Biological Sciences, Biochemistry

Structural understanding of NtcA regulation and of its coactivation by the adaptor PII/NtcA shuttling protein PipX, which connects PII regulation with gene expression regulation

Alicia Forcada-Nadal,[1,2] Sirine Bibak[3], Paloma Salinas[3], Asunción Contreras[3], Vicente Rubio[1,2,*] and José L. Llácer.[1,2,*]

[1]Instituto de Biomedicina de Valencia of the CSIC (IBV-CSIC), Valencia Spain.

[2]Group 739 at the IBV-CSIC of the Centro de Investigación Biomédica en Red en Enfermedades Raras of the Instituto de Salud Carlos III (CIBERER-ISCIII), Madrid, Spain

[3]Departamento de Fisiología, Genética y Microbiología, Universidad de Alicante, Alicante, Spain.

***Corresponding authors**

rubio@ibv.csic.es

jllacer@ibv.csic.es

Funding: Supported by grant PID 2020-116880GB-I00 of the Plan Estatal de I+D+I of the Spanish Government to JLL, grants PID2020-118816GB-I00 and PID2023-149456NB-I00 to AC and by grants BFU2017-84264-P from the Spanish Ministry of Science of the Spanish Government; and CIV20A6610 of the Fundación Ramón Areces, both of them to VR. SB was supported by a National Grant from the Algerian Ministry of Higher Education and Scientific Research.

Abstract

The CRP-FNR superfamily of transcriptional regulators includes the cyanobacterial master regulator NtcA, which orchestrates large responses of cyanobacteria to nitrogen scarcity. NtcA uses as allosteric activator 2-oxoglutarate (2OG), a signal of nitrogen poorness and carbon richness, and binds a coactivating protein (PipX) that shuttles between the signaling protein PII and NtcA depending on nitrogen richness, thus connecting PII signaling and gene expression regulation. Here, combining structural (X-ray crystallography of six types of crystals including NtcA complexes with DNA and 2OG and PipX), modelling and functional (EMSA and bacterial two-hybrid) studies, we clarify the binding and exquisite specificity of NtcA for its target DNA, its mechanisms of activation by 2OG, and its coactivation by PipX. Our crystal structures of PipX-NtcA-DNA complexes prove that PipX does not interact with DNA, although it increases NtcA-DNA contacts, and that it stabilizes the active, DNA-binding-competent conformation of NtcA. Superimposition of this complex on a very recently reported cryoEM structure of NtcA in a Transcription Activity Complex with RNA polymerase (RNAP), shows that PipX binding helps recruit RNAP by PipX interaction with RNAP, particularly with its gamma and sigma (region 4) subunits, a structural prediction supported here by bacterial two-hybrid experiments.

Keywords.

Transcriptional regulation, PipX, NtcA, CRP, FNR, RNA polymerase, nitrogen regulation, cyanobacteria, *Synechococcus elongatus*.

Introduction

Regulation of gene expression is a crucial process for adaptation of bacteria to changing environments, colonization of diverse niches and pathogenicity. Key players in those processes are transcriptional regulators. Of paramount importance among transcriptional regulators is the superfamily of these proteins known as CRP-FNR (cAMP receptor protein-fumarate/nitrate reductase regulator), which encompasses at least 21 subclasses and is represented in most bacterial phyla (Körner *et al.* 2003; Krol *et al.* 2023). The members of this superfamily are homodimers composed of an N-terminal dimerization and effector-binding domain (EBD) and of a C-terminal DNA binding domain (DBD) with a helix-turn-helix (HTH) motif. Members of this superfamily regulate a broad range of processes, including responses to hunger, oxygen, carbon monoxide, oxidative stress, nitrogen fixation, availability to metabolically active nitrogen forms and photosynthesis (Körner *et al.* 2003; Roberts 2024; Vegas-Palas *et al.* 1992).

CRP is considered the paradigm for this superfamily and has reached textbook-example status (Lawson *et al* 2004). The determination, more than 30 years ago, of the crystal structure of CRP bound to its target DNA revealed determinants involved in the specificity of DNA sequence recognition (Schultz *et al.* 1991). DNA binding involved the insertion of the αF helices of the HTH motif from each subunit of the CRP dimer into adjacent major grooves of the double-stranded DNA. CRP binding induced a significant bend on the DNA, which was found to facilitate transcription activation (Shultz *et al.* 1991, Parkinson *et al.* 1996, Chen *et al.* 2001, Chen *et al.* 2001, Napoli *et al.* 2006, Benoff *et al.* 2002).

DNA binding and transcriptional regulation mechanisms appear to be shared by the members of the CRP-FNR superfamily. However, structural snapshots for protein-DNA complexes have been obtained in just a few cases (Levy *et al.* 2008; Hall *et al.* 2016; Bonnet *et al.* 2013; Werel *et al.* 2023) and there is even less information concerning transcription activation complexes (TACs), which include the RNA polymerase (RNAP) holoenzyme and is critical for understanding transcriptional regulation. A crystal structure of the bacterial class II TAC from *Thermus thermophilus* revealed how the CRP/FNR activator TAP enhances transcription initiation through stabilizing interactions with RNAP (Feng *et al.* 2016). Cryo-electron microscopy studies of *E. coli* CRP-TAC complexes of class I (Liu *et al.* 2017) and class II (Shi *et al.* 2020) types have illustrated diverse interactions between CRP and RNAP, offering detailed views of how CRP

binding to DNA facilitates transcription by RNAP. Very recently, cryo-electron microscopy structures of *Anabaena* sp. PCC7120 TAC with the NtcA master regulator (a CRP/FNR superfamily member) have provided additional information on relevant interactions and on cooperative transcription activation by NtcA and the LysR-like regulator NtcB (Han *et al.* 2024).

NtcA, the subject of the present work, is a global nitrogen regulator conserved in cyanobacteria that controls a very large regulon mainly involved in nitrogen assimilation (Vegas-Palas *et al.* 1992; Herrero *et al.* 2001; Herrero *et al.* 2004, Su *et al.* 2005; Mitschke *et al.* 2011; Espinosa *et al*. 2014; Giner-Lamia *et al.* 2017). NtcA binds to 2-oxoglutarate (2OG), a molecule that signals the intracellular carbon/nitrogen balance (high levels when carbon is abundant and there is poorness of metabolizable nitrogen; Llácer *et al*. 2008) and induces the active conformation of NtcA, binding at a topologically equivalent site to that of cAMP in CRP (Llácer *et al.* 2010; Zhao *et al.* 2010). However, NtcA is unique within the CRP/FNR family because it is coactivated by PipX, a small protein conserved in cyanobacteria (Burillo *et al.* 2004) that shuttles between the carbon/nitrogen/energy sensor/signaling protein PII and NtcA, depending on whether nitrogen richness is high or low (Espinosa *et al.* 2006, Labella *et al* 2016; Labella *et al.* 2017; Forcada-Nadal *et al.* 2018), thus enabling control by PII of gene expression regulation in cyanobacteria (Llácer *et al.* 2010).

In previous work we determined (Llácer *et al.* 2010) a crystal structure of NtcA bound to PipX and proposed that this binding would stabilize the active form of NtcA that should be competent for binding to the DNA target sequence of NtcA (the NtcA box), an imperfect palindrome with three symmetrical specificity-conferring base pairs at 5-7 nucleotides upstream and downstream from the palindrome centre (Luque *et al.* 1994). Unhappily, no experimental structure of the DNA-NtcA complex could be obtained, and, therefore, this proposal was speculative, as it was, too, our proposal that PipX, by interacting with RNAP, could perhaps help recruit this polymerase to the NtcA binding site in DNA (Llácer *et al.* 2010). Unfortunately, the highly valuable structures of the NtcA TAC at the *nirA* promoter from Han *et al.* (2024) did not include PipX and did not directly address the specificity of DNA binding to NtcA.

Here we define the structural determinants of NtcA binding to its sites of regulation in the DNA by providing a 3 Å crystallographic structure of *Synechococcus elongatus* strain PCC7942 (from now on abbreviated as *S. elongatus*) NtcA with DNA and 2-OG bound. We also provide the crystal structure of the quaternary complex of NtcA

with PipX, 2OG and DNA and show that PipX, without contacting promoter DNA sequences, substantially enhances the binding of NtcA to its target site. We also confirm our previous proposal that PipX binding freezes NtcA in its DNA binding-competent conformation. Importantly, by modelling our DNA-NtcA-PipX complex onto the NtcA-TAC structure (PDB 8H4O) from Han et al. (2024) (a structure that did not contain PipX), and by subsequently testing interactions between PipX and relevant domains of the sigma and gamma subunits of RNAP in the bacterial two-hybrid system, we strongly support that, indeed, PipX enhances NtcA-specific transcription by contributing to the recruitment of the RNAP to the NtcA sites in DNA. Thus, our present work, in addition to defining in great structural detail the determinants of promoter selectivity of NtcA, expands the repertoire of mechanisms of transcriptional regulation, calling attention to the PipX coactivator protein, a specific regulatory feature of cyanobacteria that conceivably could also occur in other taxonomic bacterial groups by using other potential small protein coactivators.

Materials and Methods

Production and crystallization of NtcA-DNA, PipX-NtcA-DNA and of NtcA in inactive apo forms.

NtcA and PipX from *Synechococcus elongatus* (strain PCC7942) with N-terminal His_6 tags were produced in *E. coli* Rosetta (DE3) cells (from Novagen) from plasmids pTrc99A-*pipX* and pET15b-*NtcA,* respectively, and were purified essentially as reported (Espinosa *et al.* 2006, Llácer *et al.* 2010). To produce the V187E mutant form of NtcA, site-directed mutagenesis of the pET15-NtcA plasmid was carried out using the Quickchange II system (from Stratagene, La Jolla, CA), following the manufacturer's instructions, and utilizing the oligonucleotide pair 5'TCGGCTCAACGCGGGAGACAGTGACG3' and 5'CGTCACTGTCTCCCGCGTTGAGCCGATCG3'. After corroborating by DNA sequencing the introduction of the mutation, the NtcA mutant was expressed and purified as the corresponding wild-type protein.

The DNA used in the crystallization of the NtcA-2OG-DNA complex was prepared by 10-min heating at 65ºC of an equimolar mixture of the synthetic oligodesoxyribonucleotides 5'-AGCTGA**TAC**ATAAAAAT-3' and 5'-CATTTTTAT**GTA**TC-3' (from Sigma) in 5 mM Tris-HCl, pH 7.5, followed by slow cooling to 4ºC for annealing. These oligonucleotides include one half of the NtcA binding

site of the *S. elongatus glnA* promoter (consensus bases are in bold type), to produce upon annealing the following chained duplex (underlining and italics are used to differentiate individual oligonucleotides; bold-type, consensus bases in the NtcA box):

5'CATTTTTAT**GTA**TC*AGCTGA**TAC**ATAAAAAT 3'*

3'TAAAAATA**CAT**AGTCGA*CTA**TG**TATTTTTAC 5'*

Then, 0.3 mg/ml of this annealed DNA mixture and 0.85 mg/ml NtcA in buffer A (50 mM Na citrate, pH 6.5, 0.5 M NaCl, 5 mM $MgCl_2$, 50 mM L-arginine-HCl, 50 mM Na L-glutamate and 10 mM 2OG), was incubated 5 min at 22ºC, and concentrated to 7.3 mg protein/ml by ultrafiltration in an Amicon Ultracel YM-50 device. Crystallization sitting drops were prepared by mixing 0.4 µl of the protein-DNA solution and 0.4 µl of crystallization solution. The crystal used was obtained at 21ºC in a crystallization solution consisting of 0.1 M Bis-Tris, pH 6.5 and 28% polyethylene glycol (PEG) monomethyl ether 2K.

The procedure to crystallize the PipX-2OG-NtcA-DNA complex was similar to that for the NtcA-DNA complex, although the annealed DNA contained the whole perfect NtcA box (consensus bases are in bold type) in a symmetric non-cleaved 30-bp DNA duplex with sticky ends (5' protruding C and 3' overhanging G in one and the other oligonucleotides) to create the following duplex:

5'CATTTTTAT**GTA**TCAGCTGA**TAC**ATAAAAAT

TAAAAATA**CAT**AGTCGACT**ATG**TATTTTTAG5'

We first prepared the NtcA-DNA complex by incubating 10 min at 22ºC 0.6 mg/ml of this annealed DNA and 1.4 mg/ml NtcA in buffer A. Then 0.6 mg/ml PipX was added, the mixture was incubated 10 more min and then it was concentrated to ~6 mg protein/ml by centrifugal ultrafiltration. Crystallization sitting drops were prepared as above, obtaining after several weeks two types of crystals with the following crystallization solutions: for crystal type I, 0.2 M $MgSO_4$ and 20% PEG 3350; and for crystal type II 0.1 M Tris-HCl, pH 8.5, 2 M ammonium phosphate and 0.1 M $ZnCl_2$. The first crystals obtained of this complex were of type II. They diffracted X-rays at 6 Å-resolution, with 63% solvent content. The resolution was improved to 4.3 Å in one crystal by controlled dehydration for 30 min via vapor diffusion versus 50 µl of a solution of 4 M KCl. Initial phases were obtained with this crystal (see below). Then we obtained crystals of type I that diffracted X-rays up to 3.8 Å resolution.

Crystals of NtcA without 2OG or DNA were obtained in three different forms from a 6.1 mg/ml NtcA solution prepared as above except for the omission of 2OG from

all solutions and from buffer A, using as crystallization solutions, for form A1, 0.1 M Na acetate pH 4.5, 40% PEG 200; for form A2, 0.1 M Na HEPES pH 6.5, 20.5% PEG 4 K; and for form B1, 0.05 M Tris-HCl pH 8, 2 M NaCl, 2% 2-methyl-2,4-pentanediol, 10 mM $MgCl_2$ and 5% DMSO.

Data collection and structure determination.

Crystals were flash-frozen in liquid N_2. Cryoprotection of crystal II of PipX-NtcA-DNA was as indicated in the previous section. The crystal of NtcA apo form A1 did not need added cryoprotection. All other crystals were cryoprotected by harvesting them in their corresponding crystallization solutions fortified as indicated: for NtcA-DNA, 38% PEG monomethyl ether 2K; for type I crystals of PipX-NtcA-DNA, 39% PEG monomethyl ether 3.35K; and for NtcA apo forms A2 and B1, 25 % PEG 400, and 27% 2-methyl-2,4-pentanediol, respectively.

Diffraction was at 100 K in the indicated beamlines (Table 1) of the ESRF or Diamond synchrotrons. Datasets obtained for NtcA A1 and B1, at respective resolutions of 2.70 and 3.33 Å, and for PipX-NtcA-DNA I and II at 3.8 and 4.3 Å resolutions, respectively, were processed and scaled with MOSFLM and SCALA [CCP4 suite (Collaborative Computational Project Number 4, 1994)], whereas those for NtcA-DNA (3.00 Å resolution) and NtcA apo A2 (2.85 Å) were processed with XDS (Kabsch *et al.* 2010), and prepared with COMBAT (Collaborative Computational Project Number 4, 1994) for scaling with SCALA. See Table 1 for the space groups and cell dimensions for the different crystals.

Crystallographic phases for the NtcA-DNA crystal (containing also one bound molecule of 2OG per subunit) were determined by molecular replacement with MOLREP (Vagin *et al.* 2010) using as a model the *S. elongatus* NtcA dimer in its active conformation (PDB accession number 2XHK; Llácer *et. al.* 2010). The initial model was improved by rigid body and maximum likelihood restrained refinement using REFMAC (Murshudov *et al.* 2011). The phases obtained from this partially refined solution were improved by density modification using program PARROT (Zhang *et al.* 1997). An initial model for DNA was manually built using COOT (Emsley *et al.* 2010). Then, iterative rounds of model building with COOT and refinement with REFMAC were carried out. Tight non-crystallographic symmetry (NCS) restraints were applied until the final rounds of refinement, when they were gradually released. TLS (Winn *et al.* 2001) was used in the last steps of refinement, defining the TLS groups after TLSMD analysis (Painter *et*

al. 2006). All the diffraction data were used throughout the refinement processes, except the 5% randomly selected data for calculating R_{free}. The stereochemical quality of the model was analysed with RAMPAGE (Lovell *et al.* 2002). The final refined model, at 3.00 Å-resolution, contained one NtcA dimer and two chains of each oligodeoxyribonucleotide (Table 1). 2OG was modeled in the NtcA-DNA structure as seen in the active *Anabaena* structure (Zhao *et al.* 2010)

One subunit of the structure of active NtcA from *S. elongatus* was also used as a model for phasing the dataset for NtcA A1, by molecular replacement with MOLREP. Model building with COOT and refinement with REFMAC and TLS were as for the NtcA-DNA complex. One of the subunits of this final model (which contained two subunits in the asymmetric unit) was used for phasing the dataset for NtcA A2 by molecular replacement using PHASER (McCoy *et al.* 2007). A satisfactory solution was found, consisting in one dimer in the asymmetric unit. Model building and refinement were as above to a final resolution of 2.85 Å (Table 1). For NtcA B1, molecular replacement with MOLREP using a subunit of the inactive conformation found previously with NtcA from *S. elongatus* (PDB accession number 2XKP, Llácer *et al.* 2010) yielded a satisfactory solution, consisting in one subunit in the asymmetric unit, which was refined as with the other datasets, generating a dimer by application of the crystal symmetry.

Crystal II of the PipX-NtcA-DNA complex was originally assigned by XDS to space group I422. From space considerations (Matthews coefficient calculations; Collaborative Computational Project Number 4, 1994), the asymmetric unit could accommodate one NtcA dimer, one DNA molecule and two PipX molecules. Initial attempts to find the phases by molecular replacement using the structures of NtcA-DNA (see above) or of NtcA-PipX (Llácer *et al.* 2010) as searching models failed with both MOLREP and PHASER. We obtained satisfactory phasing solutions by using as searching models either the NtcA dimer alone or the NtcA-PipX complex with only one PipX monomeric molecule. In all these cases, the existence of packing problems in this I422 group for reasonably accommodating a second PipX molecule suggested the possibility of merohedral twinning which would increase the apparent symmetry of the crystal, thus requiring space-group re-evaluation. To assess the twinning of the PipX-NtcA-DNA crystal complex, analyses of the intensity statistics were performed in *phenix.xtriage*, from the *PHENIX* software package (Adams *et al.* 2010) for an I4 data set, determining the twin law governing the merohedrally twinned crystal as (*–h, k, -l*),

with a twinning fraction around 0.45 (0.43 or 0.48 for H-test or the Maximum likelihood method, respectively). In fact, the use of the dataset for space group I4 instead of I422 only reduced the R_{sym} value from 8.5% to 7.4%, which is another indication of the nearly perfect (α=0.5) merohedral twinning of the crystal. Thus, molecular replacement in space group I4 carried out with PHASER worked with a model composed of one NtcA dimer, one DNA double helical molecule and one PipX molecule, with Matthew's coefficient of 3.25 $Å^3 D^{-1}$ and 62.2% solvent content. This initial model was subjected to rigid body refinement using REFMAC, applying the twin operator (-*h*, *k*, −*l*) during each round of refinement. Then, and according to the density map, some residues were removed from the model whereas other regions were moved as rigid domains using COOT. The model was finally refined with REFMAC, using the jelly body refinement tool and applying the twin operator (-*h*, *k*, −*l*) during each round of refinement. Local non-crystallographic symmetry (NCS) restraints were automatically applied. The final refined model contained in the asymmetric unit two NtcA dimers bound to two DNA duplexes, where one NtcA dimer had two PipX molecules bound as previously reported for the NtcA-PipX dimer (Llácer *et al.* 2010). The other NtcA dimer molecule only had one PipX site canonically occupied by one PipX monomer, while the other site contacted aberrantly another PipX molecule from neighboring DNA-$NtcA_2$-$PipX_2$ complex.

Then we obtained a dataset for crystal I at 3.8 Å resolution. Analyses of the intensity statistics in *phenix.xtriage* revealed translational pseudosymmetry. Considering that the resulting symmetry for the initially assigned space group P21 can be defined as $P12_11$ by using the universal Herman-Mauguin space group symbols, molecular replacement was carried out with PHASER in this space-group, using as searching assembly the DNA-$NtcA_2$-$PipX_2$ complex of crystal II, yielding a good solution of two such assemblies in the asymmetric unit, with a Matthew's coefficient of 2.43 $Å^3 D^{-1}$ and 49.5% solvent content. After rigid body refinement with REFMAC, model building in COOT was alternated with a combined refinement using Phenix.refine (14) and REFMAC using the jelly body refinement tool. In REFMAC, local non-crystallographic symmetry (NCS) restraints were automatically applied. In spite of the limited resolution, both PipX-NtcA-DNA structures (I and II) exhibit very good geometries (Table 1).

EMSA assay.

A fluorescein (Flc)-labeled 159 bp DNA fragment containing the *glnA* gene promoter was generated by PCR-amplification with Deep Vent DNA polymerase (New

England Biolabs), the primer pair 5´-(Flc)CACAACCAGGAACTGAAGAC-3´ and 5´-(Flc)CGCCTGCAAGATTTCGTTAC-3´, and genomic DNA of *S. elongatus* as template. The amplified fragment, purified with GeneClean (MP Biomedicals), was used in gel retardation assays by first incubating 10 min at 32ºC 30 ng of this labeled DNA, in 20 µl of a solution containing 50 mM Hepes pH 8, 3 mM $MgCl_2$, 20% glycerol, 1 mM dithiothreitol, 25 µg/ml bovine serum albumin and 0.5 µg of poly[d(I-C)] (from Roche), and, when indicated, 100 ng of purified NtcA, either wild type or carrying the V187E mutation, and 3.2 mM 2OG. At the end of the incubation the mixtures were subjected to electrophoresis at 120V for 90 min at 4ºC in native 6% polyacrylamide gels in 11 mM Tris-HCl pH 8, 11 mM boric acid and 0.25mM EDTA (low ionic strength TBE buffer). Experiments were repeated independently in duplicate. Fluorescein fluorescence was visualized using a FUJIFILM- FLA-5000 imaging system.

Modelling of an NtcA-PipX-DNA-RNAP transcriptional activator complex.

To model the interactions between RNA polymerase and NtcA-bound PipX, we superimposed our PipX-NtcA-DNA complex from crystal I onto the NtcA-TAC of *Anabaena* (PDB 8H4O) (Han *et al.* 2024). Our crystal I structure comprises two NtcA-PipX-DNA complexes, with NtcA subunits labeled as A, B, D, and F. Each NtcA subunit from our complex was superimposed onto the NtcA subunit of the NtcA-TAC that is closest to the sigma subunit of RNAP. In each of the four superimpositions, the corresponding PipX coordinates were integrated to generate a PDB file. These PDB files were then analyzed using the PISA server (Krissinel *et al.* 2007) to identify interfacing residues between PipX and the RNA polymerase subunits. The results showed consistent PipX interaction sites, regardless of which NtcA subunit was superimposed. For graphical representations, the superposition of our NtcA subunit F on NtcA subunit X was used. Essentially the same results were obtained when the model was generated by superimposing on the NtcA dimer of the NtcA-TAC structure, one or the other of our two NtcA dimers in our PipX-NtcA-DNA complex of crystal I.

Bacterial two hybrid (BATCH) assays.

Strains and plasmids and oligonucleotides used for this purpose are listed in Supplementary Table S1. Cloning procedures were carried out with *E. coli* XL1-Blue, using standard techniques (Sambrook, 1989). The construction of BACTH plasmids involved, in brief, the PCR amplification of the relevant coding region, its digestion with two enzymes, and its cloning into the corresponding BACTH vectors previously cut with

the same enzymes. The correctness of all the constructs was confirmed by automated dideoxy DNA sequencing.

After transformation of *E. coli* BTH101 with pairwise combination of plasmids (50 ng) and subsequent selection on ampicillin (50 µg ml^{-1}) and kanamycin (40 µg ml^{-1}) LB plates, five clones from each plate were inoculated into 0.5 mL of LB plus antibiotics and 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) and incubated at 30 ºC for 24 hours. Interactions were assayed by dropping 3 µl of each saturated culture on M63 reporter agar plates containing 0.3 % maltose, 0.0001 % thiamine, 1 mM magnesium sulphate, 0.5 mM IPTG, and 40 µg ml^{-1} X-gal; and on MacConkey reporter plates containing 1% lactose and 0.5 mM IPTG. Reporter plates were incubated for 24 (MacConkey) or 48 hours (M63) at 30 ºC and photographs were taken at 24 h intervals.

Other methods.

Protein concentrations were routinely assayed by the method of Bradford (Bradford 1976) using a commercial reagent from Bio-Rad and bovine serum albumin as a standard. Sequence alignments were carried out with ClustalW (Thompson *et al.* 1994), using default values. Superposition of structures was carried out with programs SSM (Krissinel *et al.* 2004) and LSQKAB [CCP4 (Collaborative Computational Project Number 4, 1994)]. Contacts between NtcA and DNA were calculated using PISA (Krissinel *et al.* 2007) or program Contacts (Collaborative Computational Project Number 4, 1994). DNA helical parameters were analysed using w3DNA (Zheng *et al.* 2009). Figures of protein structures were generated using Pymol (http://pymol.sourceforge.net/).

Results

The structure of the NtcA-DNA-2OG ternary complex

We determined at 3 Å resolution the crystal structure of NtcA (Table 1 and Fig. 1A) bound to its target DNA, a palindrome called the NtcA box (consensus sequence, $GTAN_8TAC$) (Luque *et al*. 1994). For the purpose of crystallization, this box was prepared as a 30 bp DNA duplex incorporating a symmetric NtcA box derived from the first half of the corresponding box of the *S. elongatus glnA* promoter (which is strongly regulated by NtcA) (Reyes *et al*. 1997; Forcada-Nadal *et al*. 2014) and the preceding nine bases. In this way, a symmetric, synthetic, and expectedly high affinity NtcA box was

prepared (see below Fig. 1D). To facilitate DNA bending (characteristic for CRP family regulators) this DNA had in each strand a break point between the GTA and TAC parts of the box, with cohesive intervening ends to maintain duplex integrity (see Material and Methods).

In the *S. elongatus* NtcA homodimer bound to its target DNA (Fig. 1A) the DNA binding domains (DBDs) of both subunits are wrapped around by a strongly bent (~80º) DNA duplex, with the bending largely due to sharp kinks between bases -5 and -6 and 5 and 6 on each side of the dyad axis. In this structure the protein interacts directly with 20 of the 30 base pairs of the DNA duplex (Fig. 1A-D), with all the interactions except one (with R29, a residue of the effector binding domain, EBD) being mediated by the two DNA binding domains (DBDs). The major groove of the double helix, hosting in each of two successive turns a half of the palindromic signature sequence, hosts in each turn the initial two turns of the DNA binding helix (helix F) of one subunit or the other subunit of NtcA (Fig. 1B). The axes of the helices are approximately perpendicular to the direction of the DNA chain. In this way, amino acid side chains emerging radially from each F helix can interact with the surrounding bases and phosphates in each turn of the major groove (Fig 1B, C). In addition, phosphates in the minor groove of the DNA preceding and following the consensus motive make some contacts with protein amino acids (Fig 1B, D). The NtcA dimer in this complex closely resembles the NtcA dimer reported previously in the "active" 2OG-bound conformation lacking DNA from *S. elongatus* (Llácer *et al*. 2010) or from *Anabaena* sp. (Zhao *et al*. 2010) (rmsd for the superimposition of the dimers, 0.72 Å for 421 C^{α} atoms or 0.77 Å for 388 C^{α} atoms, respectively; and see superimposition in Supplementary Fig. S1A). Indeed, our present DNA-bound structure also exhibits the effector 2OG bound on both EBDs via direct and water-mediated hydrogen bonds (inset of Fig. 1A).

Our structure of DNA-bound NtcA is also highly similar to that of the NtcA dimer contained in the recently reported cryoEM structure of NtcA-TAC complex from *Anabaena*, a complex that also includes DNA (Han *et al.* 2024) (rmsd for the superimposition of the dimers, 2.46 Å for 369 C^{α} atoms) (Fig. 1E). Actually, our NtcA-DNA complex is also quite similar to the complex of CRP with its target DNA (Supplementary Fig. S1B) (rmsd for the superimposition of 353 C^{α}-atoms with the *E. coli* CRP dimer bound to DNA, 2.69 Å). The reported structures of DNA complexes of other members of the CRP-FNR family are also highly similar (Supplementary Fig. S1C) (Parkinson *et al.* 1996, Levy *et al.* 2008, Bonnet *et al.* 2013, Hall *et al.* 2016, Werel *et al.*

2023), and thus our present results add up to the existing evidence in favour that all members of the CRP transcription factor superfamily activate transcription by binding to DNA essentially in the same way.

NtcA box recognition

As already mentioned, the interactions between NtcA and DNA are mediated by residues located in the helix-turn-helix motif of NtcA, with the specificity being provided by residues of helix αF, which is the one that is inserted in the major groove of DNA (Fig. 1B). Nevertheless, there are abundant interactions with phosphate groups of the DNA (Fig 1B, D), which contribute to the stability of the binding of NtcA to the DNA, although expectedly, not to the specificity of such binding for the NtcA box. Particular stabilization is likely to be obtained from ion pairs between the positively charged side chains of R142, R197 and K207 and the phosphate groups of the DNA (Fig. 1D and Table 2). Particular mention deserves such interaction between R142 and the DNA because, unlike in other members of the CRP family, in which each subunit of the transcription factor exclusively interacts with one half of the DNA consensus sequence, in the case of R142 of NtcA there is cross-interaction with the half of the NtcA box that predominantly interacts with the other subunit of NtcA. Interestingly, the position of R142 (a residue located in the dimerization helix αC), changes depending on the binding or lack of binding of 2OG to NtcA (Llácer *et al.* 2010; and see below). Only in the presence of 2OG is R142 positioned at an appropriate distance to interact with the DNA, while in the absence of 2OG the orientation of αC is such that R142 is too far to make contact with the DNA. Therefore, the indicated (Fig. 1D) interaction of R142 with the DNA phosphate appears particularly important, possibly being a key determinant for enhancing NtcA affinity for its box when 2OG is bound (see Discussion).

As already mentioned, only residues of helix F can confer specificity to the interaction with the NtcA box, since only four residues of this helix, all of them invariant in the NtcA sequences across cyanobacterial species, R186, V187, T190, and R191, directly contact DNA bases (Table 2). In the other determined structure of NtcA bound to DNA (cryoEM structure of the NtcA-TAC of *Anabaena;* Han *et al.* 2024), at least three corresponding residues, R187, V188, and R192, also contact the DNA bases, supporting the importance of these interactions. Two residues, R186 and R191 (numbering for the *S. elongatus* NtcA sequence), form hydrogen bonds with the DNA bases, while V187 and T190 contribute with van der Waals contacts. The guanidinium group of R186 forms two

hydrogen bonds with the O6 and N7 atoms of an invariant G:C pair in the NtcA box, at 7 base pairs from the symmetry axis of the palindrome (Figs. 1B and 1D). R191 similarly makes bonds with the O6 and N7 atoms of a guanine at 3 base pairs from the symmetry axis, outside the conserved bases of the consensus sequence (Figs. 1B and 1D). This could partially explain the different affinities of NtcA for different promoters that have identical consensus sequences but that show variability in the eight central bases of the palindrome (Jiang *et al.* 2000; Vázquez-Bermúdez *et al.* 2002). Massive transcriptomics data (Mitschke *et al.* 2011) identified a G at this position only in ~50% of the putative NtcA promoters of *Anabaena sp*.

Concerning residues V187 and T190, they interact with the methyl group of a thymine at 6 base pairs from the DNA symmetry axis (Fig. 1B and 1C, and its inset). V187 is particularly notable for distinguishing NtcA from CRP, being invariant across cyanobacteria, whereas it is replaced by glutamate in CRPs. A mutation substituting V187 by glutamate (as in CRP) abolishes NtcA's ability to bind to its target DNA, as shown in electrophoretic mobility shift assays (EMSA) (Fig. 1F). This underscores the critical role of V187 in NtcA-DNA interaction. The EMSA in the absence and presence3 of 2OG suggest that NtcA's DNA binding mechanism plays a key role for binding of NtcA to its DNA box even in the absence of 2OG (Fig. 1F).

Structures of the PipX-NtcA-2OG-DNA complex.

We determined the structure of the quaternary complex of NtcA with DNA, 2OG and PipX using two crystals (crystals I and II, Table 1) of this complex grown under different conditions, and diffracting X-rays at 3.8 and 4.3 Å resolutions, respectively. Unless indicated, we will refer to the structure of crystal I (Fig. 2A-C) since it is the one obtained at the highest resolution (3.8 Å). Crystal I contained in the asymmetric unit (Fig. 2A) two complete PipX-NtcA-DNA complexes in two DNA duplexes chained by the C and G overhangs, totalizing two 30-bp-long DNA molecules hosting two NtcA dimers and four NtcA-bound PipX molecules. Despite the limited resolution, the R_{factor}/R_{free} values were quite good (Table 1) and, as illustrated in Fig. 2C, the map had excellent quality, allowing the tracing of most amino acid side chains, including those at the interfaces between the different components of the complex, what allowed quite detailed mapping of the interactions between the NtcA and DNA.

The two complexes in the asymmetric unit of crystal I are very similar (rmsd of 0.97 Å for 594 superimposed C^{α} atoms). The NtcA dimer and PipX have the same

topography and conformation as in the complex of these two components of *S. elongatus* in the absence of DNA (Llácer *et al*., 2010) (rmsd of 1.04 ± 0.02 Å for 547±1 superimposed C^{α} atoms), with preservation of the PipX-NtcA interactions observed in the PipX-NtcA complex (Llácer *et al.* 2010). PipX in this complex adopts the *flexed* conformation (Fig. 2A, B) in which the C-terminal helix (helix B runs antiparallel to the preceding helix A forming a two-helix bundle. The flexed conformation was found also when PipX is free (not bound to any partner, Forcada-Nadal *et al.* 2017), whereas the B helix was found to be extended in at least one of the three PipX molecules trapped in its complex with PII (Llácer *et al.* 2010).

DNA is bound as in the DNA-NtcA complex (Fig. 1A), and PipX does not interact with DNA (Fig. 2A). As already mentioned, the interactions of NtcA with the DNA are highly similar to those in the NtcA-2OG-DNA ternary complex, with just three extra interactions made by each NtcA subunit with each half of the palindrome (Fig. 1D, highlighted in blue; and Table 3). Because of this increase in the number of interactions, PipX might activate NtcA transcription not only by the previously proposed mechanism (Llácer *et al.* 2010) of stabilizing the "active" conformation of NtcA, but also by increasing the intrinsic affinity of the "active" NtcA dimer for the target DNA, as shown in surface plasmon resonance experiments (Forcada-Nadal *et al.* 2014).

Although the lower resolution of the diffraction data for complex II (Fig. 2D, E) does not favour detailed structural analysis (despite the observation of quite a good map for this resolution, Fig. 2E), the PipX molecules in crystal II (Fig. 2D, E) also have their C-terminal helices flexed and do not contact the DNA, and the individual components in the complexes observed in this crystal are architecturally very similar to the corresponding elements in crystal I and in the NtcA-PipX complex having no DNA bound (Llácer *et al.* 2010). For example, rmsd values for superimposition of all the PipX molecules in the complexes in crystals I, II and in the NtcA-PipX complex without DNA (Llácer *et al.* 2010) were very low (mean of 0.62±0.09 Å for 85±1 C^{α} atoms). Furthermore, the architectures of the NtcA-PipX complexes in crystals I and II are highly similar (rmsd values of 1.00-1.09 Å for 509-594 superimposed C^{α} atoms). Nevertheless, an interesting finding in crystal II was the lack of one PipX molecule: of the two NtcA dimers bound to DNA, one has two PipX molecules bound as in the complexes observed in crystal I and in the NtcA-PipX complex (Llácer *et al.* 2010), whereas the other one has one normally bound PipX molecule while the second site for PipX is occupied by the free

end of the protruding flexed helices of one of the two PipX molecules of the other dimer (Fig. 2D, labelled "aberrant binding of PipX"). Although this mode of binding may not occur in solution, the complexes in crystal II reveal that NtcA-DNA complexes with canonical (Llácer *et al.* 2010) occupation of only one PipX site might occur *in vivo*, raising the question of whether the complexes of NtcA with two PipX molecules or with one PipX molecule bound could differ in their transcription activation efficiency. As shown below, the existence of a complex with only one molecule of PipX bound would fit the possibility that one PipX molecule could interact with RNA polymerase.

Modelling of PipX -RNAP interactions in the NtcA-TAC complex

To model the potential interactions between PipX and RNA polymerase in the NtcA-TAC complex, we superimposed our DNA-NtcA-PipX complex (Crystal I) onto the NtcA-TAC structure (PDB 8H4O) (Han *et al.* 2024), following the methodology described in the Materials and Methods section. The resulting model indicated that PipX could fit into the complex without causing any steric clashes (Fig. 3A-C). The model also shows that only one PipX molecule interacts with the RNA polymerase (Fig.3A and Supplementary Fig. S2), consistent with the structural observation in our DNA-NtcA-PipX complex (crystal II), where only one PipX molecule is canonically bound to NtcA. This suggests the redundancy of the second PipX molecule for transcriptional regulation and reinforces the functional relevance of our observed structure.

The model predicts that PipX contacts through its C-terminal helices A and B with domain 4 of the sigma subunit (region 4) of RNA polymerase, as well as with the gamma subunit and, to a lesser extent, with the omega subunit (Fig. 3A-C). PISA server analysis predicts that the gamma subunit, via residues I393, R394, S395, G396, M397, and N399, could interact with helix A of PipX (potentially interacting residues of PipX, N55, R58, M59, D62, N63, and R66); that the sigma subunit region 4 could contact via residues P379, N380, S383, V384, K386, and E387 with PipX helices A (residues R54, N55 and R58) and B (residues Q86, T87, F88, and L89), whereas PipX residues R59, R66, and R70 would be positioned near residues S49, A50, and E51 of the omega subunit of the polymerase (Fig. 3C).

Functional evidence of PipX-RNAP interactions by bacterial two-hybrid assays

Functional support for PipX-RNA interaction was obtained in pilot studies with the bacterial two-hybrid (BACTH) system. This is a genetic approach to test for protein-protein interactions based on the reconstitution of the adenylate cyclase of *Bordetella*

pertussis in *E. coli* (Karimova *et al.* 1998; Battesti and Bouveret, 2012) when the T18 and T25 domains of this enzyme, each one in a chimera with an interacting partner, are placed close together as a consequence of the interaction. This approach appears not to yield false-positive results (Jerez *et al.* 2021). Our studies (Fig. 3D) using this technique provided clear support for the interaction in vivo of PipX with portions of the gamma and sigma subunits of RNAP that encompass the residues predicted by our structural modelling to participate in the interactions. PipX C-terminally fused to the T25 domain, yielded positive responses (monitored in two independent colour assays, see Materials and Methods and Fig. 3D) with the fragments of both the gamma and the sigma subunits of RNAP linked at their C-termini with the T18 domain. The fact that the results were negative when the T25 domain was linked to the N-terminus of PipX (Fig. 3D) suggests insufficient mutual approach of the linked domains when the T25 domain is N-terminal, also constituting a reassuring negative control for the positive results when the T25 domain was C-terminal to PipX. As a further control, we tested the influence of deleting helix B of PipX on the interaction with the sigma domain fragment, since the prediction from the structural modelling was that three residues of helix B participate in the interactions, in addition to 3 residues from helix A (see above). Therefore, the deletion of helix B might be expected to produce a decreased but not nil response. Preliminary tests for the strongest possible response (needed to maximize the possibility of detection of weakly positive responses upon helix B deletion), led us to use chimeras in which the T18 and T25 domains were linked to the C-termini of PipX and the sigma subunit fragment, respectively. The results, shown in the framed inset in Fig. 3D, confirmed the expectations, with the removal of the last helix ($PipX^{1-70}$) decreasing but not abolishing the color response in both assays tested. Overall, these results agree with the predictions made by the structural modelling.

Newly identified NtcA Apo forms and NtcA Activation by 2OG

NtcA exhibits structural variability depending on its active or inactive state dictated by the presence or absence of 2-oxoglutarate (2OG). Active NtcA, either alone or complexed with PipX (Llácer *et al.* 2010) and DNA (present studies), maintains an essentially constant conformation that closely resembles that of other active Crp/Fnr superfamily members, highlighting a conserved mechanism of activation (Supplementary Fig. S3).

In contrast, the reported "inactive" forms of *Anabaena* NtcA without 2OG (Zhao *et al*. 2010) (called here form A) or of *S. elongatus* NtcA with aberrantly bound 2OG (Llácer *et al.* 2010) (called here form B) show significant mutual structural differences (RMSD of 3.67 Å for 306 Cα atoms; and see below) where A resembles more the active form than form B. By determining new crystal structures of *S. elongatus* NtcA (Table 1) in the absence of 2OG, we now report three additional inactive conformations which have no 2OG bound (called "apo forms") (Fig. 4A). Two novel forms (A1 and A2, Fig. 4A) resemble but are not identical to the A form reported previously for *Anabaena* NtcA (Zhao *et al*. 2010) (Supplementary Fig. S4, leftmost two and one structures of the 1st and 2nd rows, respectively). These forms exclude that the A-type inactive form were species-specific for *Anabaena*. The third novel inactive form reported now (form B1, Fig. 4A) closely resembles the previously reported (Llácer *et al.* 2010) inactive form B (Supplementary Fig. S4, the two rightmost structures of the 1st row), but lacks the aberrantly bound 2OG found in the B form, with important differences with B in the 2OG site, particularly affecting residues 76-78, and 88, having nearly buried the empty 2OG site (Supplementary Fig. S5A). Despite their mutual similarities, forms A1 and A2 presented different crystal packing (C2221 and P212121, respectively), as was the case for form B1, which belonged to space group P6422 (Table 1).

The comparison of the various inactive conformations of NtcA (epitomized by the structures shown in Fig. 4A) with the highly constant active conformation of this transcriptional regulator hosting correctly bound 2OG (Figs. 1A and E; 2 B and E and 3A) shows that 2OG triggers a conformation exhibiting rearrangements in the EBD that propagate throughout the NtcA dimer, significantly impacting the relative positioning of the DNA-binding helices (αF) (Fig. 4B). In the active state, the distance between the DNA-interacting helices of the two subunits is short (34.8 Å), aligning perfectly for DNA binding at the two adjacent turns of the NtcA box where they sit (Figs. 1A and E and 2B and E). In contrast, these transcriptionally "inactive" forms exhibit changes of orientation and displacements of the F and F* DNA binding helices of the dimer (the asterisk signals the other subunit of the NtcA homodimer), reflected in increases in their mutual distance (to 37.6, 38.4 and 53.1 Å in A1, A2 and B1, respectively; Fig. 4A), thus preventing proper interaction of these F helices with DNA for transcriptional activation (Fig. 4B). This alteration is closely tied to shifts in the interfacial α-helices (αCs), which in active NtcA are close to the twofold axis of the dimer, forming a true coiled-coil, whereas in inactive forms they diverge more or become more parallel (Fig. 4A,B). Key to these reorientations

of the C-helices are 2OG-triggered changes in the interactions across the interface that involve the side-chains of E133 and E137 (Fig. 4 C, D and Supplementary Fig. S5B).The changing partners of these two glutamate side-chains depending on whether 2OG is bound or it is not bound, favors the reorientation of the C-helices and, secondarily, of the DNA-binding domains, which become poised in a fixed conformation fitting the two turns of the major groove of the double helix in the region of the NtcA box. A critical feature in this structural modulation is the β4-β5 hairpin (Fig. 4E). In the active conformation, the interaction between 2OG and E133 disrupts prior contacts involving R55 and Y89 from the β4-β5 hairpin, freeing the latter to adopt a DNA-compatible conformation. Furthermore, this conformational change may be key to understand the second mechanism, besides increasing promoter affinity, by which 2OG could activate NtcA-dependent transcription, as seen in other family members of class II transcriptional activators such as FNR, CooA, and CRP (Bell & Busby, 1994; Leduc *et al.* 2001; Lawson *et al*. 2004). The β4-β5 hairpin belongs to the region called AR3 ("activator region 3"), the area of the transcription factor that interacts with the sigma subunit of RNA polymerase in class II TACs. Moreover, in NtcA the conformation of the β4-β5 hairpin is also important for the interaction with PipX, since in "active" NtcA, this hairpin interacts with the C-terminal end of the C-terminal helix (helix G) of the neighboring subunit (Fig. 5A), so that the conformation of this helix changes to one suitable for interacting with PipX (Llácer *et al.* 2010).

In summary, our results show that NtcA, as other factors of the CRP-FNR superfamily, can adopt multiple transcriptionally inactive conformations (Supplementary Fig. S4) in the absence of its small-molecule allosteric effector (2OG in the case of NtcA), in contrast with the constancy of the active conformation of the transcriptional regulators of this family (Supplementary Fig. S3) (Popovych *et al.* 2009; Sharma *et al.* 2009; Gallagher *et al.* 2009; Kumar *et al.* 2010; Lanzilotta *et al.* 2000; Townsend *et al.* 2014; Seok *et al.* 2014).

Discussion

Transcription activation by NtcA and other members of the Crp/Fnr superfamily relies on their ability to specifically bind DNA in their active states, followed by subsequent interaction with RNA polymerase (RNAP) (Kuchinskas *et al.* 2006, Mazon *et al.* 2007, Agari *et al.* 2008, Llácer *et al.* 2010, Bonnet *et al.* 2013, Townsend *et al.* 2014, Volveda *et al.* 2015, Feng *et al.* 2016, Hall *et al.* 2016). NtcA and several related

factors target class II activated promoters, where key interactions occur approximately 41.5 nucleotides upstream of the transcription start site, facilitating RNAP recruitment and the formation of active transcription complexes (Valladares *et al.* 2008, Leduc *et al.* 2001, Lawson *et al.* 2004, Bell & Busby 1994). Activation typically involves a rotation of the DBD upon induction, optimizing the arrangement of DNA recognition helices for binding (Sharma *et al.* 2009, Lanzilotta *et al.* 2000). The binding of the allosteric effector and nitrogen poorness sensor 2OG is crucial for NtcA activation, as it enables specific interactions with the NtcA box that are necessary for effective RNAP recruitment.

Recently, the structure of a NtcA-TAC complex from *Anabanena* has been determined using cryo-electron microscopy (Han *et al.* 2024), providing insights into how DNA looping facilitates transcription activation by NtcA. DNA looping allows NtcA to bring together distantly located DNA-binding sites, enabling the transcription factor to effectively recruit RNA polymerase and activate transcription in a cooperative manner. Despite the overall relatively high resolution of the cryo-electron microscopy structure (3.6 Å), local resolution for NtcA and the NtcA box was rather low (worse than 5.5 Å-resolution), and therefore our present structure of NtcA bound to its target DNA at 3.0 Å resolution offers greater detail regarding the specific recognition of the NtcA box.

NtcA binding induces substantial DNA bending, which is consistently observed in both our structure and the NtcA-TAC complex (Figure 1E). The specific recognition of the NtcA box is critically dependent on four invariant amino acid residues: Arg186, Val187, Thr190, and Arg191. These residues are directly involved in binding of DNA bases, thus conferring specificity, with Arg186 and Arg191 forming hydrogen bonds with conserved guanine bases in the NtcA box, while Val187 and Thr190 also recognize NtcA box-defining DNA bases through van der Waals interactions. Thus, this precise interaction pattern is essential for the ability of NtcA to recognize and bind to its target DNA sequences. The importance of these residues is further highlighted by the fact that when Val187 is replaced by a glutamate (as found in CRP proteins), NtcA loses its ability to bind to the NtcA box, as observed in electrophoretic mobility shift assays (Fig. 1F) and in agreement with previous reports showing the inability of CRP proteins to bind to NtcA promoters (Forcada-Nadal *et al.* 2014). This mutation introduces steric and polarity clashes that prevents proper interaction with the DNA of the NtcA box, underscoring the critical role of Val187 in maintaining correct base contact.

In addition to these direct interactions, NtcA also engages in extensive contacts with the DNA phosphate backbone, involving residues such as Arg29, Met144, His175,

Thr185, and Thr188. Some of these interactions are shared with CRP, but others, particularly those mediated by Arg142, are unique to NtcA. Arg142 is especially interesting because it does not interact with the DNA site directly adjacent to its subunit but instead crosses over to interact with the DNA bound by the opposite NtcA subunit (Fig.4D). Moreover, this residue not only interacts with a phosphate group from the DNA but also forms a salt bridge with the invariant E61 of the β4-β5 turn from the regulatory domain (EBD) of the adjacent NtcA subunit.

The movement of the β4-β5 hairpin that also differentiates the active and inactive states of NtcA (Fig. 4E), is integral to its function, involving an ion-pair network centered around Arg142, connecting the DNA with dynamic elements of NtcA like helix C and the already mentioned β4-β5 hairpin (Fig. 4C, D and Supplementary Fig. S5B). The different conformations of this β hairpin depending on the activation state of NtcA (Fig. 4E) are connected to movements of the HTH DNA binding motif of the same subunit (Fig. 4D). In the inactive forms, interactions between E133 from one subunit of NtcA and R55 from the other, promote a different conformation of the β4-β5 hairpin, affecting the HTH (Figs. 4C,D and Supplementary Fig. S5B). The β4-β5 hairpin is part of the "activator region 3" (AR3), a region of the transcription factor that interacts with the sigma subunit of RNA polymerase (Bell & Busby 1994, Leduc *et al.* 2001, Lawson *et al.* 2004). In agreement with this assignation, in the NtcA-TAC complex (Han *et al.* 2024) NtcA interacts with sigma subunit region 4 (σR4) (Fig. 3A) through residues D39, E59, E62, and E63 (Fig. 3C), of which the latter three are located in the β4-β5 hairpin (Fig. 5B). Superimposition of the inactive NtcA A1 form onto the NtcA-TAC complex highlights the shift of the position of the β4-β5 hairpin upon 2OG binding, which moves closest towards the sigma subunit of RNAP, thus strengthening the interaction of NtcA with RNAP (Fig.5B) and also highlighting why the binding of 2OG promotes transcriptional activation by NtcA via promotion of effective interaction with RNAP. Interestingly, the αG helix, the C-terminal helix of NtcA which is characteristic of this protein, not being present in CRP and mediating some of the interactions of NtcA with PipX (Llácer *et al*. 2010), undergoes positional shifts upon 2OG binding to NtcA. The shifts should facilitate NtcA interaction with PipX (Fig. 5A), likely contributing to the coactivation process driven by PipX.

Determination of the PipX-NtcA-DNA complex structure has been essential in elucidating the co-activator role of PipX in the NtcA transcription system. Our structure reveals that PipX is positioned away from the DNA (Fig. 2A, B, D), and in this complex

with DNA, both the NtcA dimer and PipX retain the same conformation observed in the NtcA-PipX complex without DNA (Llácer *et al.* 2010). Despite the moderate resolution of our structure of the DNA-NtcA-PipX complex (3.8 Å), our data clearly indicate additional interactions between NtcA and DNA within the ternary complex which are absent in the DNA-NtcA binary complex (Fig. 1D), possibly explaining part of the co-activator role of PipX as an indirectly enhancer of the number of interactions between NtcA and DNA.

Another important factor in the co-activating role of PipX is its obvious stabilization of the active conformation of NtcA that is promoted by 2OG binding to this transcriptional activator (Llácer *et al.* 2010). By binding selectively to the active form, PipX shifts the equilibrium between inactive and active form in favor of the latter form, which is the one that is most competent for DNA binding and the one that, once bound to DNA, promotes transcriptional activity. Actually, each PipX molecule interacts with the EBD of one NtcA subunit and with the DNA-binding domain (DBD) of the other subunit of the dimer (Llácer *et al.* 2010), in practice stabilizing the correct orientation of this last domain and particularly of its F helix for effective DNA binding. These roles of PipX in enhancing NtcA box binding of 2OG-activated NtcA, without being directly involved in interacting with DNA is reflected in the fact that PipX is required for maximal expression of non-canonical NtcA-sensitive promoters like *glnN* (Espinosa *et al.* 2006), helping orchestrate maximal NtcA responses of cyanobacteria to conditions of nitrogen starvation.

We previously proposed (Llácer *et al.* 2010) that PipX could also help recruit RNAP to DNA-bound NtcA, thus enhancing NtcA transcriptional activation not only by increasing the affinity of NtcA for its target DNA sequences, but also by making more effective transcriptional initiation at the sites of NtcA binding. This proposal was based on the fact that significant portions of the surface of PipX, particularly its two C-terminal helices, remained exposed in the PipX-NtcA complex, something that is confirmed now in the structure of Ntca-PipX-DNA, indicating potentiality for interactions of NtcA-bound PipX with additional components of the transcriptional machinery, particularly with RNAP. This is presently supported by our model in which we superimposed our PipX-NtcA-DNA complex structure on the NtcA-TAC lacking PipX (Han *et al.* 2024), leading to excellent fitting of PipX in the supercomplex, where the RNAP C-terminal helices interact with region 4 of the sigma subunit of RNAP, as well as with the gamma subunit, and, to lesser extent, with the omega subunit. Pilot functional BACTH interaction

studies *in vivo* (Fig. 3D) support the interactions of PipX with the sigma and gamma subunits of RNAP, specifically with the regions of the gamma and sigma subunits of RNAP predicted to interact with PipX in our superposition model (Fig. 3). Interestingly, our model suggests that a single PipX molecule would be enough for these interactions, and thus for helping recruit RNAP towards NtcA DNA targets, which actually aligns with some of our structural data (crystal II) in which one of the two DNA-bound NtcA molecules only had one PipX site properly occupied by this co-activator macromolecule (Fig. 2D). Taken together, experimental crystal structures, structural modelling and BACTH results support a role of PipX not only in enhancing the ability of NtcA to bind to DNA and to stabilize the active conformation of NtcA for proper DNA binding, but also to favor the recruitment of the transcriptional machinery to the NtcA binding sites on DNA.

However, this might not be all the functions of PipX with respect to NtcA-mediated transcription. In the structure of the NtcA-TAC complex obtained by Han et al. (2024) (PDB 8H4O), NtcA deviates from the typical behavior of class II activators (Busby *et al.* 1999, Lawson *et al.* 2004) by not interacting with the α subunit of RNAP's N-terminal domain (αNTD) or C-terminal domain (αCTD). This raises the intriguing possibility that PipX might be necessary for these interactions to occur. Furthermore, PipX could engage not only in the NtcA-TAC complex but also in the complex of DNA wth NtcA-TAC-NtcB also reported by Han et al. (2024) (PDB 8H3V). The NtcB in this complex is a second transcription factor (of the LysR-type) identified as mediator of nitrite-triggered transcription in *Synechococcus*; Aichi *et al.* 1997). Again, NtcA-bound PipX could be integrated in that complex without causing steric clashes (Supplementary Fig. S6), suggesting a potential mechanism for combined regulation of certain genes, which might require the coordinated action of two transcription factors and PipX, particularly under conditions of nitrogen depletion. This coactivation role of PipX could be crucial in fine-tuning gene expression in response to environmental signals, especially in pathways where multiple regulatory proteins are involved.

The broader implications of these findings are significant for understanding transcriptional regulation in bacteria. The PipX-NtcA-DNA complex serves as a model for the CRP-FNR superfamily, illustrating how co-activators like PipX can modulate transcription factor activity. These results emphasize the importance of the PipX-NtcA interaction in transcriptional regulation, particularly under nitrogen starvation conditions in cyanobacteria. The structural details elucidated here pave the way for further research

into the regulatory mechanisms of other CRP-FNR family members that might involve still unrecognized small coactivating proteins. Ultimately, our findings might have applications in biotechnology and medicine, particularly in the development of strategies to control bacterial gene expression.

Acknowledgements.

Supported by grant PID 2020-116880GB-I00 of the Plan Estatal de I+D+I of the Spanish Government to JLL, grants PID2020-118816GB-I00 and PID2023-149456NB-I00 to AC and by grants BFU2017-84264-P from the Spanish Ministry of Science of the Spanish Government; and CIV20A6610 of the Fundación Ramón Areces, both of them to VR. SB was supported by a National Grant from the Algerian Ministry of Higher Education and Scientific Research.

Authors' contributions.

JLL and VR conceived the work. AF and JLL did most of the experimental work excepting the BACTH experiments which were done by SB and PS under AC direction. VR and JLL supervised the work and, together with AF, wrote the paper. JLL and AF draw the structural figures. All the authors read the manuscript, made suggestions concerning its contents and agreed with its contents.

References

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66:213-21.

Agari Y, Kashihara A,Yokoyama S, Kuramitsu S, Shinkai A (2008). Global gene expression mediated by *Thermus thermophilus* SdrP, a CRP/FNR family transcriptional regulator. *Mol. Microbiol.* 70:60-75.

Aichi M, Omata T (1997). Involvement of NtcB, a LysR family transcription factor, in nitrite activation of the nitrate assimilation operon in the cyanobacterium *Synechococcus sp.* strain PCC 7942. *J. Bacteriol*. 179:4671–4675.

Battesti A, Bouveret E (2012). The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods* 58:325-34.

Bell A, Busby S (1994). Location and orientation of an activating region in the *Escherichia coli* transcription factor, FNR. *Mol. Microbiol.* 11:383-390.

Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM, Ebright RH (2002). Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* 297**:**1562-1566.

Bonnet M, Kurz M, Mesa S, Briand C, Hennecke H, Grutter MG (2013). The structure of *Bradyrhizobium japonicum* transcription factor FixK2 unveils sites of DNA binding and oxidation. *J. Biol. Chem.* 288; 14238-14246.

Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.

Burillo S, Luque I, Fuentes I, Contreras A (2004). Interactions between the nitrogen signal transduction protein PII and N-acetyl glutamate kinase in organisms that perform oxygenic photosynthesis. *J. Bacteriol.* 186:3346-54.

Busby S, Ebright RH (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* 293:199-213.

Chen S, Vojtechovsky J, Parkinson G, Ebright RH, Berman HM (2001). Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: DNA binding specificity based on energetics of DNA kinking. *J. Mol. Biol.* 314:63-74.

Chen S, Gunasekera A, Zhang X, Kunkel TA, Ebright RH, Berman HM (2001). Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: alteration of DNA binding specificity through alteration of DNA kinking. *J. Mol. Biol.* 314:75-82.

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and Development of Coot. *Acta Crystallogr. D.* 66:486-501.

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.* 61:457-469.

Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F, Contreras A (2014). PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 111:E2423-30.

Feng Y, Zhang Y, Ebright RH (2016). Structural basis of transcription activation. *Science* 352:1330-3.

Forcada-Nadal A, Forchhammer K, Rubio V (2014). SPR analysis of promoter binding of Synechocystis PCC6803 transcription factors NtcA and CRP suggests cross-talk and sheds light on regulation by effector molecules. *FEBS Lett.* 588:2270-2276.

Forcada-Nadal A, Palomino-Schätzlein M, Neira JL, Pineda-Lucena A, Rubio V (2017). The PipX Protein, When Not Bound to Its Targets, Has Its Signaling C-Terminal Helix in a Flexed Conformation. *Biochemistry* 56:3211-3224.

Forcada-Nadal A, Llácer JL, Contreras A, Marco-Marín C, Rubio V (2018). The PII-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. *Front. Mol. Biosci*. 13;5:91.

Gallagher DT, Smith N, Kim SK, Robinson H, Reddy PT (2009). Profound asymmetry in the structure of the cAMP-free cAMP Receptor Protein (CRP) from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 284:8228-8232.

Giner-Lamia J, Robles-Rengel R, Hernández-Prieto MA, Muro-Pastor I, Florencio FJ, Futschik ME (2017). Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp. PCC 6803. *Nucl. Ac. Res.* 45:11800-11820.

Hall M, Grundström C, Begum A, Lindberg MJ, Sauer UH, Almqvist F, Johansson J, Sauer-Eriksson AE (2016). Structural basis for glutathione-mediated activation of the virulence regulatory protein PrfA in Listeria. *Proc. Natl. Acad. Sci. USA*. 113:14733-14738.

Han SJ, Jiang YL, You LL, Shen LQ, Wu X, Yang F, Cui N, Kong WW, Sun H, Zhou K, Meng HC, Chen ZP, Chen Y, Zhang Y, Zhou CZ (2024). DNA looping mediates cooperative transcription activation. *Nat. Struct. Mol. Biol.* 31:293-299.

Herrero A, Muro-Pastor AM, Flores E (2001). Nitrogen control in cyanobacteria. *J. Bacteriol.* 183:411-425.

Herrero A, Muro-Pastor AM, Valladares A, Flores E (2004). Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol. Rev.* 28:469-487.

Jerez C, Salinas P, Llop A, Cantos R, Espinosa J, Labella JI, Contreras A (2021). Regulatory Connections Between the Cyanobacterial Factor PipX and the Ribosome Assembly GTPase EngA. *Front. Microbiol*. 12:781760

Jiang F, Wisén 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.* 301:783-793.

Karimova G, Pidoux J, Ullmann A, Ladant D (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. USA*. 95:5752-6.

Körner 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.* 27:559-592.

Krissinel E, Henrick K (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D.* 60:2256-2268.

Krissinel E, Henrick K (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774-797.

Krol E, Werel L, Essen LO, Becker A (2023). Structural and functional diversity of bacterial cyclic nucleotide perception by CRP proteins. *Microlife* 1;4:uqad024.

Kuchinskas M, Li H, Conrad M, Roberts G, Poulos TL (2006). The role of the DNA-binding domains in CooA activation. *Biochemistry* 45:7148-53.

Kumar P, Joshi DC, Akif M, Akhter Y, Hasnain SE, Mande SC (2010). Mapping conformational transitions in cyclic AMP receptor protein: crystal structure and normal-mode analysis of Mycobacterium tuberculosis apo-cAMP receptor protein. *Biophys. J.* 98:305-14.

Labella JI, Obrebska A, Espinosa J, Salinas P, Forcada-Nadal A, Tremiño L, Rubio V, Contreras A (2016). Expanding the Cyanobacterial Nitrogen Regulatory Network: The

GntR-Like Regulator PlmA Interacts with the PII-PipX Complex. *Front. Microbiol.* 7:1677.

Labella JI, Cantos R, Espinosa J, Forcada-Nadal A, Rubio V, Contreras A (2017). PipY, a Member of the Conserved COG0325 Family of PLP-Binding Proteins, Expands the Cyanobacterial Nitrogen Regulatory Network. *Front. Microbiol*. 8:1244.

Lanzilotta WN, Schuller DJ, Thorsteinsson MW, Kerby RL, Roberts GP, Poulos TL (2000). Structure of the CO sensing transcription activator CooA. *Nat. Struct. Biol.* 7:876-880.

Lawson CL, Swigon D, Murakami KS, Darst SA, Berman HM, Ebright RH (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* 14:10-20.

Leduc J, Thorsteinsson MV, Gaal T, Roberts GP (2001). Mapping CooA-RNA polymerase interactions. Identification of activating regions 2 and 3 in CooA, the CO-sensing transcriptional activator. *J. Biol. Chem.* 276:39968-39973.

Levy C, Pike K, Heyes DJ, Joyce MG, Gabor K, Smidt H, van der Oost J, Leys D (2008). Molecular basis of halorespiration control by CprK, a CRP-FNR type transcriptional regulator. *Mol. Microbiol.* 70:151-167.

Liu B, Hong C, Huang RK, Yu Z, Steitz TA (2017). Structural basis of bacterial transcription activation. *Science* 358:947-951.

Llácer JL, Fita I, Rubio V (2008). Arginine and nitrogen storage. *Curr. Opin. Struct. Biol.* 18:673-81.

Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K, Rubio V (2010). Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc. Natl. Acad. Sc.i USA* 107:15397-15402.

Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2002.) Structure validation by Cα geometry: ϕ/ψ and Cβ deviation. *Proteins* 50:437-450.

Luque I, Flores E, Herrero A. (1994). Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J.* 13:2862-9.

Mazon H, Gábor K, Leys D, Heck AJ, van der Oost J, van den Heuvel RH. (2007). Transcriptional activation by CprK1 is regulated by protein structural changes induced by effector binding and redox state. *J. Biol. Chem.* 282:11281-11290.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. (2007) Phaser crystallographic software. *J. Appl. Cryst.* 40:658-674.

Mitschke J, Vioque A, Haas F, Hess WR, Muro-Pastor AM (2011). Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc. Natl. Acad. Sci. USA* 108:20130-20135.

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* 67:355-367.

Napoli AA, Lawson CL, Ebright RH, Berman HM (2006). Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: recognition of pyrimidine-purine and purine-purine steps. *J. Mo.l Biol.* 357:173-183.

Painter J, Merritt EA (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* 62:439-450.

Parkinson G, Wilson C, Gunasekera A, Ebright YW, Ebright RH, Berman HM (1996). Structure of the CAP-DNA complex at 2.5 angstroms resolution: a complete picture of the protein-DNA interface. *J. Mol. Biol.* 260:395-408.

Popovich 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* 106:6927-6932.

Reyes JC, Muro-Pastor MI, Florencio FJ (1997). Transcription of glutamine synthetase genes (glnA and glnN) from the cyanobacterium *Synechocystis* sp. strain PCC 6803 is differently regulated in response to nitrogen availability. *J. Bacteriol.* 179:2678-2689.

Roberts MG, Dent MR, Ramos S, Thielges MC, Burstyn JN (2024). Probing conformational dynamics of DNA binding by CO-sensing transcription factor, CooA. J *Inorg. Biochem*. 259:112656.

Sambrook, J., Fritsch, E. R., & Maniatis, T (1989). Molecular Cloning: A Laboratory Manual (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Seok, S.-H., Im, H., Won, H.-S., Seo, M.-D., Lee, Y.-S., Yoon, H.-J., Cha, M.-J., Park, J.-Y. and Lee, B.-J. (2014). Structures of inactive CRP species reveal the atomic details of the allosteric transition that discriminates cyclic nucleotide second messengers. *Acta Crystallogr. D*, 70: 1726–1742.

Schultz SC, Shields GC, Steitz TA (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253:1001-1007.

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* 106:16604-16609.

Shi W, Jiang Y, Deng Y, Dong Z, Liu B (2020). Visualization of two architectures in class-II CAP-dependent transcription activation. *PLoS Biol.* 18(4):e3000706.

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.* 33:5156-5171.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.

Townsend, P.D., Jungwirth, B., Pojer, F., Bußmann, M., Money, V.A., Cole, S.T., Pühler, A., Tauch, A., Bott, M., Cann, M.J. and Pohl, E. (2014). The crystal structures of apo and cAMP-bound GlxR from *Corynebacterium glutamicum* reveal structural and dynamic changes upon cAMP binding in CRP/FNR family transcription factors. *Plos One* 9:113265.

Vagin A, Teplyakov A (2010). Molecular replacement with MOLREP. *Acta Crystallogr. D* 66:25-25.

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.* 190:6126-6133.

Vázquez-Bermúdez MF, Flores E, Herrero A (2002). Analysis of binding sites for the nitrogen-control transcription factor NtcA in the promoters of *Synechococcus* nitrogen-regulated genes. *Biochim. Biophys. Acta* 1578:95-98.

Vega-Palas MA, Flores E, Herrero A (1992). NtcA, a global nitrogen regulator from the cyanobacterium *Synechococcus* that belongs to the Crp family of bacterial regulators. *Mol. Microbiol.* 6:1853-9.

Volbeda, A., Darnault, C., Renoux, O., Nicolet, Y. and Fontecilla-Camps, J.C. (2015). The crystal structure of the global anaerobic transcriptional regulator FNR explains its extremely fine-tuned monomer-dimer equilibrium. *Science advances* 1:1501086.

Werel L, Farmani N, Krol E, Serrania J, Essen LO, Becker A (2023). Structural Basis of Dual Specificity of *Sinorhizobium meliloti* Clr, a cAMP and cGMP Receptor Protein. *mBio.* 14(2):e0302822.

Winn MD, Isupov MN, Murshudov GN (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* 57:122-133.

Zhang KYJ, Cowtan K, Main P (1997). Combining constraints for electron-density modification. *Methods Enzymol.* 277:53-64.

Zhao MX, Jiang YL, He YX, Chen YF, Teng YB, Chen Y, Zhang CC, Zhou CZ (2010). Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proc. Natl. Acad. Sci. USA* 107:12487-12492.

Zheng G, Lu XJ, and Olson WK (2009) Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* 37 (Web Server issue): W240-W246.

Table 1. Data collection and refinement statistics

	Crystals					
	1 NtcA-DNA	**2 NtcA-PipX-DNA Crystal I**	**3 NtcA-PipX-DNA Crystal II**	**4 Apo NtcA Form A1**	**5 Apo NtcA Form A2**	**6 Apo NtcA Form B1**
Beamline (synchrotron)	BM16 (ESRF)	I04.1 (Diamond)	ID23-1 (ESRF)	ID23-1 (ESRF)	BM16 (ESRF)	BM14 (ESRF)
Wavelength, Å	0.9797	0.9173	0.9772	0.9769	0.9797	0.9762
Space group	C2	$P12_11$	I4	$C222_1$	$P2_12_12_1$	$P6_422$
Unit cell a, b,c (Å) α,β,γ (º) if ≠ 90	120.0, 47.0, 139.6 β= 98.7	113.3, 68.2, 126.3 β= 115.1	165.7, 165.7, 177.6	95.2, 95.5, 160.8	66.5, 88.9, 99.7	96.4, 96.4, 126.9
Resolution[a] (Å)	24.2-3.00 (3.16-3.00)	114.4-3.80 (4.01-3.80)	82.9-4.30 (4.53-4.30)	80.4-2.65 (2.79-2.65)	66.3-2.85 (3.00-2.85)	83.5-3.33 (3.51-3.33)
Reflections, total / unique	41,267 / 15,466	44,255 / 16,087	59,483/ 16,116	73,746 / 21,137	65,283 / 14,209	36,587/ 5,530
Completeness[a] (%)	98.5 (97.3)	92.2 (86.6)	98.5 (98.5)	98.0 (98.0)	99.0 (95.7)	100 (100)
Multiplicity[a]	2.7 (2.4)	2.8 (2.4)	3.8 (3.7)	3.5 (3.3)	4.6 (3.6)	6.6 (6.9)
I/σ[a]	7.9 (1.8)	3.0 (1.6)	4.8 (1.6)	5.8 (1.8)	12.1 (2.3)	4.0 (1.8)
R_{sym}[a] (%)	8.5 (42.9)	19.3 (39.8)	7.4 (47.2)	5.7 (40.8)	5.6 (34.1)	13.6 (41.9)
R_{cryst} / R_{free} (%)	22.4/25.4	31.6/36.8	31.3/35.6	24.7/25.7	22.5/27.0	24.9/ 28.5
rmsd bond lenght (Å) rmsd bond angles (º)	0.007 1.26	0.009 1.05	0.005 1.01	0.009 1.25	0.008 1.10	0.007 1.03
Molecules or atoms / average B-factors ($Å^2$) Polypeptide chains Protein atoms 2-Oxoglutarate DNA atoms Water molecules	 2 3343/62.5 2/34.8 1230/78.2 5/31.4	 8 9548/61.4 4/54.5 2460/74.0 0	 7 8833/- 4/- 2460/- 0	 2 3098/96.1 0 0 12/42.4	 2 3101/73.8 0 0 33/26.1	 1 1652/61.5 0 0 0
Ramachandran plot (%) Favoured Allowed Outlier	 95.5 3.6 0	 95.3 4.3 0.3	 96.5 3.1 0.4	 96.2 3.8 0	 93.1 6.9 0	 91.5 8.5 0

[a]Values in parenthesis are data for the highest resolution shell.

Table 2. NtcA-DNA contacts

Half-complex A[a]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G-7	E	NH2	R186	2.41
O6	G-7	E	NH2	R186	3.19
O6	G-7	E	NH1	R186	3.41
O4	T-6	E	NH2	R186	3.47
N7	G-3’	D	NH2	R191	3.17
O6	G-3’	D	NH1	R191	3.19

Thymine methyl interactions (< 4.2Å)

C7	T-8	E	NH1	R186	4.01
C7	T-8	E	CD	R186	3.84
C7	T-8	E	NE	R186	3.67
C7	T-8	E	CZ	R186	3.78
C7	T-6	E	CG2	T190	3.71
C7	T-2’	D	NH2	R191	4.19

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T-2’	D	N	M144	2.60
OP1	G-3’	D	OG	S184	2.95
OP2	G-3’	D	OG1	T185	2.78
OP2	G-3’	D	OG1	T188	2.14
OP1	T-8	E	NZ	K207	3.00
OP2	T-8	E	N	H175	2.75
OP2	G-7	E	NH2	R197	3.15
OP1	G-7	E	NH2	R197	2.72
OP1	A-13′	D	NH1	R29	3.63

Cross-interactions

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Átomo j	Residue (chain)	Distance i-j (Å)
OP1	A4	F	NH1	R142 (A)	3.32
OP2	A4	F	NH1	R142 (A)	2.54
OP2	A-4’	D	NH1	R142 (B)	3.02
OP1	A4	F	NH1	R142 (A)	3.32

Half-complex B[b]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G7’	C	NH2	R186	2.17
O6	G7’	C	NH1	R186	3.34
O6	G7’	C	NH2	R186	3.43
N7	G3	F	NH2	R191	3.18
O6	G3	F	NH1	R191	3.39

Thymine methyl interactions (< 4.2Å)

C7	T8’	C	CZ	R186	4.05
C7	T8’	C	NE	R186	3.77
C7	T8’	C	CD	R186	3.81
C7	T6’	C	CB	T190	4.12
C7	T6’	C	CG2	T190	3.77
C7	T6’	C	CG1	V187	4.13

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T2	F	N	M144	2.72
OP1	G3	F	OG	S184	2.65
OP2	G3	F	OG1	T185	2.89
OP2	G3	F	OG1	T188	2.06
OP2	A4	F	OG1	T185	3.43
OP1	T8’	C	N	H175	2.92

[a]Defined as the half-complex dominated by the interactions of the A subunit of NtcA.

[b]Defined as the half-complex dominated by the interactions of the B subunit of NtcA

Table 3. NtcA-DNA contacts in PipX-NtcA-DNA complex (crystal I)

Half-complex A[a]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G-3’	C	NH2	R191	3.12
O6	G-3′	C	NH1	R191	3.51
N7	G-7	K	NH2	R186	2.58
O6	G-7	K	NH1	R186	3.33

Thymine methyl interactions (< 4.2Å)

C7	T-2’	C	CZ	R191	3.56
C7	T-2’	C	NH1	R191	3.94
C7	T-2’	C	NE	R191	3.73
C7	T-2’	C	NH2	R191	3.73
C7	T-5′	C	CG2	V187	3.76
C7	T-8	K	NE	R186	3.48
C7	T-8	K	CZ	R186	3.68
C7	T-8	K	NH2	R186	3.69

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T-2′	C	N	M144	3.16
OP2	G-3′	C	OG	S184	2.65
OP2	G-3′	C	N	T185	3.05
OP2	G-3′	C	OG1	T185	2.40
OP2	G-3′	C	OG1	T188	2.53
OP2	T-8	K	N	H175	2.92
OP1	G-7	K	NH2	R197	3.41
OP2	G-7	K	NH2	R197	3.07

Half-complex B[b]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G7'	G	NH2	R186	2.62
O6	G7'	G	NH1	R186	3.39
N7	G3	I	NH2	R191	2.97
O6	G3	I	NH1	R191	3.39

Thymine methyl interactions (< 4.2Å)

C7	T8′	C	CZ	R186	3.70
C7	T8′	C	NH2	R186	3.59
C7	T8′	C	NE	R186	3.54
C7	T5	K	CG2	V187	3.63
				T190	

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T8′	C	N	H175	3.08
OP1	G7′	C	NH1	R197	3.13
OP1	G7′	C	NH2	R197	2.71
OP1	T2	K	N	M144	3.26
OP1	G3	K	OG	S184	3.06
OP2	G3	K	OG1	T188	2.51
OP2	G3	K	OG	S184	2.91
OP2	G3	K	N	T185	3.33

Cross-interactions

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP2	A-4'	C	NE	R142	3.36

Table 3 (continued)

Half-complex D[c]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G-3'	G	NH2	R191	2.75
O6	G-3'	G	NH1	R191	3.28
N7	G-7	I	NH2	R186	3.38
O4	T-4	I	NH1	R191	3.44

Thymine methyl interactions (< 4.2Å)

C7	T-2′	G	CZ	R191	4.20
C7	T-5′	G	CG2	V187	3.46
C7	T-8	I	NE	R186	3.01
C7	T-8	I	CZ	R186	3.68
C7	T-8	I	NH2	R186	3.98
C7	T-8	I	CD	R186	3.48

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T-2′	G	N	M144	3.27
OP1	G-3′	G	OG	S184	3.11
OP2	G-3′	G	OG	S184	2.92
OP2	G-3′	G	OG1	T185	2.56
OP2	G-3′	G	OG1	T188	2.44
OP2	G-3′	G	N	T185	3.39
OP2	T-8	I	N	H175	2.80
OP1	G-7	I	NH2	R197	2.69

Half-complex F[d]

Base contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
N7	G7'	G	NH2	R186	2.62
O6	G7'	G	NH1	R186	3.39
N7	G3	I	NH2	R191	2.97
O6	G3	I	NH1	R191	3.39

Thymine methyl interactions (< 4.2Å)

C7	T8'	G	CZ	R186	3.64
C7	T8'	G	NH2	R186	3.65
C7	T8′	G	NE	R186	3.48
C7	T6′	G	CG2	T190	3.43
C7	T5	I	CG2	V187	3.74
C7	T2	I	NE	R191	4.00
C7	T2	I	CZ	R191	3.83
C7	T2	I	NH2	R191	4.06

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP1	T-2′	G	N	M144	3.27
OP1	G-3′	G	OG	S184	3.11
OP2	G-3′	G	OG	S184	2.92
OP2	G-3′	G	OG1	T185	2.56
OP2	G-3′	G	OG1	T188	2.44
OP2	G-3′	G	N	T185	3.39
OP2	T-8	I	N	H175	2.80
OP1	G-7	I	NH2	R197	2.69

Cross-interactions

Phosphate contacts

Hydrogen bonds

Atom i	Base	Chain	Atom j	Residue	Distance i-j (Å)
OP2	A-4'	G	NE	R142	3.45
OP2	A4	I	NH1	R142	3.43

[a]Defined as the half-complex dominated by the interactions of the A subunit of the NtcA dimer composed of subunit A and B.

[b]Defined as the half-complex dominated by the interactions of the A subunit of the NtcA dimer composed of subunit A and B.

[c]Defined as the half-complex dominated by the interactions of the D subunit of the NtcA dimer composed of subunit D and F.

[d]Defined as the half-complex dominated by the interactions of the F subunit of the NtcA dimer composed of subunit D and F.

Figure Legends

Figure 1. The NtcA-DNA complex structure. (*A*) Overall structure of the complex in ribbons, with bound 2OG shown in spheres representation. Some secondary structure elements of NtcA are labeled. One NtcA subunit is colored bluish and the other yellow, with the DNA binding domain shown in darker hues. The inset details the bound 2OG, modeled as in the NtcA-2OG *Anabaena* structure (Zhao *et al.* 2010), encased within its electron density omit map contoured at 2.5σ. Broken lines denote hydrogen bonds with the indicated residues. W, water. (*B*) Stereo view of the NtcA-DNA interface in one half of the NtcA-DNA complex. NtcA is colored yellow (except R142 and the loop between αC and αD from the neighbor NtcA subunit, which are bluish). One DNA chain is colored green and the other chain in gray. Residue side-chains and DNA are in sticks representation, and hydrogen bonds as black broken lines. (*C*) The DNA and helix F of NtcA to highlight the complementary Van der Waals surfaces and hydrophobic nature of the interactions of the V187 side-chain and of the bases at positions 5 and 6 of the DNA. The inset illustrates the fact that V187 and T190 clamp the methyl group of the thymine base at position 6' of the box, thus endowing it with specificity. (*D*) Diagram of the NtcA-DNA contacts in the NtcA-DNA and PipX-NtcA-DNA (crystal I) structures. The DNA bases that characterize the consensus NtcA box are highlighted in bold-type. While the majority of the residues shown interact with DNA in both the NtcA-DNA and PipX-NtcA-DNA complexes, those shadowed in blue only interact in the complex that includes PipX. NH144 and NH175 denote that the interactions shown of amino acids 144 and 175 are mediated by the main-chain NH group of these residues. NtcA residues contributing exclusively non-polar contacts are labeled red. The interactions with the bases and with the phosphate groups are differentiated. Double interactions with a base are marked with a double line. R142(A) and R142(B) denote that the interactions are with Arg142 of the other subunit than the one providing all other contacts with the B or A moiety of the palindromic target DNA sequence. The black ellipse signals the twofold symmetry of the palindrome. The black squares mark sites of primary kinks in the DNA. *(E)* Superposition of the structures of NtcA-DNA (blue, this study) and NtcA-DNA from the NtcA-TAC complex of *Anabaena* (salmon; PDB file 8H40; Han *et al*. 2024) to show the great similarity of both structures. *(F)* Gel retardation assays of biotinylated DNA encompassing the *S. elongatus glnA* promoter without (left) or with (panels to the right)

100 ng of NtcA either wild-type or hosting the V187E mutation, in the absence or, when indicated, in the presence of 3.2 mM 2OG. For details, see Materials and Methods.

Figure 2. The PipX-NtcA-DNA complexes in crystals I (*A-C)* and II *(D-E),* Proteins are in ribbon representation. The different components are labelled in the same color in which they are drawn. (*A, D) C*ontents of the entire asymmetric units of the corresponding crystals. *(B)* Close-up view of the PipX-NtcA-DNA complex from the crystal at 3.8 Å resolution, with PipX in green. The allosteric effector 2OG is shown in spheres representation (C and O atoms in yellow and red, respectively). *(C)* Illustrative example of the quality of the electron density map. 2Fo-Fc map contoured at 1σ (blue-violet grid) around the sticks model for DNA (one chain in magenta, the other one in orange) and for PipX (yellow) and NtcA (one subunit cyan and the other one green). *(E)* Close-up view of one complex of crystal II to show the 2Fo-Fc electron density map at 1σ encircling the model, to illustrate the quality of the map.

Figure 3. PipX-RNA Polymerase (RNAP) interactions. For clarity, only the Gamma, Omega, and domain 4 of the Sigma subunit of RNAP are depicted in all representations. *(A)* Cartoon representation of the model created by superimposing the structure of our PipX-NtcA-DNA complex from *S. elongatus* (crystal I) over the one of the NtcA-TAC complex (PDB 8H4O) of *Anabaena*. The common elements of the two structures (DNA and NtcA dimer) were superimposed with Pymol (http://pymol.sourceforge.net/). DNA is in the same color for both structures, but NtcA dimers are in salmon for *Anabaena* and blue for *S. elongatus*. PipX molecules are blue-cyan and the parts shown of the σ, γ and ω subunits of RNAP are gold, blue and orange, respectively. *(B)* and *(C)* Close-ups of the NtcA-TAC complex with modelled PipX on the basis of the superimposition in (*A*). Proteins are shown in transparent surface representation to exhibit inside their chain folds in cartoon representation. In *(C)* the zoom is increased to show the PipX-interfacing residues (see Materials and Methods) and their positions within the model, alongside the corresponding RNAP interfacing residues. Amino acids are identified in single-letter code, and the two C-terminal helices A and B of PipX are labelled. (*D*) Experimental evidence using pilot bacterial two-hybrid (BACTH) analyses of PipX interactions with the indicated regions of the Gamma and Sigma subunits of RNAP ($\sigma^{319-399}$ and $\gamma^{374-424}$), which encompass the residues proposed in panel (*C*) of this figure to mediate the interactions. The relative position of the T25 or T18 domains of *CyaA*-encoded adenylate cyclase in the indicated chimeras is shown by placing the name of the indicated domain

before or after the name of the potentially interacting protein. Representative photographs are shown for the results in the M63-maltose-X-gal assay after 48h of incubation at 30ºC (blue-green; to the left) or on MacConkey-lactose after 24h at 30ºC (red; to the right). For further details, see Jerez et al. 2021.

Figure 4. NtcA apo structures and movements of the interfacial helix, the DNA-binding helix and the β4-β5 hairpin upon activation. (*A*) Ribbon representations of NtcA in its novel apo structures, A1, A2 and B1, colored as in Fig.1A. The distance between the DNA-binding region of the F helices of both subunits are given in Å over a double-headed broken arrow. (*B*) Movements of the interfacial and the DNA-binding helices upon activation. The twofold axis of NtcA (black ellipse) is perpendicular to the paper. Only αB, the interfacial helices (αC) and the DNA-binding helices (αF) are shown as cylinders, and are labeled in the color of the indicated inactive or active form. An asterisk is used to differentiate helices from one subunit from those of the other subunit. *Active forms of NtcA:* NtcA-DNA (red), PipX-NtcA-DNA (orange), NtcA-2OG from *S. elongatus* (cyan; PDB file 2XHK). *Inactive forms of NtcA*: A1 (grey), A2 (green), B (yellow). The C and C* helices are approximately perpendicular to the paper and thus their colors are shown only at their bases. The DNA molecule belongs to the PipX-NtcA-DNA complex and is in backbone representation. *(C, D)* Mechanism of NtcA activation based on the structures of NtcA B1 *(C)* and NtcA-DNA *(D)* from *S. elongatus*. Ribbon representations showing key residues involved in signal transmission from the 2OG binding site to the DNA-binding helices. One NtcA subunit is coloured blueish and the other one greenish, with the residues of the latter subunit marked with an asterisk. Hydrogen bonds are depicted as dashed lines. An equivalent projection of inactive form A1 is shown as supplementary material (Supplementary Fig. S5B) to show that changes in essentially the same residues although of different magnitude do occur. (*C*) Comparison of β4-β5 hairpin conformation between the active and inactive forms of NtcA. NtcA-DNA (red), PipX-NtcA-DNA (brown), PipX-NtcA (cyan; PDB file 2XKO), NtcA-2OG from *S. elongatus* (blue; PDB file 2XHK), NtcA-2OG from *Anabaena* sp. (green; PDB file 3LA2), NtcA A1 (grey), NtcA A2 (black), NtcA A from *Anabaena* sp. (orange; PDB file 3LA7), NtcA B1 (yellow) and the reported NtcA B from *S. elongatus* in inactive conformation hosting aberrantly bound 2OG (magenta; PDB file 2XKP).

Figure 5. Involvement of movements of the β4-β5 hairpin and of helix G of NtcA on the activation of this transcriptional regulator by 2OG. (*A*) αG helix movement in the PipX-

NtcA-DNA complex (αG helix in red) relatively to the inactive NtcA A1 structure (αG helix in green). The β4-β5 hairpin of NtcA in the PipX-NtcA-DNA complex is colored in orange. 2OG is shown in spheres representation. (*B*) Superposition of the NtcA Apo A1 structure from *S. elongatus* (this work) onto the NtcA subunit that is closest to the sigma subunit of RNAP in the NtcA-TAC complex (PDB 8H4O), to highlighting the conformational change in the β4-β5 hairpin of NtcA upon 2OG binding. For clarity, only the sigma subunit domain 4 (yellow) and NtcA (salmon) of NtcA-TAC are shown, also showing (in light blue) the β4-β5 hairpin of the A1 structure, whereas the β4-β5 hairpin from the NtcA-TAC complex is hot pink. Three hairpin residues that interact with sigma R4 are shown as spheres and are labelled.

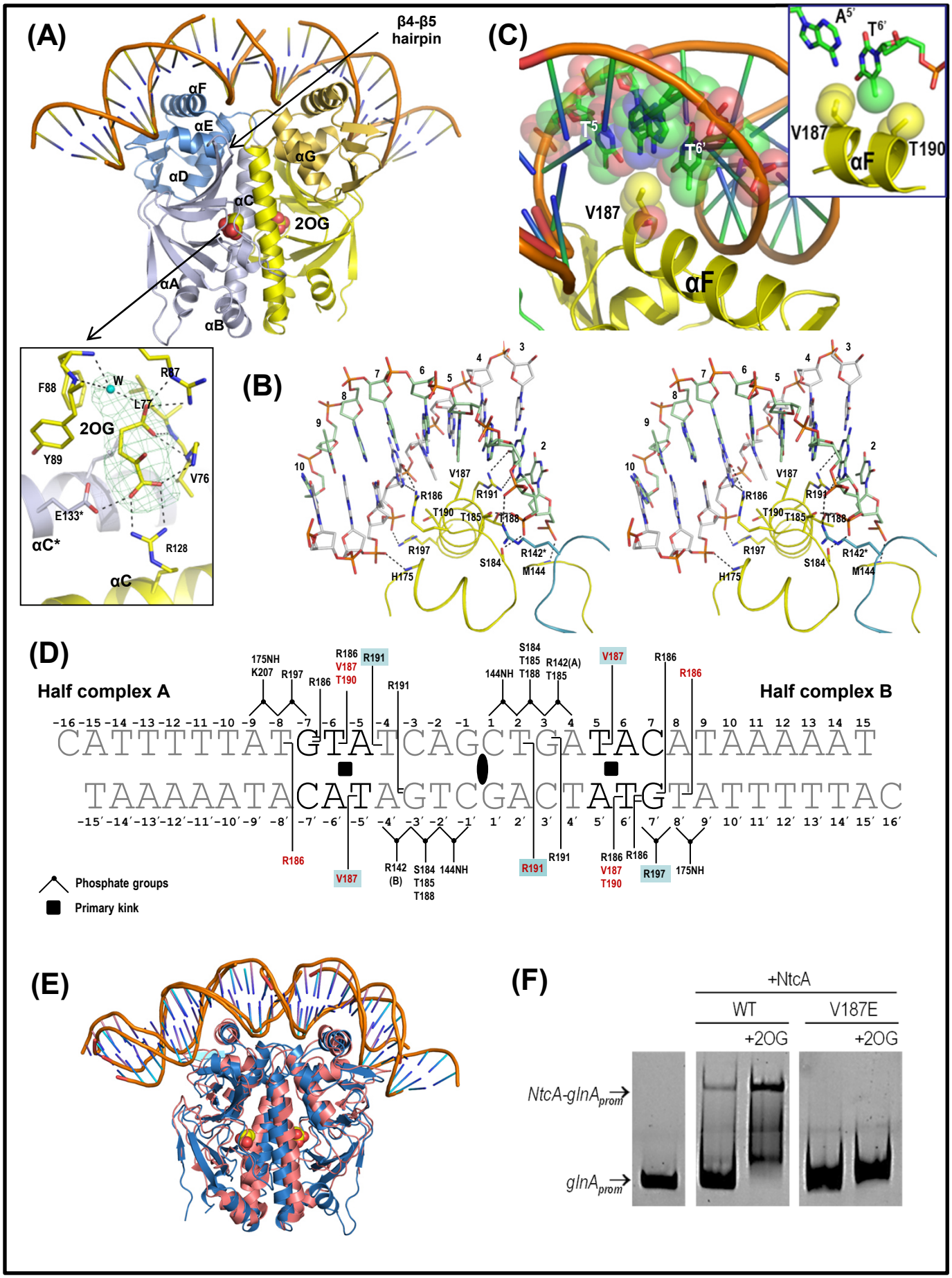
(A)
β4-β5 hairpin
αF
αE
αD
αC
αG
2OG
αA
αB
F88
W
R87
L77
2OG
Y89
V76
E133*
αC*
R128
αC
(B)
V187
R186
R191
T190
T185
T188
R197
S184
R142*
M144
H175
(C)
T5
T6'
V187
αF
A5'
T6'
V187
αF
T190
(D)
Half complex A
Half complex B
175NH
K207
R197
R186
R186
V187
T190
R191
R191
144NH
S184
T185
T188
R142(A)
T185
V187
R186
R186
CATTTTTATGTATCAGCTGATACATAAAAAT
TAAAAATACATAGTCGACTATGTATTTTTAC
R186
V187
R142
(B)
S184
T185
T188
144NH
R191
R191
R186
V187
T190
R186
R197
175NH
Phosphate groups
Primary kink
(E)
(F)
+NtcA
WT
V187E
+2OG
+2OG
NtcA-glnA prom
glnA prom

Fig.1

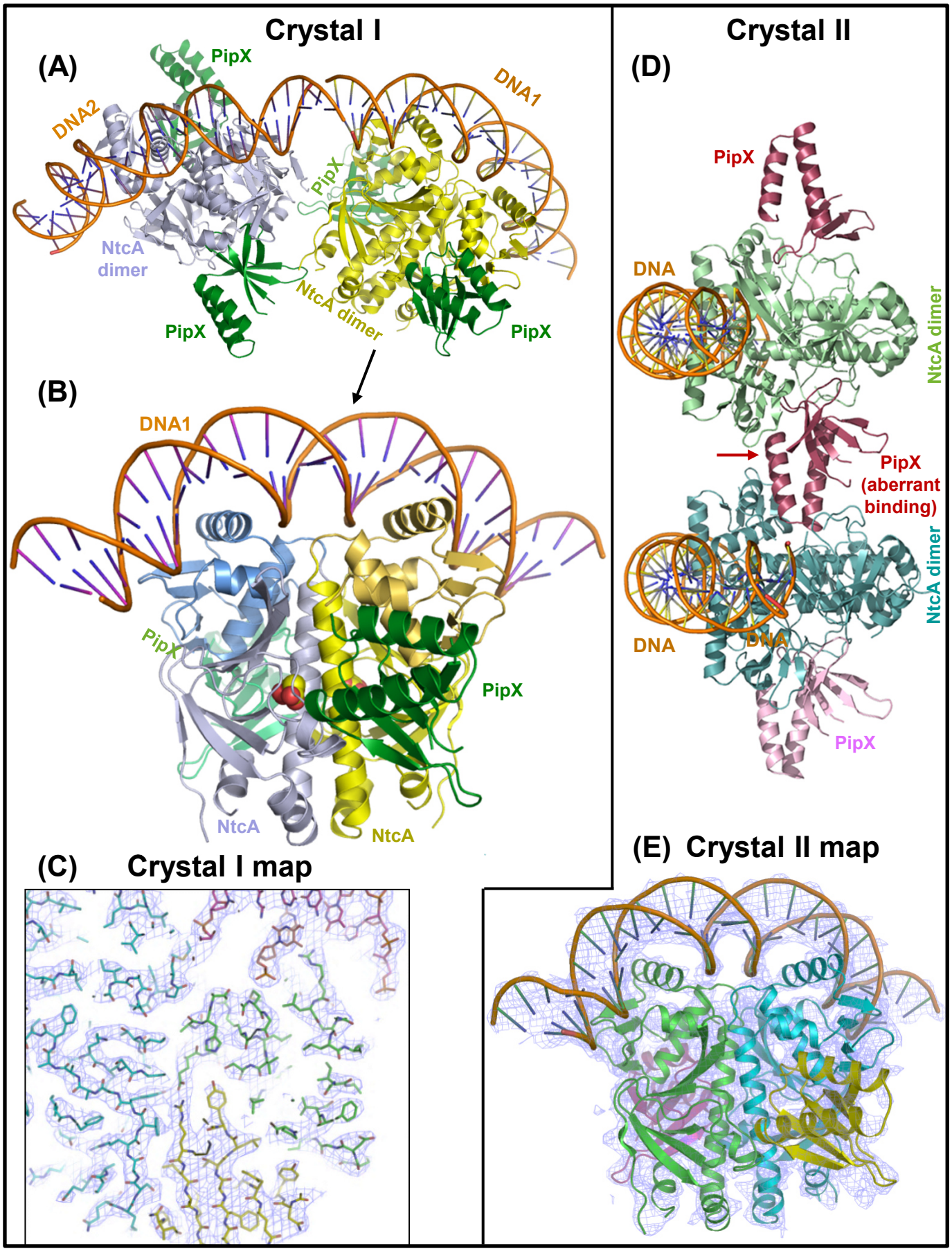
Crystal I
(A)
PipX
DNA2
DNA1
PipX
NtcA
dimer
NtcA dimer
PipX
PipX
(B)
DNA1
PipX
PipX
NtcA
NtcA
(C)
Crystal I map
Crystal II
(D)
PipX
DNA
NtcA dimer
PipX
(aberrant
binding)
NtcA dimer
DNA
DNA
PipX
(E) Crystal II map

Fig.2

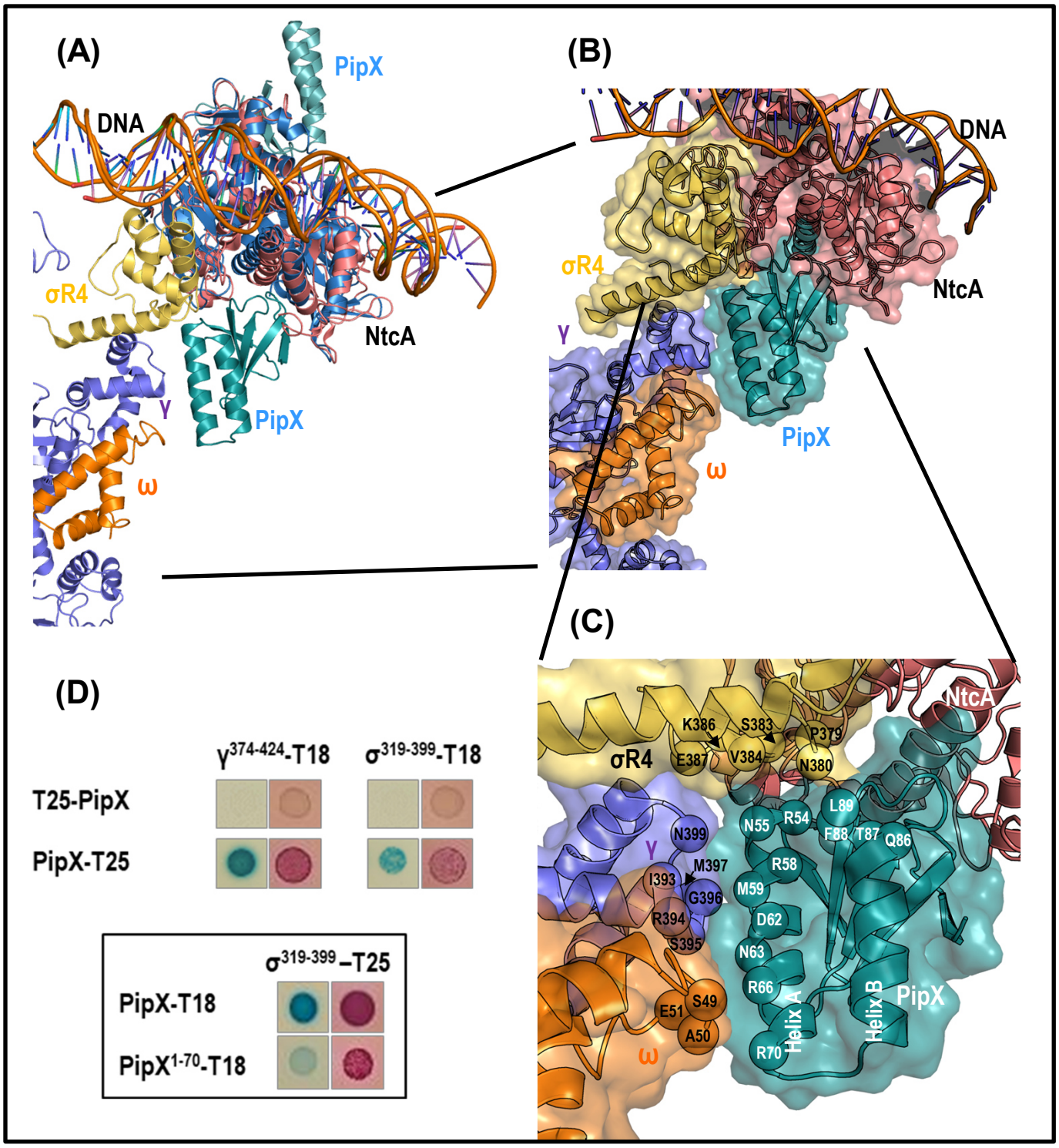
(A)
PipX
DNA
σR4
NtcA
γ
PipX
ω
(B)
DNA
σR4
NtcA
γ
PipX
ω
(C)
NtcA
K386
S383
P379
σR4
E387
V384
N380
L89
N55
R54
F88
T87
Q86
N399
γ
R58
M397
I393
M59
G396
R394
D62
S395
N63
R66
S49
E51
Helix A
Helix B
PipX
A50
R70
ω
(D)
γ374-424-T18
σ319-399-T18
T25-PipX
PipX-T25
σ319-399–T25
PipX-T18
PipX1-70-T18

Fig.3

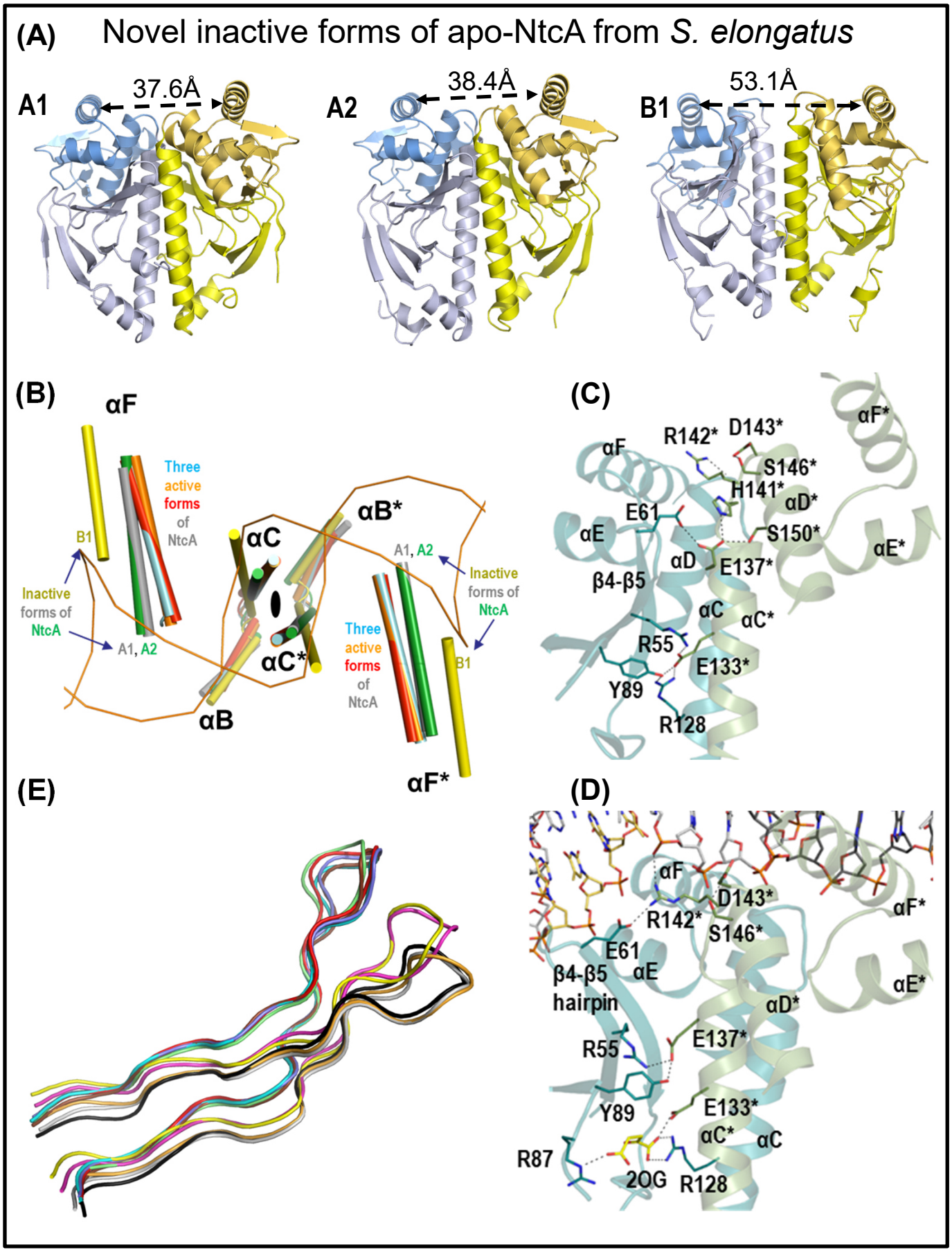
(A)
Novel inactive forms of apo-NtcA from S. elongatus
A1
37.6Å
A2
38.4Å
B1
53.1Å
(B)
αF
Three active forms of NtcA
B1
Inactive forms of NtcA
A1, A2
αC
αC*
αB
αB*
A1, A2
Inactive forms of NtcA
Three active forms of NtcA
B1
αF*
(C)
αF
R142*
D143*
αF*
S146*
H141*
αD*
E61
αE
S150*
αE*
αD
E137*
β4-β5
αC
αC*
R55
E133*
Y89
R128
(E)
(D)
αF
D143*
αF*
R142*
S146*
E61
β4-β5
αE
αE*
hairpin
αD*
R55
E137*
Y89
E133*
αC*
αC
R87
2OG
R128

Fig.4

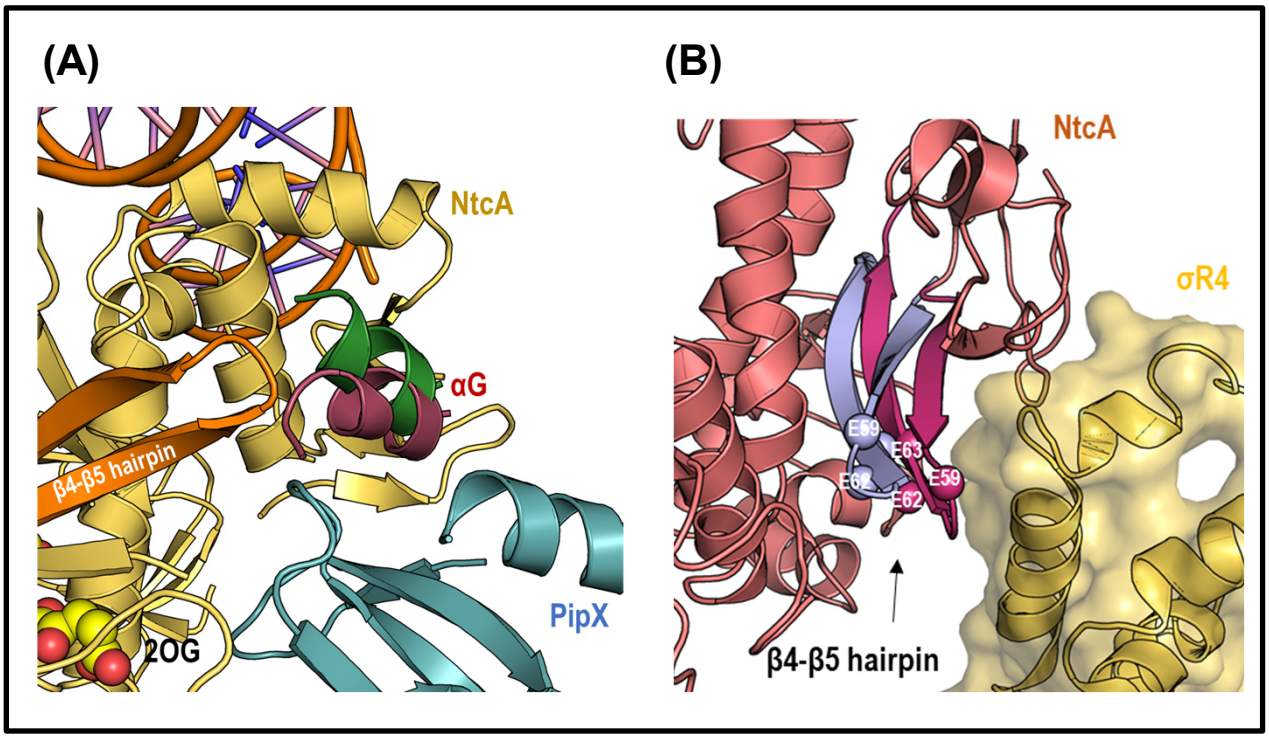
(A)
NtcA
αG
β4-β5 hairpin
PipX
2OG
(B)
NtcA
σR4
E69
E63
E59
E62
β4-β5 hairpin

Fig.5