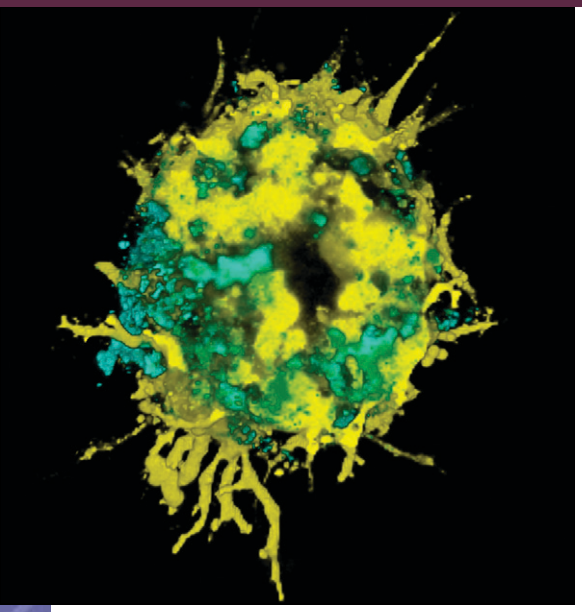
July 13, 2010 | vol. 107 | no. 28 | 12407–12734

PNAS

Proceedings of the National Academy of Sciences of the United States of America www.pnas.org



Cover image: Pictured are small blood vessels that have formed in vitro from endothelial cells. A type of vascular abnormality known as cerebral cavernous malformation (CCM) has been linked to mutations in one of three genes that contribute to neurovascular development. Joycelyn Wüstehube et al. show that *CCM1* (also known as *KRIT1*) keeps the endothelium quiescent (blue cells compared to yellow control cells) by preventing uncontrolled angiogenesis, and that *CCM1*-silenced human endothelial cells recapitulate hallmark CCM vascular irregularities in mice. See the article by Wüstehube et al. on pages 12640–12645. Image courtesy of Arne Bartol and Joycelyn Wüstehube.

From the Cover

Contents

Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate

Meng-Xi Zhao[a], Yong-Liang Jiang[a], Yong-Xing He[a], Yi-Fei Chen[b], Yan-Bin Teng[a], Yuxing Chen[a], Cheng-Cai Zhang[b,1], and Cong-Zhao Zhou[a,1]

[a]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, China; and [b]Aix Marseille Université and Laboratoire de Chimie Bactérienne, Unités Propres de Recherche 9043, Centre National de la Recherche Scientifique, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France

Edited* by Robert Haselkorn, University of Chicago, Chicago, IL, and approved May 28, 2010 (received for review February 12, 2010)

2-oxogluatarate (2-OG), a metabolite of the highly conserved Krebs cycle, not only plays a critical role in metabolism, but also constitutes a signaling molecule in a variety of organisms ranging from bacteria to plants and animals. In cyanobacteria, the accumulation of 2-OG constitutes the signal of nitrogen starvation and NtcA, a global transcription factor, has been proposed as a putative receptor for 2-OG. Here we present three crystal structures of NtcA from the cyanobacterium *Anabaena*: the apoform, and two ligand-bound forms in complex with either 2-OG or its analogue 2,2-difluoropentanedioic acid. All structures assemble as homodimers, with each subunit composed of an N-terminal effector-binding domain and a C-terminal DNA-binding domain connected by a long helix (C-helix). The 2-OG binds to the effector-binding domain at a pocket similar to that used by cAMP in catabolite activator protein, but with a different pattern. Comparative structural analysis reveals a putative signal transmission route upon 2-OG binding. A tighter coiled-coil conformation of the two C-helices induced by 2-OG is crucial to maintain the proper distance between the two F-helices for DNA recognition. Whereas catabolite activator protein adopts a transition from off-to-on state upon cAMP binding, our structural analysis explains well why NtcA can bind to DNA even in its apoform, and how 2-OG just enhances the DNA-binding activity of NtcA. These findings provided the structural insights into the function of a global transcription factor regulated by 2-OG, a metabolite standing at a crossroad between carbon and nitrogen metabolisms.

crystal structure | cyanobacteria | transcription activator | DNA recognition

2-oxoglutarate (2-OG) is the carbon skeleton for the assimilation of nitrogen. It is derived from the Krebs cycle, a highly conserved central metabolic pathway, and stands at the crossroad between carbon and nitrogen metabolism. Interestingly, increasing evidence indicates that 2-OG is not only a central metabolite but also a signaling molecule in a variety of organisms, including bacteria (1), plants (2), and animals (3). One such example is provided by studies in cyanobacteria *Anabaena* sp. PCC 7120, where it has been shown that the accumulation of 2-OG in vivo constitutes the nitrogen starvation signal and thus triggers a series of cellular responses by readjusting cellular metabolism (4). It has been proposed NtcA, a transcription factor belonging to the Crp-Fnr family (cAMP receptor protein and fumarate/nitrate reductase regulator, Pfam code PF00325), could be one of the receptors of 2-OG because 2-OG enhances the DNA-binding activity of NtcA (4, 5). However, no evidence is available to show that 2-OG is indeed a ligand for NtcA.

Transcription factors in the Crp-Fnr family are widespread among bacteria. They regulate a large spectrum of cellular activities, including carbon or nitrogen metabolism, redox or stress responses, etc., enabling bacteria to adapt to diverse environmental changes. They are usually characterized by a helix-turn-helix DNA-binding domain (DBD) at the C-terminus (6). Based on their sequences and functions, members of this family were originally classified into three groups, namely Crp, Fnr, and the nitrogen responsive regulator NtcA (7). The most widely known member is *Escherichia coli* cAMP receptor protein (CRP) or catabolite activator protein (CAP), which, in response to levels of cAMP, controls the process of catabolic repression (8). The structure of *E. coli* CAP in complex with cAMP, determined in 1981, was the first of its kind among the Crp-Fnr family members (9), and has served as the basis for understanding the molecular mechanisms of transcription factors activated by small effectors. The active form of CAP exists as a homodimer (10, 11), with each subunit containing an N-terminal cAMP-binding domain and a C-terminal DBD bridged together by a long helix (C-helix) (9). Information concerning the structure of the apoform of CAP was not available until the NMR (12) and the crystal (13) structures were reported in March 2009 and June 2009, respectively. These two structures at last made it possible to begin appreciating the precise allosteric mechanisms behind activation of CAP upon cAMP binding. Despite a discrepancy in the allosteric mechanisms deduced, both structures indicated an off-to-on state of CAP triggered by cAMP. Comparison between the NMR structure of apo-CAP (12) and the crystal structures of CAP-cAMP (14) and CAP-cAMP-DNA (15) indicated that the two C-helices are subject to a coil-to-helix transition upon cAMP binding (12). The larger dimeric interface resulting from the three regenerated turns of the C-helix decreases the relative distance between the two DNA recognition helices (F helices) from 41 Å in the apoform to 34 Å in the cAMP-bound form, which fits perfectly into a turn of DNA major groove (12). In an alternative model deduced from the crystal structure of apo-CAP (13), the F-helices are wrapped in the core of DBDs in the absence of cAMP, which results in a low DNA-binding affinity.

In contrast to the mounting structural investigations on CAPs, no information on the structure of either Fnr or NtcA groups has been reported. The representative member NtcA of one of these groups is a global nitrogen regulator found in all cyanobacteria sequenced to date, such as *Nostoc* spp. (including *Anabaena* sp. PCC 7120), *Cyanothece* sp., *Plectonema boryanum*, *Prochlorococcus* sp., *Synechococcus* sp., and *Trichodesmium* sp. (6). It regulates the expression of a large number of genes, especially those

Author contributions: Y.C., C.-C.Z., and C.-Z.Z. designed research; M.-X.Z., Y.-F.C., and Y.-B.T. performed research; M.-X.Z., Y.-L.J., Y.-X.H., Y.C., C.-C.Z., and C.-Z.Z. analyzed data; and M.-X.Z., Y.-L.J., Y.C., C.-C.Z., and C.-Z.Z. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 3LA7, 3LA2, and 3LA3).

[1]To whom correspondence may be addressed. E-mail: zcz@ustc.edu.cn or cczhang@ifr88.cnrs-mrs.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1001556107/-/DCSupplemental.

involved in nitrogen metabolism, as well as other aspects of cyanobacterial metabolism, such as carbon metabolism, photosynthesis, and stress responses (16). *Anabaena* sp. PCC 7120 and many other cyanobacteria can also fix nitrogen via a reaction sequestered in a particular cell type, the heterocysts, which are formed along the filament in response to combined nitrogen starvation (17). This unique cell differentiation strategy enables cyanobacteria to take advantage of both oxygen-sensitive nitrogen fixation (in heterocysts) and oxygen-producing photosynthesis (in vegetative cells) (5, 18). NtcA plays crucial roles at different stages of heterocyst differentiation (5). The consensus NtcA-binding site adopts a palindromic sequence pattern, TGT-(N_9 or N_{10})-ACA or GTA-(N_8)-TAC (19–21). The DNA-binding activity of NtcA is regulated by the accumulation of the metabolite 2-OG (4, 22–24). Because of the absence of 2-OG dehydrogenase in cyanobacteria (25), the 2-OG produced by the Krebs cycle serves mainly as a carbon skeleton for nitrogen assimilation through the glutamine synthetase-glutamate synthase pathway. When there is a shortage of nitrogen, 2-OG transiently accumulates in *Anabaena* (26), the increased level of which is sensed by NtcA as a signal of nitrogen starvation. This phenomenon has been mimicked by the accumulation of a nonmetabolizable analogue of 2-OG, 2,2-difluoropentanedioic acid (DFPA), which was sufficient to trigger the differentiation of *Anabaena* heterocysts (4). Moreover, the presence of 2-OG was required for the NtcA-dependent transcriptional initiation as indicated by in vitro transcription assays (27).

To date, information on the binding site of 2-OG or DFPA and on the activation mechanism of NtcA by these effectors remains undetermined. A set of preliminary X-ray crystallographic data of NtcA indicated that it might share a fold and may consequently present a mechanism of activation similar to that of CAP (28, 29). However, 2-OG is distinct from cAMP both in terms of molecular size and structural features, which theoretically should be reflected in differing effector-binding sites and activation mechanisms. We have determined the crystal structures of NtcA both in its apoform and ligand-bound forms in complex with 2-OG and DFPA, respectively. Despite the structural similarities between NtcA and CAP, comparative structural analysis showed that NtcA utilizes a rather different allosteric mechanism triggered by 2-OG. The signal from 2-OG is first sensed by the effector binding domains (EBD) before being transmitted to the F-helices in DBDs via the C-helices and the succeeding hinge regions, with no accompanying coil-to-helix transition or dramatic conformation switch of the F-helices.

Results

Crystallization and Structure Solution. Because the first structure of *E. coli* CAP was solved in 1981 (9), determining the apoform structure of CAP has proved to be a major challenge and was achieved 28 years later (12, 13), largely because of the in vitro instability and/or intrinsic flexibility of CAP in the absence of cAMP. We have encountered similar difficulties during the purification and crystallization of NtcA, which might explain why little is known about its structure and 2-OG binding mode despite preliminary X-ray crystallographic data reported by Wisen et al. in 2004 (28). Without the ligand 2-OG/DFPA or the detergent *n*-octyl-β-D-glucopyranoside (OBG), NtcA is quite unstable in solution and is prone to precipitation. However, after adding OBG to a final concentration of 0.1% (w/v), the DFPA/2OG-bound form of NtcA could reach a concentration of 20 mg/mL. We first obtained high-quality crystals of NtcA in complex with DFPA (designated NtcA-DFPA), and with a multiple wavelength anomalous dispersion (MAD) dataset of seleno-methionine labeled NtcA-DFPA (Se-NtcA-DFPA) at 2.70 Å, we solved the phase problem and determined the structure. Using this structure as the MR search model, we eventually determined the native structures of NtcA-DFPA at 2.40 Å and NtcA-2-OG at 2.60 Å.

An even bigger challenge has been solving the structure of the apoform of NtcA (apo-NtcA), crucial to illustrating the allosteric effect of 2-OG. Our previously optimized crystal of Se-Met-labeled apo-NtcA (Se-apo-NtcA) diffracted to only 3.10 Å and data showed severe anisotropy. Nevertheless, using this 3.10 Å MAD dataset, we determined the initial phases and built the partial model into the low-resolution density map with the help of the molecular replacement (MR) method using the NtcA-DFPA structure. Among the diamond-like crystals, we found several crystals of apo-NtcA of a different shape (Fig. S1), which diffracted at 1.90 Å and with lower anisotropy.

Overall Structures of NtcA. As postulated earlier (29, 30), NtcA adopts the same fold as classic CAPs, and NtcA-2-OG, NtcA-DFPA, and apo-NtcA share a similar overall structure (Fig. 1). Taking NtcA-2-OG as an example, the two subunits (A and B) in an asymmetric unit form a tight dimer with a 2-fold axis (Fig. 1*A*). Each subunit has two distinct domains, an N-terminal EBD and a C-terminal DBD, connected by a long helix from Pro116 to Ala141 (designated the C-helix, and the DNA recognition helix in DBD designated the F-helix, in accordance with those of CAP) (Fig. 1*A*). The two C-helices adopt a coiled-coil conformation and are the major contributors to the stabilization of the homodimer. The EBD adopts a β-barrel motif composed of seven antiparallel β-strands (β1–β7), within which a molecule of 2-OG is embedded. The DBD consists of a DNA-binding helix-turn-helix motif and a four-strand β-sheet (β8–β11).

NtcA-2-OG and NtcA-DFPA adopt an almost identical overall structure, whereas apo-NtcA differs considerably, especially in the EBDs (Fig. 1*B*). Vacuum electrostatic potential distribution

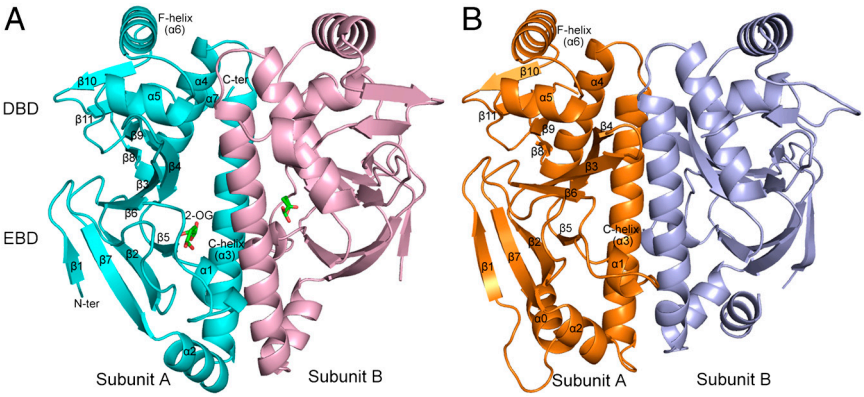


Fig. 1. Overall structures of NtcA homodimers. (*A*) NtcA-2-OG, (*B*) apo-NtcA. The two subunits are colored cyan and pink for NtcA-2-OG, and orange and light blue for apo-NtcA. The secondary structure elements are numbered sequentially and α3 and α6 are designated C-helix and F-helix, respectively. An extra α-helix at the N-terminus of apo-NtcA is designated α0. 2-OG is shown as sticks and is colored according to atom type: C, green; O, red; F, gray.

revealed a putative DNA-binding cleft within each DBD of the three structures, indicating that both the apo- and ligand-bound NtcA possess DNA-binding capacity.

The Binding Site of 2-OG and DFPA. In NtcA-2-OG, the EBD of each subunit contains a molecule of 2-OG that fits well in the electron density map (Fig. S2*A*). 2-OG binds to the protein with both direct and water-mediated hydrogen bonds (H-bonds) (Fig. 2*A*). The major contribution comes from the H-bonds between the main-chain atoms of residues Gly76, Val77, and Leu78 and the two carboxyl groups of 2-OG. The O7 and O9 of 2-OG form two H-bonds with Nα of Gly76, whereas O1 of 2-OG makes H-bonds with the Nα atoms of Val77 and Leu78. Thus, the main-chain conformation of Gly76–Leu78 provides a scaffold for embracing 2-OG. Moreover, Arg129-Nη1 and -Nη2 form two H-bonds with O10 and O9 of 2-OG, respectively, whereas Arg88-Nη1 interacts with O3 of 2-OG through a salt bridge and Glu134′-Oε2 (residues from subunit B are labeled with a prime) also forms an H-bond with O10 of 2-OG. In addition to these direct interactions, the Nα atoms of Tyr90 and Phe89 form two H-bonds with O3 of 2-OG through a water molecule.

We solved the structure of NtcA–DFPA to examine the structural basis behind DFPA being able to mimic the signaling function of 2-OG in vivo (4). As expected, it shares an almost identical overall structure with NtcA-2-OG and, moreover, DFPA adopts a binding pattern quite similar to that of 2-OG (Fig. 2*B*). Superposition of these two ligand-bound structures yields an rmsd of 0.40 Å for 384 Cα atoms of the dimer and 0.49 Å for 182 Cα atoms in the EBDs of two subunits. DFPA and 2-OG share the same binding residues with an almost identical H-bond network, except for a few slight differences. The acceptor atom O1 of 2-OG in the H-bond formed with Nα of Val77 was substituted by the atom F8 of DFPA. In addition, the H-bond between the atom F7 of DFPA and Glu134′-Oε1 acts as a counterpart of that between Nα of Gly76 and O7 of 2-OG.

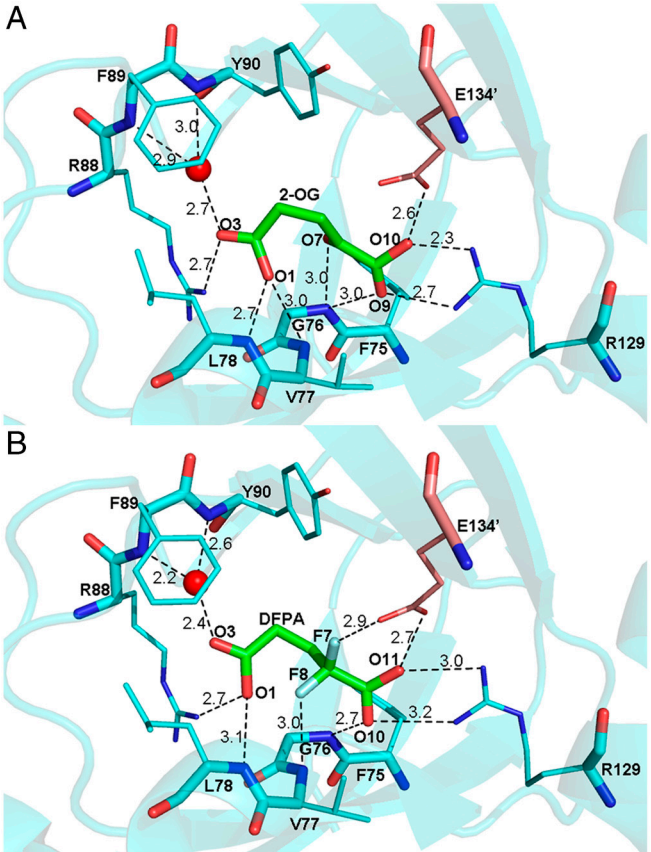


Fig. 2. The effector binding sites. (*A*) 2-OG and (*B*) DFPA. The residues are shown in cyan as thick sticks for the main chains and thin sticks for the side chains, water molecules are shown as red spheres. Residue Glu134 from the symmetric subunit is labeled with a prime and shown in pink. The O and F atoms of 2-OG and DFPA are numbered sequentially according to the PDB files (3LA2 and 3LA3). Black dotted lines denote the polar interactions with the distance in angstrom (H-bonds or salt bridges). The backbone of the protein is presented as a semitransparent cartoon.

Structural Comparison of apo-NtcA and NtcA-2-OG. Apo–NtcA shares the same topology and has a very similar overall structure to that of NtcA-2-OG. Superposition of apo-NtcA and NtcA-2-OG dimers yielded an rmsd of 2.22 Å for 360 Cα atoms. In contrast to one subunit that could be relatively well superimposed, another subunit showed considerable displacement (Fig. 3*A*). Moreover, for each subunit, most variation was found in the EBD in contrast to the minor conformational changes observed in the DBD. Superposition of subunit A of apo-NtcA and that of NtcA-2-OG results in an rmsd of 1.81 Å for 181 Cα atoms, with 0.48 Å for 73 Cα atoms of DBD compared to 2.78 Å for 91 Cα atoms of EBD. Remarkably, the binding pocket containing β5, α1, β6, and the connecting loops undergoes dramatic conformational changes upon 2-OG binding (Fig. 3*B*). In apo-NtcA, this pocket adopts an open, extended conformation that allows 2-OG to access the binding site. When 2-OG is approaching the binding pocket, as shown in Fig. 3*B*, the Cα atoms of residues Phe75, Gly76, and Val77 move toward 2-OG by 1.1, 1.6, and 2.3 Å, respectively, from one side, whereas the Cα atoms of Arg88, Phe89, and Tyr90 move toward 2-OG by 0.9, 1.2, and 1.9 Å, respectively, from the other side. Subject to the narrowing of the binding pocket, both the main-chain and side-chain atoms of Thr82–Asp87 in the surface loop between α1 and β6 are shifted. The Cα atoms of Ser86 and Asp87 are attracted, whereas the Cα atoms Thr82–Lys85 are repelled by 2-OG. In addition, upon the main-chain shift and side-chain rotation of Tyr90, the β3–β4 hairpin moves outward against the binding pocket (Fig. 3*B*). Taken together, as a result of the induced fit, residues Phe75–Leu78 and Arg88–Tyr90 are now in a better conformation to fix the carboxyl groups of 2-OG via abundant polar interactions (Fig. 2*A*).

The conformational changes at the EBDs upon 2-OG binding alter the relative orientation of the two C-helices (Fig. 3*C*), which in turn results in a larger dimeric interface. Apo–NtcA dimer has a buried interface of approximately 1700 Å^2, with the two C-helices adopting a rather parallel (at an angle of 17°) orientation. Upon 2-OG binding, the angle between the two C-helices is twisted from 17° to 23°, resulting in a tighter coiled-coil conformation and a more stabilized dimer with an interface increased to approximately 2000 Å^2 (Fig. 3*C*).

Putative Route of Signal Transmission from 2-OG to the F-Helices. How does a series of conformational changes transfer the 2-OG signal from the effector binding sites in the EBDs to the DNA-binding sites of the two F-helices in the DBDs? The 2-OG signal is first sensed by the residues at the β-barrel pocket in EBD (Fig. 3*B*). All the conformational changes occurring upon 2-OG binding together lead to the distortion of the EBDs as well as with the side-chain shifts of several residues directly linked to the C-helices (Fig. 3*C*). Among these residues, Tyr90 contributes at a particularly important level. In apo-NtcA, the hydroxyl group of Tyr90 forms an H-bond with Glu134′-Oε1 and occupies the binding pocket, whereas Glu62-Oε1 in the β4-strand forms an H-bond with Glu138′-Oε2. In NtcA-2-OG, the phenol ring of Tyr90 rotates around the peptide bond by approximately 30° and shifts the hydroxyl group by approximately 3.5 Å away from 2-OG. Meanwhile, following the breakage of the H-bond with Glu62-Oε1 in the apoform, the side chain of Glu138′ is rotated

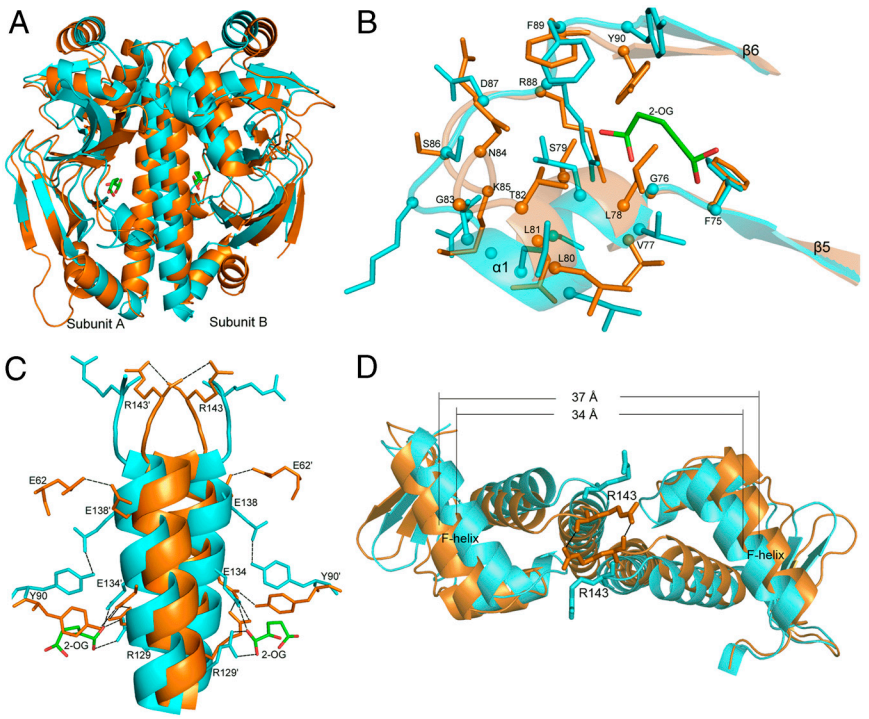


Fig. 3. Structural comparison of apo-NtcA and NtcA-2-OG. (*A*) Superposition of apo-NtcA (in orange) and NtcA-2OG (in cyan). The subunit A of apo-NtcA and NtcA-2OG superimposed well, whereas the subunit B showed greater shift. (*B*) A close-up view of the conformational changes of the binding pocket. The residues involved are highlighted as sticks. (*C*) A close-up view of the superimposed C-helices. The key residues are shown as sticks. The polar interactions are represented by black dotted lines. (*D*) A top view of the C-helices at the center to highlight the conformational change in the hinge regions. The key Arg143 residues are shown as sticks and the distance of the Cα atoms between apo-NtcA (orange) and NtcA-2-OG (cyan) are labeled.

toward 2-OG by approximately 67° and forms an H-bond with the hydroxyl group of Tyr90 (Fig. 3*C*).

Besides these 2-OG binding residues in EBD, the 2-OG signal is sensed directly and simultaneously by residues Arg129 and Glu134′, which protrude from the two C-helices (Fig. 3*B*). In apo-NtcA, the carboxyl group of Glu134′ is fixed by H-bonds with Arg129-Nη1 and -Nη2 via Oε2, and another with the hydroxyl group of Tyr90 via Oε1. In NtcA-2-OG, Glu134′ is shifted slightly toward and makes an H-bond with the O10 of 2-OG, whereas Arg129-Nη1 and -Nη2 form two H-bonds with the atoms O9 and O10 of 2-OG (Fig. 3*C*). The interactions of 2-OG with residues Arg129 and Glu134′ facilitate the adoption of a tighter coiled-coil conformation by the two C-helices.

The signal is then transmitted along the C-helices to the hinges linking the C-helices and DBDs. In apo-NtcA, the pair of C-helices adopt a relatively parallel or looser coiled-coil conformation, and the two Arg143 residues located at the hinge regions are close enough to each other to make two H-bonds between the Nη2 and main-chain oxygen atoms (Fig. 3*C*). In NtcA-2-OG, the Cα atoms of two Arg143 residues move outward relative to their counterparts in apo-NtcA by an average of 3.6 Å, accompanied by a rotation of the side chain outward by approximately 30° (Fig. 3*D*). The breaking of these two H-bonds, in addition to the conformational changes of Glu138′, Arg129, and Glu134′, contribute to a tighter coiled-coil conformation of the C-helices in the ligand-bound NtcA. As seen from the top view, the 17–23° transition of the C-helices shifts the hinge regions to a position in which they cross over slightly (Figs. 3*D* and 4). Consequently, the two DBDs move closer to each other as a rigid body, similar to that found in the DBDs of CAP (12, 13). Thus, the distance between the two F-helices is reduced from 37 Å in apo-NtcA to 34 Å in NtcA-2-OG (Figs. 3*D* and 4).

In short, the signal of 2-OG is first sensed by the binding residues in EBDs before being transmitted along the C-helices through residues Arg129, Glu134, and Glu138, further amplified by Arg143 at the hinge regions, and finally accepted by the two F-helices in the DBDs. Upon 2-OG binding, the two C-helices adopt a tighter coiled-coil conformation, which is crucial to bringing the two rigid DBDs relatively closer and providing a more adapted distance between the two F-helices for DNA recognition.

Discussion

Structural Comparison of Ligand-Bound Forms of NtcA and CAP. The structure of NtcA represents the first solved structure of this group in the Crp-Fnr family. Both NtcA and CAP are dimeric proteins with a similar overall structure. Superposition of NtcA-2-OG against *E. coli* CAP-cAMP [Protein Data Bank (PDB) ID 1G6N) (14) yields an rmsd of 2.83 Å for 351 Cα atoms, and reveals that the DBD and EBD adopt a similar relative position in both structures (Fig. S3). In the structure of CAP-cAMP-DNA, the two F-helices are inserted into successive DNA major grooves and make extensive contacts with the bases and the sugar-phosphate backbone (15, 31). In NtcA-2-OG, the two F-helices (Arg187–Lys200) possess numerous positively charged residues Arg187, Arg192, Arg198, and Lys200, which allow, in

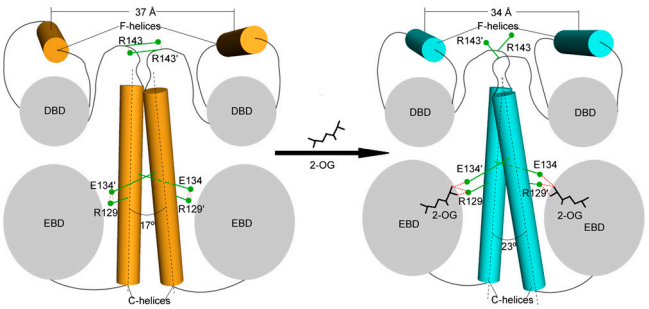


Fig. 4. A model illustrating the mechanism of allosteric control of NtcA.

the 2-OG or DFPA-binding state, a conformation more suitable to recognize successive DNA major grooves.

Based on the sequence-structure alignment, the binding sites of 2-OG in NtcA clearly correspond to those of cAMP in CAP, and even share several binding residues (Fig. S4). However, the binding pattern of 2-OG is quite different from that of cAMP, due to significant differences in both their molecular size and structural features. It is notable that residues Ser83 and Ala84 in CAP are substituted by the aromatic residues Phe89 and Tyr90 in NtcA, resulting in the occupancy of the cAMP-binding pocket. Moreover, Ser128 in CAP is substituted by Glu134′ in NtcA, which is located on the C-helix of the neighboring symmetric subunit (Fig. 5). Glu134′ forms an H-bond with the carboxyl group of 2-OG and Ser128 forms an H-bond with the nitrogen atom of the cAMP purine moiety. The side chain of Glu134′ in NtcA also occupies the corresponding space for the cAMP purine moiety in CAP.

Common and Distinct Features of the Allosteric Mechanisms Between CAP and NtcA. In both NtcA and CAP, the effector signal is transmitted from the EBD to the DBD via the C-helix bridge. A series of conformational changes induce the F-helices to adopt a more suitable distance and orientation for DNA binding, which then activates the transcription of target genes. During the signal transmission, the individual DBDs of both NtcA and CAP move as a rigid body resulting in a DNA major groove-favoring conformation of the two F-helices.

Besides these common features, several structural properties distinguish NtcA from CAP by their different associated activation mechanisms. According to the NMR structure, the binding of cAMP extends the C-helix by three turns and concomitantly forces the F-helices to rotate by 60° and adopt a correct orientation and a perfect distance for DNA recognition (12). However, the crystal structure suggests that the binding of cAMP extends the C-helix by 1.5 turns and causes the buried F-helices in apo-CAP to be exposed to the surface of the DBDs in cAMP-CAP (13). In contrast, upon 2-OG binding, neither of the two individual C-helices of NtcA undergoes any dramatic conformational change, but instead shows a shift toward each other with a tighter coiled-coil conformation resulting from the breakage and regeneration of H-bonds donated by residues Arg129, Glu134, Glu138, and Arg143. The significant change in the relative orientation between the two C-helices shortens the distance of the two parallel F-helices from 37 to 34 Å, achieved without rotation.

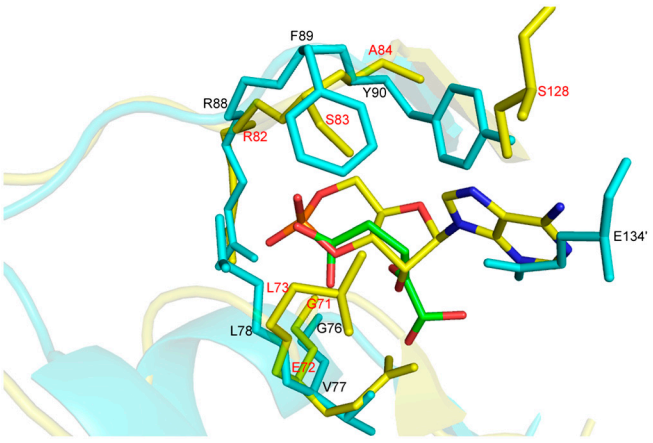


Fig. 5. Structural comparison of the effector binding sites of NtcA and CAP. NtcA and CAP (PDB ID 1G6N) are colored cyan and yellow, respectively. The residues involved are numbered in black for NtcA-2-OG and red for CAP-cAMP. The molecule 2-OG is colored as in Fig. 1, whereas cAMP is colored according to atom type: C, yellow; O, red; N, blue.

In addition, in both apo-NtcA and NtcA-2-OG, a positively charged cleft made by residues exposed to the surface of the DBDs provides a DNA-binding patch. This indicates that, in contrast to apo-CAP (32), both apo-NtcA and NtcA-2-OG possess a DNA-binding capacity. Why is the enhancement of DNA-binding affinity of NtcA by 2-OG not as strong as that of CAP by cAMP? As shown in the present NtcA structures, the binding of 2-OG shifts the two F-helices to a distance more adapted for DNA binding and consequently increases DNA-binding affinity. Even in the absence of 2-OG, the conformation of the two F-helices allows NtcA to bind DNA with relatively weaker affinity. This observation agrees well with the fact that the apoform of NtcA can indeed bind to DNA in vitro at a dissociation constant of 52 nM, slightly higher than that of the 2-OG-bound form at 40 nM, as found in *Synechococcus* PCC 7942 (22, 23). The fine tuning of NtcA by 2-OG might have evolved from evolutionary pressure whereby cyanobacteria needed NtcA to regulate the expression of a much wider spectrum of genes than does CAP in *E. coli*. In fact, NtcA not only activates the expression of genes involved in nitrogen metabolism upon the 2-OG accumulation induced by nitrogen deprivation, but it also controls many genes involved in other functions such as photosynthesis and oxidative stress response (33–35). In addition to functioning as an activator, in some cases NtcA acts as a repressor (36). The structure-function relationship of NtcA may therefore have been fashioned during evolution by the necessity to coordinate a diversity of cellular functions, not simply involving an off-to-on transition in the case of CAP, but by using a more subtle change of the relative distance between the two F-helices.

Numerous studies have been carried out on the DNA-binding affinity of NtcA proteins from cyanobacteria *Anabaena* PCC 7120 (4, 21, 24, 29) and *Synechococcus* PCC 7942 (20, 22, 23). Based on the present crystal structures, we performed a sequence alignment of 17 representative NtcA proteins (Fig. S5). The 2-OG binding site and the key residues undergoing conformational changes upon 2-OG binding are almost identical. Moreover, the F-helix is also highly conserved. Therefore, the allosteric mechanism proposed in the present work for the *Anabaena* NtcA should be valid for all other cyanobacterial NtcA proteins.

More investigations are needed to fully understand the complex mechanism behind the 2-OG-tuning of NtcA likely involving other known or unknown factors. Using *Synechococcus* NtcA as a bait, the PII-interacting protein X (PipX) was preyed by two-hybrid approaches, and the interaction between NtcA and PipX in the presence of 2-OG was also shown by the surface plasmon resonance method (37, 38). Furthermore, there is evidence suggesting that PipX acts as a positive regulator in NtcA-dependent transcriptional activation under the high levels of 2-OG caused by nitrogen deficiency. On the basis of these results, it was proposed that the transcription of NtcA-dependent promoters could be activated by the PipX–NtcA complex (37). An attempt to solve the complex structure of PipX–NtcA is in progress to test this hypothesis.

Materials and Methods

Protein Expression, Purification, and Crystallization. The coding region of NtcA gene was cloned into pET28b (Novagen) as previously described (23). Both the wild-type and Se-Met labeled apo-NtcA were expressed, purified and crystallized by the procedures reported by Wisen et al. (28). For ligand-bound forms, 15 mg/mL of protein with 5 mM 2-OG/DFPA was mixed with an equal volume of the reservoir solution containing 0.05 M $(NH_4)_2SO_4$ and 25% pentaerythritol ethoxylate (3/4, EO/OH) equilibrated against the reservoir solution. All crystals were grown at 16 °C using the hanging drop vapor diffusion technique.

Data Collection and Processing. All crystals were flash frozen in liquid nitrogen using the cryoprotectant of their respective reservoir solutions, and the data were all collected at 100 K in a nitrogen stream. The native data of NtcA-2-OG and NtcA–DFPA were collected on a MarResearch 345 image-plate detector using a Rigaku MM007 X-ray generator (1.5418 Å) at the School of Life

Sciences, University of Science and Technology of China (USTC). The data of Se-NtcA-DFPA were collected at the Beijing Synchrotron Radiation Facility and those of native apo-NtcA and Se-apo-NtcA were collected at the Shanghai Synchrotron Radiation Facility. The data was integrated and scaled with the program HKL2000 (39).

Structure Determination and Refinement. The selenium sites from the data of Se-NtcA-DFPA were determined using the program SHELXD (40). The initial phases were calculated with the program phenix.solve implemented in the PHENIX program suite. The program phenix.resolve was used to perform solvent flattening and automatic model building. The resultant model was subsequently used as the search model for determining the structures of NtcA-2-OG and NtcA–DFPA by the MR method with the program MOLREP (41) implemented in CCP4i (42). The initial model was refined by REFMAC5 (43) and rebuilt interactively with the programs PHENIX (44) and COOT (45) until the free R-factor converged. The final models were refined to the highest resolution shown in Table S1, and both were checked using the programs MOLPROBITY (46) and PROCHECK (47).

The resulting electron density map of Se-apo-NtcA at 3.10 Å was found to be quite poor. Using the DBD and the C-helix part of the NtcA–DFPA structure as a search model, we determined the phase by MR in combination with MAD phasing. With the aid of the selenium sites and residues of bulky side-chains for the residue type assignment, we successfully built the initial model of apo-NtcA using COOT and Autobuild implemented in PHENIX (44). This model was then used as a search model against the 1.90 Å data of the native apo-NtcA. All structure figures were prepared with PyMOL (48).

ACKNOWLEDGMENTS. We thank Rui Bao at Tongji University and Xiao-Xiao Ma at USTC for data collection. We are grateful to all the developers of the CCP4 Suite, ESPript, TreeView, and PyMOL, and to Dr. Ling Peng for providing DFPA. We thank Shanghai Synchrotron Radiation Facility (SSRF) and Beijing Synchrotron Radiation Facility (BSRF). This work was supported by the Ministry of Science and Technology of China (Projects 2006CB910202 and 2006CB806501), the National Natural Science Foundation of China (Program 30870490), and the Agence Nationale de la Recherche (Physico-Chimie du Vivant program) in France.

1. Ninfa AJ, Jiang P (2005) PII signal transduction proteins: Sensors of alpha-ketoglutarate that regulate nitrogen metabolism. *Curr Opin Microbiol* 8(2):168–173.
2. Feria Bourrellier AB, et al. (2010) Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. *Proc Natl Acad Sci USA* 107(1):502–507.
3. Hebert SC (2004) Physiology: Orphan detectors of metabolism. *Nature* 429(6988): 143–145.
4. Laurent S, et al. (2005) Nonmetabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci USA* 102(28):9907–9912.
5. Herrero A, Muro-Pastor AM, Valladares A, Flores E (2004) Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev* 28(4):469–487.
6. Korner H, Sofia HJ, Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* 27(5):559–592.
7. Fischer HM (1994) Genetic regulation of nitrogen fixation in rhizobia. *Microbiol Rev* 58(3):352–386.
8. Zubay G, Schwartz D, Beckwith J (1970) Mechanism of activation of catabolite-sensitive genes: A positive control system. *Proc Natl Acad Sci USA* 66(1):104–110.
9. McKay DB, Steitz TA (1981) Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* 290(5809):744–749.
10. Anderson WB, Schneider AB, Emmer M, Perlman RL, Pastan I (1971) Purification of and properties of the cyclic adenosine 3′,5′-monophosphate receptor protein which mediates cyclic adenosine 3′, S-monophosphate-dependent gene transcription in *Escherichia coli*. *J Biol Chem* 246(19):5929–5937.
11. Riggs AD, Reiness G, Zubay G (1971) Purification and DNA-binding properties of the catabolite gene activator protein. *Proc Natl Acad Sci USA* 68(6):1222–1225.
12. Popovych N, Tzeng SR, Tonelli M, Ebright RH, Kalodimos CG (2009) Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc Natl Acad Sci USA* 106(17):6927–6932.
13. Sharma H, Yu S, Kong J, Wang J, Steitz TA (2009) Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proc Natl Acad Sci USA* 106(39):16604–16609.
14. Passner JM, Schultz SC, Steitz TA (2000) Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. *J Mol Biol* 304(5):847–859.
15. Schultz SC, Shields GC, Steitz TA (1991) Crystal structure of a CAP-DNA complex: The DNA is bent by 90 degrees. *Science* 253(5023):1001–1007.
16. Herrero A, Muro-Pastor AM, Flores E (2001) Nitrogen control in cyanobacteria. *J Bacteriol* 183(2):411–425.
17. Zhang C-C (1996) Bacterial signalling involving eukaryotic-type protein kinases. *Mol Microbiol* 20(1):9–15.
18. Zhang C-C, Laurent S, Sakr S, Peng L, Bedu S (2006) Heterocyst differentiation and pattern formation in cyanobacteria: A chorus of signals. *Mol Microbiol* 59(2):367–375.
19. Ramasubramanian TS, Wei TF, Golden JW (1994) Two *Anabaena* sp. strain PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific genes. *J Bacteriol* 176(5):1214–1223.
20. Luque I, Flores E, Herrero A (1994) Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* 13(12):2862–2869.
21. Jiang F, Wisen S, Widersten M, Bergman B, Mannervik B (2000) Examination of the transcription factor NtcA-binding motif by in vitro selection of DNA sequences from a random library. *J Mol Biol* 301(4):783–793.
22. Tanigawa R, et al. (2002) Transcriptional activation of NtcA-dependent promoters of *Synechococcus* sp. PCC 7942 by 2-oxoglutarate *in vitro*. *Proc Natl Acad Sci USA* 99 (7):4251–4255.
23. Vazquez-Bermudez MF, Herrero A, Flores E (2002) 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus* glnA promoter. *FEBS Lett* 512(1–3):71–74.
24. Chen H, et al. (2006) Studying the signaling role of 2-oxoglutaric acid using analogs that mimic the ketone and ketal forms of 2-oxoglutaric acid. *Chem Biol* 13(8):849–856.
25. Stanier RY, Cohen-Bazire G (1977) Phototrophic prokaryotes: The cyanobacteria. *Annu Rev Microbiol* 31:225–274.
26. Li JH, Laurent S, Konde V, Bedu S, Zhang C-C (2003) An increase in the level of 2-oxoglutarate promotes heterocyst development in the cyanobacterium Anabaena sp. strain PCC 7120. *Microbiology* 149:3257–3263.
27. Valladares A, Flores E, Herrero A (2008) Transcription activation by NtcA and 2-oxoglutarate of three genes involved in heterocyst differentiation in the cyanobacterium *Anabaena sp.* strain PCC 7120. *J Bacteriol* 190(18):6126–6133.
28. Wisen S, Sjogren T, Olin B, Mannervik B (2004) Purification, crystallization and preliminary X-ray data of the transcription factor NtcA from the cyanobacterium *Anabaena* PCC 7120. *Acta Crystallogr D* 60(Pt 5):923–925.
29. Wisén S, Bergman B, Mannervik B (2004) Mutagenesis of the cysteine residues in the transcription factor NtcA from *Anabaena* PCC 7120 and its effects on DNA binding in vitro. *Biochim Biophys Acta* 1679(2):156–163.
30. Wisén S (2003) Characterization of the DNA-binding properties of the cyanobacterial transcription factor NtcA. PhD thesis (Uppsala University, Sweden).
31. Parkinson G, et al. (1996) Structure of the CAP–DNA complex at 2.5 Å resolution: A complete picture of the protein–DNA interface. *J Mol Biol* 260(3):395–408.
32. Dai J, Lin SH, Kemmis C, Chin AJ, Lee JC (2004) Interplay between site-specific mutations and cyclic nucleotides in modulating DNA recognition by *Escherichia coli* cyclic AMP receptor protein. *Biochemistry* 43(28):8901–8910.
33. Su Z, Olman V, Mao F, Xu Y (2005) Comparative genomics analysis of NtcA regulons in cyanobacteria: Regulation of nitrogen assimilation and its coupling to photosynthesis. *Nucleic Acids Res* 33(16):5156–5171.
34. Cheng Y, et al. (2006) A pair of iron-responsive genes encoding protein kinases with a Ser/Thr kinase domain and a His kinase domain are regulated by NtcA in the Cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 188(13):4822–4829.
35. Lopez-Gomollon S, et al. (2007) Cross-talk between iron and nitrogen regulatory networks in *Anabaena* (*Nostoc*) sp. PCC 7120: identification of overlapping genes in FurA and NtcA regulons. *J Mol Biol* 374(1):267–281.
36. Jiang F, Mannervik B, Bergman B (1997) Evidence for redox regulation of the transcription factor NtcA, acting both as an activator and a repressor, in the cyanobacterium *Anabaena* PCC 7120. *Biochem J* 327(Pt 2):513–517.
37. Espinosa J, Forchhammer K, Burillo S, Contreras A (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Mol Microbiol* 61(2):457–469.
38. Espinosa J, Forchhammer K, Contreras A (2007) Role of the *Synechococcus* PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. *Microbiology* 153(Pt 3): 711–718.
39. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol Crystallogr* 276:307–326.
40. Schneider TR, GM S (2002) Substructure solution with SHELXD. *Acta Crystallogr D* 58:1772–1779.
41. Vagin A, Teplyakov A (1997) MOLREP: An automated program for molecular replacement. *J Appl Crystallogr* 30:1022–1025.
42. Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 50(5):760–763.
43. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
44. Adams PD, et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* 58:1948–1954.
45. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
46. Davis I, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–383.
47. Laskowski R, Macarthur M, Moss D, Thornton J (1993) Procheck—a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
48. DeLano W (2002) The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA) http://www.pymol.org.

Supporting Information

Zhao et al. 10.1073/pnas.1001556107

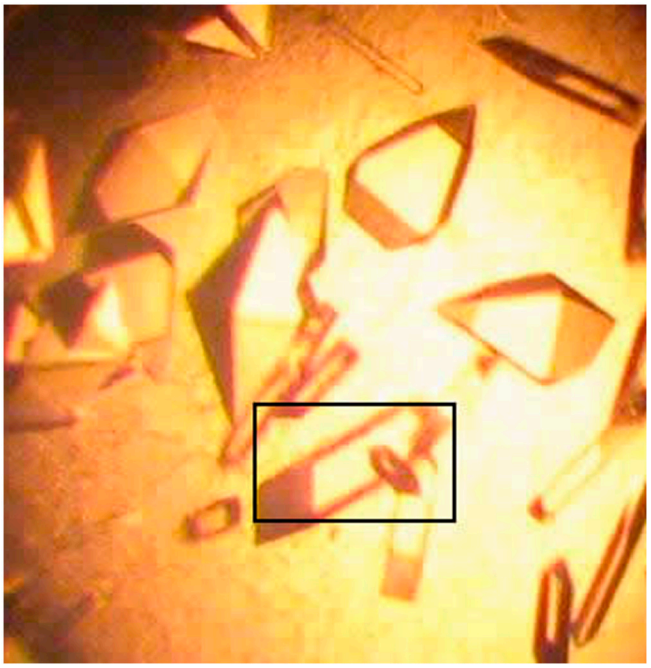


Fig. S1. The shape of the crystal of apo-NtcA in the rectangle is different from that of the diamond-like crystals.

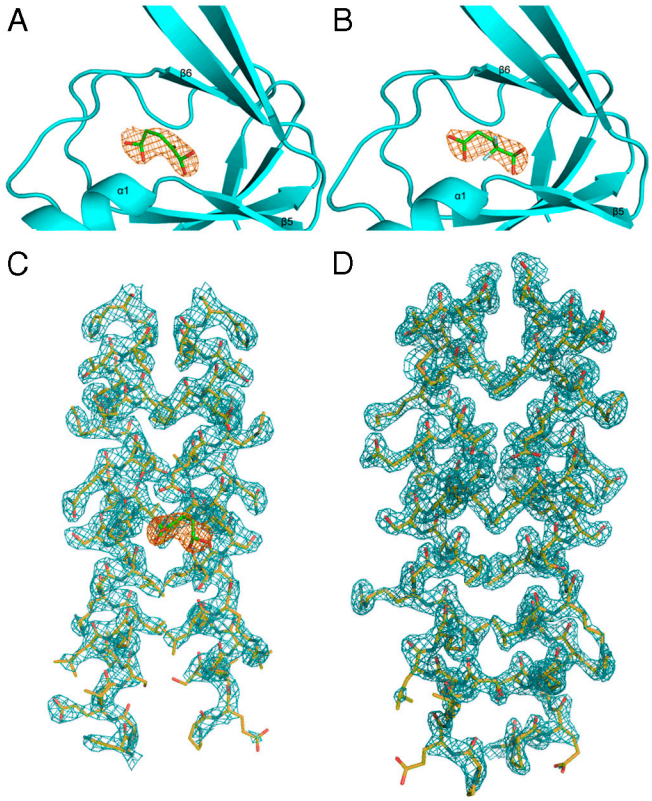


Fig. S2. The 2Fo-Fc electron density map for 2-OG/DFPA and NtcA-2-OG /apo-NtcA, contoured at 1.5σ. The 2Fo-Fc electron density maps for (*A*) 2-OG and (*B*) 2,2-difluoropentanedioic acid are shown as orange mesh. A stereoview of the 2Fo-Fc electron density map for (*C*) NtcA-2-OG and (*D*) apo-NtcA centered on the C-helices. The maps are shown in blue mesh for protein and orange mesh for 2-OG.

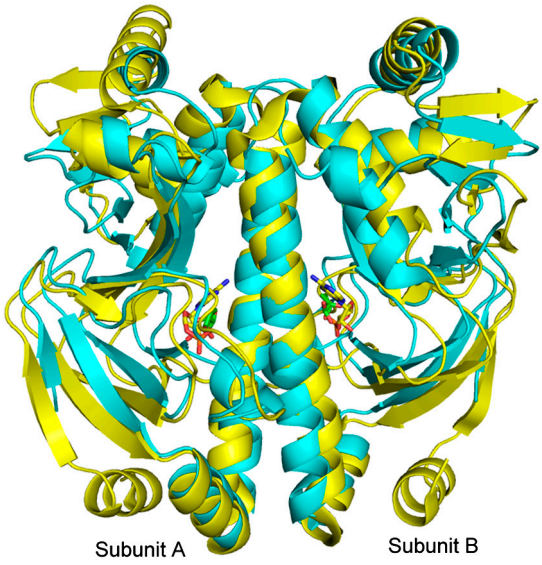


Fig. S3. Structural comparison of NtcA-2-OG (cyan) and cAMP-CAP (yellow; Protein Data Bank ID 1G6N). The molecules of 2-OG and cAMP are shown as sticks.

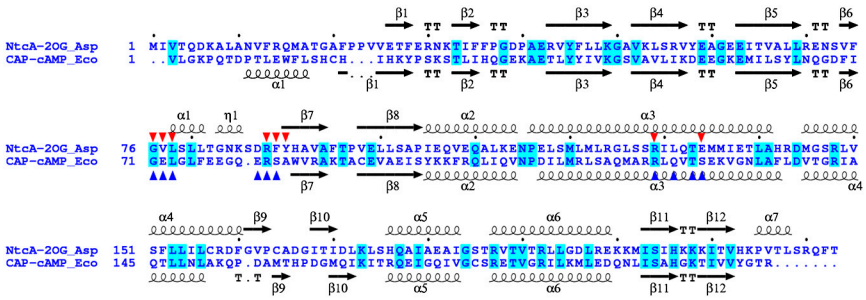


Fig. S4. Structure-based sequence alignment of NtcA and *E. coli* CAP. The binding site of 2-OG in NtcA and that of cAMP in CAP are indicated by red and blue triangles, respectively. The residues marked in yellow are the main residues different between the effector binding sites in NtcA and CAP.

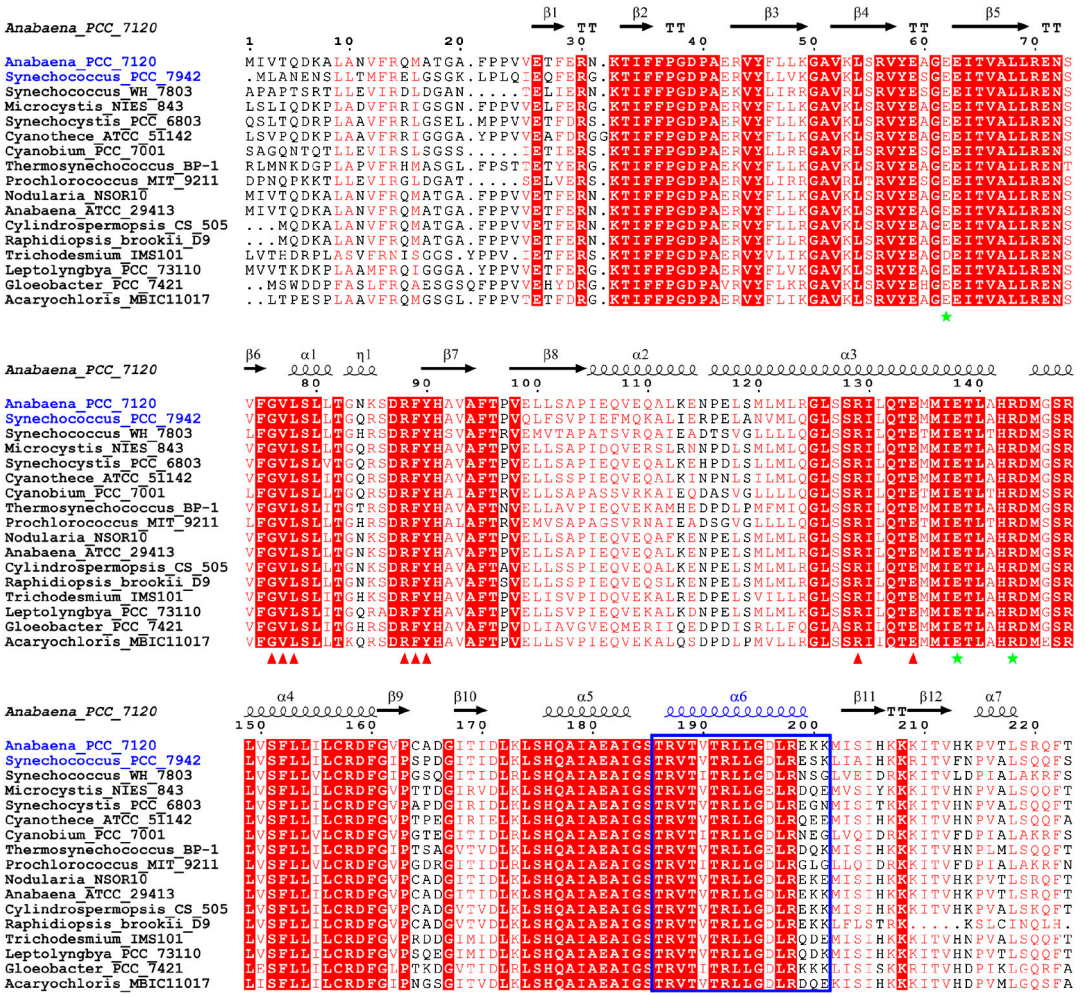


Fig. S5. Structure-based sequence alignment of 17 representative NtcA proteins from cyanobacteria. The 2-OG binding site was marked with the red triangles and the key residues undergoing conformational changes upon 2-OG binding were marked with green stars. All these sites are almost identical. The highly conserved F-helix is boxed in a blue frame.

Table S1. Crystal parameters, data collection, and structure refinement

Molecule	Se-NtcA-DFPA			NtcA-DFPA	NtcA-2-OG	Se-apo-NtcA			apo-NtcA
Data collection	*Peak*	*Inflection*	*Remote*			*Peak*	*Inflection*	*Remote*	
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$C222_1$
Unit cell dimensions									
a (Å)	67.44	67.43	68.09	68.03	68.14	67.47	67.51	68.31	67.51
b (Å)	69.26	69.26	69.92	70.08	70.26	67.47	67.51	68.31	85.62
c (Å)	149.29	149.23	150.68	149.73	149.99	160.07	160.15	161.85	171.96
α, β, γ (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00
Resolution range (Å)	50.00–2.70 (2.75–2.70) *	50.00–2.70 (2.75–2.70)	50.00–2.70 (2.75–2.70)	20.45–2.40 (2.53–2.40) *	23.73–2.60 (2.74–2.60)	50.00–3.10 (3.21–3.10)	50.00–3.10 (3.21–3.10)	50.00–3.10 (3.21–3.10)	50.00–1.90 (1.93–1.90)
Unique reflections	36,968 (1,852)	36,910 (1,839)	37,908 (1,811)	27,164 (3,632)	22,720 (3,261)	7,131 (592)	7,141 (588)	7,305 (568)	39,299 (1,931)
Completeness (%)	99.9 (99.5)	99.9 (99.7)	99.8 (97.5)	94.7 (88.3)	99.6 (99.4)	97.9 (84.7)	98.0 (83.8)	97.1 (79.7)	99.0 (98.2)
I/σ (I)	16.6 (1.7)	17.3 (1.8)	14.3 (1.3)	18.4 (3.1)	13.3 (2.6)	30.7 (2.6)	32.2 (2.0)	32.5 (2.2)	18.2 (4.5)
R_{merge} † (%)	7.8 (53.2)	7.4 (52.6)	8.5 (62.9)	4.1 (34.2)	7.9 (69.7)	9.7 (42.6)	8.6 (42.4)	8.9 (39.8)	9.0 (53.4)
Average redundancy	4.1 (3.8)	4.1 (3.9)	4.1 (3.4)	3.6 (3.6)	5.3 (5.3)	17.1 (10.9)	18.7 (10.7)	18.6 (10.3)	10.7 (7.9)
Structure refinement									
Resolution range (Å)				20.45–2.40	23.73–2.60				50.00–1.90
R-factor ‡/R-free § (%)				21.1/25.3	22.2/26.9				21.5/26.4
Number of protein atoms				3,119	3,080				3,277
Number of water atoms				160	120				170
Rmsd ¶ bond length (Å)				0.011	0.013				0.013
Rmsd bond angle (°)				1.326	1.386				1.410
Mean B factors (Å^2)				37.75	25.96				40.81
Ramachandran plot ‖ (residues, %)									
Most favored (%)				96.42	94.07				96.57
Additional allowed (%)				3.58	5.93				3.43
Outliers (%)				0	0				0
PDB entry				3LA3	3LA2				3LA7

*The values in parentheses refer to statistics in the highest bin.

†$R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for its unique reflection; Summations are over all reflections.

‡R-factor $= \sum_h |Fo(h) - Fc(h)| / \sum_h Fo(h)$, where Fo and Fc are the observed and calculated structure-factor amplitudes, respectively.

§R-free was calculated with 5% of the data excluded from the refinement.

¶Rmsd from ideal values.

‖Categories were defined by Molprobity.