# Regulation of the Transcriptional Activator NtrC1: Structural Studies of the Regulatory and AAA$^+$ ATPase Domains

Seok-Yong Lee[1,2], Armando De La Torre[2,3], Dalai Yan[4], Sydney Kustu[4], B. Tracy Nixon[5], David E. Wemmer[1,2,6,7]

[1] Graduate Group in Biophysics, University of California, Berkeley, CA. 94720, USA; [2] Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd. Berkeley, CA. 94720, USA; [3] Department of Molecular Cellular Biology, University of California, Berkeley, CA. 94720, USA; [4] Department of Plant and Microbial Biology, University of California, 111 Koshland Hall, Berkeley, CA. 94720, USA; [5] Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA. 16802, USA; [6] Department of Chemistry, University of California, Berkeley, CA. 94720, USA

[7] Corresponding author

E-MAIL dewemmer@lbl.gov; FAX (510) 486-6059.

**Running Title:** Structural studies of NtrC1

[*Keywords*: AAA$^+$ ATPase; crystal structure; sigma-54; transcriptional activator; response regulator; two component system]

**Abstract**

Transcription by $\sigma^{54}$ RNA polymerase depends on activators that contain ATPase domains of the $AAA^{+}$ class. These activators, which are often response regulators of two-component signal transduction systems, remodel the polymerase so it can form open complexes at promoters. Here we report the first crystal structures of the ATPase domain of an activator, the NtrC1 protein from the extreme thermophile *Aquifex aeolicus*. This domain alone, which is active, crystallized as a ring-shaped heptamer. The protein carrying both the ATPase and adjacent receiver domains, which is inactive, crystallized as a dimer. In the inactive dimer one residue needed for catalysis is far from the active site and extensive contacts among the domains prevent oligomerization of the ATPase domain. Oligomerization, which completes the active site, depends on surfaces that are buried in the dimer and hence on a rearrangement of the receiver domains upon phosphorylation. A motif in the ATPase domain known to be critical for coupling energy to remodeling of polymerase forms a novel loop that projects from the middle of an α-helix. The extended, structured loops from the subunits of the heptamer localize to a pore in the center of the ring and form a surface that could contact $\sigma^{54}$.

**Introduction**

Activators of the alternative $\sigma^{54}$-holoenzyme form of RNA polymerase catalyze an ATP-dependent conformational change in the polymerase that allows it to form transcriptionally productive open complexes at promoters (Rombel et al. 1998; Buck et al. 2000; Xu and Hoover 2001; Zhang et al. 2002). Remodeling depends on previous binding of the polymerase to a promoter in a closed complex. The activators, which are usually composed of regulatory, oligomerization/ATPase, and DNA binding domains, respond to different environmental signals

and are unusual in binding to DNA enhancer-like sequences between 100 and 150 bp upstream of promoters. In many cases their activity is controlled by phosphorylation of an aspartate residue in the regulatory domain, which induces a conformational change that allows the central ATPase domain to oligomerize. Oligomerization is essential for ATP hydrolysis and productive contact with $\sigma^{54}$-holoenzyme. Although the structures of isolated regulatory and DNA binding domains have been solved (Volkman et al. 1995; Kern et al. 1999; Pelton et al. 1999; Stock et al. 2000; Meyer et al. 2001; Park et al. 2002a), there is not yet a structure for the oligomerization/ATPase domain of an activator of $\sigma^{54}$-holoenzyme and hence many questions remain unanswered: What are the molecular bases for regulation of the central ATPase domain by the regulatory domain? How do the activators, which are dimers in their inactive state, form active oligomers, which appear to be hexamers or octamers? How is ATP hydrolysis by the activators coupled to the conformational change in $\sigma^{54}$-holoenzyme that allows DNA melting at promoters? We have begun structural studies of NtrC1 (GI: 2983588) from the hyperthermophilic bacterium *Aquifex aeolicus* to try to answer some of these questions. Although this protein was classified as an NtrC family member on the basis of its high amino acid similarity (59%) to the well-studied nitrogen regulatory protein NtrC from *Salmonella enterica serovar typhimurium*, the genes regulated by NtrC1 and the environmental signals controlling its function have not been identified (Deckert et al. 1998). We here describe in both its inactive and active states the structure of the ATPase domain of NtrC1. We compare the structure of the active ATPase to those of other ATPases associated with various cellular activities ($AAA^+$) proteins, namely: N-ethylmaleimide sensitive factor (NSF), which dissociates a complex of SNARES (soluble NSF attachment protein [SNAP] receptors) to allow further cycles of membrane fusion (Lenzen et al. 1998; Yu et al. 1998); p97, which is thought to have an

analogous role in disassembling SNARE complexes; heat shock locus protein U (HslU), which completes the unfolding of partially unfolded proteins and transfers them to the protease HslV (Bochtler et al. 2000); RuvB, which promotes DNA branch migration at Holliday junctions during genetic recombination; and the clamp loader complex, which opens the sliding clamp ring and loads it onto DNA to mediate processive movement of DNA polymerase (Neuwald et al. 1999; Vale 2000; Jeruzalmi et al. 2001a; Jeruzalmi et al. 2001b; Ogura and Wilkinson 2001).

## Results

### *Overall architecture of the protein structures*

The structure of the joined N-terminal regulatory and central ATPase domains ($NtrC1^{RC}$) of NtrC1 from *Aquifex aeolicus* and the structure of the isolated central domain ($NtrC1^{C}$) were determined to resolutions of 2.4 Å and 3.1 Å, respectively by x-ray crystallography (Table 1). Both were in their ADP-bound form. $NtrC1^{RC}$ crystallized with a dimer in the asymmetric unit, whereas $NtrC1^{C}$ crystallized with two heptamers (Fig. 1). This change allows us to infer the fundamental process by which the protein is activated, the details of which are discussed below.

### *Structure of the inactive dimer*

$NtrC1^{RC}$ is a dimer in which the regulatory domain of each monomer is adjacent to the central domain of its dimeric partner (Fig. 1A). The regulatory domain, often called a receiver domain in the context of two-component signal transduction, has the expected $(\alpha/\beta)_5$ fold and contains the conserved active site residues grouped together as they were in many previous structures of homologs (Nixon et al. 1986; Stock et al. 2000). However $NtrC1^{R}$ has a substantial extension of C-terminal helix 5 (to residue 135, about 10 residues longer than the equivalent helix in *S.*

*typhimurium* NtrC), and the extensions from the two monomers form a pair of crossed helices that are major dimerization determinants of $NtrC1^{RC}$. The dimerization interface between the two receiver domains and the extended helices is very similar to that seen for a homologous $\sigma^{54}$ activator, dicarboxylate transport regulator D (DctD) of *Sinorhizobium meliloti* (Meyer et al. 2001; Park et al. 2002a).

The central ATPase domain (residues 143-387) of NtrC1 is composed of an $\alpha/\beta$ subdomain that is typical of P-loop NTPases and an $\alpha$-helical subdomain that is a distinctive feature of $AAA^+$ ATPases (Neuwald et al. 1999). The five $\beta$-strands in the central ATPase domain of $NtrC1^{RC}$ could be superimposed on those of NSF domain 2 (D2), considered a prototype structure of a $AAA^+$ ATPase (Neuwald et al. 1999), with an r.m.s.d of 1.0 Å between the aligned segments (proteins displayed side-by-side, Fig. 2A). It is known that these five $\beta$-strands are structurally conserved among $AAA^+$ ATPases (Bochtler et al. 2000). Although the tilts and lengths of the $\alpha$ helices in the $\alpha/\beta$ subdomain are quite different in the two cases, the overall structural folds are similar.

There are two notable insertions in NtrC1 that have no counterpart in NSF D2 (Fig. 2). These two ordered structures are in helix 8 (H8) and after helix 9 (H9). The $\beta$-hairpin loop after H9 contacts the insertion in H8. There are $\beta$-hairpin loops after H9 in two other $AAA^+$ ATPases, HslU and RuvB (Bochtler et al. 2000; Han et al. 2001; Yamada et al. 2001). The insertion in H8 forms an unusual structural motif, an ordered loop that projects out of the middle of this helix (Fig. 2B). At the base it is connected to H8 by two extended strands (containing two H-bonds typical of anti-parallel $\beta$ sheets). The segments of H8 before and after the insertion have a similar helical axis but are separated by one helical turn as a result of the insertion. Comparison to structures of the corresponding regions of other $AAA^+$ ATPases shows that the two helical

segments of H8 in NtrC1 are normally a single helix (*e.g.* H8 from NSF D2 and H4 from the δ subunit of the DNA clamp loader). The insertion begins after the second turn of H8 (AEL) and is followed by residues (FELA) that complete the helix. The portion of the inserted sequence corresponding to the ordered loop has been referred to as the GAFTGA motif in NtrC1 and other $\sigma^{54}$ activators (Fig. 2B).

Several interesting features stabilize the unusual GAFTGA-containing structural motif. First, the GAFTGA loop itself is composed of two back-to-back type-II β-turns. Second, there are two glycine residues (G210 and G225) in the extended strands that flank the GAFTGA loop and allow for the distortion of the main chain configuration of H8 required to accommodate the insertion. These are highly conserved among $\sigma^{54}$ activators. Third, there are two conserved phenylalanine residues (F209 and F227) before and after the conserved glycine residues. These pack against L208 and other neighboring hydrophobic residues to support the structure of H8 (Fig. 2, B and C). Two proposed roles for the novel GAFTGA structural motif --energy coupling and oligomerization-- are discussed below.

*Structure of the heptamer rings*

The central ATPase domain of NtrC1 alone ($NtrC1^{C}$) which is competent for assembly, ATP hydrolysis, and transcriptional activation (B.T. Nixon, manuscript in preparation), formed a ring-like heptamer (Fig. 1B). There were two such heptamers in the P1 unit cell, which were stacked on each other in a head-to-head fashion with an offset of ~30Å (not shown). The diameter and height of the rings are ~124Å and ~40 Å, respectively. The fold of the ATPase domain in the rings is identical to that in the dimer, except for a subdomain reorientation discussed in detail below. Calculation of rotation angles relating the seven protomers in the heptamer indicated that

the heptamer rings are not perfectly symmetrical. Four of the protomers (chains D, E, F, G in Fig. 1B) are related to each other by rotations of 51–55°, and the other three protomers (A, B, C) are related to each other by rotations of 53-55°, but the rotation angles between these two groups is 47°. There are small gaps between the groups of three (A, B, C) and four (D, E, F, G) protomers. Whether this is due to crystal packing is unclear. The rotation angles and spacings between protomers are quite similar in the two rings.

*Interfaces involved in formation of dimers and heptamers*

In general, $AAA^+$ ATPases function as oligomeric rings (Vale 2000). Nature has used three ways to create such rings: 1) regulated assembly of oligomers from monomers or dimers; 2) formation of stable rings that remain oligomerized; or 3) formation of a ring from one polypeptide chain containing six $AAA^+$ ATPase domains (Vale 2000). Activators of $\sigma^{54}$-RNA polymerase fall in the first category. The formation of oligomers by these activators is often regulated by phosphorylation of receiver domains in response to environmental signals. Receiver domains can control oligomerization either positively (in which case phosphorylation is required for activity) (Drummond et al. 1990; Weiss et al. 1992) or negatively (in which case phosphorylation or deletion of the receiver domain is required to relieve inhibition) (Simms et al. 1985; Huala et al. 1992; Gu et al. 1994). Using an *in vivo* transcription assay and an *in vitro* ATPase assay we have shown that deletion of the receiver domain of *A. aeolicus* NtrC1 increases both transcriptional activation and ATPase activity several hundred fold (B. T. Nixon, manuscript in preparation). Thus, regulation of the activity of NtrC1 is negative and is analogous to that of *S. rhizobium* DctD rather than *S. typhimurium* NtrC (Weiss et al. 1992; Lee et al. 1994; Meyer et al. 2001).

The extensive dimeric interface between receiver domains of NtrC1$^{RC}$ is very similar to one observed previously for the receiver domain of DctD (henceforth DctD$^{R}$) (Meyer et al. 2001). In both cases the interface involves contacts between H4, β5, and H5 of the two monomers, in addition to contacts between the crossed extensions of H5. Nixon and colleagues refer to this extension as a "linker" because not all σ$^{54}$ activators with N-terminal receiver domains have such a structured extension between their receiver and output domains (e.g. *S. typhimurium* NtrC does not) (Keener and Kustu 1988; Wootton and Drummond 1989), whereas some σ$^{54}$ activators (and other proteins) with different sorts of regulatory domains apparently do [e.g. the Xylene (Xyl) R protein from *Pseudomonas putida*] (Garmendia and de Lorenzo 2000; Meyer et al. 2001; O'Neill et al. 2001). Many mutations screened for increasing transcriptional activation by intact DctD in the absence of its physiological phosphodonor resulted in weakening of the extensive dimer interface in DctD$^{R}$, indicating that this interface is important for maintaining the off state of the ATPase domain (Meyer et al. 2001). Activated DctD$^{R}$ formed a very different dimer(s) that did not contain the crossed helices found in the off-state dimer (Park et al. 2002a; Park et al. 2002b).

As discussed above, contacts mediated by the crossed linkers of NtrC1$^{RC}$ appear to be major dimerization determinants. In support of this view, amino acid substitutions predicted to cause destabilization of contacts between them --based on findings with DctD$^{R}$-- increased transcriptional activation by NtrC1$^{RC}$ *in vivo* (B. T. Nixon, manuscript in preparation). However, there are also regions of contact between central domains and the receiver domains (including linkers) of opposite monomers in the dimer and between the two central ATPase domains of NtrC1$^{RC}$ (Fig. 1A). Strand β2 and helix H8 contact the linker in the partner monomer, and both the insertion in H8 and the β-hairpin after H9 contact the C-terminal parts of H2 and H3 in this

opposite receiver domain (Fig. 1A). Contacts between the ATPase domains are primarily in the α-helical subdomain. The dimerization interface spans about 7100 $Å^2$ (3550 $Å^2$ per monomer).

The amount of buried surface area in the heptamer of $NtrC1^C$ is very similar to that in the dimer of $NtrC1^{RC}$. One surface of the ATPase domain that participates in forming the heptamer overlaps with the surface that participates in forming the dimer (Figs. 3, 6). In the heptamer, protomers pack front to back so that there are then two surfaces in each ATPase domain participating in intermolecular interactions. This increases the intermolecular contact surface between ATPase domains, which would drive oligomerization once phosphorylation had altered the conformation of the receiver and linker domains. The novel structural motif inserted into H8 and H8 itself also participate in oligomerization. In the $NtrC1^{RC}$ dimer, part of the β hairpin after H9 and the insertion in H8 contact the receiver domain and thus are buried, whereas in the $NtrC1^C$ heptamer, these regions are involved in contacts with other protomers (Figs. 1, 3).

*Subdomain motions between dimer and heptamer and their effect on the nucleotide binding pocket*

Comparison of the central ATPase domain in the inactive dimer ($NtrC1^{RC}$) to that in the assembled oligomer ($NtrC1^C$) shows a hinge motion between the α/β subdomain and the α-helical subdomain, even though both forms have the same nucleotide (ADP) bound (Fig. 4A). In HslU an analogous hinge motion between these subdomains was observed when the nucleotide bound to the hexamer was changed (Bochtler et al. 2000; Sousa et al. 2000; Wang et al. 2001a; Wang et al. 2001b). The rotations between these subdomains in NtrC1 are about 10° counterclockwise in going from dimer to heptamer (Fig. 4A), with the rotation axis at the midpoint between H7 and residue 308 and running parallel to H7. The α-helical subdomain,

along with H6, opens relative to the $\alpha/\beta$ subdomain, leading to a flatter molecule that may have a preference for heptameric packing.

A close view of the nucleotide-binding pocket shows that there are some rearrangements of ADP and the residues that contact it in going from dimer to heptamer (Fig. 4B). Because the domain rotation axis is located near the N-terminus of H7, these do not involve the Walker A (GXXXGKE; also called the P-loop) or Walker B (DEXX) motifs, or the two phosphates of ADP. However, residues R357 and K360, which are located in the long helical region referred to by Neuwald et al as Sensor II (Neuwald et al. 1999), and the ribose and adenine base of ADP do show significant changes. The ADP in both $NtrC1^{RC}$ and $NtrC1^{C}$ has the *anti* configuration, as was the case for p97 (Zhang et al. 2000) and several other $AAA^{+}$ proteins (Ogura and Wilkinson 2001).

*Correlation of structure with genetic and biochemical studies of $\sigma^{54}$ activators*

Structures of NtrC1 allow us to map amino acid substitutions identified in genetic studies of other $\sigma^{54}$ activators. Mutational changes that affect ATP binding and/or hydrolysis are shown in Figure 5A. Most of those that abolish ATP binding are located in the Walker A (P-loop) region or the sensor II region (e.g. G354 and R357: numbering in NtrC1 from *A. aeolicus*) (Perez-Martin and de Lorenzo 1996; Gao et al. 1998; Rombel et al. 1999; Chaney et al. 2001; Lew and Gralla 2002; Zhang et al. 2002). In NtrC1 the side-chain of R357, which has been called the sensor II residue, contacts the $\gamma$-phosphate of ADP, implying that this residue plays an important role in binding the phosphate region of the nucleotide. G354 is located in the loop preceding the sensor II helix. This glycine, which is highly conserved among $\sigma^{54}$ activators, may orient the sensor II helix so that R357 can contact the $\gamma$-phosphate of the nucleotide (Figs. 4B, 5A). Amino

acid substitutions that prevent ATP hydrolysis without disrupting binding are located in the Walker B motif, at or around R293, which was predicted to be a catalytic arginine, or in residues of H8 just preceding the GAFTGA insertion (e.g. A206, which is an S in many other activators, and E207) (Rombel et al. 1998; Rombel et al. 1999; Chaney et al. 2001; Zhang et al. 2002). D238 in Walker B, which is thought to be required for divalent cation binding (Rombel et al. 1999), interacts with $Mg^{2+}$, although $Mg^{2+}$ was seen in only one monomer of the $NtrC1^{RC}$ dimer and was not seen in the lower resolution structure of the $NtrC1^{C}$ heptamer. This $Mg^{2+}$ also interacts with ADP (Fig. 4B). R293, which is highly conserved among $\sigma^{54}$ activators, is located close to ADP bound by the neighboring protomer in the heptamer and could probably contact the $\gamma$-phosphate of ATP if it were present (Fig. 5A). R293 is far from the nucleotide in the $NtrC1^{RC}$ dimer (Fig. 1A). It has been proposed that the catalytic arginine for the $AAA^{+}$ superfamily is one close to R293 (R299) (Neuwald et al. 1999).

The GAFTGA portion of the insertion into H8 is highly conserved among $\sigma^{54}$ activators and has been shown to be involved in coupling the energy available from ATP hydrolysis to a conformational change of $\sigma^{54}$-holoenzyme (Wang et al. 1997; Gonzalez et al. 1998; Rombel et al. 1998; Chaney et al. 2001; Lew and Gralla 2002; Zhang et al. 2002). Mutations that change the GAFTGA sequence often abolish transcriptional activation but do not affect ATP binding or hydrolysis. Buck and colleagues recently showed that the GAFTGA loop contacts $\sigma^{54}$ when the ATPase domain of phage shock protein F (PspF) is bound to ADP-$AlF_x$ (Chaney et al. 2001; Bordes et al. 2003), a nucleotide analogue which mimics ATP in the transition state for hydrolysis. Notably, the GAFTGA loop is localized in the central pore region of the heptamer structure of $NtrC1^{C}$ (Fig. 1B). In this oligomer structure, the inserted loops can form an extended binding surface for $\sigma^{54}$ and presumably play a role in mediating the energy coupling that allows

NtrC1 to change the conformation of $\sigma^{54}$-RNA polymerase at promoters. One of the mutational amino acid substitutions in the GAFTGA region of *S. typhimurium* NtrC (which corresponds to G218K in NtrC1) increases non-specific DNA-binding by about 10-fold (North et al. 1996). This lesion provides evidence that the GAFTGA loop is near the DNA, perhaps near the distorted DNA in the -12 promoter region to which $\sigma^{54}$ is initially bound (Guo et al. 1999; Guo et al. 2000; Chaney et al. 2001). As mentioned above, some lesions that decrease ATP hydrolysis affect residues in H8 just prior to the GAFTGA insertion. These residues may be involved in communication between active sites and sites of strong contact with the sigma factor. They have also been implicated in a contact with $\sigma^{54}$ that does not depend on the presence of nucleotide and occurs even in (inactive) dimers of DctD (Lee and Hoover 1995; Wang et al. 1997; Kelly and Hoover 2000), and they appear to play a role in oligomerization of *S. typhimurium* NtrC and the ATPase domain of PspF (Li et al. 1999; Chaney et al. 2001; Lew and Gralla 2002; Zhang et al. 2002).

In *S. typhimurium* NtrC mutational amino acid substitutions that apparently favor oligomerization at low concentrations of the phosphorylated protein occur throughout the central ATPase domain (Li et al. 1999; Yan and Kustu 1999). In particular, such substitutions have been found in the GAFTGA loop (residues corresponding to A215C and G218C of *A. aeolicus* NtrC1) and in and around R293.

**Discussion**

It has been postulated that transcription by $\sigma^{54}$ holoenzyme confers one notable advantage over transcription by $\sigma^{70}$-holoenzyme, the capacity to vary transcriptional efficiency at a given promoter over a wide range (Rombel et al. 1998; Buck et al. 2000; Zhang et al. 2002). This wide

range of control--not usually accessible to $\sigma^{70}$-holoenzyme at a single promoter--is achieved by the use of $AAA^{+}$ activators that couple energy to the process of open complex formation. Rather than melting the DNA themselves, activators remodel $\sigma^{54}$-RNA polymerase bound to promoters in transcriptionally silent closed complexes so that it can form open complexes. The structures of *A. aeolicus* $NtrC1^{C}$ (active) and $NtrC1^{RC}$ (inactive) allow us to return to the questions posed about activators in the introduction. We begin with the related questions of how receiver domains of $\sigma^{54}$ activators, which are a large subclass of their regulatory domains, control the activity of the adjacent ATPase domains and how the dimeric forms of the activators form active oligomers.

*Oligomerization of NtrC1*

Remarkably, receiver domains can control the output of their adjacent ATPase domains either positively or negatively (Simms et al. 1985; Drummond et al. 1990; Huala et al. 1992; Weiss et al. 1992; Gu et al. 1994). This is the case even though activators such as *S. meliloti* DctD (negatively regulated) and *S. typhimurium* NtrC (positively regulated) have 38% identity and 53% sequence similarity over their entire lengths. In the case of *A. aeolicus* NtrC1 control is negative and appears to be very similar to that in DctD. Contacts between the receiver domain and linker of $NtrC1^{RC}$ and regions of the central ATPase domain of the opposite monomer in the dimer effectively bury one of the two surfaces required for the central domain to form heptamers (Figs. 1A, 3, 6). In addition, the central ATPase domains themselves are held in a front-to-front orientation in the dimer, whereas they must reorient in a front-to-back orientation to form the heptamer (Fig. 6). Hence, in their crossed state the long helices at the ends of the receiver domains (linkers) prevent oligomerization. Deletion of the receiver domain and linker yields an

active oligomeric ATPase in which the active site for ATP hydrolysis has been completed and a surface has been created for interaction with $\sigma^{54}$.

When phosphorylation alters the conformation of the receiver domains of NtrC1 the dimerization interfaces between the receiver and ATPase domains are presumably disrupted. Where phosphorylated receiver domains are located in oligomers of full-length NtrC1 is not known, but the rearrangement seen in active $DctD^{R}$ indicates that the linker helices are no longer crossed in activated dimers and that the ATPase domains of this protein, were they attached, would be free to repack in the front to back orientation required for oligomer formation (Fig. 6) (Park et al. 2002a; Park et al. 2002b). The presence of two contact surfaces between ATPase domains in the heptamer, one of which involves many residues that form the single contact surface in the dimer, would be sufficient to drive oligomerization (Figs. 3, 6). Use of the same contact surface in the "on" and "off" states provides a mechanism for minimizing both unwanted interactions and the free energy cost of switching between states.

Although we cannot prove that the active form of NtrC1 is a heptamer rather than *e.g.* a hexamer or an octamer, we know of no reason why crystallization would have favored a heptamer if this were not the case. A homoheptamer (ClpB), a heteropentamer (DNA clamp loader), and a mixture of homoheptamer and homohexamer (HslU) have been observed among $AAA^{+}$ ATPases (Rohrwild et al. 1997; Kim et al. 2000; Miyata et al. 2000; Jeruzalmi et al. 2001a). Because we can infer only the outlines of energy coupling **(**see below**),** our current interpretations do not depend on knowing the precise number of protomers in the active oligomer. Neither the inactive construct from NtrC1 ($NtrC1^{RC}$) nor the active construct ($NtrC1^{C}$) carries the C-terminal DNA-binding determinants of the protein. Where these are located and how they affect oligomerization and function of the intact protein remain to be determined.

*Novel structural motif and energy coupling*

The GAFTGA motif in NtrC1 and other activators of $\sigma^{54}$-RNA polymerase plays a direct role in contact with the polymerase and coupling of energy to a change in its conformation (Chaney et al. 2001; Zhang et al. 2002; Bordes et al. 2003). This sequence motif is part of an insertion with respect to other AAA+ ATPases (Fig. 2). The insertion forms an unusual structural motif consisting of two extended strands and an ordered loop (GAFTGA itself) projecting from the middle of H8 in the $\alpha/\beta$ subdomain of the ATPase. Although a somewhat similar insertion was observed in a helix of actin (PDB entry 1ijj), in that case the inserted loop is shorter and is not connected to the helix with strands (Bubb et al. 2002). In addition, the helical segments before and after the loop have different helical axes. To our knowledge, the unusual structural motif containing the GAFTGA loop has not been seen before.

The ordered GAFTGA loop is found in the central pore of the $NtrC1^{C}$ heptamer. Loops from different subunits form a surface that could contact $\sigma^{54}$ and mediate the energy coupling between NtrC1 and $\sigma^{54}$-holoenzyme that allows the polymerase to initiate transcription. The only residue shown to be essential for contact between the polymerase and a short GAFTGA-containing peptide was T217 (Bordes et al. 2003), which is in the middle of the loop (Fig. 2).

There are a number of other $AAA^{+}$ ATPases in which functionally important sequences are located in the region of H8 in NtrC1 and are found in the pore at the center of the oligomer. One of these is HslU, which contains a long insertion (I-domain) after H8 (Bochtler et al. 2000) that binds to partially unfolded proteins. HslU completes the unfolding process by means of a hinge motion between the $\alpha/\beta$ and helical subdomains that occurs as a result of ATP hydrolysis (Wang et al. 2001a; Wang et al. 2001b). A second example is the DNA clamp loader $\gamma$-complex,

for which the structure of the $AAA^+$ domain has been determined both with and without its target, the $\beta$ protein (clamp) (Jeruzalmi et al. 2001a; Jeruzalmi et al. 2001b). In the complex between the two proteins, H4 of the catalytic $\delta$-subunit of the clamp loader (which corresponds to H8 of NtrC1) undergoes a drastic conformational change, as does the $\beta$ subunit (Jeruzalmi et al. 2001a; Jeruzalmi et al. 2001b). Finally, a comparison of the structures of NSF D2 and p97 leads to the proposal that H8 in p97 relocates upon ATP hydrolysis, again through a domain motion involving the $\alpha/\beta$ and helical subdomains (Zhang et al. 2000). Recently, it was shown using cryo-electron microscopy that there are dramatic changes in the central pore of the p97 oligomer through the ATPase cycle (Rouiller et al. 2002). Taken together with the structure of $NtrC1^C$, these other structures support the view that the inserted GAFTGA loop in H8 of NtrC1, which is implicated in binding $\sigma^{54}$, will undergo changes in position and/or orientation as ATP is hydrolyzed. These changes are presumably driven by subdomain reorientations in the ATPase domain of NtrC1 and are coupled to a conformational change in $\sigma^{54}$ RNA polymerase, but the detailed mechanism of energy coupling remains to be determined.

*Conserved regions of $\sigma^{54}$ activators*

Morett and colleagues noted that activators of $\sigma^{54}$-holoenzyme have 7 regions of conserved sequence in their ATPase domains, which they designated C1 through C7 (Morett and Segovia 1993). Together with the analysis of Neuwald et al (Neuwald et al. 1999), the structure of the $NtrC1^C$ oligomer leads to a better understanding of the roles of these conserved regions (Fig. 5B). The C1 region is the Walker A motif. Region C2, which was previously of unknown

function, clearly participates in contacts between the protomers of the NtrC1$^{C}$ heptamer, as does region C5. They contact one another and a number of mutational amino acid substitutions in and around these regions appear to increase oligomerization of phosphorylated *S. typhimurium* NtrC (Li et al. 1999). Region C3 carries the GAFTGA motif and region C4 is the Walker B motif. Region C6 carries the putative sensor I residue (N280), the sensor I region, the catalytic arginine, which has also been called an arginine finger, and the hinge between the two subdomains. Although Neuwald et al (Neuwald et al. 1999) predicted that the catalytic arginine was conserved R299, there is genetic evidence that it may be conserved R293: A lesion in *S. typhimurium* NtrC that changed the corresponding residue to cysteine yielded a protein that bound nucleotide and oligomerized normally but failed to hydrolyze ATP (Rombel et al. 1999). In the oligomer either R299 or R293 presumably completes the active site of an adjacent protomer (Figs. 1, 5A). Region C7 carries the sensor II residue (R357) and the sensor II helix and is the only conserved region located in the helical subdomain of the ATPase.

*Conclusions*

Unlike the case for many other AAA and AAA$^{+}$ ATPases (Vale 2000), $\sigma^{54}$ activators do not bind tightly to the polymerase, their target, in the presence of ATP (or ADP). Tight binding, which has been achieved for PspF and the ATPase domain of *Klebsiella pneumoniae* nitrogen fixation protein A (NifA) (Chaney et al. 2001) depends on the presence of ADP-AlFx, an analogue of ATP in the transition state for hydrolysis. In addition, it requires that activator and polymerase be mixed before this analogue is added, implying that the proteins must first associate weakly. Under normal circumstances contacts between activators and polymerase are apparently short-lived, allowing for very high levels of transcription from $\sigma^{54}$-dependent promoters. Together

with additional structures in the presence of various nucleotides, the detailed biochemical and genetic analyses that have been made of $\sigma^{54}$ activators should facilitate progress in defining the mechanochemical action of these unusual transcriptional regulators and of $AAA^{+}$ proteins generally.

*Coordinates*

Coordinates and reflection data for $NtrC1^{RC}$ and $NtrC1^{C}$ have been deposited at the Protein Data Bank under accession codes 1NY5 and 1NY6, respectively.

## Materials and Methods

*Protein expression and purification*

The *Aquifex aeolicus* genomic DNA was a gift from R. Huber. Both $NtrC1^{RC}$ (*Aquifex aeolicus*, GI # 2983588, residues 1-387) and $NtrC1^{C}$ (residues 121-387) were cloned from the genomic DNA into pET21a vectors (NOVAGEN). A selenomethionine derivative of the $NtrC1^{RC}$ protein was obtained by growing *E. coli* B834 (DE3) carrying plasmid SJS1244 in M9 minimal medium supplemented with selenomethionine (Leahy et al. 1994). $NtrC1^{C}$ protein was expressed using *E. coli* BL21 (DE3) with Rosetta.pLysS (to provide rare tRNAs). $NtrC1^{RC}$ and $NtrC1^{C}$ were purified in a similar manner. Cells were harvested and disrupted by sonication. Cell lysate in 50 mM Tris buffer (pH 8.2), 1 mM $MgCl_2$, 1 mM ATP, 1 mM tris-2-carboxy-ethylphosphine (TCEP) was heated at 75-80°C for 25 min and the supernatant was dialysed against 50 mM Tris buffer (pH 8.0) and 5 mM EDTA overnight. The protein was further purified on a Q-sepharose column, and then a gel filtration column. The purified protein was extensively dialysed against 5 mM ammonium bicarbonate (2 liter chamber, 3 buffer changes for about 24 h) and then

lyophilized. Lyophilized $NtrC1^{RC}$ was redissolved in 50 mM Tris (pH 8.0), 20 mM ATP, 10 mM TCEP, 5 mM $MgCl_2$, 5% glycerol (vol/vol). $NtrC1^{C}$ was dissolved in the same buffer except without glycerol. The pH of TCEP and ATP solutions was adjusted to 7.0 before mixing with the protein solution and the final pH of the protein solution was 8.0. Protein solutions were filtered with a 0.2 μm filter before use in crystallization trials. Gel filtration chromatography (Superdex200, 10/30) and dynamic light scattering confirmed the dimeric state of $NtrC^{RC}$ and the higher oligomeric state (M.W. ~220 kDa) of $NtrC1^{C}$. The oligomeric state of $NtrC1^{C}$ is dependent on salt concentration with oligomer stabilized at low salt concentration (data not shown).

*Crystallization and structure determination*

Crystals of $NtrC1^{RC}$ were grown at room temperature using the hanging drop vapor diffusion method. Six μl of protein solution (10 mg/ml) was mixed with an equal volume of well solution containing 0.6-0.7 M $NaH_2PO_4/KH_2PO_4$, 100 mM citric acid (pH 6.5), 10 mM imidazole, and 3% (v/v) methanol. Crystals appeared after 1 day and grew to approximately 0.2 x 0.2 x 1.0 mm (hexagonal rod shape) after one week. Crystals were transferred to well solution plus 10% glycerol for 1 min and then transferred to well solution plus 20% glycerol for 5 min, after which they were flash frozen and stored in liquid nitrogen until being used for data collection. These crystals had the space group $P3_221$, with unit cell dimensions a=b=94.8Å and c=195.0Å with two molecules in the asymmetric unit. SAD (Single wavelength Anomalous Diffraction) data were collected at the beam line 5.0.2 of the Advanced Light Source (ALS) at Lawrence Berkeley National Lab. The data were processed and scaled with DENZO and SCALEPACK from the HKL program suite (Otwinowski and Minor 1997). Heavy atom sites (Se) were found with

SOLVE at 3 Å (Terwilliger 2000). Eleven out of 12 SeMet sites were found and 10 sites were related by two-fold rotations that are not related to crystallographic symmetry. Phasing was calculated with SHARP (de La Fortelle and Bricogne 1997) up to 3Å resolution and was improved and extended by RESOLVE (Terwilliger 2000) using two-fold non-crystallographic symmetry (NCS) averaging and solvent flattening. Model building was done using O (Jones et al. 1991). Using partial models, two separate NCS operators for both regulatory and central ATPase domains were calculated and two domain NCS-averaging was performed using DM (Wang 1985). This improved the map quality significantly and allowed us to complete the model building. The GAFTGA loop that is inserted into H8 was not initially visible but the connecting strands were clear in the very early stage of the experimental map before NCS averaging. The loop was built after completion of the rest of the chain. The structure was refined with CNS (Brunger et al. 1998). Anisotropic B factors and bulk solvent corrections as well as the cross validation methods were applied throughout the refinement. NCS restraints were applied only in the initial refinement stage and were released after the R-working and R-free factors dropped to 33%/35%. Water and solvent picking was performed after the $R_{working}$ and $R_{free}$ dropped to 25%/28.5% using CNS. PROCHECK was used to monitor geometric parameters (Laskowski et al. 1993).

Crystals of $NtrC1^{C}$ were grown at room temperature using the hanging drop vapor diffusion method. Two μl of protein solution (20 mg/ml) was mixed with an equal volume of well solution containing (pH 6.5), 50 mM diammonium tartrate and 8% (vol/vol) PEG 3350. Crystals grew to approximately 0.2 x 0.2 x 0.2 mm (diamond shape) after one day. Crystals were transferred to well solution plus 10% glycerol for 2 min and then transferred to well solution plus 25% glycerol for 5 min, after which they were flash frozen and stored in liquid nitrogen until

being used for data collection. These crystals had the space group P1, with unit cell dimensions a= 106.8Å, b=108.3Å, c=110.0Å, a=70.3°, b=85.9° and g=73.3°, with fourteen molecules in the unit cell. Data were collected at the beam line 8.3.1 of the Advanced Light Source (ALS). The data were processed and scaled with MOSFLM and SCALA using the Elves interface (Holton, manuscript in preparation). Seven two-fold NCS peaks were found (at about 5σ) using the self-rotation function of Molrep (Vagin and Isupov 2001). Locked-cross-rotation was performed using the self–rotation solutions by Molrep at several different resolution ranges. Three sets of solutions (ten solutions each) were found at several different resolution ranges and sphere radius (25Å-40Å). Solutions that were not located to form rings were discarded and the correct solutions were confirmed by the decrease of $R_{free}$ in the initial refinement with CNS. Eleven solutions were found in this way and the remaining three molecules were placed manually in the $2F_o$-$F_c$ electron density map. After all fourteen solutions had been found, two-domain rigid body refinement (α/β subdomain and α-helical subdomain) improved $R_{working}$/ $R_{free}$ factor significantly. Refinement was done with CNS.

**Acknowledgements**

We thank Donghae Shin for excellent advice in crystallography, Ho S. Cho for help with sample preparation, David King for help in characterizing NtrC1 by mass spectrometry, the beam line staff at the ALS for help with data collection, and Seth Rubin for critical reading of the manuscript. This work was supported by National Institute of Health grants (GM62163) to D.W. and (GM38361) to S. K., and National Science Foundation grant MCB9727745 to B. T. N.

**References**

Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D., and Huber, R. 2000. The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature* **403**: 800-805.

Bordes, P., Wigneshweraraj, S. R., Schumacher, J., Zhang, X., Chaney, M., and Buck, M. 2003. The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: Identifying a surface that binds $\sigma^{54}$. *Proc. Natl. Acad. Sci. U S A* **100**: 2278-2283.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. M., Rice, L. M., Simonson, T., and Warren, G. L. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 905-921.

Bubb, M. R., Govindasamy, L., Yarmola, E. G., Vorobiev, S. M., Almo, S. C., Somasundaram, T., Chapman, M. S., Agbandje-McKenna, M., and McKenna, R. 2002. Polylysine induces an antiparallel actin dimer that nucleates filament assembly: crystal structure at 3.5-A resolution. *J. Biol. Chem.* **277**: 20999-21006.

Buck, M., Gallegos, M.-T., Studholme, D. J., Guo, Y., and Gralla, J. D. 2000. The bacterial enhancer-dependent $\sigma^{54}$ ($\sigma^{N}$) transcription factor. *J. Bacteriol.* **182**: 4129-4136.

Chaney, M., Grande, R., Wigneshweraraj, S. R., Cannon, W., Casaz, P., Gallegos, M. T., Schumacher, J., Jones, S., Elderkin, S., Dago, A. E., Morrett, E., and Buck, M. 2001. Binding of transcriptional activators to $\sigma^{54}$ in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. *Genes Dev.* **15**: 2282-2294.

de La Fortelle, E. and Bricogne, G. 1997. *Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods*. Academic Press Inc, San Diego.

Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**: 353-358.

Drummond, M. H., Contreras, A., and Mitchenall, L. A. 1990. The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. *Mol. Microbiol.* **4**: 29-37.

Gao, Y., Wang, Y. K., and Hoover, T. R. 1998. Mutational analysis of the phosphate-binding loop of *Rhizobium meliloti* DctD, a $\sigma^{54}$-dependent activator. *J. Bacteriol.* **180**: 2792-2795.

Garmendia, J. and de Lorenzo, V. 2000. The role of the interdomain B linker in the activation of the XylR protein of *Pseudomonas putida*. *Mol. Microbiol.* **38**: 401-410.

Gonzalez, V., Olvera, L., Sobernon, X., and Morett, E. 1998. In vivo studies on the positive control function of NifA: a conserved hydrophobic amino acid patch at the central domain involved in transcriptional activation. *Mol. Microbiol.* **28**: 55-67.

Gu, B., Lee, J. H., Hoover, T. R., Scholl, D., and Nixon, B. T. 1994. *Rhizobium meliloti* DctD, a $\sigma^{54}$-dependent transcriptional activator, may be negatively controlled by a subdomain in the C-terminal end of its two-component receiver module. *Mol. Microbiol.* **13**: 51-66.

Guo, Y., Lew, C. M., and Gralla, J. D. 2000. Promoter opening by $\sigma^{54}$ and $\sigma^{70}$ RNA polymerases: sigma factor-directed alterations in the mechanism and tightness of control. *Genes Dev.* **14**: 2242-2255.

Guo, Y., Wang, L., and Gralla, J. D. 1999. A fork junction DNA-protein switch that controls promoter melting by the bacterial enhancer-dependent sigma factor. *EMBO J.* **18**: 3736-3745.

Han, Y. W., Iwasaki, H., Miyata, T., Mayanagi, K., Yamada, K., Morikawa, K., and Shinagawa, H. 2001. A unique β-hairpin protruding from AAA+ ATPase domain of RuvB motor protein is involved in the interaction with RuvA DNA recognition protein for branch migration of Holliday junctions. *J. Biol. Chem.* **276**: 35024-35028.

Huala, E., Stigter, J., and Ausubel, F. M. 1992. The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. *J. Bacteriol.* **174**: 1428-1431.

Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. 2001a. Crystal structure of the processivity clamp loader gamma (γ) complex of *E. coli* DNA polymerase III. *Cell* **106**: 429-441.

Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M., and Kuriyan, J. 2001b. Mechanism of processivity clamp opening by the τ subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* **106**: 417-428.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47 (Pt 2)**: 110-119.

Keener, J. and Kustu, S. 1988. Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: Roles of the conserved amino-terminal domain of NTRC. *Proc. Natl. Acad. Sci. U S A* **85**: 4976-4980.

Kelly, M. T. and Hoover, T. R. 2000. The amino terminus of *Salmonella enterica serovar typhimurium* $\sigma^{54}$ is required for interactions with an enhancer-binding protein and binding to fork junction DNA. *J. Bacteriol.* **182**: 513-517.

Kern, D., Volkman, B. F., Luginbuhl, P., Nohaile, M. J., Kustu, S., and Wemmer, D. E. 1999. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* **402**: 894-898.

Kim, K. I., Cheong, G. W., Park, S. C., Ha, J. S., Woo, K. M., Choi, S. J., and Chung, C. H. 2000. Heptameric ring structure of the heat-shock protein ClpB, a protein-activated ATPase in *Escherichia coli*. *J. Mol. Biol.* **303**: 655-666.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283-291.

Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P., and Hendrickson, W. A. 1994. Crystallization of a fragment of human fibronectin - introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. *Protein-Struct. Funct. Genet.* **19**: 48-54.

Lee, J. H. and Hoover, T. R. 1995. Protein crosslinking studies suggest that *Rhizobium meliloti* C4-dicarboxylic acid transport protein D, a $\sigma^{54}$-dependent transcriptional activator,

interacts with $\sigma^{54}$ and the beta subunit of RNA polymerase. *Proc. Natl. Acad. Sci. U S A* **92**: 9702-9706.

Lee, J. H., Scholl, D., Nixon, B. T., and Hoover, T. R. 1994. Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of Rhizobium meliloti DCTD, a $\sigma^{54}$-dependent transcriptional activator. *J. Biol. Chem.* **269**: 20401-20409.

Lenzen, C. U., Steinmann, D., Whiteheart, S. W., and Weis, W. I. 1998. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**: 525-536.

Lew, C. M. and Gralla, J. D. 2002. New roles for conserved regions within a $\sigma^{54}$-dependent enhancer-binding protein. *J. Biol. Chem.* **277**: 41517-41524.

Li, J., Passaglia, L., Rombel, I., Yan, D., and Kustu, S. 1999. Mutations affecting motifs of unknown function in the central domain of nitrogen regulatory protein C. *J. Bacteriol.* **181**: 5443-5454.

Meyer, M. G., Park, S., Zeringue, L., Staley, M., McKinstry, M., Kaufman, R. I., Zhang, H., Yan, D., Yennawar, N., Yennawar, H., Farber, G. K., and Nixon, B. T. 2001. A dimeric two-component receiver domain inhibits the $\sigma^{54}$-dependent ATPase in DctD. *Faseb J.* **15**: 1326-1328.

Miyata, T., Yamada, K., Iwasaki, H., Shinagawa, H., Morikawa, K., and Mayanagi, K. 2000. Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J. Struct. Biol.* **131**: 83-89.

Morett, E. and Segovia, L. 1993. The $\sigma^{54}$ bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* **175**: 6067-6074.

Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**: 27-43.

Nixon, B. T., Ronson, C. W., and Ausubel, F. M. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. *Proc. Natl. Acad. Sci. U S A* **83**: 7850-7854.

North, A. K., Weiss, D. S., Suzuki, H., Flashner, Y., and Kustu, S. 1996. Repressor forms of the enhancer-binding protein NtrC: Some fail in coupling ATP hydrolysis to open complex formation by $\sigma^{54}$-holoenzyme. *J. Mol. Biol.* **260**: 317-331.

O'Neill, E., Wikstrom, P., and Shingler, V. 2001. An active role for a structured B-linker in effector control of the $\sigma^{54}$-dependent regulator DmpR. *EMBO J.* **20**: 819-827.

Ogura, T. and Wilkinson, A. J. 2001. AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells* **6**: 575-597.

Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**: 307-326.

Park, S., Meyer, M., Jones, A. D., Yennawar, H. P., Yennawar, N. H., and Nixon, B. T. 2002a. Two-component signaling in the AAA+ ATPase DctD: binding $MG_2^+$ and $BeF_3^-$ selects between alternative dimeric states of the receiver domain. *Faseb J.* **16**: 1964-1966.

Park, S., Zhang, H., Jones, A. D., and Nixon, B. T. 2002b. Biochemical evidence for multiple dimeric states of the *Sinorhizobium meliloti* DctD receiver domain. *Biochemistry* **41**: 10934-10941.

Pelton, J. G., Kustu, S., and Wemmer, D. E. 1999. Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. *J. Mol. Biol.* **292**: 1095-1110.

Perez-Martin, J. and de Lorenzo, V. 1996. ATP binding to the $\sigma^{54}$-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. *Cell* **86**: 331-339.

Rohrwild, M., Pfeifer, G., Santarius, U., Muller, S. A., Huang, H. C., Engel, A., Baumeister, W., and Goldberg, A. L. 1997. The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nat. Struct. Biol.* **4**: 133-139.

Rombel, I., North, A., Hwang, I., Wyman, C., and Kustu, S. 1998. The bacterial enhancer-binding protein NtrC as a molecular machine. *Cold Spring Harb. Symp. Quant. Biol.* **63**: 157-166.

Rombel, I., Peters-Wendisch, P., Mesecar, A., Thorgeirsson, T., Shin, Y. K., and Kustu, S. 1999. MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. *J. Bacteriol.* **181**: 4628-4638.

Rouiller, I., DeLaBarre, B., May, A. P., Weis, W. I., Brunger, A. T., Milligan, R. A., and Wilson-Kubalek, E. M. 2002. Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. *Nat. Struct. Biol.* **9**: 950-957.

Simms, S. A., Keane, M. G., and Stock, J. 1985. Multiple forms of the CheB methylesterase in bacterial chemosensing. *J. Biol. Chem.* **260**: 10161-10168.

Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S., and McKay, D. B. 2000. Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* **103**: 633-643.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* **69**: 183-215.

Terwilliger, T. C. 2000. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **55 (Pt 4)**: 849-861.

Vagin, A. A. and Isupov, M. N. 2001. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D Biol. Crystallogr.* **57**: 1451-1456.

Vale, R. D. 2000. AAA proteins. Lords of the ring. *J. Cell Biol.* **150**: F13-F19.

Volkman, B. F., Nohaile, M. J., Amy, N. K., Kustu, S., and Wemmer, D. E. 1995. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. *Biochemistry* **34**: 1413-1424.

Wang, B. C. 1985. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**: 90-112.

Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H., Seong, I. S., Lee, C. S., Chung, C. H., and Eom, S. H. 2001a. Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure (Camb)* **9**: 177-184.

Wang, J., Song, J. J., Seong, I. S., Franklin, M. C., Kamtekar, S., Eom, S. H., and Chung, C. H. 2001b. Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU. *Structure (Camb)* **9**: 1107-1116.

Wang, Y. K., Lee, J. H., Brewer, J. M., and Hoover, T. R. 1997. A conserved region in the $\sigma^{54}$-dependent activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. *Mol. Microbiol.* **26**: 373-386.

Weiss, D. S., Klose, K. E., Hoover, T. R., North, A. K., Porter, S. C., Wedel, A. B., and Kustu, S. 1992. Prokaryotic transcriptional enhancers. In *Transcriptional Regulation* (ed. S. McKnight and K. Yamamoto), pp. 667-694. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Wootton, J. C. and Drummond, M. H. 1989. The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng.* **2**: 535-543.

Xu, H. and Hoover, T. R. 2001. Transcriptional regulation at a distance in bacteria. *Curr. Opin. Microbiol.* **4**: 138-144.

Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. 2001. Crystal structure of the Holliday junction migration motor protein RuvB from Thermus thermophilus HB8. *Proc. Natl. Acad. Sci. U S A* **98**: 1442-1447.

Yan, D. and Kustu, S. 1999. "Switch I" mutant forms of the bacterial enhancer-binding protein NtrC that perturb the response to DNA. *Proc. Natl. Acad. Sci. U S A* **96**: 13142-13146.

Yu, R. C., Hanson, P. I., Jahn, R., and Brunger, A. T. 1998. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat. Struct. Biol.* **5**: 803-811.

Zhang, X., Chaney, M., Wigneshweraraj, S. R., Schumacher, J., Bordes, P., Cannon, W., and Buck, M. 2002. Mechanochemical ATPases and transcriptional activation. *Mol. Microbiol.* **45**: 895-903.

Zhang, X., Shaw, A., Bates, P. A., Newman, R. H., Gowen, B., Orlova, E., Gorman, M. A., Kondo, H., Dokurno, P., Lally, J., Leonard, G., Meyer, H., van Heel, M., and Freemont, P. S. 2000. Structure of the AAA ATPase p97. *Mol. Cell* **6**: 1473-1484.

**Table 1. Data collection, Phasing, Refinement Statistics**

| Crystal | $NtrC1^{RC}$(dimer) | $NtrC1^{C}$(heptamer) |
|---|---|---|
| Space group | $P3_221$ | $P1$ |
| Cell parameters (Å) | a = b = 94.76, c=195.01 | a=106.79,b=108.26,c=110.02 |
| | | $\alpha$=70.25, $\beta$=85.90, $\gamma$= 73.27 |
| Source | ALS 5.0.2 | ALS 8.3.1 |
| Data and Phasing | | |
| Resolution (Å) | 50.0 – 2.4 | 62.02 – 3.1 |
| Wavelength (Å) | 0.9793 | 1.0 |
| Completeness (%) | 99.4 (83.4) | 98.0 (97.1) |
| $R_{sym}$ (%)[a] | 8.0 (41.5) | 6.1 (42.1) |
| $R_{cullis}$ (ano)[b] | 0.38 | -- |
| Figure of merit[c] | 0.52 | -- |
| Refinement Statistics | | |
| Resolution (Å) | 42.6-2.4 | 20.0-3.1 |
| No. of reflections | 36719 | 78375 |
| No. of atoms | protein/ADP/solvent | |
| | 6192/269/54 | 27677/378/0 |
| No. of molecules/AU | 2 | 14 |
| $R$ factors (%)[d] | $R_{work}/R_{free}$, 21.3/25.7 | 26.6/32.9 |
| R.m.s.d of bonds[e] | length/angle, 0.007 Å /1.4° | 0.01 Å / 1.6° |

[a] $R_{sym} = \sum |I_i - \langle I_i \rangle| / \sum I_i$, where $\langle I_i \rangle$ is the average intensity of symmetry equivalent reflections. Numbers in parenthesis are statistics for the highest resolution bin. Same for completeness.
[b] $R_{cullis}$ (ano) = $\sum | \Delta_{ano}(obs) - \Delta_{ano}(cal) | / \Delta_{ano}(obs)$ for acentric reflections, where $\Delta_{ano}(obs)$ is the anomalous difference.
[C] Both figure of merit and $R_{cullis}$ were calculated with reflections up to 3.0 Å
[d] $R$ factor = $\sum |F(obs) - F(calc)| / \sum F(obs)$.
[e] R.m.s.d = root-mean-square-deviations.

**Figure Legends**

**Figure 1.** Overall structures of $NtrC1^{RC}$ and $NtrC1^{C}$. **a)** Structure of $NtrC1^{RC}$. $NtrC1^{RC}$ forms a dimer in the asymmetric unit. The monomers are colored gray and gold. The GAFTGA and β-hairpin insertions into the central ATPase domain are colored blue and green, respectively. ADP and the catalytic arginine residue (R293) are shown using a ball-and-stick representation. R293 cannot contact the nucleotide in the inactive dimer. **b)** Structure of the $NtrC1^{C}$ heptamer. The left side is a top view illustrating how the protomers pack in the context of the heptamer. Each protomer is labeled with a letter, A through G. The GAFTGA insertion (blue) and β-hairpin (green) comprise the pore in the heptamer. The Sensor II helix is colored cyan. On the right side is a side view.

**Figure 2.** The GAFTGA insertion forms an unusual structural motif. **a)** Comparison of the central ATPase domain of $NtrC1^{RC}$ (left) and NSF D2 (right). The α/β and α-helical subdomains are colored gray and light blue, whereas the helix carrying the GAFTGA motif and the β-hairpin insertion in NtrC1 are colored blue and green, respectively. The putative Sensor I residue (N280) and the Sensor II residue (R357) of NtrC1 are shown using a ball-and-stick representation. Note that ADP is bound in $NtrC1^{RC}$ but ATP is bound in NSF D2. Arrows indicate the insertions in NtrC1 relative to NSF D2. **b)** A close view of the GAFTGA insertion. Two H-bonds, typical of anti-parallel β-strands, are shown as dashed lines. Sequence alignments among $\sigma^{54}$ activators and other $AAA^{+}$ ATPases are shown with the conserved sequences highlighted. The sequences that form the last turn of H8 in other $AAA^{+}$ ATPases align well with those of $\sigma^{54}$ activators. **c)** OMIT map of the GAFTGA region including H8 (residues 203-232) with density contoured at 3.5σ.

**Figure 3.** Surface representations of both the inactive $NtrC1^{RC}$ dimer and the active $NtrC1^{C}$ heptamer. a) Surface plot of $NtrC1^{RC}$ dimer. The monomers are colored gray and green. b) The

central domain of one monomer in the inactive dimer. Buried regions of contact with the other central domain in the dimer are shown in gray. c) Same as b) but rotated 90°. The common region that is buried in both dimers and heptamers is colored red. d) Surface plot of $NtrC1^{C}$ heptamer. Two protomers from the heptamer are colored gold and blue. e) One protomer from the active heptamer. The contact regions with the preceding and following protomers in the heptamer are colored gray and blue, respectively. f) Same as e) but rotated 90°.

**Figure 4.** Subdomain motion in the central ATPase domain between inactive $NtrC1^{RC}$ and active $NtrC1^{C}$. **a)** The hinge motion between the α/β subdomain and the α-helical subdomain. Residues 160 to 306 of both inactive and active NtrC1 are superimposed. The central domain of the inactive $NtrC1^{RC}$ (dimer) is colored blue and that of the $NtrC1^{C}$ (heptamer) is colored green. The red dot indicates the rotation axis, which is perpendicular to the plane of the paper. **b)** The nucleotide binding sites for both $NtrC1^{RC}$ (blue) and $NtrC1^{C}$ (green) are shown. R293 (orange) from the adjacent protomer is located near the nucleotide-binding site in the active $NtrC1^{C}$ heptamer, whereas it is located on the opposite side in the inactive $NtrC1^{RC}$ dimer (see Fig. 1A). The single $Mg^{2+}$ present in the dimer is indicated by a green dot.

**Figure 5.** Correlation of structure with genetic and biochemical studies. **a)** Mutations that affect the function of $\sigma^{54}$ activators are mapped onto the structure of the central ATPase domain. Mutations that affect ATP binding, ATP hydrolysis, and transcriptional activation are colored red, green, and blue, respectively. Residues shown in ball-and-stick representation are: E174 and R357 (red); E207, L208, F209, D237, D238 and R293 (different subunit) (green); and A215, F216, T217, and A219 (blue). **b)** The seven conserved regions (C1 through C7) of the central ATPase domain of $\sigma^{54}$ activators are shown. These regions cover residues: 167-174 (C1); 191-

198 (C2); 205-219 (C3); 232-242 (C4); 251-256 (C5); 273-310 (C6); 352-361 (C7). The significance of each region is discussed in the text.

**Figure 6.** Oligomerization of NtrC1. Surface representations of NtrC1 changing from a dimer to a heptamer. In brackets, the surface representations for the receiver and central domains are separated to illustrate how they adopt different conformations upon phosphorylation. The dashed lines indicate the connections between the receiver and central domains. The asterisk indicates a transient state. For the phosphorylated receiver domain, we used the structure of $BeF_3^-$-$DctD^R$ (PDB code:1L5Y), the activated form (see text).

Fig. 1

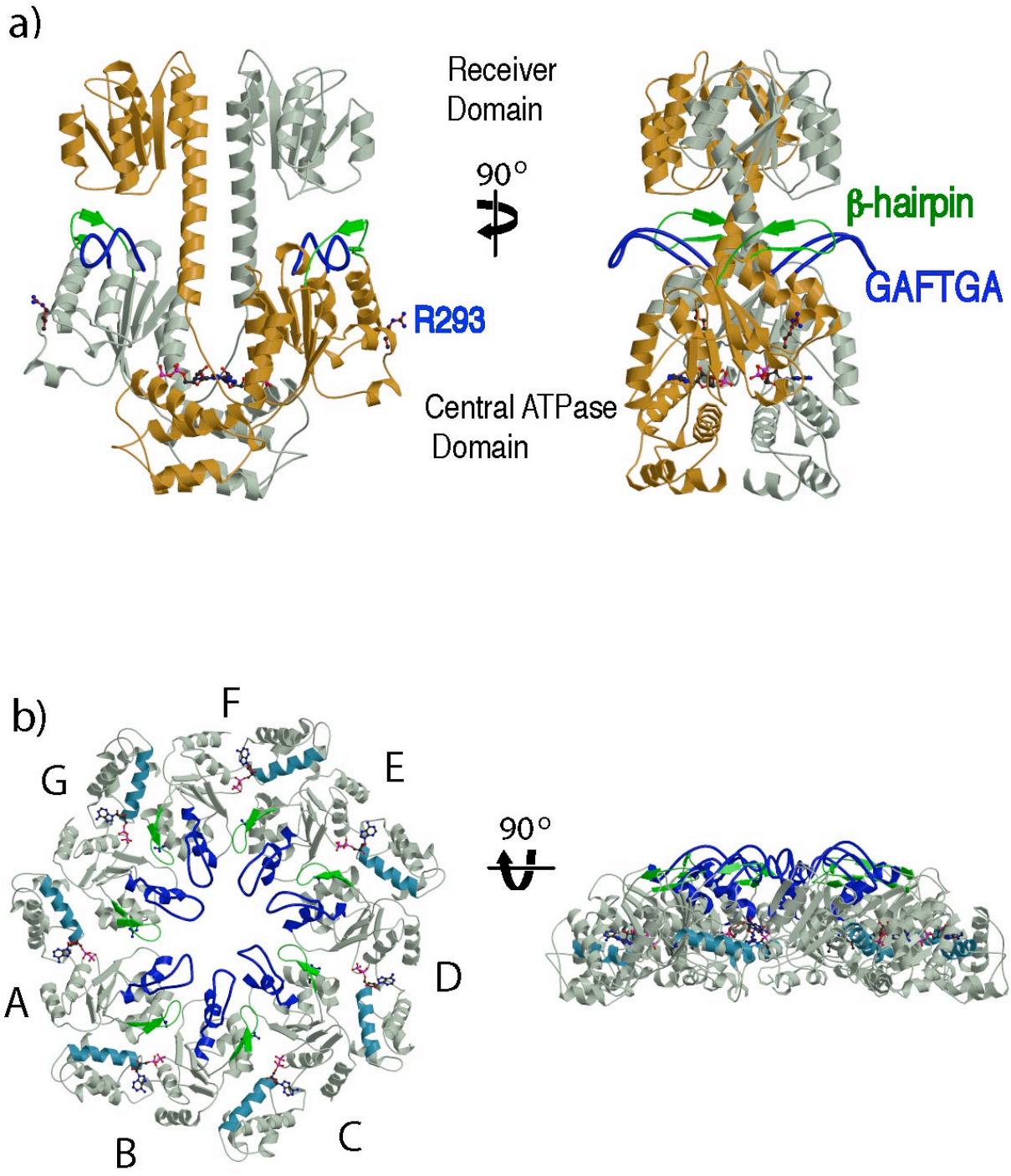

a)
Receiver
Domain
90°
β-hairpin
GAFTGA
R293
Central ATPase
Domain
b)
F
G
E
90°
D
A
B
C

Fig. 2

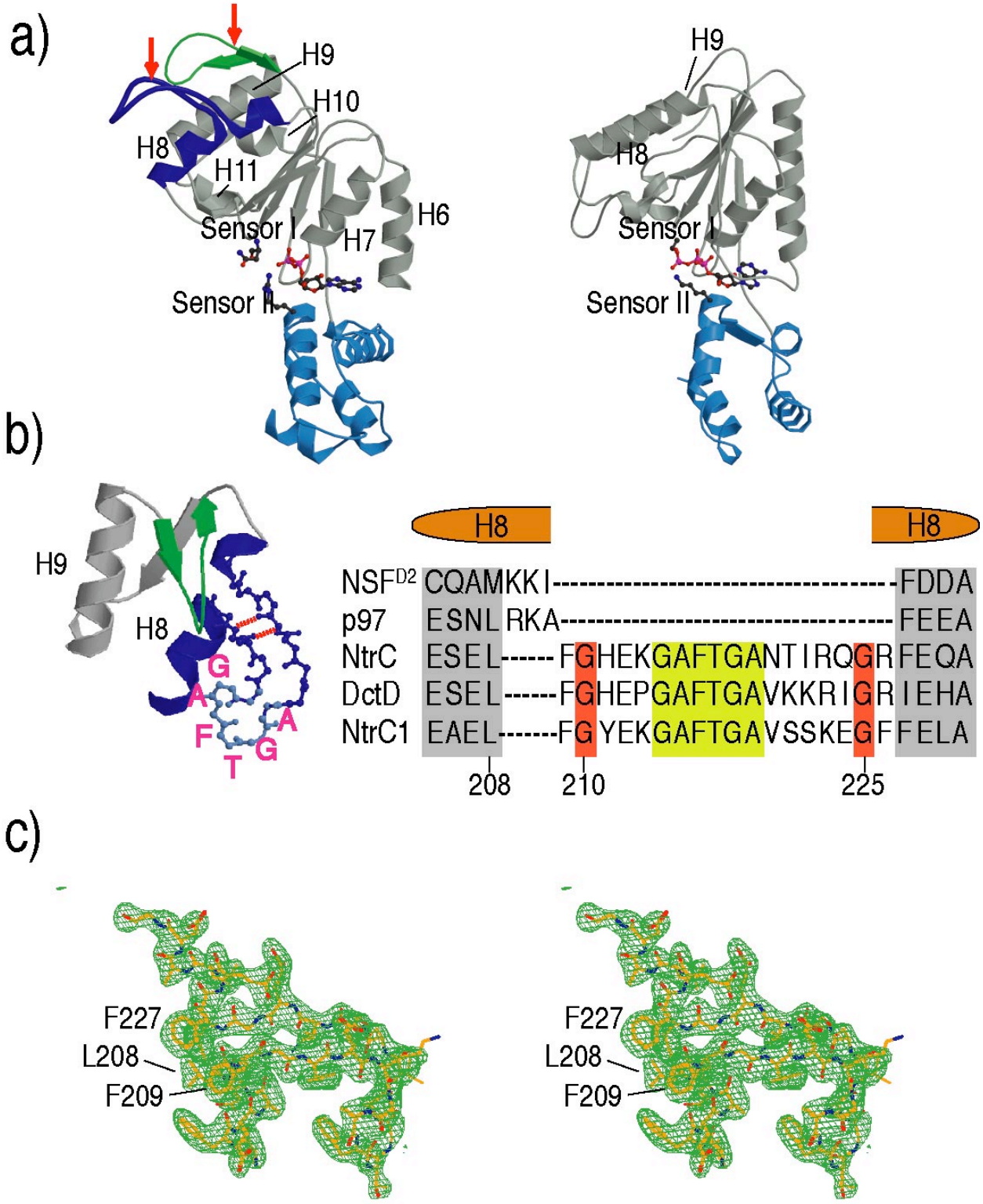

a)
H9
H10
H8
H11
Sensor I
H7
H6
Sensor II
H9
H8
Sensor I
Sensor II
b)
H9
H8
G
A
A
F
G
T
H8
H8
NSF D2 CQAMKKI----------FDDA
p97 ESNL RKA----------FEEA
NtrC ESEL------FGHEKGAFTGANTIRQGRFEQA
DctD ESEL------FGHEPGAFTGAVKKRIGRIEHA
NtrC1 EAEL-------FGYEKGAFTGAVSSKEGFFELA
208
210
225
c)
F227
L208
F209
F227
L208
F209

Fig. 3

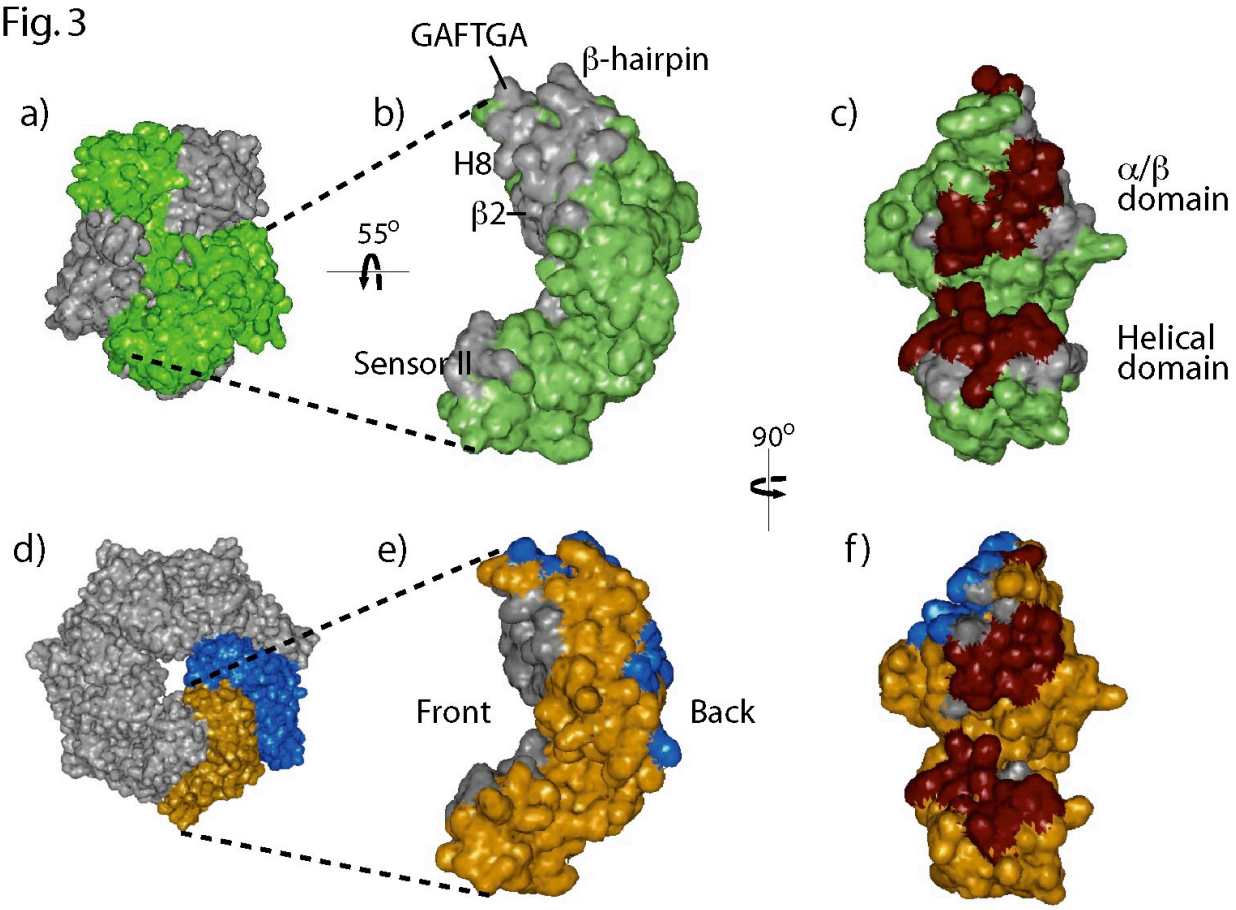

a)
b)
GAFTGA
β-hairpin
H8
β2
55°
Sensor II
c)
α/β
domain
Helical
domain
90°
d)
e)
Front
Back
f)

Fig. 4

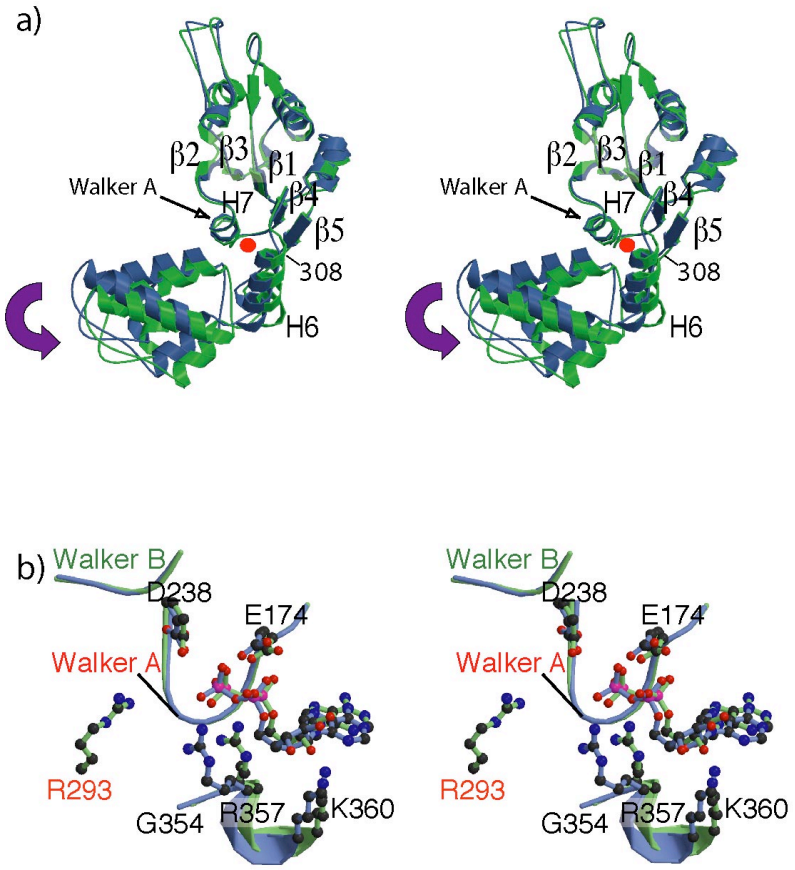
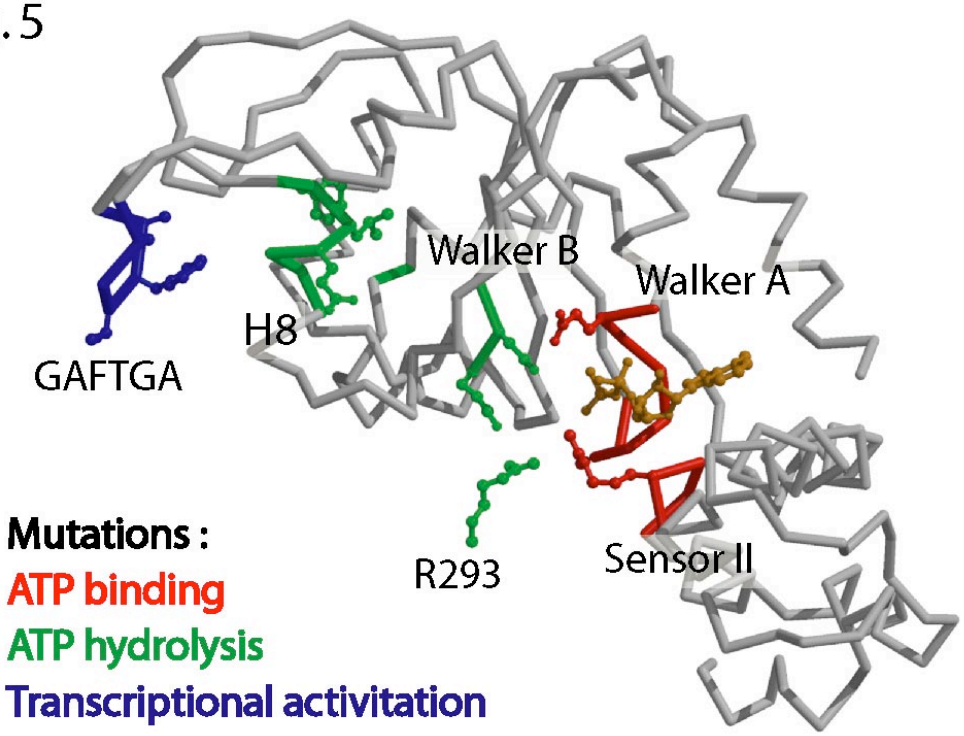
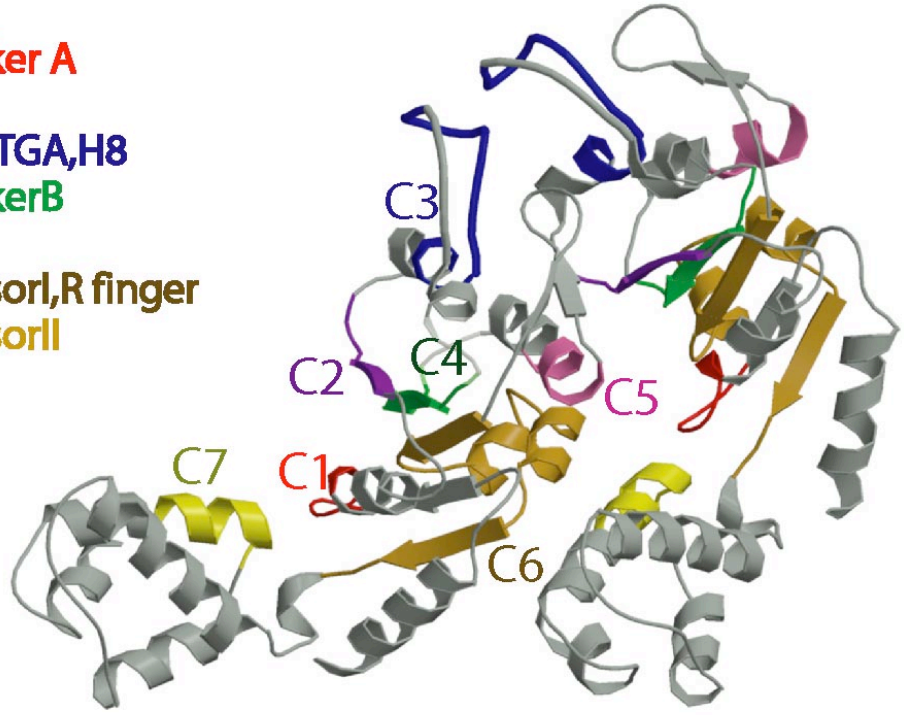

a)
β2
β3
β1
Walker A
H7
β4
β5
308
H6
β2
β3
β1
Walker A
H7
β4
β5
308
H6
b)
Walker B
D238
E174
Walker A
R293
G354
R357
K360
Walker B
D238
E174
Walker A
R293
G354
R357
K360

Fig. 5

a)


Walker B
Walker A
H8
GAFTGA
R293
Sensor II
Mutations :
ATP binding
ATP hydrolysis
Transcriptional activitation

b)


C1:Walker A
C2
C3:GAFTGA,H8
C4:WalkerB
C5
C6:SensorI,R finger
C7:SensorII
C3
C2
C4
C5
C7
C1
C6

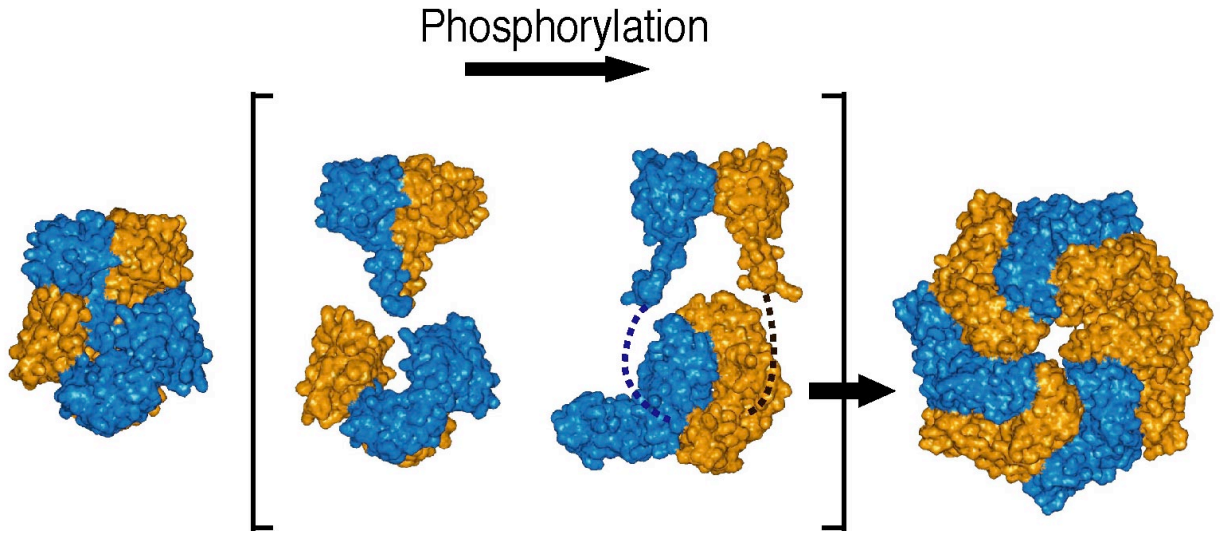



Fig. 6