, Förster B, Rourke L, Howitt SM, Price GD. Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. PLoS ONE. 2014; 9:e115905. [PubMed: 25536191]

69. Erhardt H, et al. Disruption of individual nuo-genes leads to the formation of partially assembled NADH:ubiquinone oxidoreductase (Complex I) in *Escherichia coli*. Biochim Biophys Acta BBA - Bioenerg. 2012; 1817:863–871.

70. Hrdy I, et al. Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature. 2004; 432:618–622. [PubMed: 15577909]

71. Xu Y, et al. GlnK, a PII-homologue: structure reveals ATP binding site and indicates how the T-loops may be involved in molecular recognition. J Mol Biol. 1998; 282:149–165. [PubMed: 9733647]

72. Jiang P, et al. Structure/function analysis of the PII signal transduction protein of *Escherichia coli*: genetic separation of interactions with protein receptors. J Bacteriol. 1997; 179:4342–4353. [PubMed: 9209053]

73. Bonatto AC, Souza EM, Pedrosa FO, Yates MG, Benelli EM. Effect of T- and C-loop mutations on the *Herbaspirillum seropedicae* GlnB protein in nitrogen signalling. Res Microbiol. 2005; 156:634–640. [PubMed: 15950123]

74. Shetty ND, Reddy MCM, Palaninathan SK, Owen JL, Sacchettini JC. Crystal structures of the apo and ATP bound *Mycobacterium tuberculosis* nitrogen regulatory PII protein. Protein Sci. 2010; 19:1513–1524. [PubMed: 20521335]

75. Xu Y, et al. The structures of the PII proteins from the cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803. Acta Crystallogr D Biol Crystallogr. 2003; 59:2183–2190. [PubMed: 14646076]

76. Fokina O, Chellamuthu VR, Zeth K, Forchhammer K. A novel signal transduction protein P(II) variant from *Synechococcus elongatus* PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. J Mol Biol. 2010; 399:410–421. [PubMed: 20399792]

77. Fokina O, Chellamuthu VR, Forchhammer K, Zeth K. Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. Proc Natl Acad Sci U S A. 2010; 107:19760–19765. [PubMed: 21041661]

78. Yildiz Ö, Kalthoff C, Raunser S, Kühlbrandt W. Structure of GlnK1 with bound effectors indicates regulatory mechanism for ammonia uptake. EMBO J. 2007; 26:589–599. [PubMed: 17203075]

79. Rajendran C, et al. Crystal structure of the GlnZ-DraG complex reveals a different form of PII-target interaction. Proc Natl Acad Sci. 2011; 108:18972–18976. [PubMed: 22074780]

80. Xu Y, Carr PD, Huber T, Vasudevan SG, Ollis DL. The structure of the PII–ATP complex. Eur J Biochem. 2001; 268:2028–2037. [PubMed: 11277925]

81. Heeswijk WC, van Westerhoff HV, Boogerd FC. Nitrogen assimilation in *Escherichia coli*: putting molecular data into a systems perspective. Microbiol Mol Biol Rev. 2013; 77:628–695. [PubMed: 24296575]

82. Ninfa AJ, Atkinson MR. PII signal transduction proteins. Trends Microbiol. 2000; 8:172–179. [PubMed: 10754576]

83. Dror RO, et al. Signal transduction Structural basis for nucleotide exchange in heterotrimeric G proteins. Science. 2015; 348:1361–1365. [PubMed: 26089515]

84. Hubbell WL, Cafiso DS, Altenbach C. Identifying conformational changes with site-directed spin labeling. Nat Struct Mol Biol. 2000; 7:735–739.

85. Albe KR, Butler MH, Wright BE. Cellular concentrations of enzymes and their substrates. J Theor Biol. 1990; 143:163–195. [PubMed: 2200929]

86. Overbeek R, et al. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 2005; 33:5691–5702. [PubMed: 16214803]

87. Jensen LJ, et al. STRING 8–a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 2009; 37:D412–416. [PubMed: 18940858]

88. Sievers F, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2014; 7:539–539. [PubMed: 21988835]

89. Dereeper A, et al. Phylogenyfr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008; 36:W465–W469. [PubMed: 18424797]

90. Gibson DG, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009; 6:343–345. [PubMed: 19363495]

91. Malakhov MP, et al. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J Struct Funct Genomics. 2004; 5:75–86. [PubMed: 15263846]

92. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

93. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

94. Bricogne, G., et al. BUSTER version XYZ. Cambridge, United Kingdom: Global Phasing Ltd; 2011.

95. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

96. Hubbell WL, Froncisz W, Hyde JS. Continuous and stopped flow EPR spectrometer based on a loop gap resonator. Rev Sci Instrum. 1987; 58:1879–1886.

Research Highlights

- A protein related to nitrogen PII proteins is conserved near α-carboxysome genes.

- Crystal structures of CPII are elucidated with and without potential ligands.

- ADP and AMP binding drive major conformational changes in the protein trimer.

- Bicarbonate, a substrate of carboxysomes, enhances ADP and AMP binding.

- A unique clade of PII proteins, herein named CPII, is regulated by inorganic carbon.
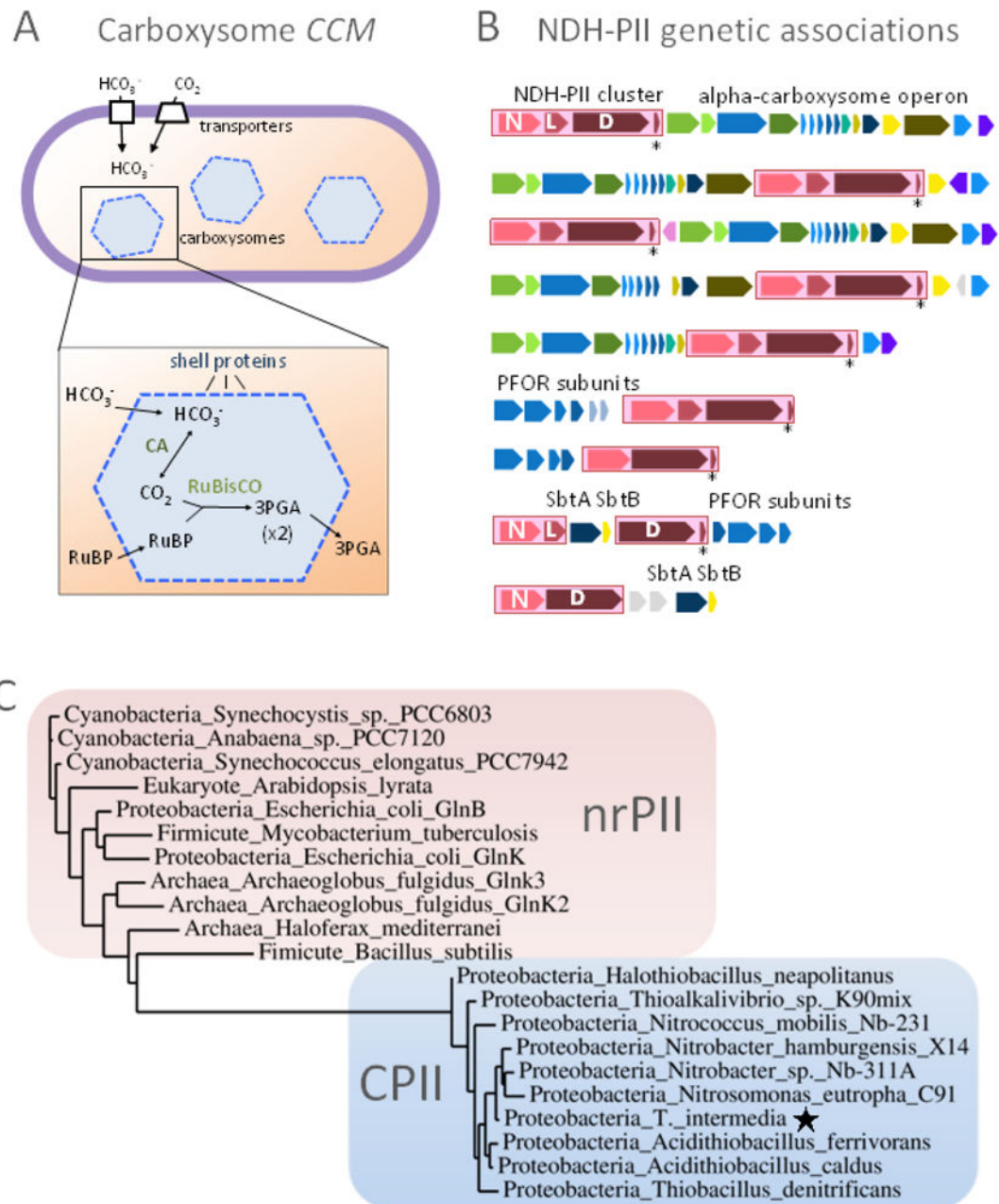
Figure 1. Conserved genomic associations for CPII, a novel type of PII regulatory protein in bacteria

A. A schematic diagram of the CO_2 concentrating mechanisms in bacteria including the carboxysome. For comparison to panel A, the α-carboxysome operons encode the RuBisCO and carbonic anhydrase (CA) enzymes along with multiple paralogs of the carboxysome shell proteins (shown as small blue genes in panel A). **B.** The CPII gene (denoted by an asterisk) occurs as part of a conserved NDH-PII gene cluster near the following: α-carboxysome operons, pyruvate:ferredoxin oxidoreductase (PFOR) genes, and putative bicarbonate transporter systems SbtAB. PFOR genes are represented in shades of blue. The putative gene for SbtA is colored dark blue and appears adjacent SbtB, colored yellow.

Representative gene clusters were selected for illustration among a large number of diverse cases. Operon diagrams were drawn from the following organisms, from top to bottom: *Thiomonas intermedia, Acidithiobacillus caldus, Nitrosomonas eutropha, Halothiobacillus neapolitanus, Nitrococcus mobilis, Hydrogenivirga, Sulfurihydrogenibium yellowstonense, Sulfurihydrogenibium azorense* and *Roseiflexus castenholzii*. "N" marks the position of the NDH-5 homolog, "L", the low complexity protein, "D" the DUF2309 protein and, asterisk (*), the location of CPII. In the last case (*Roseiflexus castenholzii*), the NHD-5 and DUF2309 are encoded near SbtA and SbtB homologs without a low complexity "L" or CPII protein. **C.** A phylogenetic tree of representative protein sequences of PII and PII-like sequences. CPII forms a cluster separate from nrPII proteins. A star (★) marks the protein described in this work from *Thiomonas intermedia*.
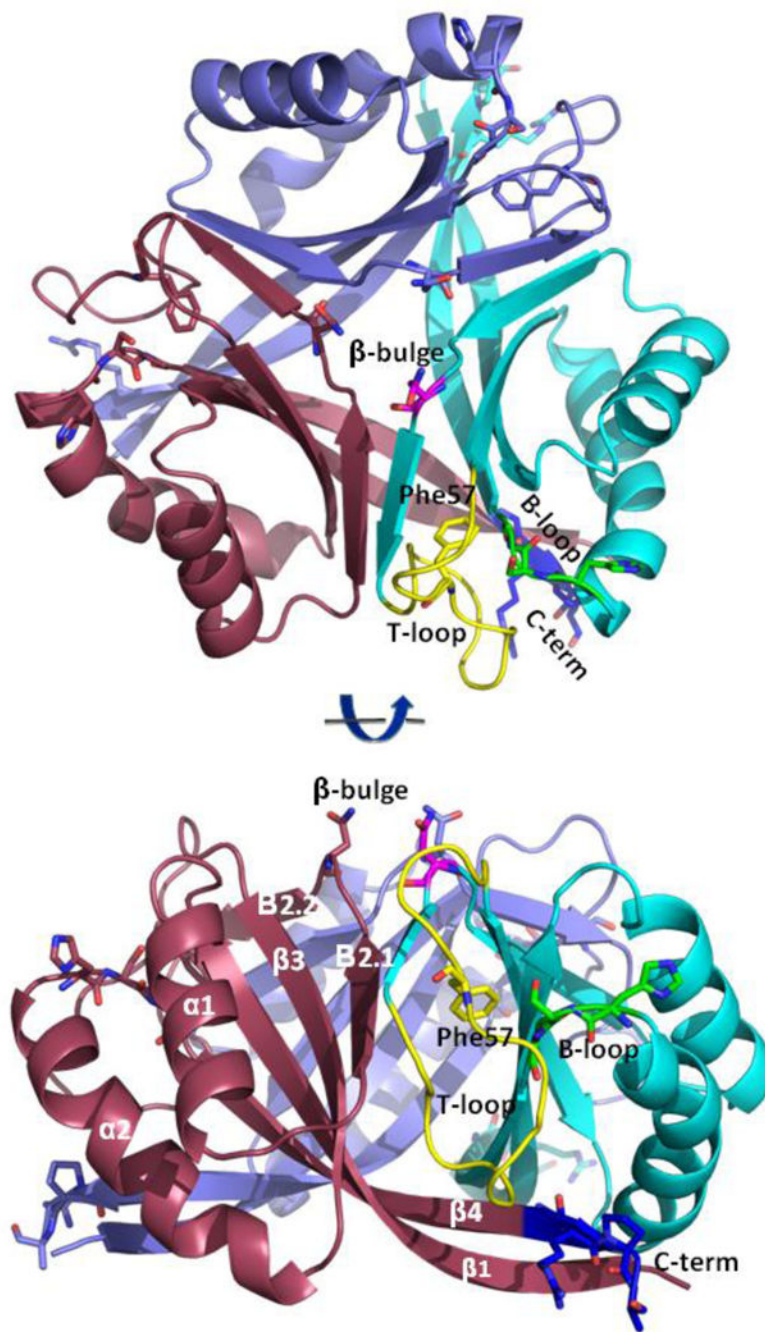
Figure 2. Crystal structure of Apo-CPII from *Thiomonas intermedia*
A. A schematic diagram of the CPII monomer in an illustrative extended configuration to emphasize the secondary structure and key tertiary interactions in the folded state. Key secondary structure motifs are labeled. **B.** The apo-CPII trimer (PDB ID 5D4L) is shown as cartoons from the top and side views, with select motifs and residues labeled. Phe-57 is indicated by an asterisk (*).
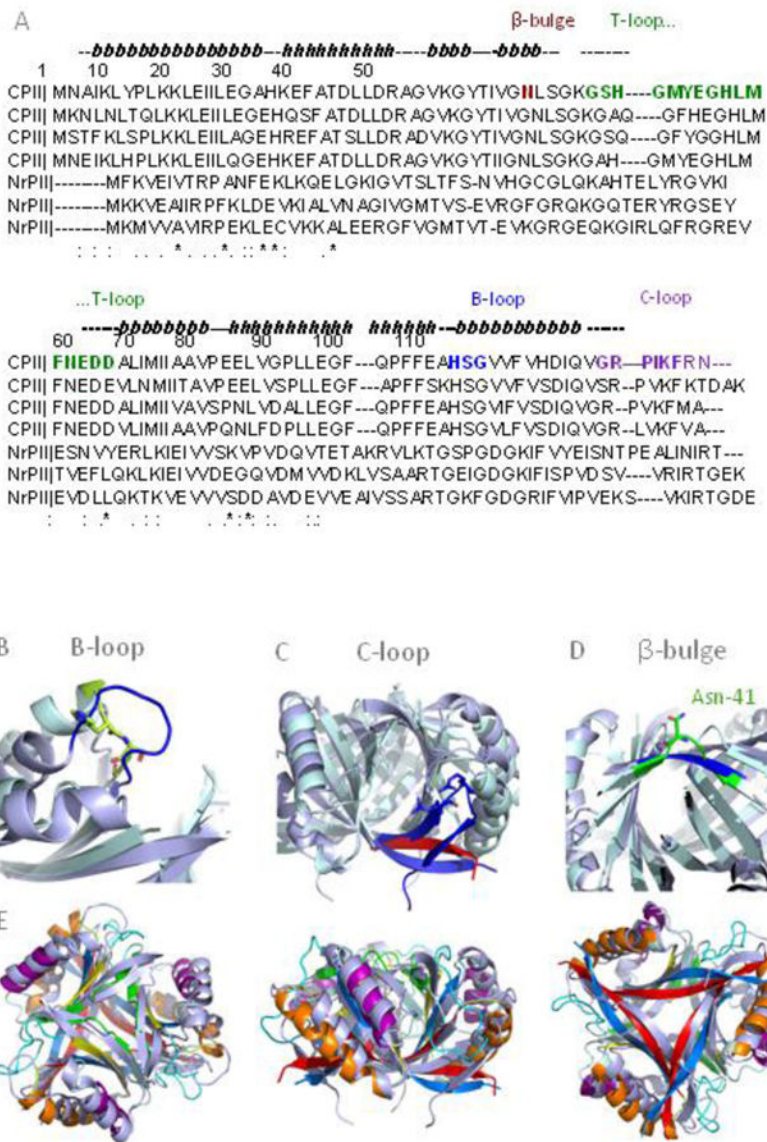
Figure 3. Comparison of secondary and tertiary structures of CPII and nrPII
A. Protein sequence alignments of select CPII and nrPII proteins. For the *Thiomonas intermedia* CPII sequence, secondary structural motifs are shown as 'b' for beta-sheet, 'h' for alpha-helix and '-' for coil. Conserved motifs of the CPII T-loop, B-loop and C-loop are colored green, blue and purple, respectively. From top to bottom, the sequences correspond to CPIIs from *Thiomonas intermedia* K12, *Halothiobacillus neapolitanus* C2, *Nitrosomonas eutropha* C91, *Nitrobacter* sp. Nb-311a, nrPIIs from *Bacillus subtilis*, *Synechocystis* sp PCC 6803, *Archaeologlobus fulgidu*s. **B–E.** Apo-CPII (PDB 5D4L) is aligned to nrPII of *Synechococcus* sp. PCC 6803 (PDB 1QY7). **B.** CPII's B-loop with conserved residues GSH shown in sticks is significantly shorter than the nrPII B-loop. **C.** The C-loop in nrPII is shown in dark blue. No analogous C-loop is present in apo-CPII. **D.** Asn-41 in CPII contributes to the β-bulge, which is generally not present in nrPII proteins. **E.** Two consecutive 90° rotations of CPII in superposition with nrPII.
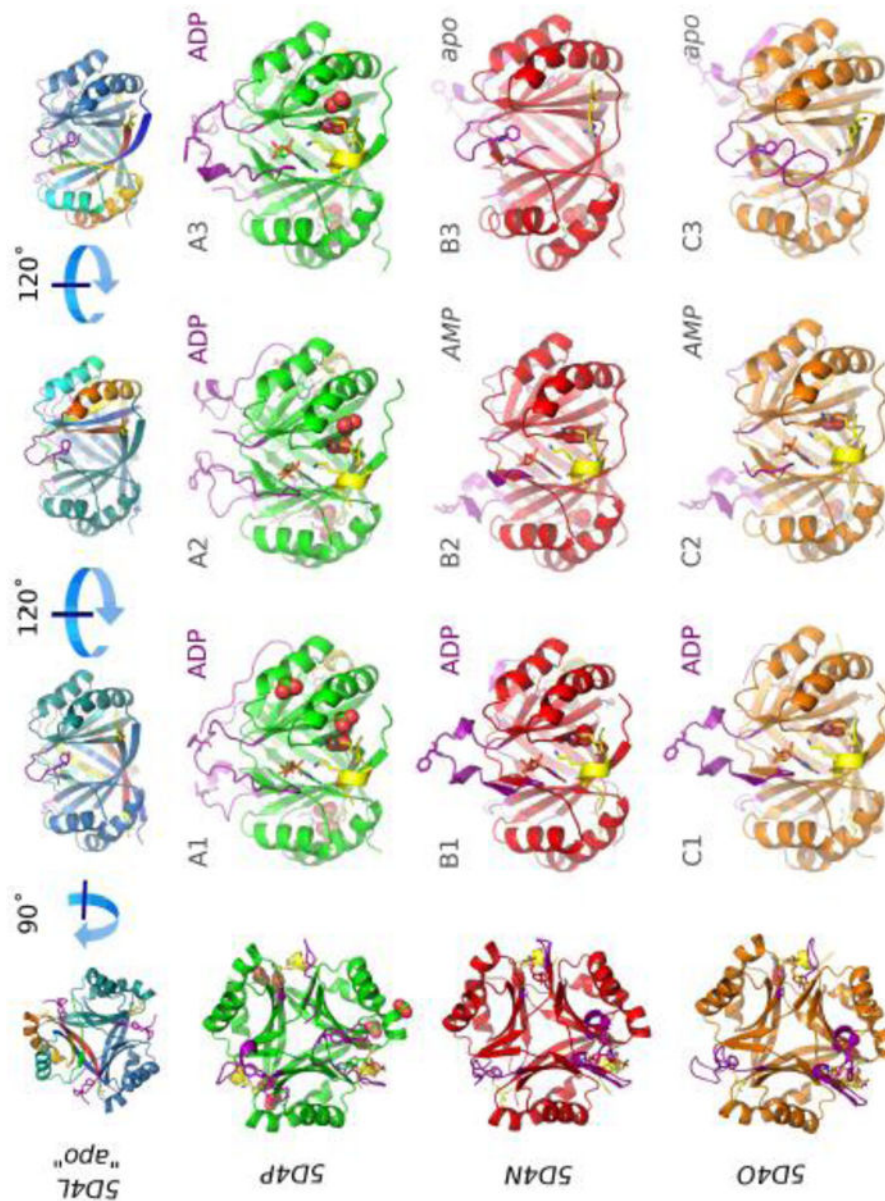
Figure 4. Nucleotide binding is associated with T-loop and C-loop conformational changes in the CPII trimer

Crystal structures are arranged in rows, viewed from the top, and from the side at three 120° rotations, showing all three nucleotide binding clefts, which are sometimes different within the same trimer. In the top row, apo-CPII exemplifies the respective orientations shown in each column. T-loops are colored purple with Phe-57 residues shown in stick form. C-loops are colored yellow, with Lys-105 and Arg-102 shown in stick form. Bicarbonate and acetate are shown as spheres. The following three rows show the diversity of subunit and loop conformations in CPII in multiple distinct crystal structures with ligands present in some or all of the available binding sites. Panels **A1–3** show the three ADP- and bicarbonate-bound clefts of the CPII trimer (PDB 5D4P). Note that all three T-loops are elevated above the plane of the CPII trimer, and all three C-loops are stabilized in an α-helical conformation.

B1–3 show the nucleotide binding clefts in PDB structure 5D4N. The clefts shown in B1, B2 and B3 are occupied by ADP and acetate, AMP and acetate, and no ligands, respectively. **C1–3** show the three nucleotide binding clefts in PDB structure 5D4O, bound to ADP and bicarbonate, AMP and bicarbonate, and no ligand, respectively. The nucleotides are shown in stick representation.
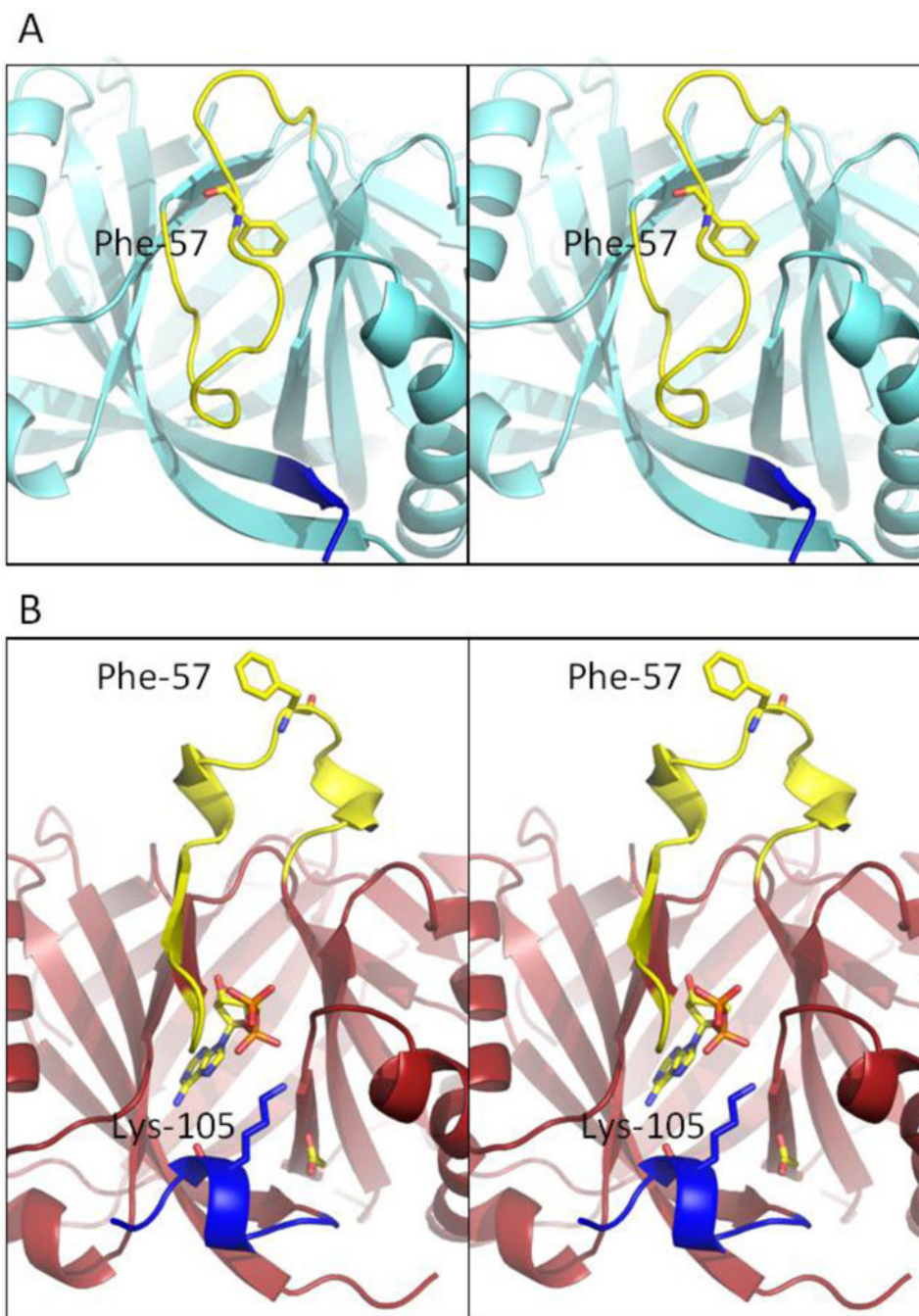
Figure 5. Conformational changes in the T-loop and C-loop upon ligand binding
T-loops are colored yellow with Phe-57 shown in stick form. C-loop residues are colored
dark blue with Lys-105 shown in stick form. A. Stereo view of apo-CPII (PDB 5D4L). B.
Stereo view of ADP- and acetate- bound CPII (PDB 5D4N).
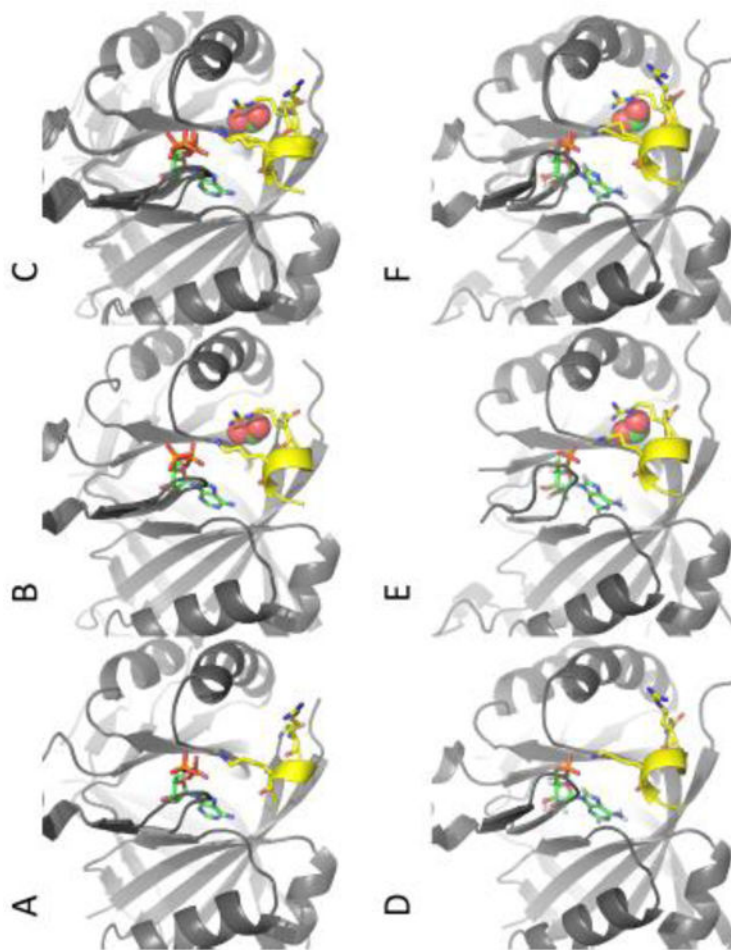
Figure 6. Bicarbonate and acetate interact with Arg-102, a conserved residue in the C-terminal helix motif in CPII

A. The ADP-bound cleft (PDB 5DRK) (without bicarbonate) shows Arg-102 in the C-terminal helix motif (yellow) extending towards the solvent. **B.** The ADP and bicarbonate-bound cleft in PDB structure 5D4O shows Arg-102 positioned for interaction with bicarbonate. **C.** Overlay of panels A and B. **D.** The AMP-bound cleft in PDB structure 5DS7 (without bicarbonate) shows Arg-102 pointed towards solvent. **E.** In the AMP and bicarbonate-bound cleft (PDB 5D4O), Arg-102 is elevated towards the interior of the cleft, coordinating bicarbonate. **F.** Overlay of panes D and E. For all panels, the C-terminal segment (residues G101–R107) is colored yellow. R102, K105 and nucleotides are shown in sticks. Bicarbonate is shown in spheres.

CIPII-R1 EPR spectra with increasing nucleotide concentrations
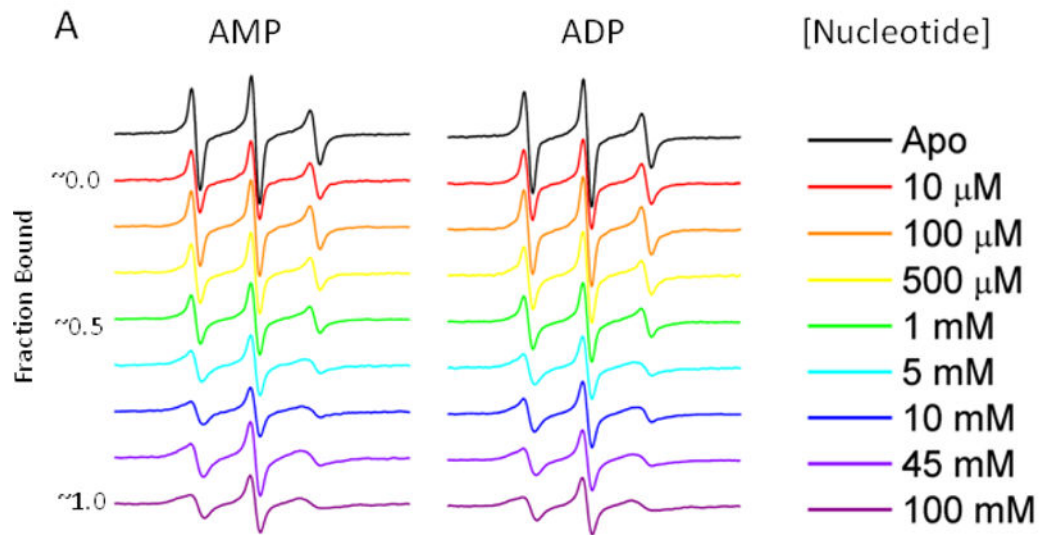
CPII-R1 nucleotide affinity is dependent on bicarbonate (HCO$_3^-$)
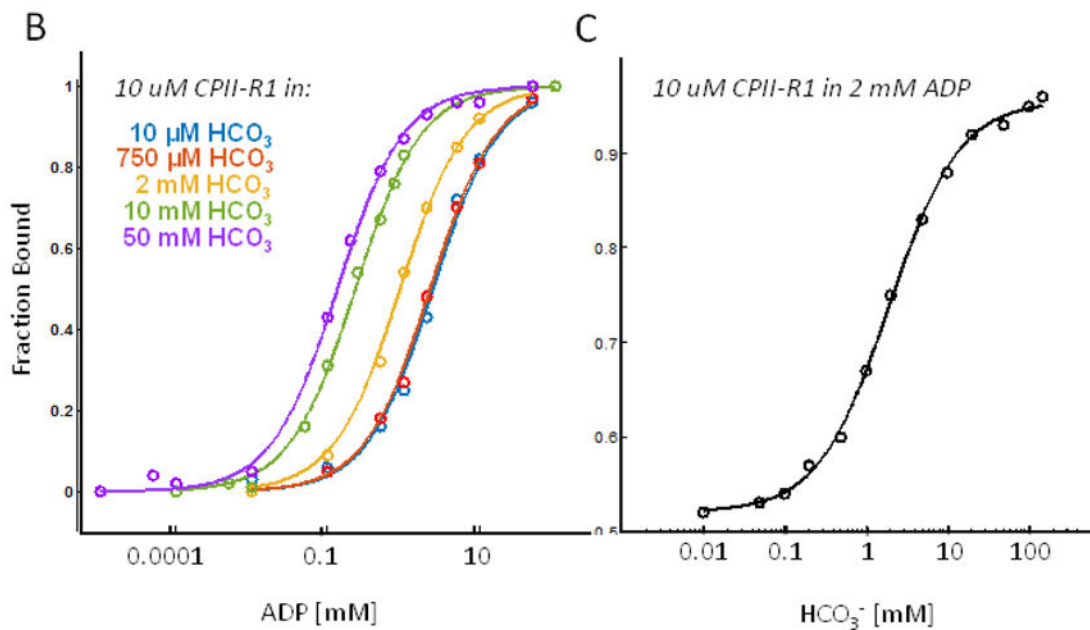
Figure 7. The affinity of CPII-R1 for ADP and AMP is influenced by bicarbonate
A. Continuous Wave (CW) EPR spectra of spin labeled CPII (C108R1) titrated with ADP and AMP. Incubation of 10 μM CPII C108R1 with increasing concentrations of ADP or AMP drives an unstructured-to-structured transition at the protein's C-terminus; consistent with the formation of a C-terminal α-helix seen in the nucleotide bound crystal structure. All spectra have been normalized and the structural transition can be seen as a broadening of the line-shape profile as nucleotide concentration increases. Line-shape broadening reflects restricted local motion of the spin labeled residue C109-R1 as the nucleotide binds and the

C-terminus becomes structured. **B.** Titrating 10 μM CPII-R1 with ADP under various concentrations of sodium bicarbonate reveals a bicarbonate induced increase in binding affinity of CPII for ADP. A representative set of titrations are plotted to illustrate this shift under 10 μM (green), 5 mM (yellow), 10 mM (blue), and 50 mM (orange) sodium bicarbonate conditions. Experiments were performed at room temperature (298 K). **C.** 10 μM CPII-R1 was titrated with sodium bicarbonate in the presence of 2mM ADP (the approximate K_D of CPII and ADP). Increasing the concentration of bicarbonate stabilizes the ADP bound conformation of CPII and increases the binding affinity of CPII for ADP.
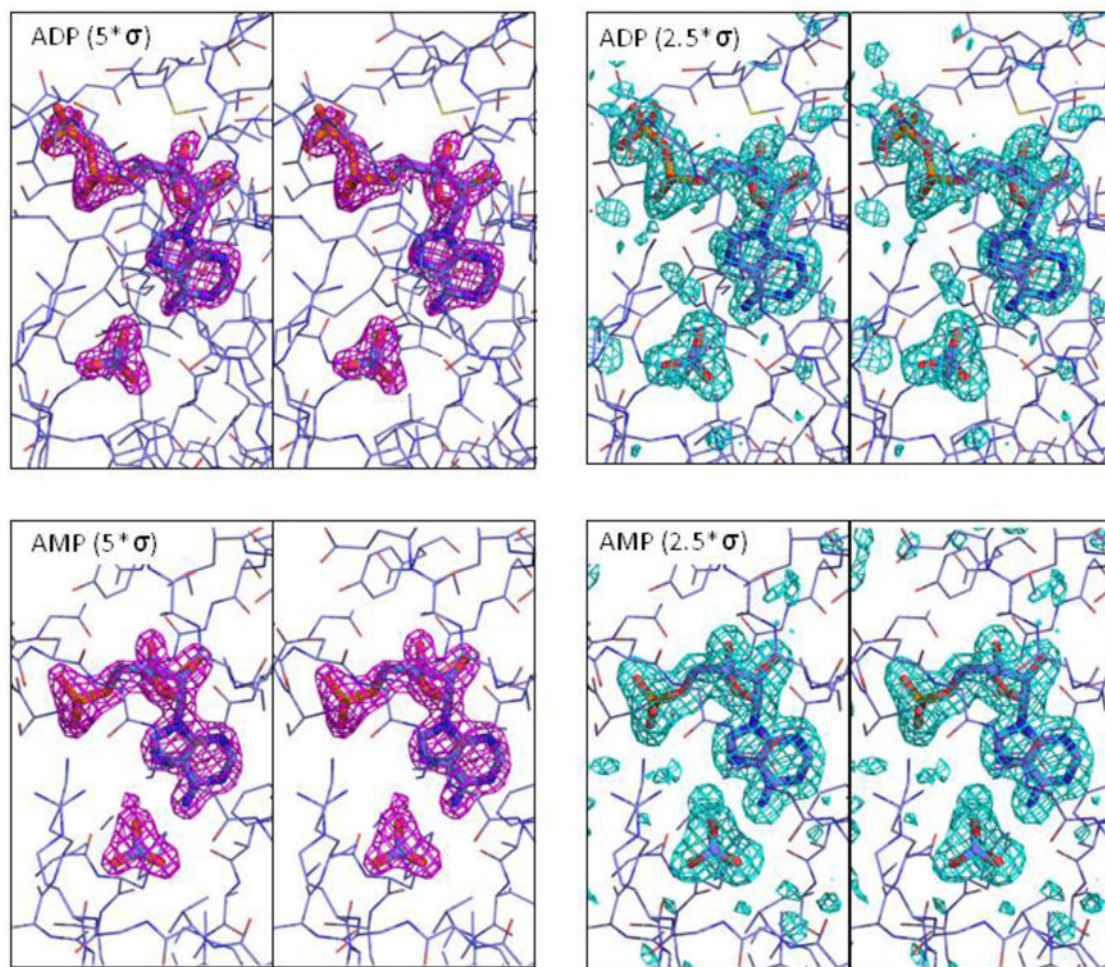
Figure 8. Electron density maps reveal features that can be unambiguously modeled by ADP, AMP and bicarbonate

Difference electron density maps are shown calculated as 'omit maps', with atoms belonging to the ligands removed from the phase calculation. Different binding sites (in crystal structure PDB ID 5D4O) showing occupancy of bicarbonate and either ADP (top) or AMP (bottom) are illustrated in stereo view at contour levels of 5 (left) and 2.5 (right) standard deviations above the mean density value.

Table 1

X-ray Diffraction Data and Atomic Refinement Statistics

PDB ID	5D4L	5D4N	5D4O	5DRK	5DS7	5D4P
Space group	P3	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell: a b c (Å)	82.7,82.7,31.5	51.5, 75.4, 79.6	51.4, 75.5, 80.0	51.9, 75.7, 80.3	51.5, 75.6, 79.9	49.8, 76.2, 87.9
Unit cell: α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 90	90, 90, 90	90, 90,90	90, 90, 90
Resolution (Å)	71.6 − 2.3	54.7 − 1.6	54.9 − 1.8	55.0 − 2.0	54.9 − 2.3	43.3 − 2.2
Data completeness (%)	97.4 (91.9)	99.7 (97.5)	99.3 (94.1)	99.70 (96.1)	99.7 (94.6)	97.6 (91.9)
Total reflections	121103 (7666)	267556 (19026)	191688 (11606)	160664 (10357)	160664 (10357)	98900 (6227)
Unique reflections	10491 (747)	41577 (2984)	29590 (2037)	13077 (910)	13077 (910)	17217 (1177)
Multiplicity	11.5 (10.3)	6.4 (6.4)	6.5 (5.7)	14.2 (13.1)	12.3 (11.4)	5.7 (5.3)
Rmerge	0.13	0.04	0.05	0.11	0.09	0.08
I/σ(I)	17.0 (3.2)	21.4 (1.9)	19.2 (2.7)	22.8 (3.5)	20.5 (3.1)	15.0 (3.0)
Refinement program	PHENIX	BUSTER-TNT	PHENIX	PHENIX	PHENIX	PHENIX
R, R-free	0.177, 0.234	0.211, 0.234	0.192, 0.229	0.230, 0.256	0.2439, 0.245	0.191, 0.247
R-free test set (reflections)	525, or 5.3%	2095, or 5.3%	2959, or 11.1%	1068, or 5.0%	694, or 5.0%	1720, or 11.1%
Bonds (Å) r.m.s.d. from ideal	0.008	0.010	0.007	0.013	0.018	0.008
Angles (°) r.m.s.d. from ideal	1.23	1.11	1.26	1.64	1.72	1.28
Non-hydrogen atoms						
Protein chains/atoms	2/1474	3/2242	3/2305	3/2128	3/2198	3/2341
ADP molecules	0	1	1	1	0	3
AMP molecules	0	1	1	1	3	0
Bicarbonate molecules	0	0	2	1	0	7*
Acetate molecules	0	2	0	0	0	0
Waters molecules	22	103	100	7	63	45
Number of residues	194	298	302	284	291	312
Wilson B-factor (Å²)	28.9	30.0	28.0	37.4	27.5	34.2
Ave. B-factor (Å²)	34.0	35.6	33.0	44.8	33.9	43.0
Ramachandran (% favored)	184 (97.9%)	276 (100%)	287 (99.0%)	263 (98.1%)	276 (98.9%)	300 (98.7%)
Ramachandran (% outliers)	1 (0.5%)	0 (0%)	0 (0%)	1 (0.4%)	0 (0%)	0 (0%)

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

*
Only three of the bicarbonates occupy positions believed to be biologically relevant (i.e. in the protein binding cleft). Others appear to represent fortuitous binding within the context of the crystal state.