Cyanobacteria, blue-green algae, are the most abundant autotrophs in aquatic environments and form the base of the food chain by fixing carbon and nitrogen into cellular biomass. To compensate for the low selectivity of Rubisco for CO₂ over O₂, cyanobacteria have developed highly efficient CO₂-concentrating machinery of which the ABC transport system CmpABCD from Synechocystis PCC 6803 is a component. Here, we have described the structure of the bicarbonate-binding protein CmpA in the absence and presence of bicarbonate and carbonic acid. CmpA is highly homologous to the nitrate transport protein NrtA. CmpA binds carbonic acid at the entrance to the ligand-binding pocket, whereas bicarbonate binds in nearly an identical location compared with nitrate binding to NrtA. Unexpectedly, bicarbonate binding is accompanied by a metal ion, identified as Ca²⁺ via inductively coupled plasma optical emission spectrometry. The binding of bicarbonate and metal appears to be highly cooperative and suggests that CmpA may co-transport bicarbonate and calcium or that calcium acts as a cofactor in bicarbonate transport.

Cyanobacteria are the most abundant microorganisms in aquatic environments and play a key role in the global carbon cycle (1). It is estimated that these photosynthetic microbes are responsible for ~50% of carbon fixation in the oceans. Over their 2.7-billion-year existence, cyanobacteria had to adapt to a changing gaseous environment where the levels of CO₂ declined and O₂ increased (2). Because O₂ can compete with CO₂ for binding to the carbon-fixing enzyme Rubisco (3), cyanobacteria evolved the most effective CO₂-concentrating mechanism (CCM)² that allows them to concentrate CO₂ levels around Rubisco up to 1000-fold. The CCM involves the import and accumulation of inorganic carbon as HCO₃⁻ in the cytoplasm and subsequent conversion to CO₂ in the protein microcompartment called the “carboxysome” via carbonic anhydrase.

One component of this CCM machinery in Synechocystis PCC 6803 is the cmpABCD operon that encodes a high affinity bicarbonate ABC transporter that is induced under low CO₂ conditions (4). This transporter is composed of four polypeptides, a high affinity solute-binding lipoprotein (CmpA), an integral membrane permease (CmpB), a cytoplasmic ATPase (CmpD), and an ATPase/solute-binding fusion protein (CmpC) that regulates transport (Fig. 1). The CmpABCD transporter is the highest affinity bicarbonate transporter of cyanobacteria (4). This affinity is predominantly conferred by the binding of bicarbonate to CmpA (K₅ = 5 µM) (5). CmpA is anchored to the periplasmic face of the cytoplasmic membrane via a lipid anchor attached to a conserved cysteine (6). The closest known homolog of CmpA is NrtA, the solute-binding protein of the nitrate-specific NrtABCD transporter that is 48% identical and 61% similar in amino acid sequence (7). We recently published an analysis (8) of the 1.6-Å structure of NrtA complexed with nitrate to elucidate the molecular determinants of nitrate specificity. From this structure, it seemed likely that the nitrate versus bicarbonate specificity was mainly due to the replacement of a lysine in the nitrate coordination sphere in NrtA with a glutamate in CmpA.

To better compare and contrast these two important ABC transport systems, the x-ray structure of CmpA has been determined in three different states, complexed with H₂CO₃ (carbonic acid) at pH 5.0, bound with HCO₃⁻ (bicarbonate) at pH 8.0, and in the absence of ligands at pH 8.0. The “C-clamp” structure of CmpA is remarkably homologous to that of NrtA (8) with ligands binding in the cleft between the two domains. At pH 5.0, nearly all of the bicarbonate is in the fully protonated, carbonic acid form. Under these conditions, the ligand is not bound deep inside the cleft, as was observed in NrtA, but rather at the entrance to the ligand-binding region. At pH 8.0, almost all of the dissolved inorganic carbon is in the bicarbonate form. In this case, the ligand is found deep inside the cleft of the C-clamp and, unexpectedly, is bound concomitantly with a calcium ion. Indeed, bicarbonate binds to CmpA if (and only if) calcium is also present. These and other results suggest that calcium and bicarbonate bind in...
Structure of a Cyanobacterial Bicarbonate Transport Protein

Cloning of cmpA from Synechocystis PCC 6803—The solute-binding domain of CmpA (residues 27–452) from Synechocystis PCC 6803 was cloned from genomic DNA. The cmpA gene was PCR-amplified with Platinum Pfx DNA polymerase (Invitrogen) according to the manufacturer’s instructions and standard cycling conditions. The forward primer 5′-GGGAA-TTCCATATGGCGGCAATCCCCCCGAT-3′ included an Ndel restriction site, and the reverse primer 5′-CCGCC-TCGAGTTAGACTTTTTTGATGCGAACACTTTTCGAG-3′ included an Xhol restriction site for cloning into pET-28a. The PCR product was purified with the QIAquick PCR purification kit (Qiagen) followed by digestion with both Ndel and Xhol at 37 °C overnight. The gene was separated from digestion by-products on a 1.0% agarose gel and purified by the QIAquick Gel purification kit (Qiagen). The purified cmpA fragment was then ligated into the expression vector pET-28a (Novagen). The pET-28a (Novagen) vector was previously modified such that the thrombin cleavage site was replaced with a recombinant tobacco etch virus (rTEV) cleavage site. *Escherichia coli* DH5α cells were transformed with the ligation mixture and then plated onto LB medium supplemented with 30 μg/ml kanamycin. Individual colonies were selected and cultured overnight, and plasmid DNA was extracted with the QIAprep Spin Miniprep kit (Qiagen). Positive clones were sequenced by Applied Science. Cells were lysed on ice by four cycles of sonication (45 s) separated by 3 min of cooling. Cellular debris was removed by centrifugation at 4 °C for 30 min at 30,000 × g. The clarified lysate was loaded onto a 5-ml HiTrap HP cartridge (GE Healthcare) charged with Ni²⁺ and pre-equilibrated with buffer A. After loading, the column was washed with ~40 ml of 90% buffer A/10% buffer B (25 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole, pH 8.0) followed by gradient elution of the protein from 10 to 100% buffer B. Protein-containing fractions were pooled based on SDS-PAGE and dialyzed against 20 mM HEPES (pH 7.5) with 200 mM NaCl overnight at 4 °C. For removal of the His tag, rTEV protease was added to dialyzed CmpA at a molar ratio of rTEV:CmpA of ~1:50 CmpA, and the mixture was incubated for 4 h at ~23 °C and then 16 h at 4 °C. The rTEV protease contained an N-terminal His₈ tag, which allowed the separation of cleaved CmpA from both His-tagged CmpA and rTEV via Ni²⁺ affinity chromatography, as described above. Cleaved CmpA was dialyzed overnight at 4 °C against 20 mM HEPES, pH 7.5, with 20 mM NaCl and then loaded onto a 5-ml HiTrap Q HP anion exchange cartridge (GE Healthcare) pre-equilibrated with the dialysis buffer. After loading, the column was washed with HEPES buffer until the A₂80 fell below ~0.1 OD, and then the inoculation of 2 × 4-liter baffled flasks containing TB (Terrific Broth) medium supplemented with 30 μg/ml kanamycin and 30 μg/ml chloramphenicol. The cells were grown at 37 °C with aeration to an A₆₀₀ of ~0.4, at which time the temperature was lowered to 22.5 °C for the remainder of the experiment. Thirty minutes after lowering the temperature, protein expression was induced by the addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. The cells were allowed to grow for an additional 16 h before harvesting by centrifugation. The cell paste was frozen in liquid nitrogen and stored at −80 °C.

Expression of the Selenomethionine-labeled Protein in *E. coli*—Selenomethionine-substituted protein was expressed via the methionine inhibitory pathway (9). Rosetta(DE3)pLysS (Novagen) cells were transformed with the pET28a-CmpA plasmid and plated onto LB medium supplemented with kanamycin. After ~16 h growth at 37 °C, the colonies were harvested from the plates and used for the inoculation of 2 × 4-liter baffled flasks containing M9 (42 mM Na₂HPO₄, 136 mM KH₂PO₄, 58 mM NaCl, 53.5 mM NH₄Cl) medium supplemented with 0.015 mM thiamine, 30 μg/ml kanamycin, and 30 μg/ml chloramphenicol. Cultures were grown at 37 °C to an A₆₀₀ of ~0.4 before the temperature was adjusted to 22.5 °C for the remainder of the growth. Cultures were grown to an A₆₀₀ of ~0.8 before each flask was supplemented with 200 mg each of L-lysine, L-threonine, and L-phenylalanine and 100 mg each of L-leucine, L-isoleucine, L-valine, and L-selenomethionine. After 30 additional min of growth, the cells were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside and allowed to grow for 16 h. Cultures were harvested by centrifugation, and the cell paste was frozen in liquid nitrogen for storage at −80 °C.

**Protein Purification**—Native and selenomethionine-substituted CmpA were purified in an identical manner. Approximately 10 g of frozen cells were thawed in 50 ml of cold buffer A (25 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole, pH 8.0) with one Complete EDTA-free protease inhibitor tablet (Roche Applied Science). Cells were lysed on ice by four cycles of sonication (45 s) separated by 3 min of cooling. Cellular debris was removed by centrifugation at 4 °C for 30 min at 30,000 × g. The clarified lysate was loaded onto a 5-ml HiTrap Q HP cartridge (GE Healthcare) charged with Ni²⁺ and pre-equilibrated with buffer A. After loading, the column was washed with ~40 ml of 90% buffer A/10% buffer B (25 mM NaH₂PO₄, 300 mM NaCl, 300 mM imidazole at pH 8.0) followed by gradient elution of the protein from 10 to 100% buffer B. Protein-containing fractions were pooled based on SDS-PAGE and dialyzed against 20 mM HEPES (pH 7.5) with 200 mM NaCl overnight at 4 °C. For removal of the His tag, rTEV protease was added to dialyzed CmpA at a molar ratio of rTEV:CmpA of ~1:50 CmpA, and the mixture was incubated for 4 h at ~23 °C and then 16 h at 4 °C. The rTEV protease contained an N-terminal His₈ tag, which allowed the separation of cleaved CmpA from both His-tagged CmpA and rTEV via Ni²⁺ affinity chromatography, as described above. Cleaved CmpA was dialyzed overnight at 4 °C against 20 mM HEPES, pH 7.5, with 20 mM NaCl and then loaded onto a 5-ml HiTrap Q HP anion exchange cartridge (GE Healthcare) pre-equilibrated with the dialysis buffer. After loading, the column was washed with HEPES buffer until the A₂80 fell below ~0.1 OD, and then the
protein was eluted with a gradient of 20–500 mM NaCl in 20 mM HEPES, pH 7.5. Protein-containing fractions were pooled based on SDS-PAGE and dialyzed against 20 mM HEPES (pH 7.5) with 100 mM NaCl. For crystallization, CmpA was concentrated to \( \frac{1}{11011} \) mg/ml based on an extinction coefficient of \( \frac{1}{18528} \) ml/cm \( \mu g \) as calculated by the program Protean (DNastar, Inc., Madison, WI).

Crystallization of Native and Selenomethionine-labeled CmpA—In the first set of crystallization experiments, 25 mM Na2CO3 was added to CmpA prior to crystallization. Potential crystallization conditions were screened using the Hampton Screen at both room temperature and 4 °C via the hanging drop method of vapor diffusion. Single crystals were observed at room temperature from 2M ammonium sulfate. Crystallization conditions were refined, and large single crystals of CmpA (native and selenomethionine-substituted) complexed with carbonate were obtained by streak seeding of the hanging drop crystals into batch plates containing 11 mg/ml CmpA, 50 mM succinate, pH 5.0, 1.1–1.5 M ammonium sulfate, and 12 mM Na2CO3. Four additional structures of CmpA were determined at pH 8.0 using the following combinations of additives, NaHCO3, EGTA, NaHCO3, EGTA, and no additives. These crystals were obtained by streak seeding the Na2CO3-containing native crystals into batch plates containing 11 mg/ml CmpA, 50 mM HEPPS, pH 8.0, 1.4–1.6 M ammonium sulfate, and 5 mM EGTA and/or 15 mM NaHCO3 where applicable. Crystals grew to dimensions of \( \frac{1}{0.5} \times \frac{0.3}{0.15} mm \) in 2–3 weeks. All of the various CmpA crystals were triclinic, with similar unit cell dimensions as noted in Tables 1 and 2. The solvent content of the crystals was 50%, with one molecule in the asymmetric unit.

High Resolution X-ray Data Collection—Both the native and selenomethionine-substituted protein crystals were flash-frozen in the same manner. Briefly, crystals were harvested from the batch plates and soaked for several hours in a synthetic mother liquor composed of 100 mM succinate (pH 5.0) or 100 mM HEPPS (pH 8.0), 1.5–1.7 M ammonium sulfate, and 300 mM NaCl. The additives EGTA, NaHCO3, or Na2CO3 were included in all synthetic mother liquor and cryoprotectant solutions, when applicable, at 5 mM above the concentrations used for crystallization. Crystals were serially transferred to a final cryoprotectant solution containing

### Table 1

| Data collection | P1 | Se-Met CmpA | P1 |
|----------------|----|-------------|----|
| Space group    | P1 | P1          | P1 |
| Cell dimensions| a, b, c (Å) \( \frac{44.67, 49.62, 57.86}{87.41, 80.52, 75.29} \) | a, b, c (Å) \( \frac{44.69, 49.62, 57.96}{87.31, 80.27, 75.26} \) | a, b, c (Å) \( \frac{44.77, 49.73, 57.84}{87.44, 80.37, 75.16} \) |
| α, β, γ (°)    | 64 (7.3) | 63 (7.6) | 6.3 (10.7) |
| L/σL           | 23.0 (1.85) | 23.0 (1.73) | 16.6 (7.85) |
| Completeness (%)| 89.1 (92.3) | 89.5 (90.7) | 94.4 (89.2) |
| Redundancy     | 2.5 (2.4) | 2.5 (2.4) | 2.5 (2.5) |
| Peak Inflection| 0.97899 | 0.97915 | 0.97132 |
| Remote         | 50-2.18 (2.26-2.18) | 50-2.19 (2.27-2.19) | 50-2.17 (2.25-2.17) |
| Wavelength     | 0.97915 | 0.97915 | 0.97132 |
| Resolution (Å) | 50-2.18 (2.26-2.18) | 50-2.19 (2.27-2.19) | 50-2.17 (2.25-2.17) |

### Table 2

| Data collection and refinement statistics (molecular replacement) |
|-----------------------|--------------------|--------------------|--------------------|
| PDB accession ID      | CmpA, pH 5.0 H2CO3 | CmpA, pH 8.0 Apo | CmpA, pH 8.0 HCO3−/Ca2+ (EGTA) |
| Space group           | P1                 | P1                 | P1                 |
| Cell dimensions       | a, b, c (Å) \( \frac{44.54, 49.64, 57.78}{87.33, 80.50, 75.31} \) | a, b, c (Å) \( \frac{44.66, 49.68, 57.57}{87.57, 81.20, 75.20} \) | a, b, c (Å) \( \frac{44.79, 49.57, 57.74}{87.67, 81.23, 75.24} \) |
| α, β, γ (°)           | 7.7 (7.0)          | 3.6 (21.9)         | 4.2 (19.6)         |
| Resolution (Å)        | 40-1.6 (1.66-1.6)  | 25.6-1.35 (1.45-1.35) | 24.6-1.35 (1.44-1.35) |
| Rmerge or Rref (°)    | 22.9 (14.1)        | 29.8 (60.0)        | 24.9 (53.0)        |
| Completeness (%)      | 92.4 (89.4)        | 99.2 (98.1)        | 98.9 (98.3)        |
| Redundancy            | 2.6 (2.2)          | 7.2 (3.7)          | 6.5 (3.6)          |
| No. reflections       | 56,952             | 103,049            | 102,903            |
| Rwork or Rfree (%)    | 17.5/19.2          | 17.9/19.3          | 20.8/22.3          |
| No. atoms             | 3123               | 3127               | 3139               |
| Protein               | 4                  | 0                  | 5                  |
| Water                 | 415                | 411                | 444                |
| B-factors             | 11.9               | 9.2                | 11.0               |
| Protein               | 28.8               | N/A                | 13.8               |
| Water                 | 22.3               | 19.2               | 19.9               |
| Root mean square deviations | 0.005         | 0.005              | 0.005              |
| Bond lengths (Å)      | 1.3                | 1.3                | 1.3                |
| Bond angles (°)       | 1.3                | 1.3                | 1.3                |
100 mM succinate, pH 5.0, or 100 mM HEPES, pH 8.0, 300 mM NaCl, 1.5–1.7 M ammonium sulfate, 21–25% ethylene glycol, and the appropriate amount of additive. The crystals were then flash-cooled by submersion in liquid nitrogen. X-ray data sets from the selenomethionine-substituted crystals and native CmpA complexed with Na₂CO₃ were collected on a “SBC3” charge-coupled detector at the Structural Biology Center 19-BM beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL). The x-ray data were collected and processed using a Bruker Proteum charge-coupled detector system. X-ray data on the remaining three complexes (EGTA, EGTA, and NaHCO₃ added, and no additives) were collected by Dr. Matthew Benning at Bruker AXS, Inc. in Madison, WI. The data were collected on a Microstar generator (running at 2.7 kW) with Helios optics. The detector was a Pt-135 charge-coupled device. Data reduction was performed using the programs SAINT and SADABS from the Proteum2 software suite (Bruker AXS Inc., 2006).

X-ray Structural Analyses—The structure of CmpA was solved via multwavelength anomalous dispersion phasing with x-ray data collected from the selenomethionine-substituted protein crystals (Table 2). The software package SOLVE (11) was used to determine and refine the positions of the selenium atoms, and the program RESOLVE (12) was then used to perform solvent flattening and initial protein model building. RESOLVE successfully built 356 of 452 residues for CmpA, and the remaining residues were added manually using the graphics package O (13). Alternate cycles of maximum likelihood refinement with crystallography NMR software and manual model building reduced the R-factor to <20% for all five models. Relevant data collection and refinement statistics are summarized in Table 2. All five models yielded similar Ramachandran analyses; 90% of residues were in the most favored regions, 9.1% in the additionally allowed regions, 0.9% in the generously allowed regions, and 0% in the disallowed regions. In all five models, three residues are in the generously allowed regions of the backbone conformation, Asn-150, Glu-270, and Tyr-319. However, the high resolution electron density for these residues is unambiguous and consistent with the model. The peptidic carbonyl oxygen of Asn-150 hydrogen bonds to Asn-152 or Glu-270, and both are involved in Ca²⁺ binding. Similarly, Glu-270 is directly involved in Ca²⁺ binding and is in a random coil region connecting two β-strands that line the back of the solute-binding pocket. Tyr-319 is part of a surface loop that connects two α-helices.

FIGURE 2. a, stereo ribbon diagram of the CmpA crystal structure at pH 5.0 with carbonic acid. The protein is gradiently colored blue to red as the chain extends from the N to the C terminus. The view is of the front of the C-clamp, which opens to the ligand-binding cleft. Carbonic acid is depicted as spheres. b, stereo ribbon diagram of the CmpA crystal structure at pH 8.0 with Ca²⁺ and bicarbonate, with the identical view and coloring as in a. Bicarbonate and Ca²⁺ are depicted as spheres. Note the difference in the position of carbonic acid and bicarbonate in the ligand-binding cleft.
**Structure of a Cyanobacterial Bicarbonate Transport Protein**

Methodology for Metals Determination—Determination of the calcium and nickel content in the protein was performed using inductively coupled plasma optical emission spectrometry (ICP-OES). A Thermo Jarrell-Ash IRIS-AP spectrometer was used. Calcium measurements were taken at wavelengths of 393.366 and 396.847 nm, and nickel content was measured at 231.604 nm. The sample was run after 10-fold dilution with de-ionized water. Duplicate reagent blanks were also run. Instrument detection limits for both calcium and nickel were <1 ng/ml. A semiquantitative scan for other elements present was performed using inductively coupled plasma mass spectrometry (ICP-MS). A ThermoElemental PQ Excel ICP-MS instrument was used. This determination revealed only the presence of iron (≈3.5 µg/ml) and zinc (≈0.2 µg/ml) at appreciable concentrations.

**RESULTS**

The overall molecular structure of CmpA with carbonic acid is shown in Fig. 2. CmpA is an α-β protein and belongs to the periplasmic-binding protein superfamily (14). Similar to the related nitrate-binding protein NrtA, CmpA is composed of two domains (I and II) organized in a C-clamp shape. The core of each domain consists of a five-stranded mixed β-sheet flanked by α-helices. Domain I, which is slightly larger than domain II, is composed of three different segments of the polypeptide chain, 54–150, 267–321, and 350–426. This domain contains β1–β4 and β10 arranged as a mixed β-sheet, flanked by α1–α4, α8–α10, and α12–α14. Similarly, domain II is composed of several portions of the polypeptide chain, 151–266, 322–349, and 427–452. The fold of domain II is similar to that of domain I with β5–β9 forming a mixed β-sheet flanked by α4–α7, α11, and α15. In addition, domain II contains two antiparallel β-strands, 11 and 12, that are adjacent to the core motif. The two domains of CmpA are connected by random coil elements and flanked on either side of the solute-binding cleft by α-helices.

Based on the order of the β-strands in each domain (21354), CmpA belongs to class II of the periplasmic-binding protein superfamily (14), which also includes the oxyanion-binding proteins specific for sulfate, phosphate, and molybdate (15–19). However, CmpA, similar to NrtA, is much larger (~100 amino acids) than these oxyanion-binding proteins. The last 100 residues of CmpA are mostly in an α-helical conformation with the exception of the 2-stranded antiparallel β-sheet in the C-terminal domain. These terminal residues form α-helices that wrap around the back of the structure and cradle the two α/β domains. As observed in NrtA, the surface charge distribution on the face opposite to the ligand-binding cleft of CmpA is almost entirely composed of acidic residues. Other solute-binding proteins (e.g. zinc, iron, sulfate, and phosphate) have a similar charge distribution and, during our analysis of the NrtA structure, led us to suggest that such a negative charge may facilitate delivery of solute to the transmembrane pore by limiting unproductive interactions of the solute-binding protein with the phospholipid membrane.

The major difference between the CmpA and NrtA structures is that the CmpA clamp is in a more open conformation than that of NrtA. This suggests that CmpA, at pH 5.0 ($R_{work} = 17.5\%$, $R_{free} = 19.2\%$), is in the unliganded form. However, electron density consistent with a CO3 molecule was observed near, but not in, the solute-binding cleft (Figs. 2–4).
At pH 5.0, nearly all of the solute is in the protonated, carbonic acid form. Therefore, it appears that carbonic acid does not bind in the expected ligand-binding pocket and, in turn, does not induce closure of the C-clamp. Interestingly, the carbonyl oxygen of carbonic acid is within 2.6 Å of the N\(^{\text{2}}\) atom of Trp-245 but does not hydrogen bond with any other residue in the protein. Bicarbonate is the predominant carbon species in the aquatic environment of *Synechocystis* and binds with a dissociation constant of \(\sim 5 \mu M\). Therefore, it is unlikely that this carbonic acid/CmpA interaction occurs in nature. However, it is also possible that the position of carbonic acid represents an initial interaction between bicarbonate and the exterior loops of the C-terminal domain to facilitate delivery of solute to the binding pocket. This kind of “preloading” of ligand to exterior loops of the transport protein was suggested in the case of the zinc-binding protein ZnuA (20).

To determine the structure of CmpA with its authentic ligand, bicarbonate, the pH of the crystallization medium was increased to pH 8.0 (\(R_{\text{work}} = 17.9\%\), \(R_{\text{free}} = 19.3\%\)). At this pH, nearly all of the dissolved CO\(_2\) is in the bicarbonate form. It is important to note that, at this pH, bicarbonate is observed in the ligand-binding cleft whether or not it was added to the crystallization mother liquor. In both cases, bicarbonate is wedged in the binding cleft created by the juxtaposition of the two globular domains of CmpA and causes the C-clamp to close over the bound ligands (Figs. 2–4). Unexpectedly, a metal ion is also observed adjacent to the bicarbonate and bound by several residues from both the N- and C-terminal domains. There is no previous evidence suggesting that bicarbonate binding and transport by CmpA is metal-dependent. The metal ion is coordinated by five amino acid side chains and one bicarbonate oxygen in a distorted octahedral arrangement. The axial ligands are provided by side chain oxygens of Glu-270.
Structure of a Cyanobacterial Bicarbonate Transport Protein

(2.3 Å), Glu-271 (2.4 Å), Asn-152 (2.5 Å), and a carboxyl oxygen (2.5 Å) from bicarbonate. The octahedral geometry of the metal coordination is completed by side chain oxygens from Glu-70 and Gln-198, which are within 2.3 and 2.4 Å of the metal, respectively. The average nonbonded distance (2.4 Å) immediately eliminated a water molecule at the center of the coordination sphere and was instead highly suggestive of a metal.

Both the chemistry and geometry of the metal-oxygen coordination sphere initially suggested that the bound metal was either Ca$^{2+}$ or Na$^+$ (21, 22). It is important to note that this putative metal was not observed in the structure at pH 5.0, and the nature of the metal/protein interactions are unlikely to be affected by the shift from pH 5.0 to 8.0. To ascertain which is the more likely metal bound to CmpA, both metals were modeled into the density and unrestrained B-value refinement was performed in crystallography NMR software. Ca$^{2+}$ refined to a B-value of 17.1 Å$^2$, nearly identical to the average B-value of the surrounding atoms (16.1 Å$^2$). In contrast, when Na$^+$ was refined at this position, the resulting B-value was, as expected, a below average value of 6.3 Å$^2$. The contention that the bound metal was a Ca$^{2+}$ ion was further supported by analysis performed using the Metalloprotein Data Base and browser (Scripps Research Institute). Although both Na$^+$ and Ca$^{2+}$ are observed to have similar ligand-metal distances, it is not as common to find Na$^+$ ions coordinated by six ligands, and when it does occur, the ligating moieties are primarily water molecules. In contrast, Ca$^{2+}$ is most commonly observed octahedrally coordinated by oxygens contributed from side chains such as Glu, Asp, Asn, and Gln. Indeed, ICP-OES revealed that such are essentially those predicted from sequence homology (8). There are, however, some differences between CmpA and NrtA (8). There are, however, some differences between these two states are found in residues 189–194 that comprise a surface loop proximal to the binding cavity (Fig. 4a). This loop includes Thr-192 that hydrogen bonds with bicarbonate and moves 1.3 Å into the ligand-binding cleft upon solute binding. This movement of Thr-192 creates a hydrogen-bonding donor to the O3 of bicarbonate (Fig. 4b).

To further examine this metal-bicarbonate interaction, CmpA crystals were grown at pH 8.0 in the presence of 10 mM EGTA and an additional 20 mM NaHCO$_3$. The resulting 1.35-Å structure of CmpA refined with an R$_{work}$ and R$_{free}$ of 20.9 and 22.2%, respectively. Unexpectedly, CmpA was observed in the “closed” conformation with both Ca$^{2+}$ and bicarbonate in the binding cleft. The coordinates from this structure and those of CmpA with bicarbonate and Ca$^{2+}$, obtained without crystallization additives, are nearly identical, with a root mean square deviation of 0.2 Å over all 402 residues. Together, these results suggest that metal and bicarbonate bind to CmpA in a strongly cooperative manner.

**DISCUSSION**

CmpA is a member of the periplasmic binding protein superfamily and consists of two $\alpha/\beta$ domains organized as a C-clamp with solute occupying the cleft created between the two domains. Unlike some other members of this family, CmpA does not undergo a dramatic conformational change upon solute binding. This suggests that the binding cavity is relatively rigid and that solute binding does not require extensive rearrangement of the ligating residues.

The closest homologue of CmpA is the nitrate-binding protein NrtA that shares 48 and 61% amino acid sequence identity and similarity, respectively. We previously reported the 1.6-Å crystal structure of NrtA from *Synechocystis* PCC 6803 complexed with nitrate (8). The structures of CmpA and NrtA can be superimposed with a root mean square deviation of 1.1 Å for 344 C-α atoms (Fig. 4a). The most significant difference in the backbone of these proteins lies in residues 220–226 (see loop noted with a *star* in Fig. 4a). In the structure of NrtA, this hydrophobic loop blocks solvent access to the nitrate-binding cavity. In contrast, the solute-binding cavity of CmpA remains slightly more open with bicarbonate partially exposed to solvent. The residues in CmpA that contact the bound bicarbonate are essentially those predicted from sequence homology between CmpA and NrtA (8). There are, however, some differences in how CmpA and NrtA bind their respective ligands. Nitrate and bicarbonate occupy the same general position, but

---

*3 N. M. Koropatkin and T. J. Smith, unpublished results.*
nitrates are buried ~1.5 Å deeper in the NrtA binding cleft and partially occupies the space taken by the Ca\(^{2+}\) atom in CmpA. In both structures, the negative charge on bicarbonate and nitrate is polarized at the O1 oxygen of each anion, but this charge is stabilized differently in each protein. In NrtA, the N\textsuperscript{ε} of Lys-269 is located within 2.8 Å of the O1 oxygen, stabilizing the negative charge (Fig. 4c). In CmpA, this residue is replaced by Glu-271 (Fig. 4b). In our analysis of the NrtA crystal structure, we predicted that the exchange of Lys for Glu was necessary to provide a hydrogen bonding acceptor to the protonated oxygen of bicarbonate. This is, in fact, the case in CmpA with 0.6 residues in the coordination sphere and the average metal-oxygen distance (2.4 Å). This was clarified by ICP-OES analysis that identified calcium as the major metal ion associated with the protein with 0.6 ± 0.04 mol of Ca\(^{2+}\) bound/mol of CmpA. Because none of the purification or crystallization reagents contained significant amounts of calcium, it is possible that CmpA acquired Ca\(^{2+}\) during expression in E. coli and carried it through subsequent purification steps. This has been observed for some periplasmic metal-binding proteins (e.g. ZnuA, FutA1) that bind their ligand with a K\textsubscript{d} of <10\(^{-6}\) M. Concomitant binding of both metal and anion has been observed with other periplasmic-binding proteins. Synergistic Fe\(^{3+}\) binding with phosphate or CO\(_3\)\(^{-}\) has been observed for both the Hae- 

**Structure of a Cyanobacterial Bicarbonate Transport Protein**

because calcium may be used to either increase the binding affinity or selectivity of bicarbonate over the other anions such as nitrate. In this way, calcium could be a "coenzyme" that facilitates transport without being an important transported nutrient itself. It is interesting to note that only the fresh water strains of *Synechocystis* have the *cmp* operon and that fresh water has lower levels of bicarbonate and calcium (0.96 and 0.38 mmol/kg, respectively) compared with ocean water (2.39 and 10.3 mmol/kg, respectively) (30, 31). Although it is tempting to speculate that the calcium bicarbonate complex is transported into the cell, there is currently no evidence that suggests that CmpABCD is involved in such a process. Indeed, there is some evidence that calcium is transported in bacteria, however, via a number of genes unrelated to the *cmpABCD* operon (32). Further studies are clearly necessary to understand the role of calcium in the facilitated transport of bicarbonate.

**REFERENCES**

1. Whitton, B. A., and Potts, M. (eds) (2000) *The Ecology of Cyanobacteria: Their Diversity in Time and Space*, Kluwer Academic Publishers, Norwell, MA
2. Gould, S. J. (2001) *The Book of Life: An Illustrated History of the Evolution of Life on Earth*, 2nd Ed., W. W. Norton & Co., New York
3. Badger, M. R., and Price, G. D. (2003) *J. Exp. Bot.* 54, 609–622
4. Omata, T., Price, G. D., Badger, M. R., Okamura, M., Gohta, S., and Ogawa, T. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 13571–13576
5. Maeda, S., Price, G. D., Badger, M. R., Enomoto, C., and Omata, T. (2000)
Structure of a Cyanobacterial Bicarbonate Transport Protein

6. Omata, T., and Ogawa, T. (1986) Plant Physiol. 80, 525–530
7. Omata, T. (1995) Plant Cell Physiol. 36, 207–213
8. Koropatkin, N. M., Pakrasi, H. B., and Smith, T. J. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 9820–9825
9. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) J. Mol. Biol. 229, 105–124
10. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
11. Terwilliger, T. C., and Berendzen, J. (1999) Acta. Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
12. Terwilliger, T. C. (2000) Acta. Crystallogr. Sect. D Biol. Crystallogr. 56, 965–972
13. Jones, T. A., Zou, J.-Y., and Cowan, S. W. (1991) Acta. Crystallogr. Sect. A 47, 110–119
14. Dwyer, M. A., and Hellinga, H. W. (2004) Curr. Opin. Struct. Biol. 14, 495–504
15. Wang, Z., Choudhary, A., Ledvina, P. S., and Quiocio, F. A. (1994) J. Biol. Chem. 269, 25091–25094
16. Vyas, N. K., Vyas, M. N., and Quiocio, F. A. (2003) Structure (Lond.) 11, 765–774
17. Pfugrath, J. W., and Quiocio, F. A. (1988) J. Mol. Biol. 200, 163–180
18. Lawson, D. M., Williams, C. E., Mitchenall, L. A., and Pau, R. N. (1998) Structure (Lond.) 6, 1529–1539
19. Hu, Y., Rech, S., Gunsalus, R. P., and Rees, D. C. (1997) Nat. Struct. Biol. 4, 703–707
20. Banerjee, S., Wei, B., Bhattacharyya-Pakrasi, M., Pakrasi, H. B., and Smith, T. I. (2003) J. Mol. Biol. 333, 1061–1069
21. Harding, M. M. (2000) Acta. Crystallogr. Sect. D Biol. Crystallogr. 56, 857–867
22. Harding, M. M. (2001) Acta. Crystallogr. Sect. D Biol. Crystallogr. 57, 401–411
23. Bruns, C. M., Anderson, D. S., Vaughan, K. G., Williams, P. A., Nowalk, A. J., McRee, D. E., and Mietzner, T. A. (2001) Biochemistry 40, 15631–15637
24. Guo, M., Harvey, I., Yang, W., Coghill, L., Campopiano, D. J., Parkinson, J. A., MacGillivray, R. T., Harris, W. R., and Sadler, P. J. (2003) J. Biol. Chem. 278, 2490–2502
25. Shouldice, S. R., Skene, R. J., Dougan, D. R., McRee, D. E., Tari, L. W., and Schryvers, A. B. (2003) Biochemistry 42, 11908–11914
26. Bruns, C. M., Nowalk, A. J., Arvai, A. S., McTigue, M. A., Vaughan, K. G., Mietzner, T. A., and McRee, D. E. (1997) Nat. Struct. Biol. 4, 919–924
27. Nowalk, A. J., Tencza, S. B., and Mietzner, T. A. (1994) Biochemistry 33, 12769–12775
28. Homann, P. H. (2002) Photosynth. Res. 73, 169–175
29. Zhao, Y., Shi, Y., Zhao, W., Huang, X., Wang, D., Brown, N., Brand, J., and Zhao, J. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 5744–5748
30. Livingston, D. A. (1963) in Data of Geochemistry (Fleischer, M., ed) 6th Ed., USGS Prof. Pap. 440-G
31. Holland, H. D. (1978) The Chemistry of the Atmosphere and Oceans, John Wiley & Sons, Inc., New York
32. Berkelman, T., Garret-Engele, P., and Hoffman, N. E. (1994) J. Bacteriol. 176, 4430–4436