doi:10.1016/j.jmb.2005.08.003 J. Mol. Biol. (2005) 353, 242–255

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Negative Regulation of AAA+ ATPase Assembly by Two Component Receiver Domains: A Transcription Activation Mechanism that is Conserved in Mesophilic and Extremely Hyperthermophilic Bacteria

Michaeleen Doucleff[1,2], Baoyu Chen[3], Ann E. Maris[1] David E. Wemmer[1,2], Elena Kondrashkina[4] and B. Tracy Nixon[3]*

[1]*Physical Biosciences Division Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA*

[2]*Department of Chemistry University of California Berkeley, CA 94720, USA*

[3]*Department of Biochemistry and Molecular Biology, The Pennsylvania State University University Park, PA 16802 USA*

[4]*BioCAT at APS/Argonne National Lab, Illinois Institute of Technology, 9700 South Cass Ave, Argonne, IL 60439, USA*

**Corresponding author*

Only a few transcriptional regulatory proteins have been characterized in extremely hyperthermophilic organisms, and most function as repressors. Structural features of the NtrC1 protein from the hyperthermophilic bacterium *Aquifex aeolicus* suggested that this protein functions similarly to the σ^{54}-polymerase activator DctD of *Sinorhizobium meliloti*. Here, we demonstrate that NtrC1 is an enzyme that hydrolyzes ATP to activate initiation of transcription by σ^{54}-holoenzyme. New structural data, including small-angle solution scattering data and the crystal structure of the phosphorylated receiver domain, show that NtrC1 uses a signal transduction mechanism very similar to that of DctD to control assembly of its AAA+ ATPase domain. As for DctD, the off-state of NtrC1 depends upon a tight dimer of the receiver domain to repress oligomerization of an intrinsically competent ATPase domain. Activation of NtrC1 stabilizes an alternative dimer configuration of the receiver domain that is very similar to the on-state dimers of the DctD and FixJ receiver domains. This alternative dimer appears to relieve repression of the ATPase domain by disrupting the off-state dimerization interface along the helical linker region between receiver and ATPase domains. Bacterial enhancer binding proteins typically have two linker sequences, one between N-terminal regulatory and central ATPase domains, and one between the central ATPase and C-terminal DNA binding domains. Sequence analyses reveal an intriguing correlation between the negative regulation mechanism of NtrC1 and DctD, and a structured N-terminal linker and unstructured C-terminal one; conversely, the very different, positive mechanism present in NtrC protein occurs in the context of an unstructured N-terminal linker and a structured C-terminal one. In both cases, the structured linkers significantly contribute to the stability of the off-state dimer conformation. These analyses also raise the possibility that a structured linker between N-terminal regulatory and central output domains is used frequently in regulatory proteins from hyperthermophilic organisms.

Keywords: NtrC1; sigma54; two component signal transduction; hyperthermophile; gene regulation

Introduction

Most bacterial and archaeal proteomes contain a diverse array of regulatory proteins that function as repressors or activators of gene transcription. Studies of the regulatory proteins found in extremely hyperthermophilic bacteria and archaea have only recently begun. Most of the proteins

Abbreviations used: ORF, open reading frame; CCD, charge-coupled device; UAS, upstream activating sequence; HSQC, heteronuclear single quantum coherence; SAXS, small-angle X-ray scattering; WAXS, wide-angle X-ray scattering.

E-mail address of the corresponding author: btn1@psu.edu

that have been characterized function as transcriptional repressors, e.g. the arginine repressor protein (ArgR) in the bacterium *Thermotoga neapolitana*[1] and maltose-specific gene regulator (TrmB) in the archaeon *Thermococcus litoralis*.[2] To date the only activator characterized is the Ptr2 protein of the archaeon *Methanococcus jannaschii*, which acts, in part, by recruiting TATA binding protein to the promoter.[3] Here we describe enzymatic and structural properties of NtrC1, a transcriptional activator from the bacterium *Aquifex aeolicus*. The results that we present extend previous structural studies[4] in which NtrC1 crystallized as a dimer in the inactive state and as a heptameric ring in the active state. *Aquifex aeolicus* grows optimally at 85 °C but can survive at 96 °C. NtrC1 is one of six putative activators of σ^{54}-holoenzyme in this organism. The σ^{54} activator proteins are widespread among bacteria and regulate the expression of genes required for diverse physiological processes. For example, NtrC of *Escherichia coli* controls transcription for genes required for nitrogen metabolism, and DctD of *Sinorhizobium meliloti* controls transcription for genes required for the transport of dicarboxylates. NtrC1 was named after NtrC because the proteins share 60% sequence similarity;[5] however, NtrC1's physiological role in the cell is unknown. Sequence homology and previous molecular structures show that NtrC1 has three domains: an N-terminal receiver domain (Pfam accession number PF00072), a central AAA+ ATPase domain (Pfam PF00158), and a C-terminal DNA binding domain (Pfam PF02954). Henceforth these will be referred to as R (regulatory or receiver), C (central) and D (DNA-binding) domains. Two segments of sequence that connect the domains identified by homology are termed linkers, with L_1 denoting the segment that joins the R and C domains, and L_2 the segment joining the C and D domains.

Here, we show directly that NtrC1 is a response regulator that uses its AAA+ ATPase domain to stimulate transcription of an open reading frame (ORF) of unknown function by σ^{54}-holoenzyme. We provide further evidence for two distinct signal transduction mechanisms between R and C domains of σ^{54}-dependent transcriptional activators, which were first revealed by comparing the function of NtrC from enteric bacteria to that of DctD from *S. meliloti*.[6] Our present functional and structural data for the R domain of NtrC1, including solution scattering data and the crystal structure for its phosphorylated state, indicate that NtrC1 uses the same mechanism as DctD to regulate oligomerization of its C domain and hence AAA+ ATPase and activation of transcription. Sequence analyses suggest that this mechanism occurs in other two-component response regulators, and that hyperthermophilic microbes use a structured linker to join the receiver and output domains more frequently than mesophilic counterparts.

Results

The AAA+ ATPase domain of NtrC1 is intrinsically competent

Removing the receiver domain and linker L_1 from DctD strongly derepresses its ATPase domain.[7] We found the same to be true for NtrC1. In a heterologous, *in vivo* system containing the *dctA-lacZ* reporter gene in *E. coli* cultures grown at 43 °C, removal of the receiver domain and L_1 linker (the C-L_2-D construct) increased transcription of NtrC1 by 80-fold (Figure 1(a)). The reporter gene contained an upstream activating sequence (UAS) specific for DctD, rather than for NtrC1. Fortuitous binding to the UAS may have enhanced the *in vivo* activation by the C-L_2-D construct of NtrC1. Further removing linker L_2 and the DNA-binding domain (leaving just C) resulted in a fragment that displayed only sixfold greater transcription activity than seen for the full-length protein. Adding back the receiver domain and linker L_1 to the C domain (construct R-L_1-C) repressed its activity, reducing it to the background level that was seen for the full-length protein.

Even in the absence of DNA, the purified C-L_2-D fragment of NtrC1 also exhibited high *in vitro* ATPase activity, hydrolyzing 300 μmol of ATP per minute per μmol of monomer at an optimal temperature of 82 °C (Figure 1(b)). The full-length NtrC1 (R-L_1-C-L_2-D construct) had no detectable ATPase activity, unless it was incubated with beryllofluoride, BeF_3^- (Figure 1(c)). BeF_3^- has been shown to bind to the active site aspartate, making an excellent analogue of aspartyl-phosphate,[8] and fully activating many isolated receiver domains,[9] including that of NtrC1 (see below). Full-length NtrC1 had maximal activity in 7 mM Be^{2+} + 42 mM F^-, with about 25% of the ATPase activity of the truncated protein. Higher concentrations (at a 1:6 ratio) were strongly inhibitory for the full-length protein (Figure 1(c)) but only modestly for the truncated C-L_2-D protein (data not shown). This is likely because high fluoride ion concentration favors formation of BeF_4^{2-} over BeF_3^-, and only the latter activates receiver domains. The C-L_2-D fragment, missing the receiver domain, does not require activation.

Repression of the AAA+ ATPase domain of NtrC1 depends upon the integrity of the receiver-L_1 linker dimer interface

Mutations in DctD that partially activated DctD *in vivo* clustered in residues of the R-L_1 dimer interface. Two of these substitutions, K121E and E122K, were shown biochemically to reduce the stability of the off-state dimer.[10] Analogous substitutions targeting the dimer interface (Figure 2(a)) were introduced into the *ntrC1* gene and assayed in the heterologous *in vivo* transcription assay described above. The amount of activity was quite

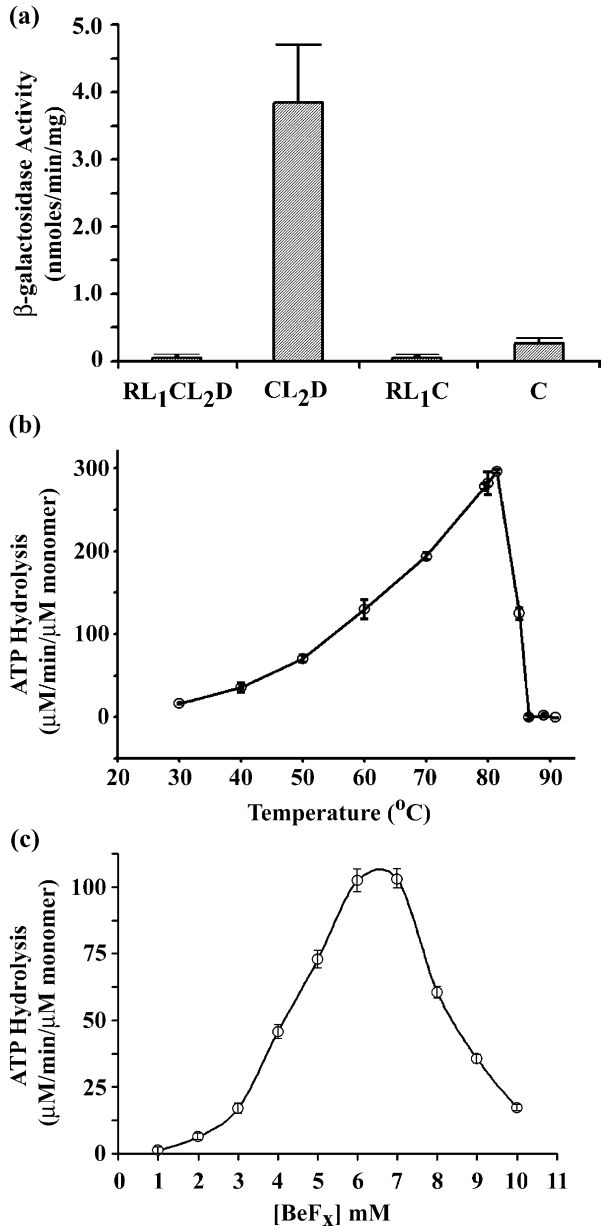

Figure 1. Assays of NtrC1 function. (a) Transcriptional activities of NtrC1. The indicated fragments were assayed in *E. coli* cells (cultured at 43 °C) containing the plasmid-borne reporter gene (*dctA-lacZ*) and transcription activator genes (T7*pro-ntrC1* constructs indicated). (b) ATPase activity *in vitro*, determined for the C-L_2-D fragment shown as a function of temperature. (c) ATPase activity of full-length NtrC1 is shown as a function of the concentration of the activating agent BeF_3^-.

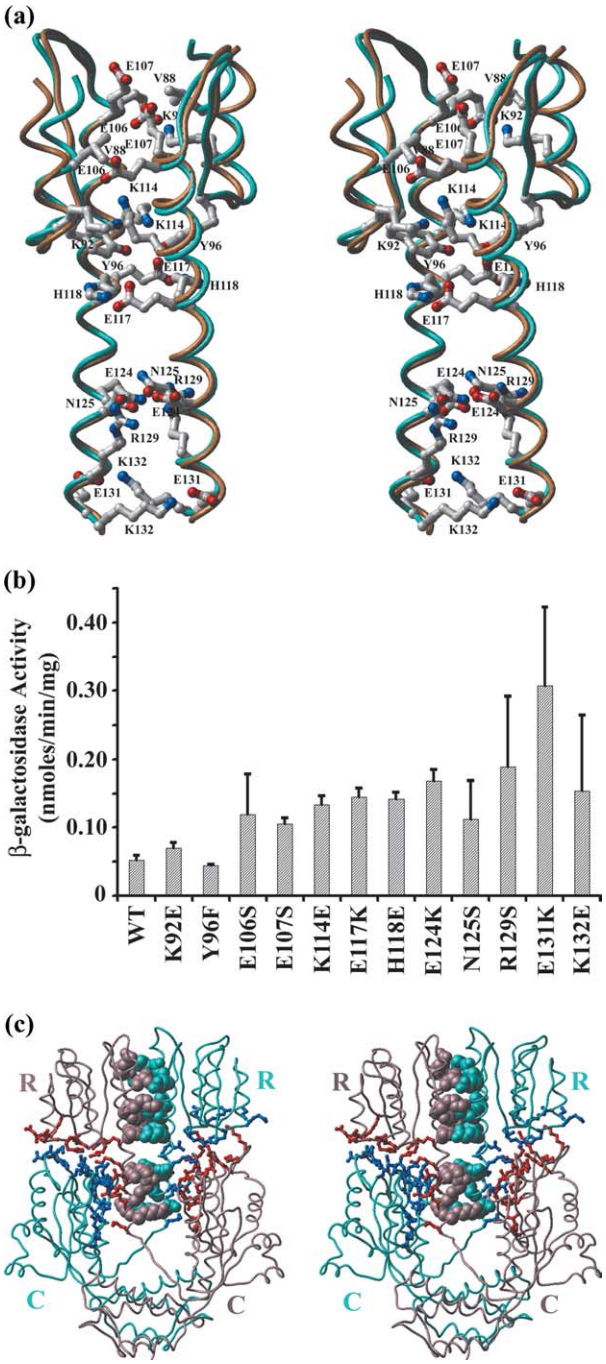

Figure 2. Mutagenesis of the NtrC1 off-state dimer interface. (a) Superpositioning of the off-state dimer interfaces of NtrC1 R-L_1 and DctD R-L_1 emphasizing their structural homology. Residues of NtrC1 involved in key interactions at the dimerization interface are modeled as sticks. (b) Except for V88, all of the interface residues indicated in (a) were altered and the resulting proteins subject to the *in vivo* transcription analysis as in (a). (c) The crystal structure of the off-state dimer of NtrC1 R-L_1-C fragment is drawn to illustrate the contacts between receiver domains studied above (CPK spheres) and additional contacts between the R-L_1 portion and the ATPase domain that were not studied (sticks).

low, probably because of the heterologous nature of the assay in which the reporter gene, core RNA polymerase, and sigma factors are all non-native (from *S. meliloti* and *E. coli*, not *A. aeolicus*). In addition, because the assays were performed in *E. coli* (no transcriptional assay is available in *A. aeolicus*), the assays could not be performed at a temperature optimal for NtrC1 ATPase activity (~70 °C).

Although the activities were low, it was still sufficient to reveal partial derepression of the central ATPase domain for substitutions E107S, K114E, E117K, H118E, E124K, R129S and E131K, and possibly for E106S, N125S and K132E, but not for substitutions K92E and Y96F (Figure 2(b)). Thus, as for DctD, the ability of the NtrC1 receiver domain and L_1 linker to repress assembly of its AAA+ ATPase depends upon interactions of residues in the off-state dimer interface, predominantly ones in the dimer interface formed by α-helix 5. The published structure of the NtrC1 fragment R-L_1-C shows additional stabilizing contacts between the R-L_1 portion of one subunit and the ATPase domain of the second subunit, and contacts between the two ATPase domains (Figure 2(c)). These contacts were not addressed in this mutagenesis study.

Activation of the NtrC1 and DctD R-L_1 proteins induces similar structural changes in solution and in crystals

^1H–^{15}N correlation spectra revealed that both NtrC1 and DctD R-L_1 proteins were fully converted to an active conformation by 5 mM $MgCl_2$, 5 mM $BeCl_2$ and 30 mM NaF (data not shown). Small and wide-angle X-ray scattering data were collected for both protein fragments in the presence of 5 mM Mg^{2+}/BeF_3^-. Essentially super-imposable scattering profiles were obtained for mid to high Q values, providing strong evidence that the activated structures of DctD R-L_1 and NtrC1 R-L_1 are highly similar in solution (Figure 3). Guinier plots of low Q data and *ab initio* structure modeling of combined low and high Q data indicated that under these solution conditions NtrC1 R-L_1 is a dimer of dimers, while DctD R-L_1 is a dimer; these dimer structures are very similar to the corresponding crystal structure presented below for phosphorylated NtrC1 R-L_1 and previously for the beryllofluoride-bound R-L_1 fragment of DctD[6] (detailed presentation and interpretation of these data, beyond the scope of this article, are being published elsewhere).

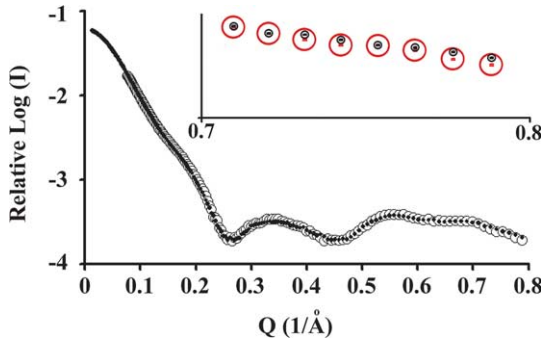

Figure 3. SAXS/WAXS data for activated R-L_1 fragments. Scattering data for the BeF_3^--NtrC1 R-L_1 fragment (scaled by a factor of 3.8; open circles) were plotted together with scattering data for the BeF_3^--DctD R-L_1 fragment (filled circles). The inset magnifies the indicated portion of the scattering curve and displays the error bars, which fall well within the symbols used to represent the data points.

The crystal structure determined for phosphorylated NtrC1 R-L_1 fragment contains two protomers in the asymmetric unit, as a non-crystallographic dimer. The structure, using diffraction data to 3.03 Å, was refined to an R_{work} of 25.1% and R_{free} of 28.3% (structure statistics are shown in Table 1). The phosphoryl groups on the conserved Asp51 residues were clearly evident in omit maps, with signal at the phosphate atoms greater than 15σ (Figure 4(a)). The configuration of conserved residues in the active site (D7, D8, D51, T79, and K101) in both protomers was similar to that seen in other activated receiver domains (such as Mg^{+2}/BeF_3^- – DctD,[6] Mg^{+2}/BeF_3^- – CheY,[11] and phosphorylated-FixJ),[12] except that the hydrogen bonding between the phosphoryl group and the amine on K101 was not apparent. Weak electron density around the CE and NZ atoms indicates that the end of this side-chain is only partially ordered and not completely constrained by hydrogen bonding with the phosphoryl group.

The phosphorylated NtrC1 R-L_1 domain forms a distinctly different dimer structure (Figure 4(b)) from that seen in the inactive NtrC1 R-L_1-C structure (Figure 2(a) and (c)). In the phosphoprotein structure, the dimeric interface involves primarily α-helix 4 and β-strand 5 (900 $Å^2$) rather than α-helix 4, β-strand 5, and α-helix 5 (2065 $Å^2$) seen in the inactive structure.[4]

NtrC1 binds with high affinity to tandem sites in a putative enhancer element

A ten kilobase-pair region flanking the *ntrC1* gene

Table 1. Crystallography data and refinement statistics

Data collection	
Space group	$P3_121$
Unit cell *a*, *b*, *c* (Å)	91.52, 91.52, 130.94
Total observations	84,596
Unique structure factors	12,451
Resolution range (Å)	17–3.03
R_{sym}[a,b]	4.3 (41.6)
Completeness[b] (%)	98.1 (97.6)
Average I/sigma[b]	12.4 (3.4)
Refinement	
Number molecules in asymmetric unit	2
Number protein atoms	2126
Refinement resolution (Å)	17–3.03
R_{cryst}[c]/R_{free}[d]	25.1/28.3
r.m.s.d. bond lengths (Å)	0.0092
r.m.s.d. bond angles (deg.)	1.60
Ramachandran	
Favorable region (%)	81.3
Allowed, generously allowed regions (%)	16.6, 2.1

[a] $R_{sym} = \Sigma_i|I_i - \langle I_i\rangle|/\Sigma_i I_i$, where $\langle I_i\rangle$ is the average intensity of symmetry equivalent reflections.

[b] Values for the highest resolution shell (3.19–3.03 Å) are given in parentheses.

[c] $R_{cryst} = 100\ \Sigma|F(obs) - F(calc)|/\Sigma\ F(obs)$. No σ cutoff was used.

[d] R_{free} calculated for 10% of randomly chosen unique reflections that were excluded from the refinement.

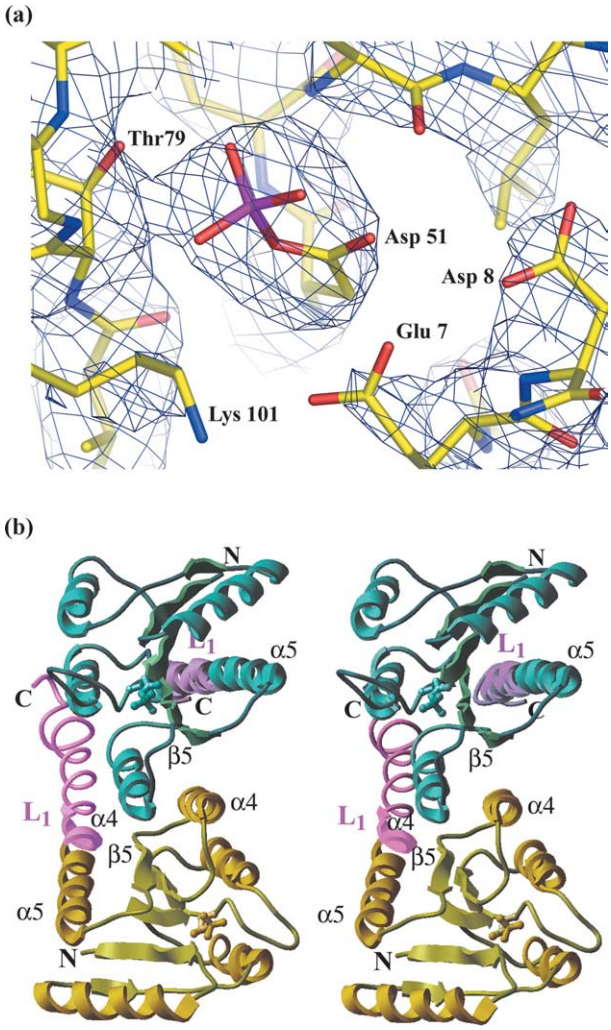

Figure 4. Crystal structure of the phospho-R-L_1 fragment of NtrC1. (a) A $2F_o - F_c$ omit map (contoured at 1.2σ) of the NtrC1 receiver domain active site made with the phosphoryl group on Asp51 removed from the model. (b) Stereo-pair of the phosphorylated NtrC1 R-L_1 structure shows the juxtaposition of subunits in the homodimer. Subunits are colored gold and cyan except for the L_1 linker segments, which are rendered in shades of magenta. The side-chain and phosphoryl group (sticks) are shown for residue D51.

of *A. aeolicus* was scanned for σ^{54}-dependent promoters. A single high-scoring site was identified, suggesting that transcription starts 100 base-pairs downstream from the end of the *ntrC1* coding region (Figure 5(a)). This suggests that the annotation of this region of the genome is incorrect. Coding region Aq_1119 begins at position −47 relative to the predicted transcription start site; however, 14 base-pairs downstream of the predicted transcription start site is an ATG codon that is preceded by a potential ribosome-binding site. If the latter is the correct translation start site, this would reduce the length of ORF Aq_1119 by 21 amino acid residues. Neighboring partial and nearly complete dyad symmetric elements were noted to span base-pairs −99 to −62 relative to the putative promoter, potentially a regulatory protein binding site. This segment is predicted to encode a strong bend in the DNA, as was previously seen to be a feature of the *dctA* UAS of *Rhizobium leguminosarum*.[13]

A 319 base-pair fragment corresponding to residues −251 to +68 of the putative Aq_1119 promoter was used in gel mobility-shift assays to evaluate *in vitro* binding. A clear shift was observed that saturated with 10 nM or less NtrC1 dimer (data not shown). DNaseI footprinting was performed under the same conditions, revealing sequence-specific binding to the tandem dyad repeats (Figure 5(b)). Quantitative analyses of the footprints (Figure 5(c)) revealed that the promoter proximal site B, which has stronger dyad symmetry, also has higher apparent affinity than the promoter distal site A (apparent K_D values (and 67% confidence intervals) were 0.7 (0.5, 1.0) and 4.8 (4.1, 5.8) nM, respectively). Additional studies with variant binding sites will be required to determine if these binding events have either positive or negative cooperativity.

Linker sequence motifs correlate with activation mechanism

A combination of structural and biochemical data available for NtrC *versus* DctD and NtrC1 indicate that the linker regions in the proteins contain sequences that correlate with specific aspects of function. First, the L_1-linker between R and C domains in NtrC seems to serve as a flexible tether.[14] The NtrC receiver domain expressed alone is monomeric. However, in the off states of DctD and NtrC1 the receiver domains are dimeric (or dimers of dimers) and the N-terminal part of L_1 is a helical extension of α-helix 5 of the receiver domain. α-Helix 5 forms a pseudo coiled coil (crossed helical pair) that is a major component of the receiver domain dimerization interface, and this interface is responsible for holding the C domains in an inactive, face to face configuration.[4] Combined results of secondary structure and coiled coil prediction programs correctly identify this element in the sequences of DctD and NtrC1, and indicate that it is not present in NtrC (Figure 6). The pseudo coiled coil structure is key to the regulatory function in DctD and NtrC1, and its presence appears to be a predictor of regulatory mechanism.

The second region of sequence that correlates with regulatory mechanism is in L_2, the linker between the end of the ATPase domain and the start of the DNA-binding domain. The DNA-binding domains of DctD, NtrC1 and NtrC all contain predicted C-terminal helix-turn-helix motifs that are expected to be responsible for the specific interactions in DNA binding. NMR data on the L_2-D portion of NtrC showed that the N-terminal part of L_2 (residues D380 to P399) is a very flexible random coil, whereas the C-terminal part of L_2 is structured, forming two helices that are an integral part of the intermolecular interface in the dimeric DNA-binding domain. This helical segment is

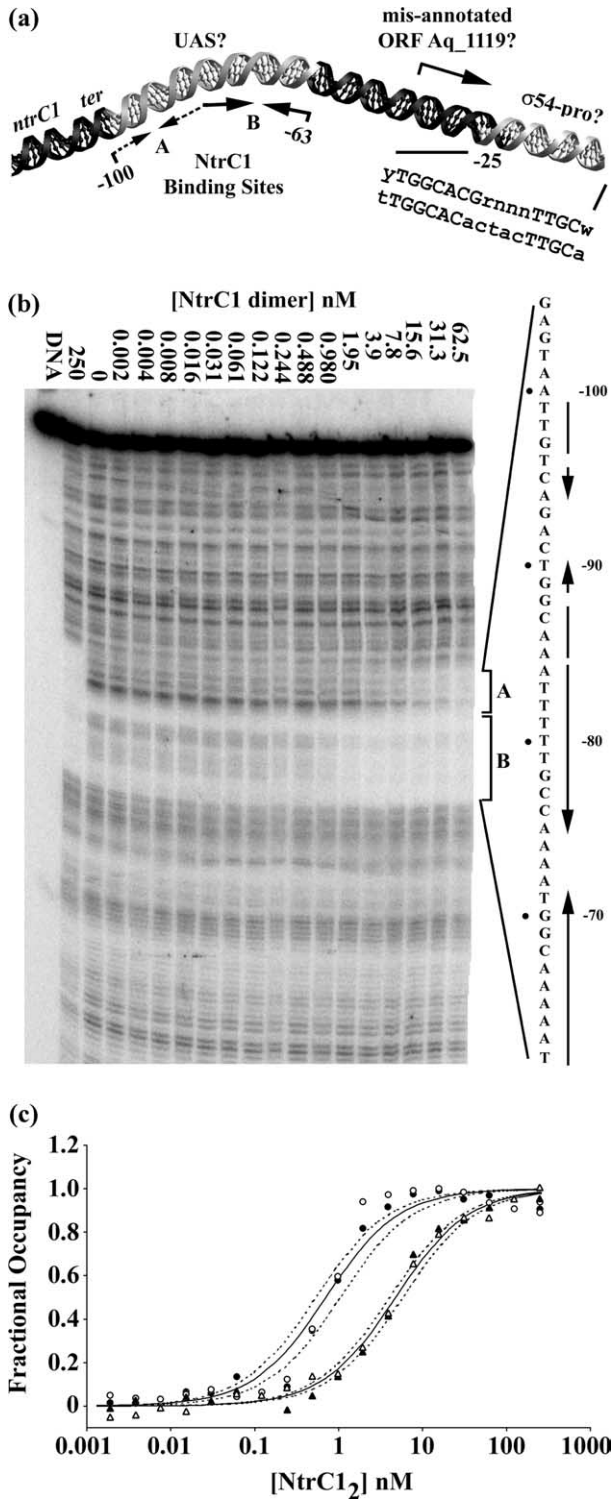

Figure 5. Putative *A. aeolicus* target gene regulated by NtrC1. (a) Termination of the *ntrC1* gene coding region is immediately followed by tandem binding sites for NtrC1 (labeled A and B) and a strong match to the σ^{54}-consensus promoter. Numbering is relative to this putative promoter, which resides a few base-pairs downstream of the annotated beginning of ORF Aq_1119. Not shown are putative ribosome-binding and translation-initiation sites located a few base-pairs downstream of the transcription start site. (b) DNaseI footprint of NtrC1 binding to the region diagrammed in (a). (c) Quantification of the binding reactions in (b), modeled as two independent sites. Filled and open symbols denote two separate experiments (site A, triangles; site B, circles).

homologous to the factor for inversion stimulation protein, Fis, in both sequence[15] and structure.[16] These two helices in L_2 are predicted well by secondary structure prediction programs. In stark contrast, for DctD and NtrC1 the L_2 region is much shorter and is predicted to have a non-structured segment, followed by a single helix that has no homology to the dimerization helices of NtrC/Fis (Figure 6). Consistent with secondary structure analysis, the DNA-binding domain of NtrC1 is not a strong dimer like NtrC (A.E.M. & D.E.W., unpublished results). Thus, the nature of L_2, recognizable from sequence, is a predictor for structure and dimerization of the DNA-binding domain. Many sequence-specific DNA-binding proteins in bacteria, including the σ^{54} activators, dimerize and recognize pseudo-palindromic DNA sequences. Either the receiver domain or the DNA-binding domain can act as a primary dimerization determinant, through recognizable sequence motifs as discussed above. There is only a need for one dimerization element in the protein, and hence, not surprisingly, the presence of the pseudo coiled coil in L_1 is anticorrelated with the presence of the extended linker and extra helix in L_2. For this reason, joint analysis of the two linker sequences should provide the most reliable predictor of the behavior in other activators.

Discussion

NtrC1 is an ATPase that stimulates transcription

Structural studies of NtrC1 (from a hyperthermophile) were initiated to complement work on NtrC (from a mesophile). Homology of NtrC1 to NtrC throughout the three functional domains, and the presence of a σ^{54} homolog in *A. aeolicus*, strongly suggested that NtrC1 is a transcriptional activator. Structures of the R-L_1-C and C domains of NtrC1 reported recently[4] indicated that its regulatory mechanism is distinct from that of NtrC. The positions of the receiver domains in the R-L_1-C structure strongly suggested a mechanistic analogy to DctD. The biochemical and structural studies presented here verify this analogy.

Prior studies revealed two distinct signal transduction mechanisms for regulating oligomerization of the AAA+ ATPase domain that is present in σ^{54} regulators. In "positive" regulation as seen with NtrC of enteric bacteria, phosphorylation of the receiver domain brings about a new interaction with the ATPase domain, leading to assembly of the active oligomer. In contrast, in "negative" regulation as seen with DctD, the "off-state" receiver domains stabilize an inactive state of the ATPase. Phosphorylation of the receiver domains leads to a rearrangement that releases the ATPases so that they can assemble the

oligomer. Figure 7(a) shows a schematic illustration of the two mechanisms emphasizing their positive and negative natures. This difference in mechanism leads to a fundamental difference in the behavior of the proteins when the receiver domains are deleted: this form of NtrC is not active under any conditions, whereas the corresponding form of DctD has been shown to be constitutively active. The structure of the R-L_1-C construct of NtrC1 clearly showed how the receiver domains could hold the ATPase domains in an inactive conformation, very likely the same mechanism as in DctD. In DctD amino acid changes that destabilize the pseudo coiled coil interface between the inactive state receiver domains lead to partial activity. Both solution (B.T.N., unpublished results) and crystal structures of the BeF_3^--activated DctD receiver domain indicated that the dimer rearranged upon activation. The data presented in Figures 1 and 2 show that the behavior of NtrC1 is exactly analogous to that of DctD. Genetic deletion of the receiver domain leads to constitutive transcriptional activity from a σ^{54} promoter driving a reporter gene. Furthermore, amino acid changes analogous to those made in the pseudo coiled coil of DctD lead to weak but clear transcriptional activity. These observations verify that NtrC1 is negatively regulated, and that the pseudo coiled coil region does play a significant role in maintaining the off state.

Our results show clearly that NtrC1, like DctD, has an intrinsically competent central domain with constitutively active ATPase activity *in vitro* and transcriptional activity *in vivo*. Thus, as predicted from sequence analysis and recent structural studies, the NtrC1 receiver domain and L_1 linker are responsible for negatively regulating the activity of the central AAA+ ATPase domain. Further, the truncated NtrC1 fragment (C-L_2-D) has a similar ATPase activity at 82 °C (300 μmol of ATP per minute per μmol of monomer) as the ATPase activity at 37 °C of the truncated DctD fragment (220 μmol of ATP per minute per μM monomer in the absence of DNA and 320 μmol of ATP per minute per μmol of monomer in the presence of DNA).[17,18] In contrast, the NtrC protein has significantly lower ATPase activity than NtrC1 and DctD. When activated, NtrC hydrolyzes ~4 μmol of ATP per minute per μmol of monomer in the absence of DNA[19] and ~20 μmol per minute per μM monomer in the presence of DNA.[20]

The active receiver domains of NtrC1 and DctD are structurally similar

The crystal structure of phosphorylated NtrC1 R-L_1 shown in Figure 4 again reinforces the similarity of NtrC1 to DctD, and essentially removes ambiguity about the structural nature of the activated dimer of DctD R-L_1. The crystal lattice of the Mg^{2+}/BeF_3^--bound DctD RL_1 fragment contained four different dimer configurations.[6] Solution structures derived from small-angle scattering data (that we report elsewhere) are inconsistent with two of the four possible dimer configurations, indicating that one of the remaining two structures is the physiologically relevant one. The phosphorylated NtrC1 dimer superimposes well with only one of the Mg^{2+}/BeF_3^--DctD dimers (r.m.s.d. ~ 1.3 Å for a superposition of the first 133 amino acid residues of active DctD and NtrC1), as well as to the dimer of phosphorylated FixJ R fragment (Figure 7(b)).

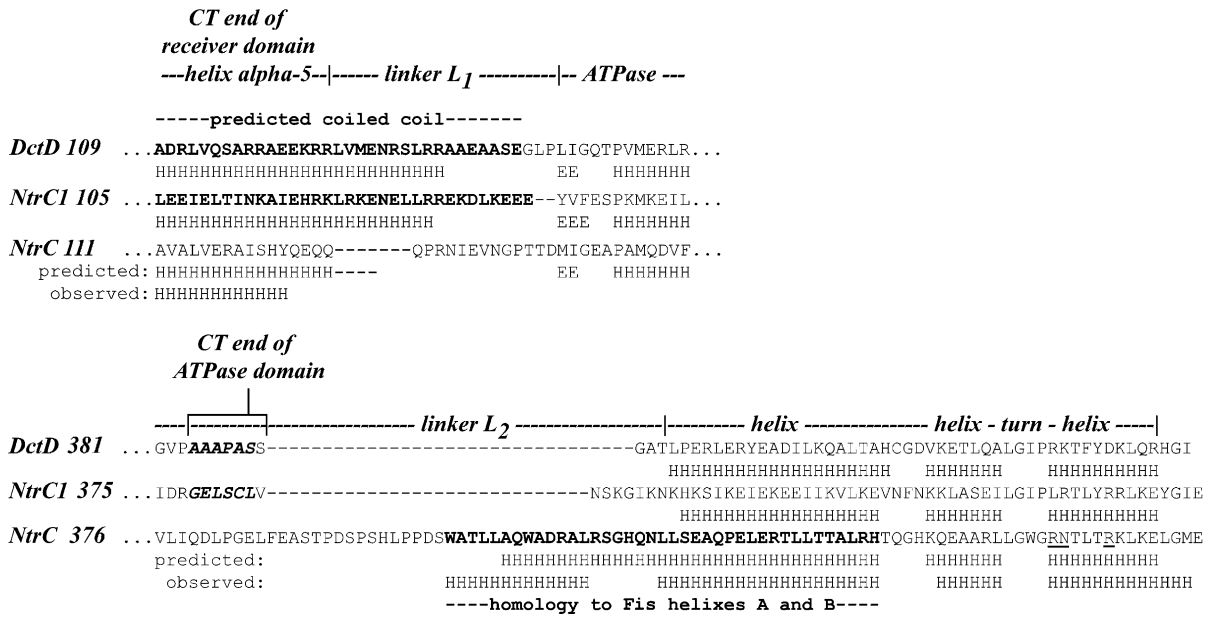

Figure 6. Sequence alignment and secondary structure prediction (H, helical; E, extended; –, random coil) for the linker regions of DctD, NtrC1, and NtrC. The top sequence shows the extended α-helix 5 and predicted coiled coils (bold residues) for linker L_1 of DctD and NtrC1, whereas in NtrC this region is unstructured. The bottom alignment shows the structured linker L_2 between the ATPase and the DNA-binding domains (in the NMR structure that defines regions of NtrC, the underlined residues of the last helix were substituted with alanine).

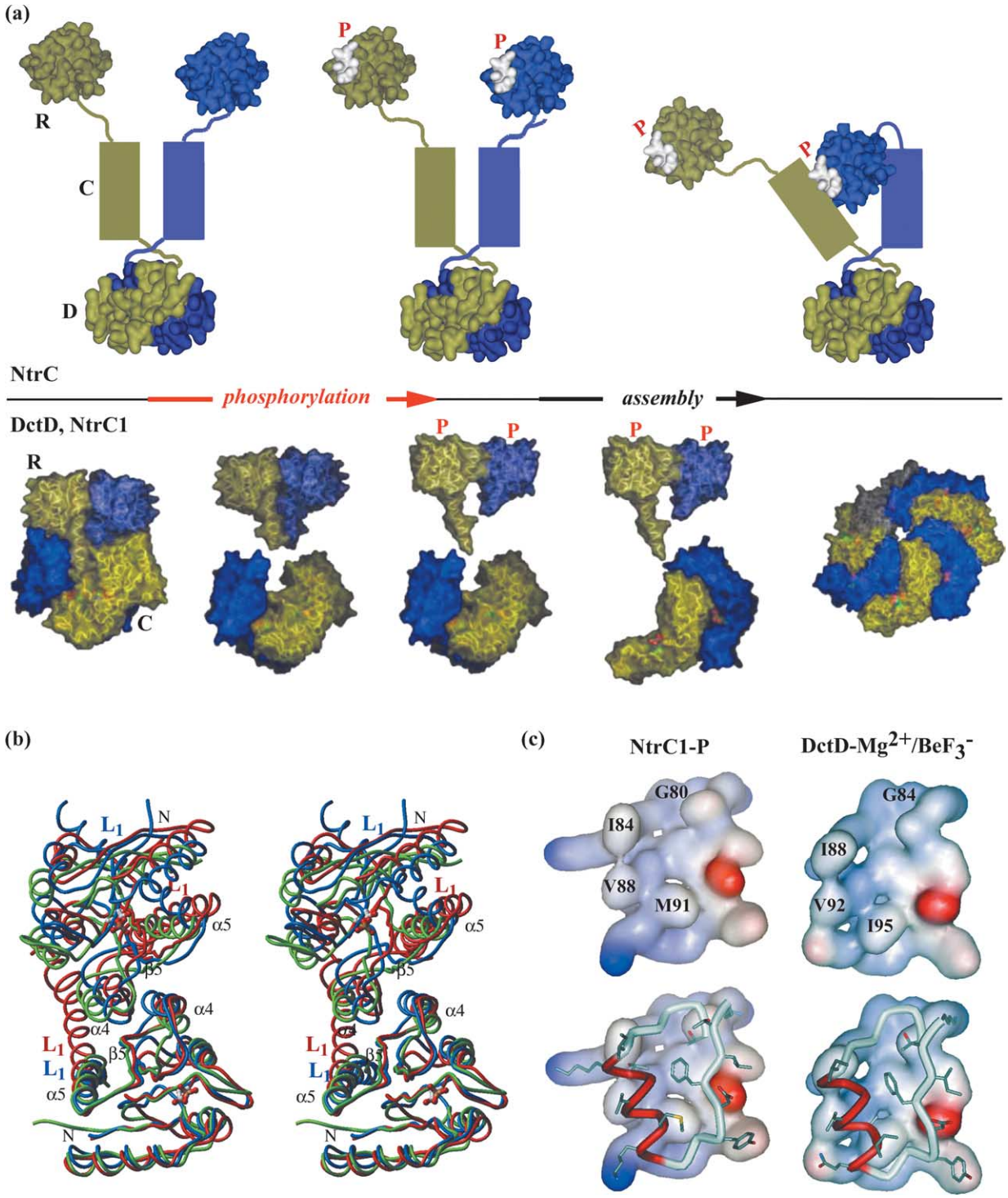

Figure 7. Different roles of linkers L_1 and L_2 in regulatory proteins. (a) Models for the transition from off to on-states following phosphorylation are shown for NtrC (top) and DctD or NtrC1 (bottom). In activated NtrC the DNA-binding domain is hidden underneath the hexamer ring. For DctD and NtrC1, no information is available to define the position of DNA-binding domains. Models were built using published structures: for NtrC fragments R (off-state PDB 1KRW, on-state 1KRX) and L_2-D (PDB entry 1NTC); and for NtrC1 and DctD fragment R of NtrC1 (this work), and fragments R-L_1-C (1NY5) and L_1-C (1NY6) of NtrC1, and R-L_1 of DctD (off 1L5Y, on 1L5Z). (b) Stereo-pair oriented as in Figure 4(b) showing the structural alignment of the phospho-NtrC1 R-L_1 homodimer (red), the BeF_3^--bound DctD R-L_1 homodimer (blue), and the phospho-FixJ R homodimer (green). For clarity, the phosphoryl group (sticks) is shown only for the NtrC1 fragment, and the L_1 linker segments are labeled. Only the bottom subunits were used to superimpose the structures to emphasize the similar juxtapositions of the second subunits in the dimer configurations (see Materials and Methods). (c) Surface representations of the homodimer interface of a single subunit of the phospho-NtrC1 R-L_1 crystal structure (left) compared with the BeF_3^--bound DctD R-L_1 crystal structure (right). Coloring indicates charge, and labeling of key residues of the T**G**X**G**X**Hyd**X_3**Hyd**X_2**Hyd** motif focuses attention on the upward opening of the β4-α4 loop that exposes the hydrophobic clefts for insertion of the hydrophobic knobs from the apposing subunit.

The rearrangement of the NtrC1 receiver domain upon activation has features related to those seen in other proteins of this class.[4,6,11,12,21,22] A key element is a change in position of the conserved Thr (residue 79 in NtrC1) to move the side-chain OH toward the active site to hydrogen bond to the phosphate. This movement changes the packing around the T79 side-chain methyl as well, and the conserved aromatic on β-strand 5 (F98 in NtrC1) then moves to a more deeply buried position that results in a repositioning of the top of helix 4, the top of strand β-strand 5, the loop between β-strand 5 and α-helix 5, and translation of α-helix 5 down along its helical axis. These rearrangements perturb the original dimerization interface between the two receiver domains, thus destabilizing it. These changes are very similar to those deduced for DctD.[6]

Molecular details of the subunit interfaces for these activated receiver domain homodimers are strikingly similar (Figure 7(c)). The activated interface is created by opening a hydrophobic pocket at the N-terminal end of α-helix 4 and β-strand 5. A specific T-G-X-G-X-Hyd-X_3-Hyd-X_2-Hyd motif, in which I84, V88 and M91 are the three hydrophobic residues (Figure 7(c)), described in more detail below), is frequently used in this interface. Conservation in this region has been discussed previously in the context of FixJ.[12]

As summarized in Figure 7(a), the relative orientations of the L_1 helices between the receiver domain and central ATPase domain in the homodimer change drastically upon phosphorylation. In the inactive dimer of NtrC1 R-L_1-C, the linker segments (the C-terminal end of α-helix 5) form a pair of crossed helices, which we found to be crucial in repressing ATPase activity. In the active dimer of NtrC1 R-L_1, these helical segments are no longer in contact but rather are pointed away from each other. In one protomer, the end of the L_1 segment makes crystal packing contacts with another symmetry-related molecule. Electron density is present for the entire length of this linker region. In the other protomer, the L_1 segment is not restrained by crystal packing contacts, and residues after amino acid residue 125 have extremely weak electron density and, hence, are difficult to model. The lack of electron density in this region suggests that the linker segment is intrinsically flexible in the phosphorylated NtrC1 R-L_1 dimer. It will be interesting to determine if the L_1 linker segment is still flexible when the central ATPase is attached to its C terminus.

NtrC1 probably regulates Aq_1119

Genes are tightly packed in the *A. aeolicus* genome.[5] The location of the putative UAS for NtrC1 is consistent with such packing because it is found immediately downstream of the termination codon of the *ntrC1* gene. The promoter proximal edge of the NtrC1 binding sites is at −62 relative to the putative promoter. Based on studies of the NtrC protein of mesophilic bacteria, this putative enhancer-binding element for NtrC1 appears to be located too close to the σ^{54} promoter. In previous studies,[23] the two high-affinity sites of the NtrC UAS functioned quite well when the proximal edge was placed −74 to −124 base-pairs upstream of the *glnA* promoter but gave only 25% and <10% function when it was placed at positions −69 and −63. If the sites we have identified immediately downstream of the *ntrC1* gene are *bona fide* promoter and enhancer elements, then other features, such as the higher ATPase activity and putative intrinsic bend in the DNA, might compensate for the apparent disadvantage arising from the close proximity. Also, high temperature may impart additional flexibility in DNA, which may facilitate looping between the UAS bound by NtrC1 and the promoter bound by σ^{54}-RNA polymerase. At this point it is unclear whether NtrC1 regulates the expression of just one gene in the *A.aeolicus* genome or more that have not been identified. There are no other predicted σ^{54}-dependent promoters that contain this apparently regulatory DNA-binding element.

Sequences in L_1 and L_2 allow classification of regulatory mechanism

Sequence comparisons used in conjunction with the available structural data reveal two sequence motifs that distinguish between the positive and negative regulatory mechanisms for σ^{54} activators. First, the L_1 region of NtrC is unstructured: its sequence is consistent with this, yielding no secondary structure prediction. The equivalent regions of DctD and NtrC1 are helical extensions of the C-terminal α-helix 5 of the receiver domain, forming pseudo coiled coil in the context of the dimer. Programs do correctly predict the secondary structure and the nature of the structure for the L_1 regions of these proteins (Figure 6). The L_2 region connecting the C and D domains similarly shows differences in length and predicted secondary structure that correlate with dimerization of the DNA-binding domains. The presence of dimerization interactions in the L_1 and L_2 regions is anticorrelated, which makes sense in light of the mechanisms of regulation in these proteins. The sequence of NtrC4 from *A. aeolicus*, a homolog of NtrC1, gives no helix or coiled coil prediction in L_1, but does have a longer L_2 region with predicted helical structure. Preliminary data are consistent with the prediction that the receiver domain from this protein is monomeric, whereas the DNA-binding domain is clearly a stable dimer.

It is interesting to examine linker sequences that join receiver and output domains in other σ^{54} activators with two-component response regulator receiver domains to classify their mechanisms. Such linker regions also appear to be relevant for other σ^{54}-dependent activators that are regulated by other types of domains. For example, in one subgroup of activator proteins, such as XylR and DmpR from *Pseudomonas*, the AAA+ ATPase is controlled by a unique regulator domain that binds and responds

directly to small effector molecules.[24] Mutational studies with both DmpR and XylR demonstrate that the integrity of the coiled coil linker between the regulator and ATPase domains, as in DctD, is critical for repressing ATPase and transcriptional activity in the absence of the signaling molecule.[25,26]

A. aeolicus has six predicted σ^{54}-dependent activators: three controlled by two-component receiver domains, two by GAF domains (Pfam accession number PF01590), and one by two PAS domains (Pfam accession number PF00989). Four of the six activators are predicted to have coiled coil or pseudo coiled coil linkers (Figure 8), as seen in DctD and NtrC1. It was previously noted that only about 4.5% of all two-component response regulators are predicted to have such structured linkers.[10] However, in that survey the majority of proteins sampled were from mesophilic bacteria. When we analyze response regulators from extremely hyperthermophilic organisms we find that about 10% (14/148) are predicted to have such structured linkers joining receiver and Pfam domains HD, HD-GYD, OmpR, PAS, GAF, HisKA, HATPase, HtH-ICLR, diacid-rec, HtH-AraC, GGDEF and trans_reg_C as well as uncharacterized ones.

An additional region that seems to have functionally relevant conserved sequences is in the α-helix 4 and β-strand 5 surfaces of the receiver domains. Phosphorylation of the NtrC and DctD/NtrC1 receiver domains leads to significantly different changes in this signaling surface. These differences underlie the two distinct mechanisms of linking phosphorylation with ring assembly of the central ATPase domain. In NtrC, activation causes a register shift in α-helix 4 creating a new binding surface for interaction with the N-terminal portion of the central ATPase domain (residues 134–160).[21,27,28] Recent studies using SAXS/WAXS and negative staining electron microscopy (B.T.N. & S. De Carlo, unpublished results) show that the receiver domain packs against the ATPase domain on the outside of a hexamer ring. In DctD, activation induces smaller changes in the signaling surface of the receiver domain that lead to a new homodimer interface. The repositioning of the loops between β-strand 3 and α-helix 4, and α-helix 4 and β-strand 5 of the receiver domain opens a hydrophobic pocket that permits the parallel stacking of hydrophobic side-chains from two subunits to form a stable on-state dimer.[6] These features appear to depend upon specific glycine and hydrophobic residues that immediately follow the highly conserved threonine or serine of the triggering mechanism (residues T83–I95 of DctD). The motif:

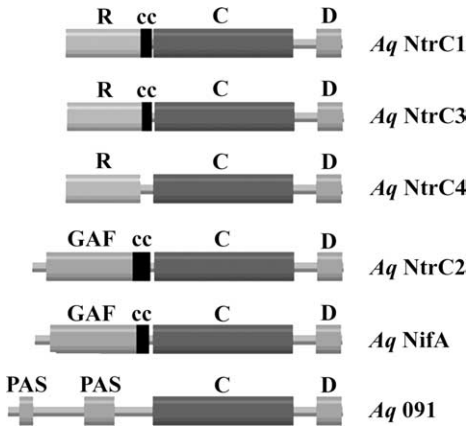

Figure 8. Predicted domain architecture for the six σ^{54}-dependent activators of *A. aeolicus*. Four of the six regulators are predicted to use a coiled coil linker L_1 (cc) to regulate assembly of the central AAA+ ATPase (C) domains. L_1 variously joins two component receiver (R) or GAF regulatory domains with the C domains. C-terminal helix-turn-helix DNA binding domains (D) are also shown.

is present in the NtrC1 receiver domain, and in those of 134 response regulators of the 5676 that were in the Pfam database at the time of this study (2.36% of the total. An additional 197 sequences (3.5% of the total) match the motif if the first or second glycine is allowed to be alanine, forming the more relaxed motif. Among these regulators are NtrC1 of *A. aeolicus*; DctD, FixJ and ActR of *S. meliloti*; DctD of *R. leguminosarum*; DctR and RegA of *Rhodobacteria*; FixJ and NodW of *Bradyrhizobium japonicum*; FixJ of *Azotobacter* (NtrX has serine instead of threonine in position 1); PgtA and SilR of *Salmonella typhimurium*; CusR of *Escherichia coli*; PilR of *Pseudomonas aeruginosa*; LuxO of *Vibrio*; SisR of *Borrelia bergdorferi*; CutR of *Streptomyces coelicolor A3*; and, many unstudied open reading frames. This and previous studies[6,12] show that activation stabilizes essentially identical dimer configurations of the receiver domains of NtrC1, DctD and FixJ.

These sequence analyses support the hypothesis that NtrC1 and DctD function similarly, and suggest that signal transduction occurs in a like manner in other regulatory proteins (of unknown structure) that also possess the appropriate sequence motifs. A helical L_1 linker connecting the regulator domain to its output domain predicts an off-state configuration similar to DctD and NtrC1. The presence of

T G H_{64} **G** X I_{54} X X X V_{96} X X M_{50} or relaxed: **T** G_{54} H_{40} G_{87} X I_{41} X X X V_{80} X X M_{37}

Y_{20}	V_{37}	I_4	L_{22}			A_{46}	Y_{20}	A_{13}	V_{37}	I_{19}	L_{36}
F_4	L_5		V_{20}				F_{13}		L_{15}		I_{14}
	M_5		I_7								V_{11}

subscripts represent percentage of sequences with that residue, X is any amino acid

this helix also suggests a tight structural coupling between regulatory and output domains in the off-state, as seen previously for DctD and now for NtrC1. In the case of σ^{54} transcriptional activators, this helix appears to be indicative of negative regulation by the regulatory domain. The presence of the motif T-G-X-Hyd-G-X-Hyd-X-X-X-Hyd-X-X-Hyd in α-helix 4 of a receiver domain suggests an on-state dimer similar to DctD, FixJ, and NtrC1.

Conclusions

Here, we present one of the first functional studies of a transcriptional activator from a hyperthermophilic organism. Our results demonstrate the conservation or lateral transfer between mesophilic and hyperthermophilic bacteria of a signal transduction mechanism used to control the functional state of σ^{54} transcriptional activators. The receiver domain and linker L_1 of NtrC1 from the hyperthermophilic *A. aeolicus* negatively regulates the central AAA+ ATPase with the same basic mechanism used by the receiver domain of DctD from the mesophilic *S. meliloti*. Despite inheriting a similar interaction between receiver and ATPase domains, the DNA binding specificities have evolved to target genes affecting different physiological needs of the respective cells. Characteristic features of both the off and on-states of the conserved mechanism can be used to predict aspects of signal transduction for additional two-component response regulators and other bacterial enhancer-binding proteins that are not controlled by two-component receiver domains.

Materials and Methods

Molecular biology

Gel mobility-shift assays and DNaseI footprints were performed largely as described by Sojda et al.[13] Oligonucleotides were used to amplify a 319 base-pair fragment from bases 781,309 to 781,628 of the *A. aeolicus* genome that was cloned into the SmaI site of the vector pUC18. Ends created by digestion with the BamHI restriction endonuclease at a site 241 base-pairs upstream of a putative σ^{54}-dependent promoter were labeled using the Klenow fragment of *E. coli* DNA polymerase. Protein and DNA were prebound for 45 min at 37 °C in buffer containing Tris–HCl (20 mM, pH 8.2), KCl (200 mM), and glycerol (12%, v/v) prior to either (a) loading in native 5% (w/v) gels for electrophoresis in TBE buffer, or (b) digestion with 250-fold diluted DNaseI (Promega) for 30 s at 37 °C. Digestions were quenched by the addition of 700 μl of stop solution (645 μl of ethanol; 5 μl of tRNA, 1 mg/ml; 50 μl of saturated ammonium acetate). DNA was precipitated at −80 °C overnight, washed with 70% (v/v) ethanol, dried, dissolved in loading buffer and electrophoresed in denaturing 6% polyacrylamide gels. Labeled DNA fragments were visualized using phosphorimaging (Molecular Dynamics). Non-linear, global regression analyses were performed using NONLIN as described.[29] *In vivo* transcription assays used a translational fusion of *S. meliloti dctA* and *E. coli lacZ* that was carried on the broad host range vector pRK290.[30] The reporter gene was introduced into the BL21/DE3/pLysS strain of *E. coli* (Novagen) for testing transcription activation by NtrC1 and recombinant versions of it that were expressed at 43 °C in overnight cultures at background levels (no IPTG added) from pET21a (Novagen) constructions. Cells were harvested by centrifugation and lysed by sonication. The extracts were then cleared by centrifugation and subjected to β-galactosidase assays as described.[30] Mutations were introduced into DNA using a commercial kit (QuickChange, Stratagene), and the entire gene for each construct was confirmed by sequence of at least one strand (Penn State Nucleic Acid Facility). To make the C-L_2-D construct, an NdeI site was placed before residue 137 and joined with the pET vector to remove the receiver domain and linker L_1. To make the R-L_1 and R-L_1-C constructs, a TAA stop codon was introduced before residue 137 and after residue 389, respectively. The following amino acid substitutions were created: K92E (AAA→GAA), Y96F (TAC→TTT), E106S (GAA→AGC), E107S (GAA→AGC), K114E (AAG→GAA), E117K (GAA→AAA), H118E (CAC→GAA), E124K (GAA→AAA), N125S (AAC→AGC), R129S (AGA→AGC), E131K (GAA→AAA) and K132E (AAG→GAA). For ATPase assays, 135 μl reactions containing 2 μM NtrC1 protein diluted into reaction buffer (50 mM Hepes (pH 6.9), 8 mM magnesium acetate, 100 mM potassium acetate) were pre-incubated for 3 min at 20 °C to 85 °C prior to the addition of 15 μl of ATP (4 mM final concentration, containing trace amounts of [γ-^{32}P]ATP, NEN-Perkin Elmer). At various times, 5 μl samples were quenched in 20 μl of 2 M HCl, to which 20 μl of chloroform and 13 μl of 2 M NaOH plus 3 M Tris base were added to return the solution to a neutral pH. The samples were then vortex mixed by and phase-separated by centrifugation at 10,000*g* for 5 min. After samples were collected, 1 μl aliquots were spotted onto a PEI cellulose TLC plate and, after drying, developed to separate ATP and Pi (buffer: 0.7 M boric acid, 0.4 M K_2HPO_4). Results were quantified using a phosphorimager (Molecular Dynamics).

Protein purification

Proteins were expressed in 20 l fermentations in a modified terrific broth containing yeast extract (40 g/l), Tryptone (20 g/l), glycerol (30 g/l), $NaH_2PO_4 \cdot H_2O$ (8 g/l), K_2HPO_4 (7 g/l) and antifoam polypropylene glycol 2000 (0.5 ml/l) or in 1 l cultures of BioExpress medium (Cambridge Isotopes Lab) for ^{15}N-labeling. Cells, resuspended in 20 mM Tris (pH 8.0), 500 mM NaCl, were lysed by sonication and cleared by centrifugation. The supernatant was heated for 30 min at 75 °C and cleared by centrifugation. Proteins were further purified by chromatography on Co-NTA, Q-Sepharose or S-Sepharose resins. NtrC1 proteins were either used immediately, dialyzed to 5 mM ammonium bicarbonate before subsequent lyophilization and rehydration when needed (R-L_1), or thawed from aliquots containing glycerol (30%; C-L_2-D, R-L_1-C, R-L_1-C-L_2-D) that were stored at −80 °C. Purity of proteins was assessed by SDS-PAGE. All proteins were centrifuged and passed through 0.1 μm pore size Anotop filter (Whatman) before data analysis.

^{1}H–^{15}N HSQC spectroscopy

Activation of NtrC1 R-L_1 by BeF_3^- was followed by

^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectra. Data were collected on a Bruker DRX-600 spectrometer and analyzed with the software NMRPipe.[31] NtrC1 R-L_1 (~750 μM monomer equivalents) was fully activated by 1 mM $BeCl_2$ and 7 mM NaF.

Static X-ray and dynamic light-scattering

Small and wide-angle X-ray (SAXS or WAXS) experiments were conducted on the Biophysics Collaborative Access Team (BioCAT) undulator beamline 18-ID at the Advanced Photon Source, Argonne National Laboratory.[32] DctD and NtrC1 R-L_1 solutions contained approximately 10 mg/ml of protein in 20 mM Tris (pH 8.0), 5 mM $MgCl_2$, 10 mM Tris(2-carboxyethyl)-phosphine (TCEp) (pH 7.75) with or without 5 mM $BeCl_2$ and 40 mM NaF. The NtrC1 R-L_1 solutions were incubated overnight at room temperature to ensure full reduction of disulfide bonds, and a BeF_3^- titration (from 10 μM to 10 mM BeF_3^-) was followed by SAXS to ensure full activation. For collection of scattering data, samples were exposed to focused X-rays (12 keV and 2×10^{13} photons/s flux) for an average of 0.58(±0.08) s. Two-dimensional scattering patterns were obtained by using a 5 cm×9 cm charge-coupled device (CCD) detector[33] at a specimen to detector distance of 2.78 m (SAXS) or 0.24 m (WAXS). Radiation damage was minimized by including 10 mM TCEP in the sample solutions and by pumping them through a 1.5 mm quartz capillary at 12.5 μl/s.[34] Recordings were repeated five times. Scattering intensity profiles over the Q range from 0.007 $Å^{-1}$ to 1.400 $Å^{-1}$ were calculated from radial averaging of the 2D scattering patterns using macros written by the APS staff for IGOR Pro (WaveMetrics, Inc). Scattering profiles from protein plus buffer and from buffer alone were scaled using incident flux values integrated over the exposure time. Protein scattering profiles were then obtained by subtracting the average of five buffer profiles from the average of five profiles of buffer plus protein. The PRIMUS program[35] was used to merge SAXS and WAXS scattering data.

Crystallography

Lyophilized powder of the NtrC1 R-L_1 fragment (residues 2–133 plus C-terminal His-tag KLAAAL EHHHHHH) was dissolved in 5% (v/v) glycerol, 5 mM MgCl, 5 mM acetyl phosphate, and 10 mM TCEP (pH 7.75) to a final protein concentration of 35 mg/ml. To activate the receiver domain, the protein solution was incubated at 37 °C for 1 h. Crystals were grown at 4 °C using the hanging-drop technique. An equal volume of the protein solution was mixed with the reservoir solution (8.5–10% (v/v) PEG-3350, 350 mM NH_4Cl, 5% glycerol; 125 mM citric acid, pH 5.5) and left at 4 °C for about 1 h. The mixture was centrifuged for 5 min, and 2–5 μl drops were placed on siliconized cover-slips. Hexagonal rod-like crystals appeared after three days. Crystals were incubated overnight sequentially in reservoir solution plus 10, 15, 20, 25, and 30% glycerol, then flash-frozen in liquid nitrogen. These crystals had space group $P3_121$ with unit cell dimensions $a=b=91.52$ Å, $c=130.94$ Å; $\alpha=90°$, $\beta=90°$, $\gamma=120°$ and two molecules in the asymmetric unit. Data were collected at beam line 8.3.1 of the Advanced Light Source (1.0° rotation per image 1.1 Å wavelength). Data were processed and scaled with MOSFLM[36] and the CCP4 program suite[37] using the ELVES interface.[38] Phases were initially obtained by molecular replacement with CNS[39] using the first 124 amino acid residues of the inactive NtrC1 R-L_1-C monomer (PDB accession number 1NY5). Refinement was performed with CNS, using 2-fold non-crystallographic symmetry in the initial stages. Electron density for the linker domain L_1 (amino acid residues 124–133) and first eight residues of the His-tag (modeled as a poly-alanine backbone) was visible in only one monomer. This region was modeled in O,[40] and the aspartyl phosphate residues on both protomers were built with CNS.

Sequence analyses

Secondary structures and coiled coils were predicted using the programs PredictProtein,[41] Coils,[42] PerCoil,[43] and MultiCoil.[44] The σ^{54}-dependent activators listed in the Pfam database[45] at the time of the study were manually sorted to identify the signal-surface motif TGXGXHydX_3HydX_2Hyd (X=any amino acid, Hyd=hydrophobic), or relaxed versions of it in which one of the two glycine residues was allowed to be alanine. Pair-wise sequence alignments were created using BLAST[46] as implemented at the NCBI website†. Promoter searches were conducted using SEQSCAN, as implemented in PROMSCAN‡.

Protein Data Bank accession codes

Coordinates have been deposited in the RCSB Protein Data Bank, accession code 1ZY2.

Acknowledgements

We thank Dr Seok-Yong Lee for help with crystallography, and Professor Sydney Kustu for continued encouragement and advice. This work was supported, in part, by grants from the National Science Foundation (grant MCB-9727745 to B.T.N.), the Huck Institute of the Pennsylvania State University (grant INNOV-TN to B.T.N.), the DOE (Office of Biological and Health Effects Research and subcontract no. 6712513 from the Lawrence Berkeley National Laboratory to B.T.N.), the National Institutes of Health (grant GM62163 to D.E.W.), and by a University of California Fellowship (M.D.). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-ENG-38. BioCAT is a National Institutes of Health-supported Research Center, grant RR-08630. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Science Division, of the U.S. Department of Energy under contract no. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

† http://www.ncbi.nlm.nih.gov
‡ http://www.promscan.uklinux.net/

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.08.003

References

1. Morin, A., Huysveld, N., Braun, F., Dimova, D., Sakanyan, V. & Charlier, D. (2003). Hyperthermophilic Thermotoga arginine repressor binding to full-length cognate and heterologous arginine operators and to half-site targets. *J. Mol. Biol.* **332**, 537–553.
2. Lee, S.J., Engelmann, A., Horlacher, R., Qu, Q., Vierke, G., Hebbeln, C. *et al*. (2003). TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Biol. Chem.* **278**, 983–990.
3. Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M. & Geiduschek, E.P. (2003). Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc. Natl Acad. Sci. USA*, **100**, 5097–5102.
4. Lee, S.Y., DeLaTorre, A., Yan, D., Kustu, S., Nixon, B.T. & Wemmer, D.E. (2003). Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev.* **17**, 2552–2563.
5. Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E. *et al*. (1998). The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature*, **392**, 353–358.
6. Park, S., Meyer, M., Jones, A. D., Yennawar, H., Yennawar, N. & Nixon, B. T. (2002). 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. Epub.
7. Gu, B., Lee, J.H., Hoover, T.R., Scholl, D. & Nixon, B.T. (1994). *Rhizobium meliloti* DctD, a σ54-dependent transcriptional activator, may be negatively controlled by a subdomain in the C-terminal end of its two-component receiver module. *Mol. Microbiol.* **13**, 51–61.
8. Cho, H., Wang, W., Kim, R., Yokota, H., Damo, S., Kim, S.H. *et al*. (2001). BeF(3)(−) acts as a phosphate analog in proteins phosphorylated on aspartate: structure of a BeF(3)(−) complex with phosphoserine phosphatase. *Proc. Natl Acad. Sci. USA*, **98**, 8525–8530.
9. Yan, D., Cho, H.S., Hastings, C.A., Igo, M.M., Lee, S.Y., Pelton, J.G. *et al*. (1999). Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators. *Proc. Natl Acad. Sci. USA*, **96**, 14789–14794.
10. Meyer, M.G., Park, S., Zeringue, L., Staley, M., McKinstry, M., Kaufman, R.I. *et al*. (2001). A dimeric two-component receiver domain inhibits the sigma54-dependent ATPase in DctD. *FASEB J*. 15: 1326–1328. Epub.
11. Lee, S.Y., Cho, H.S., Pelton, J.G., Yan, D., Berry, E. & Wemmer, D.E. (2001). Crystal structure of activated CheY: comparison with other activated receiver domains. *J. Biol. Chem.* **276**, 16425–16431.
12. Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P. *et al*. (1999). Conformational changes induced by phosphorylation of the FixJ receiver domain. *Struct. Fold. Des.* **7**, 1505–1515.
13. Sojda, J.I., Gu, B., Lee, J., Hoover, T.R. & Nixon, B.T. (1999). A rhizobial homolog of IHF stimulates transcription of dctA in *Rhizobium leguminosarum* but not in *Sinorhizobium meliloti*. *Gene*, **238**, 489–500.
14. Bantscheff, M., Weiss, V. & Glocker, M.O. (1999). Identification of linker regions and domain borders of the transcription activator protein NtrC from *Escherichia coli* by limited proteolysis, in-gel digestion, and mass spectrometry. *Biochemistry*, **38**, 11012–11020.
15. Morett, E. & Bork, P. (1998). Evolution of new protein function: recombinational enhancer Fis originated by horizontal gene transfer from the transcriptional regulator NtrC. *FEBS Letters*, **433**, 108–112.
16. Pelton, J.G., Kustu, S. & Wemmer, D.E. (1999). Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. *J. Mol. Biol.* **292**, 1095–1110.
17. Lee, J.H., Scholl, D., Nixon, B.T. & 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.
18. Xu, H., Gu, B., Nixon, B.T. & Hoover, T.R. (2004). Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a σ54-dependent transcriptional activator. *J. Bacteriol.* **186**, 3499–3507.
19. Widdick, D., Farez-Vidal, E., Austin, S. & Dixon, R. (1998). Properties of a mutant form of the prokaryotic enhancer binding protein, NtrC, which hydrolyses ATP in the absence of effectors. *FEBS Letters*, **437**, 70–74.
20. North, A.K., Weiss, D.S., Suzuki, H., Flashner, Y. & Kustu, S. (1996). Repressor forms of the enhancer-binding protein NtrC: some fail in coupling ATP hydrolysis to open complex formation by sigma54-holoenzyme. *J. Mol. Biol.* **260**, 317–331.
21. Hastings, C.A., Lee, S.Y., Cho, H.S., Yan, D., Kustu, S. & Wemmer, D.E. (2003). High resolution solution structure of the beryllofluoride-activated NtrC receiver domain. *Biochemistry*, **42**, 9081–9090.
22. Lewis, R.J., Brannigan, J.A., Muchova, K., Barak, I. & Wilkinson, A.J. (1999). Phosphorylated aspartate in the structure of a response regulator. *J. Mol. Biol.* **294**, 9–15.
23. Reitzer, L.J., Movsas, B. & Magasanik, B. (1989). Activation of *glnA* transcription by nitrogen regulator I (NRI)-phosphate in *Escherichia coli*: evidence for a long-range physical interaction between NRI-phosphate and RNA polymerase. *J. Bacteriol.* **171**, 5512–5522.
24. Shingler, V. (1996). Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism. *Mol. Microbiol.* **19**, 409–416.
25. Garmendia, J. & 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.
26. O'Neill, E., Wikstrom, P. & Shingler, V. (2001). An active role for a structured B-linker in effector control of the sigma54-dependent regulator DmpR. *EMBO J.* **20**, 819–827.
27. Lee, J., Owens, J.T., Hwang, I., Meares, C. & Kustu, S. (2000). Phosphorylation-induced signal propagation in the response regulator NtrC. *J. Bacteriol.* **182**, 5188–5195.
28. Hwang, I., Thorgeirsson, T., Lee, J., Kustu, S. & Shin, Y.K. (1999). Physical evidence for a phosphorylation-

dependent conformational change in the enhancer-binding protein NtrC. *Proc. Natl Acad. Sci. USA*, **96**, 4880–4885.
29. Scholl, D. & Nixon, B.T. (1996). Cooperative binding of DctD to the *dctA* UAS of *Rhizobium meliloti* is enhanced in a constitutively active truncated mutant. *J. Biol. Chem.* **271**, 26435–26442.
30. Ledebur, H. & Nixon, B.T. (1992). Tandem DctD binding sites of the *Rhizobium meliloti dctA* UAS are essential for optimal function despite a 50 to 100-fold difference in affinity for DctD. *Mol. Microbiol.* **6**, 3479–3492.
31. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multi-dimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
32. Fischetti, R.F., Stepanov, F., Rosenbaum, G., Barrea, R., Black, E., Gore, D. *et al*. (2004). The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the Advanced Photon Source. *J. Synchrotron Radiat.* **11**, 399–405.
33. Phillips, W.C., Stewart, A., Stanton, M., Naday, I. & Ingersoll, C. (2002). High-sensitivity CCD-based X-ray detector. *J. Synchrotron Radiat.* **9**, 36–43.
34. Fischetti, R.F., Rodi, D.J., Mirza, A., Irving, T.C., Kondrashkina, E. & Makowski, L. (2003). High-resolution wide-angle X-ray scattering of protein solutions: effect of beam dose on protein integrity. *J. Synchrotron Radiat.* **10**, 398–404.
35. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H. & Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallog.* **36**, 1277–1282.
36. Powell, H.R. (1999). The Rossmann Fourier auto-indexing algorithm in MOSFLM. *Acta Crystallog. sect. D*, **55**, 1690–1695.
37. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
38. Holton, J. & Alber, T. (2004). Automated protein crystal structure determination using ELVES. *Proc. Natl Acad. Sci. USA*, **101**, 1537–1542.
39. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W. *et al*. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
40. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
41. Rost, B. (1996). PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.* **266**, 525–539.
42. Lupas, A., Van Dyke, M. & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
43. Berger, B., Wilson, D.B., Wolf, E., Tonchev, T., Milla, M. & Kim, P.S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl Acad. Sci. USA*, **92**, 8259–8263.
44. Wolf, E., Kim, P.S. & Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6**, 1179–11789.
45. Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S. *et al*. (2004). The Pfam protein families database. *Nucl. Acids Res.* **32**, D138–D141.
46. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Edited by K. Morikawa

(Received 13 June 2005; received in revised form 1 August 2005; accepted 3 August 2005)
Available online 31 August 2005