The Pennsylvania State University

The Graduate School

Department of Biochemistry and Molecular Biology

STRUCTURAL STUDIES ON *SINORHIZOBIUM MELILOTI* DCTD RELATED TO ATP BINDING AND ACTIVATION

A Thesis in

Biochemistry, Microbiology, and Molecular Biology

by

Sungdae Park

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2002

We approve the thesis of Sungdae Park

Date of Signature

B. Tracy Nixon Associate Professor of Biochemistry and Molecular Biology Thesis Advisor Chair of Committee	
Richard J. Frisque Professor of Molecular Virology	
C.-P. David Tu Professor of Biochemistry and Molecular Biology	
Frank Pugh Associate Professor of Biochemistry and Molecular Biology	
A. Daniel Jones Senior Scientist	
Robert A. Schlegel Professor of Biochemistry and Molecular Biology Head of the Department of Biochemistry And Molecular Biology	

ABSTRACT

Sinorhizobium meliloti DctD activates transcription at the *dctA* promoter by catalyzing the transformation of closed complex of $E\sigma^{54}$ (RNA polymerase/sigma-54) into open ones in response to the phosphorylation status of DctB, a C_4-dicarboxylate sensor. DctD is σ^{54}-dependent transcriptional activator, and consists of three functional domains: an N-terminal, two-component receiver domain; a central AAA^+ ATPase domain; and a C-terminal DNA binding domain. Since the DctB and DctD proteins form a two-component regulatory system, it is easy to hypothesize a general outline of signal transduction for these proteins. But detailed molecular mechanisms underlying two-component signal transduction have only recently begun to emerge from research in several labs.

As a step toward understanding molecular events involved in the DctB/DctD signaling system in *Sinorhizobium meliloti*, I have investigated its several aspects. Tryptic digestion experiments of DctD with or without ATP showed that unphosphorylated 44 kD has the ability to bind ATP, and, upon ATP binding, undergoes conformational changes, which is registered in changes in the pattern of tryptic digestion. By modeling the decay of the 44 kD peptide at varying amounts of ATP as a combination of two first order processes; one with and one without ATP being bound, it was possible to estimate an equilibrium dissociation constant of approximately 0.7 mM for ATP. Subsequent pre-steady state kinetic studies of mant-ATP binding to DctD gave the rough estimation of an equilibrium dissociation constant of 0.87 mM. The results of both

experiments suggest that the receiver domain does not play its role simply by preventing the AAA^+ ATPase domain from binding ATP.

Crystallographic studies on the receiver domains of DctD and its variant, E121K, revealed another important feature of this signaling system. The crystal structure of E121KNL obtained with or without Mg^{2+} and BeF_3^{-1} showed that, in the active state mimicked by Mg^{2+}- BeF_3^{-1}, E121KNL has changes induced in the active site similar to those seen in FixJN~P, Spo0AN~P and CheY-BeF_3^{-1}, and assumes a dimer interface dramatically different from that of the off-state E121KNL.

Along with the results of previous crystallographic and genetic studies, the following 'working model' for signal transduction in DctD can be proposed. In the 'off-state', the receiver domain and coiled-coil linker form a dimer that inhibits oligomerization of the AAA^+ ATPase domain. In this conformation the receiver domain cannot be phosphorylated or bind Mg^{2+} and BeF_3^{-1}. The binding of Mg^{2+} and BeF_3^{-1} stabilizes an alternative dimeric conformation in which the α4-β5-α5 interface is replaced with an α4-β5 interface, thereby repositioning the phosphorylation site (Asp55) in the subunits by ~20 Å. Reorienting the receiver domains relieves inhibition of the AAA^+ ATPase domain by at least allowing or even stimulating it to oligomerize and activate transcription at the *dctA* promoter. Since ~4.5 % of two-component receiver domains are predicted to have coiled-coil linkers found in DctD, switching between alternate dimeric states as seen in DctD may be a common mechanism used by a subgroup of two-component regulatory systems.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

- **AAA** – ATPases associated with a variety of cellular activities
- **AAA^+** – superfamily of AAA
- **ADP** – adenosine 5′-diphosphate
- **ATP** – adenosine 5′-triphosphate
- **BeF** – beryllium fluoride
- **BSA** – bovine serum albumin
- **CAPS** – 3-[cyclohexylamino]-1-propanesulfonic acid
- **CCP4** – Collaborative Computational Project 4
- **CD** – circular dichroism
- **CheY~P** – phosphorylated CheY
- **$DctD_{2\text{-}162}$** – DctD receiver domain plus adjacent linker
- **$DctD_{2\text{-}384}$** – DctD receiver domain plus AAA^+ ATPase domain
- **DctDNL** – DctD receiver domain plus linker
- **DEAE-cellulose** – diethylaminoethyl-cellulose
- **DNA** – deoxyribonucleic acid
- **$E\sigma^{54}$** – sigma 54 plus the core enzyme complex of RNA polymerase
- **E121KNL** – DctD receiver domain plus linker that has E121K substitution
- **E121KNL-BeF_3^{-1}** – E121K receiver domain plus linker complexed with beryllium trifluoride
- **EBPs** – enhancer binding proteins
- **EDTA** – ethylenediamine-N,N,N′,N′-tetraacetic acid
- **Fe-BABE –** ((S)-1-[*p*-(bromoacetamido)benzyl]-EDTA-Fe)
- **FixJN** – FixJ receiver domain without linker
- **FixJN~P** – phosphorylated FixJ receiver domain without linker
- **HEPES** – N-(2-Hydroxyethyl)piperazine-N′-2-ethanesulfonic acid
- **HSQC** – heteronuclear single quantum coherence
- **IHF** – integration host factor

- **IPTG** – isopropyl-β-D-thiogalactopyranoside
- **K122ENL** – DctD receiver domain plus linker that has K122E substitution
- **K_d** – equilibrium dissociation constant
- **Mant-ATP** – 2′(3′)-O-(N-methylanthraniloyl)-ATP
- **NCS** – noncrystallographic symmetry
- **NMR** – nuclear magnetic resonance
- **NTA** – nitrilo-tri-acetic acid
- **NtrCN** – NtrC receiver domain without linker
- **NtrCN~P** – phosphorylated NtrC receiver domain without linker
- **ONPG** – o-nitrophenyl-β-D-galactopyranoside ADP
- **PCR** – polymerase chain reaction
- **PEG** – polyethylene glycol
- **PVDF** – polyvinylidene fluoride
- **RMSD** – root mean square deviation
- **SDS-PAGE** – sodium dodecyl sulfate polyacrylamide gel electrophoresis
- **Spo0AN** – Spo0A receiver domain without linker
- **Spo0AN~P** – phosphorylated Spo0A receiver domain without linker
- **TCA** – tricarboxylic acid
- **TCST** – two-component signal transduction
- **Tris** – Tris(hydroxymethyl)aminomethane
- **UAS** – upstream activation sequence

Acknowledgments

I first would like to cordially express my gratitude to my advisor, Dr. B. Tracy Nixon, for his guidance, patience, and encouragement. I am honored to have had the opportunity to work with him. I would also like to thank all my other committee members for taking time to serve on the committee and their valuable comments on my thesis. I would also like to thank all my family for their unending support and prayer throughout all my life. And I must thank my wife, Hee Jung, who has always stood by my side to comfort and encourage me.

Above all, I must give all credits to my Lord, Jesus Christ, who has orchestrated everything behind all.

INTRODUCTION

All living organisms require nitrogen as an indispensable element, most of which is bound up in amino acids and nucleotides. Most living organisms are capable of assimilating ammonia into the amido group of glutamine, which is the direct source of nitrogen for the building blocks of macromolecules and other biological compounds. Since no other pathway for the formation of glutamine other than the glutamine synthetase reaction exists, ammonia can be considered to be an essential metabolite; if an environmental source of ammonia is lacking, ammonia must be derived by the catabolism of nitrogenous compounds such as amino acids or urea or from the reduction of atmospheric dinitrogen (Stock et al., 1989). The ultimate reservoir of nitrogen for all living organisms is present in air, 78% of which is molecular nitrogen. However, only relatively few species, all of which are prokaryotes, can convert atmospheric nitrogen into ammonia through the process called biological nitrogen fixation.

Nitrogen fixation is observed in a large number of species belonging to both prokaryotic kingdoms, *Archaebacteria* and *Eubacteria*. As well as free-living diazotrophs(e.g., *Klebsiella*, *Azotobacter*, and *Rhodobacter*), many bacteria living in symbiotic associations with their host plant(s) have the ability to fix dinitrogen. The latter group of symbiotic diazotrophs, collectively known as rhizobia, includes species of three rhizobial genera *Rhizobium*, *Bradyrhizobum*, and *Azorhizobium* (Fisher, 1994).

Rhizobia are Gram-negative rods that exist either as free-living soil bacteria or as endosymbionts called bacteroids in the nodules of a variety of legumes as well as nonlegumes. Under conditions of nitrogen limitation, rhizobia stimulate their host

plant(s) to develop nodules, which the bacteria infect and inhabit. Though symbiosis is a species-specific process as suggested in Table 1, the degree of host specificity varies greatly among the rhizobia. Ultimately, the two organisms establish a metabolic cooperation which is mutually beneficial: the plants reduce carbon dioxide into sugars during photosynthesis and supply their catabolic intermediates to the bacteroids to be used as the energy source for nitrogen fixation, and, in return, the bacteroids provide the plant with reduced nitrogen. The overall process is summarized briefly in Figure 1. The symbiosis between these two organisms is exploited in agriculture to enrich the soil nitrogen as well as to provide plants for grazing. Since the expense of producing ammonia industrially for use in fertilizers increases with the cost of energy supplies, nitrogen fixation has been the subject of intense study.

1. Plants-Rhizobia interaction

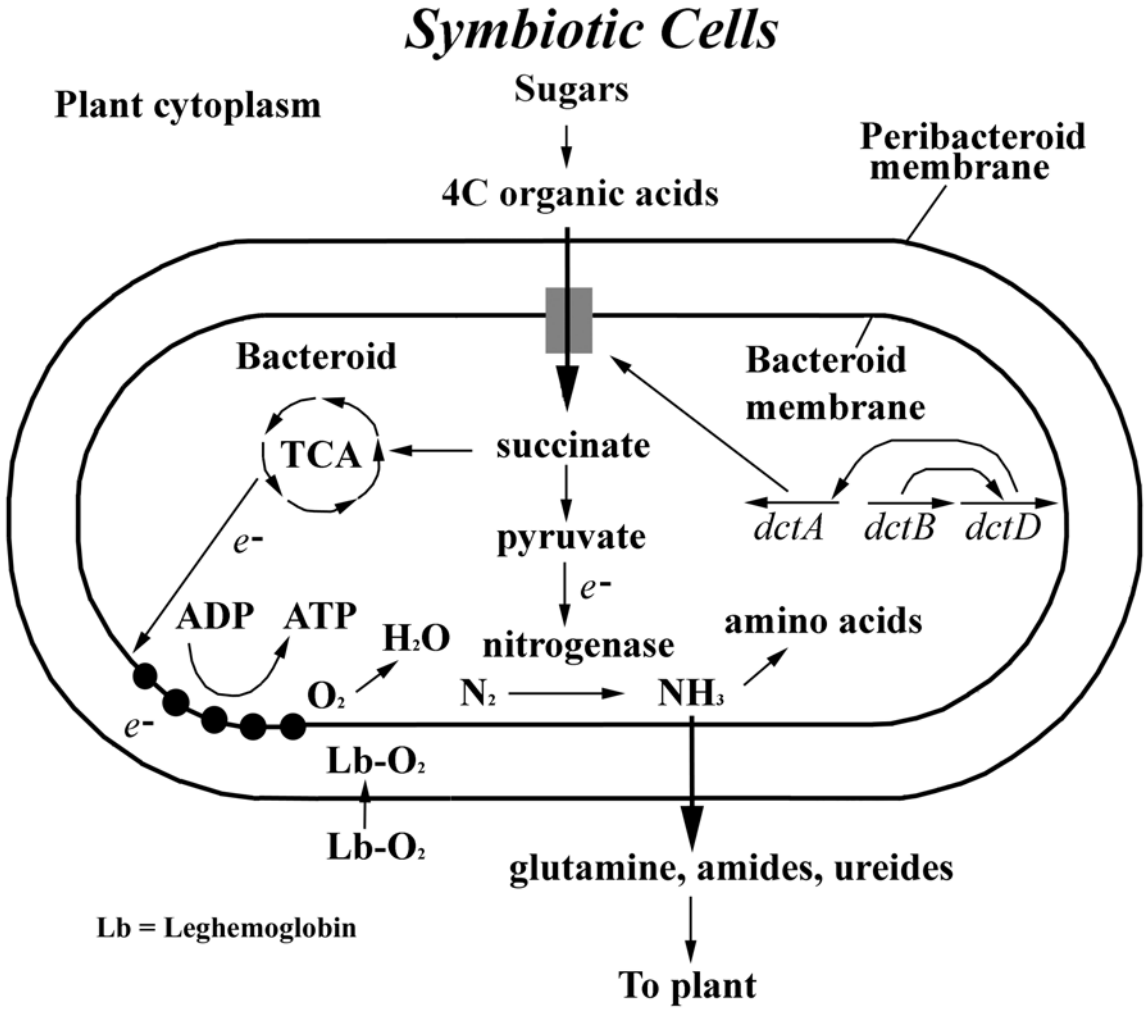
Root infection by rhizobia is a multistep process that is initiated by a molecular dialog in a zone of soil surrounding the roots of plants called the rhizosphere. Rhizobia respond by positive chemotaxis to plant root exudates such as amino acids, dicarboxylic acids, and flavonoids and move toward localized sites on the roots of plants, mainly leguminous plants (Rhijn et al., 1995). Flavonoids or isoflavonoids secreted by the host plants induce the expression of a number of nodulation (nod) genes in the cognate rhizobial bacteria. The products of nod genes are enzymes involved in the biosynthesis of species specific, substituted lipooligosaccharides called Nod factors (Fisher, 1994).

Table 1. **Rhizobium-plant associations**

Rhizobium	Host plant(s)
Sinorhizobium meliloti	Alfalfa
Rhizobium leguminosarum	
biovar viciae	Pea, vetch
biovar trifolli	Clover
biovar phaseoli	Bean
Rhizobium fredii	Soybean
Bradyrhizobium japonicum	Soybean
Rhizobium loti	Lotus
Azorhizobium caulinodans	Sesbania (stem nodulating)
Rhizobium NGR234	Tropical legumes, parasponia (non-legume)
Bradyrhizobium spp.	Parasponia (non-legume)

Figure 1. Schematic diagram of the major metabolic reactions and nutrient exchanges in the bacteroid.

Being enclosed within a peribacteroid membrane derived from the host cytoplasmic membrane, bacteroids are totally dependent on the plant for energy sources for N_2 fixation. The C_4-dicarboxylates are transported into the bacteroid via DctA, and incorporated into the TCA cycle as the C_4-dicarboxylates such as succinate, malate, and fumarate. These serve as electron donors for ATP synthesis and the ultimate source of electrons for the reduction of N_2. N_2 is reduced into NH_3 by O_2-sensitive nitrogenase in a microaerobic environment maintained by leghemoglobin. Ammonia, the first stable product of nitrogen fixation, is transported into the plant, and then assimilated into glutamine, amino acid amides, and urea derivatives called ureides. The figure is excerpted and modified from *Biology of Microorganisms*, 6^{th} edition (Brock & Madigan).


Symbiotic Cells
Plant cytoplasm
Sugars
4C organic acids
Peribacteroid membrane
Bacteroid
Bacteroid membrane
TCA
succinate
pyruvate
dctA
dctB
dctD
e-
e-
ADP
ATP
H2O
O2
nitrogenase
amino acids
N2
NH3
e-
Lb-O2
Lb-O2
glutamine, amides, ureides
Lb = Leghemoglobin
To plant

The expression of nod genes is regulated through the activity of NodD, which is a DNA binding transcriptional activator associated with the bacterial cytoplasmic membrane. NodD does this by binding to the promoters of inducible nod genes, which are characterized by a long (about 50 base pairs), highly conserved sequence called the "nod box". This "nod box" is located upstream of *nodABC*, *nodEFG* and *nodH* gene clusters (Fisher et al., 1989). Structural nod genes are classified as either common nod genes or host specific nod genes. Common nod genes encode proteins that are thought to be responsible for producing the common features of different lipooligosaccharides produced by various rhizobia. The host specific nod genes appear to be responsible for attaching identifying features to these lipooligosaccharides (Rhijn et al, 1995). These signal compounds, which are produced by induced rhizobial cells, elicit the curling of corresponding plant root hairs and division of meristematic cells, eventually leading to the formation of root nodules. Rhizobia, which are attracted to host plants by chemotaxis, attach to root hairs and start to infect plant tissue inside a host-derived infection thread that progressively penetrates into the root cortex. Subsequently, bacterial cells are released into plant cells, where they further divide and differentiate physiologically and also morphologically into bacteroids that are competent for nitrogen fixation (Fisher, 1994).

2. Nitrogen Fixation

The reduction of nitrogen to ammonia is an exergonic reaction:

$N_2 + 3H_2 \rightarrow 2NH_3 \qquad \Delta G^{o'} = -33.5$ kJ/mol

The N≡N triple bond, however, is very stable with a bond energy of 942 kJ/mol. Therefore nitrogen fixation has an extremely high activation energy (Lehninger et al., 1993). In view of the high energy demands of nitrogen fixation (up to 40 mol of ATP may be required *in vivo* for the reduction of 1 mol of dinitrogen to ammonia (Hill, 1992).) and marked oxygen sensitivity of nitrogenase, it makes perfect sense that all diazotrophs studied so far have evolved sophisticated regulatory mechanisms that allow a tight control of nitrogenase synthesis. The expression of nitrogen fixation genes is controlled by a series of regulatory genes organized in a hierarchical manner. Their concerted action enables the bacteria to sense optimal environmental conditions required for nitrogen fixation and to transmit this information to the level of gene expression (Fisher, 1994).

Symbiotic nitrogen fixation genes can be divided into *nod*, *nif*, and *fix* genes. The nod gene products are required for the early steps in nodule formation. The fix genes encode proteins that do not have homologous counterparts in free-living diazotrophs while the *nif* genes are structurally homologous to free-living diazotrophs. Among the 9 *nif* genes identified so far in rhizobial species, NifA encodes a transcriptional activator required for the transcription of other *nif* genes and some fix genes. The *nifHDK* genes encode the nitrogenase structural proteins: *nifH* for the Fe protein subunit, *nifD* for the MoFe protein α subunit, and *nifK* for the MoFe protein β subunit, respectively.

The products of the *nifE*, *nifN*, and *nifB* genes are responsible for the synthesis of the FeMo cofactor; however, the exact biochemical functions of the respective proteins are not known. The fix genes include *fixABCX*, which appear to be involved in electron transport to dinitrogenase. In free-living diazotrophs such as *K. pneumoniae*, environmental nitrogen and oxygen conditions are critical signals for the regulation. The transcription of *nifA* is regulated by NtrC, a σ^{54}-dependent transcriptional activator that is involved in the general nitrogen regulation of the cell. NtrB and NtrC are one of two component regulatory systems that respond to low intracellular nitrogen levels. They carry out this by activating transcription of genes for nitrogen fixation and nitrogen assimilation, *nifA* and *glnA*, respectively. NifA is then capable of autogenous activation of the *nifLA* promoter. The product of *nifL* prevents *nif*A-mediated activation of *nifLA* and other *nif* genes in the presence of intermediate levels of ammonia or too much oxygen, which will inactivate nitrogenase irreversibly. On the other hand, expression of nitrogen fixation genes in symbiotic bacteroids is regulated predominantly in response to the cellular oxygen levels. The lack of a strict nitrogen control mechanism probably relates to the physiological conditions under which symbiotic nitrogen fixation takes place. In the bacteroids of *Sinorhizobium meliloti*, the *nifA* gene is not activated by NtrC, and no *nifL* homolog has been found. Rather, the *nifA* gene is regulated by oxygen tension through the products of *fixLJ*, which form a two component regulatory system: FixL, a hemeprotein, senses the local oxygen level by altering the phosphorylation status of FixJ, a transcriptional activator. Control by oxygen occurs at two levels. Low oxygen conditions stimulate the autophosphorylation activity of FixL and repress the phosphatase activity of FixL-phosphate. (Fisher, 1994)

3. C_4-Dicarboxylate Transport System of Rhizobia

Requirement of C_4-Dicarboxylate Transport

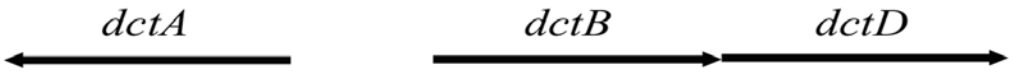
As described above, nitrogen fixation is a highly energy-consuming process due to high activation energy. In leguminous root nodules, the host plant supplies the bacteroids with the products of photosynthesis, which is in turn converted into ATP and reducing power to fuel nitrogen fixation and the maintenance of bacteroid functions. The role of ATP in this process appears to be catalytic rather than thermodynamic. Since nitrogen fixation is an exergonic reaction, it is thermodynamically favorable one: However, it is almost kinetically inert because of the high activation barrier. ATP is used to lower this high activation energy mainly by providing binding energy. In the reaction carried out by denitrogenase reductase, both ATP binding and ATP hydrolysis cause protein conformational changes for overcoming the high activation energy of nitrogen fixation. (Lehninger et al, 1993).

But how can bacteroids provide themselves with a large amount of energy for nitrogen fixation? Many studies have indicated that a C_4-dicarboxylate such as succinate, fumarate, and malate are the main energy sources provided by the host plant to meet this requirement of high energy for nitrogen fixation. Mutant Rhizobium strains that are lacking functional succinate dehydrogenase are unable to form effective nodules (Finan et al., 1981; Gardiol et al., 1982). Mutants defective in transporting succinate are incapable of fixing nitrogen (Ronson et al., 1981; Finan et al., 1981; and Yarosh et al., 1989). It was also shown that these mutants failed to grow on fumarate or malate. In

addition, revertants of these mutants recovered the ability to grow on all these three dicarboxylates, which suggests that these compounds are transported by a common system. This transport system is inducible both in free-living cells and in symbiotic bacteroids in that transport levels are increased in the presence of C_4-dicarboxylates while decreased in cells grown in glucose, arabinose, or sucrose (Ronson et al., 1981; Finan et al., 1981; and Yarosh et al., 1989).

Organization of *dct* Genes

The genes responsible for C_4-dicarboxylate transport were first identified from *Rhizobium leguminosarum* (Ronson et al., 1984). Four genes were found to be required and were designated *dctA*, *dct*B, *dctC*, and *dctD*. This was done by complementation of C_4-dicarboxylate transport mutants with a pLAFR1 cosmid library made from DNA of *R. leguminosarum*. The plasmid pLAFR1 is a broad host range vector that can replicate in rhizobia as well as *E. coli*. Subsequent DNA sequence analyses showed that *dctC* and *dctD* actually represent a single gene, now called *dctD* (Ronson et al., 1987). The *dct* locus has been shown to be on pSymb, a 1.7 Mb plasmid that is one of two megaplasmids in *S. meliloti* (Honeycutt et al., 1993; Watson et al., 1988). Sequence analyses of the *dct* genes from *S. meliloti* and *R. leguminosarum* revealed that *dct* genes are arranged in the order *dctA*, *dctB*, and *dctD*, with *dctA* transcribed divergently from *dctB* and *dctD*. (Bolton et al., 1986; Engelke et al., 1987; Watson et al., 1988; Jiang et al., 1989; and Yarosh et al., 1989).



It was also found in this work that *dctB* and *dctD* form an operon in which the two cotranscribed genes are separated by 3 bp in *S. meliloti* and 4 bp in *R. leguminosarum*.

The transcription start sites of both species were determined by primer extension analyses (Ledebur et al., 1990). It shows that *dctA* transcription starts 10 to 11 bp downstream of a σ^{54}promoter. However, the transcription start site for *dctBD* has not been determined. Additional analysis showed that the product of *ntr*A (also called *rpoN*), a fourth, unlinked gene, is also required for a functional C_4-dicarboxylate transport system. The *ntrA* encodes the alternative sigma factor, called σ^{54} or σ^{N}, which is required for transcription of genes whose products play divergent physiological roles (Ronson et al., 1987b). Subsequent sequence analyses have shown that the *dctA* promoter has the -24 and -10 regions typical of σ^{54}-dependent promoter, instead of the more common -35 and -10 boxes (Ronson and Astwood, 1985; Jiang et al., 1989).

The Eσ^{54} system has several other unique features that are uncommon among bacterial RNA polymerases. Upon binding, Eσ^{54} forms a stable and closed promoter complex. To be converted into a transcriptionally active open promoter complex, Eσ^{54} is absolutely dependent upon an activator protein bound at an upstream activation sequence (UAS), enhancer-like elements in prokaryotes. Eσ^{54} can be regarded as a "defective" holoenzyme in that it can initiate transcription only when in concert with a second trans-acting transcriptional regulator (Kutsu et al., 1991; Morett et al., 1993). The unique properties of σ^{54} are also evident in its primary structure. This protein is the only sigma factor that does not have sequence similarity to any of the other known prokaryotic sigma factors. It has several functional motifs that are normally found in eukaryotic transcriptional factors such as a leucine zipper, an acidic region, and a glutamine-rich region (Sasse-Dwight et al., 1990). Another feature of σ^{54} is its ability to interact with the promoter DNA in the absence of core RNA polymerase (Buck et al., 1992). This

leads to the suggestion that *dctA* transcription requires the σ^{54}-holoenzyme system ($E\sigma^{54}$), an upstream activating sequence, and its cognate transcriptional activator.

An upstream activation sequence was found to be required for optimal transcription of *dctA* (Ledebur et al., 1990) This sequence consists of two tandem dyad axes of symmetry, or inverted repeats, that are located 100 base pairs upstream of the *dctA* transcription start site. These sites are designated the A site (promoter distal) and the B site (promoter proximal). They are separated by 33 base pairs, which is about three helical turns of the duplex DNA apart (Jiang et al., 1989). The A site is much less symmetric than the B site.

```
-156          Site A                              Site B       -101
  CATGGTGCATGTTTTCGCCCAGGACGCCAGCACTTCTGTGCGGAAATCCGCACATA
      - ->         <- -                 - ------->  <------- -
```

From quantitative DNase I footprint experiments, Scholl showed that a dimer of DctD binds these two sites cooperatively and has ~20 fold higher affinity for the B site than for the A site (Scholl & Nixon, 1996).

Regulation of *dctA* Gene Expression

Most bacteria live in environments that frequently change. Accordingly, they have various adaptive responses ranging from rapid transient changes in mobility to long-term global reorganizations of cell morphology. Signals from within the cell and from the external environment are used to control the cellular activities that mediate a particular type of response.

Such regulations often involve two types of proteins: a histidine protein kinase, and its associated response regulator. Because of the central importance of these two components, these signal transduction systems have been designated “two-component”

systems. But this might be misleading because these systems generally contain additional signal transduction components and in some cases more than one kinase or response regulator or both (Stock et al., 1989). The two-component signal transduction is ubiquitous in bacteria, mediating a number of responses to environmental changes. In the spring of 2000, there were 1148 known two-component response regulators classified into 37 groups based on the associated output functions. Among them, DctB and DctD belong to a group that regulates functional states of σ^{54}-dependent transcriptional activators.

Control of transcription of σ^{54}-dependent promoters appears to be accomplished mainly via modulation of the activity of activator proteins. Each σ^{54}-dependent activator allows σ^{54}-holoenzyme to initiate transcription in response to a distinct physiological signal such as limitation of combined nitrogen (NtrC), low oxygen tension (NifA) and availability of dicarboxylic acids external to the cell (DctD).

The DctB protein deduced from *dctB* sequence of *Sinorhizobium meliloti* contains 622 amino acid residues, along with two transmembrane spanning regions on the basis of its amino acid sequence. The highly conserved C-terminus shows homology to the sensor class of the two-component system such as CheA, EnvZ, FixL, NtrB, PhoR, and VirA. All these proteins share a conserved COOH-terminal region that consists of more than 200 amino acids. DctB is able to phosphorylate DctD in vitro (Giblin et al., 1995). In addition, the truncated form of DctB, which has a putative C-terminus histidine kinase domain only, shows autokinase activity as well as the ability to phosphorylating DctD in vitro (Meyer et al., 2001).

Sequence analysis of *dctD* suggested that DctD consists of three functional domains: a N-terminal two-component receiver domain; a central ATPase domain; and a C-terminal DNA binding domain (Ronson et al., 1987a) (Fig. 2). The modular nature of DctD is supported by several lines of evidence: maintenance of transcriptional competence of the central domain (Huala et al., 1992); formation of functionally active chimeric protein (B. Gu, B. S. Wang, M Liang and B. T. Nixon, unpublished); and studies of a series of *Bal*31 deletion mutants of the N-terminal and C-terminal domains of DctD (Gu et al., 1994). The functional boundary between amino terminal and central domains is between residues 144 and 149, and the boundary between the central and carboxy terminal domains lies between residues 382 and 384 (Gu et al., 1994). The amino-terminal domain regulates the ATPase domain in response to the signal from its cognate sensor protein, DctB. Removal of the N-terminal regulatory domain of DctD makes DctD constitutively active (Gu et al., 1994). In addition, a sharp increase in the level of transcription after deletion from residue 106 up to residue 143 suggests that this small C-terminal portion of the N-terminus is required to keep the ATPase domain silent (Gu et al, 1994). The central domain is directly responsible for nucleotide hydrolysis and activation of transcription. The carboxy-terminal domain contains a helix-turn-helix DNA-binding motif, and confers upon DctD its specificity for DNA binding. Homology among σ^{54}-dependent transcriptional activators has been found in the central and C-terminal domains with highest in the central domain. Binding of DctD to the two sites in the *dctA* UAS is heterogeneous.

Figure 2. A) Domain structure of DctD of *S. meliloti*. The N-terminus of DctD is the receiver domain that has a putative site of phosphorylation by DctB. The central domain contains AAA^+ module and ATPase activity for transcriptional activation at the dctA promoter. The C-terminus contains a helix-turn-helix motif responsible for DNA binding at the UAS.

B) Amino acid sequence of DctD of *S. meliloti*. The three residues that are highly conserved in the receiver domains of response regulators are boxed. Residue D55 is the site for phosphorylation. Residues T83 and F102 are underlined, as these positions are always threonine or serine and an aromatic amino acid, respectively, that play crucial roles in signaling. The central ATPase domain has ATP binding motifs called Walker A (also called P-loop) and B. The carboxyl-terminal domain contains a helix-turn-helix motif characteristic of DNA binding proteins. Each motif is presented in bold letters underlined. A more detailed description of this domain was recently provided by Neuwald et al., who first described the AAA^+ superfamily of ATPases based on elaborate sequence analyses (Neuwald et al., 1999).

A

B

```
  1 MSAAPSVFLI DDDRDLRKAM QQTLELAGFT VSSFASATEA LAELSADFAG

 51 IVISDIRMPG MDGLALFGKV LALDPDLPMI LVTGHGDIPM AVQAIQDGAY
 Phosphorylation site
101 DFIAKPFAAD RLVQSARRAE EKRRLVMENR SLRRAAEAAS EGLPLIGQTP

151 AMERLRQTLK HIADTDVDVL VAGETGSGKE VVATLLHQWS RRRTGNFVAL
               ATP-binding motif (Walker A or P-loop)
201 NCGALPETVI ESELFGHEPG AFTGAVKKRI GRIEHASGGT LFLDEIEAMP
                                 ATP-binding motif (Walker B)
251 PATQVKMLRV LEAREITPLG TNLTRPVDIR VVAAAKVDLG DPAARGDFRE

301 DLYYRLNVVT LSIPPLRERR DDIPLLFSHF LARASERFGR EVPAISAAMR

351 AYLATHSWPG NVRELSHFAE RVALGVEGNL GVPAAAPASS GATLPERLER

401 YEADILKQAL TAHCGDVKET LQALGIPRKT FYDKLQRHGI NRADYVERAG
                          Helix-turn-helix
451 PGRPNAISKT
```

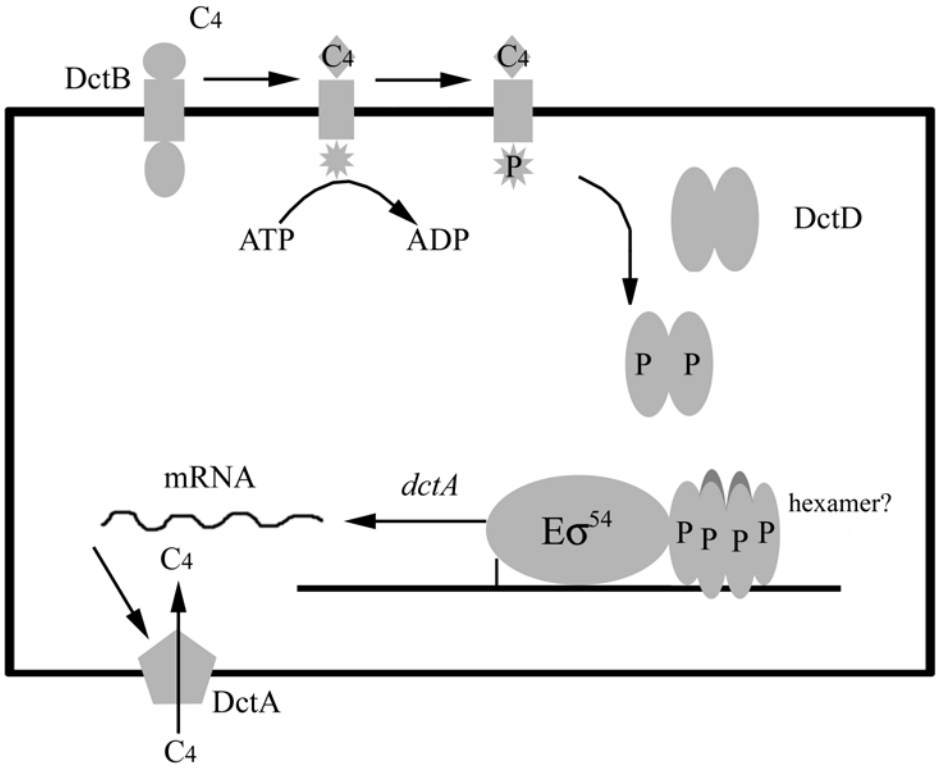
The intrinsic free energy of binding to the A site is smaller than that of binding to the B site. However, the presence of cooperativity makes occupancy isotherms of the two sites nearly identical (Scholl et al,, 1996).

Based upon these observations as well as comparisons with the NtrB/NtrC system, I propose a working model for DctB/DctD signal transduction involving a cascade of reactions responding to a particular environmental signal (Fig. 3). In its basal state, DctD binds ATP with subsequent conformational change. DctD binds cooperatively to two tandem sites of the *dct*A UAS. The DctD then interacts with σ^{54}-holoenzyme bound to its cognate promoter to form a transcriptionally inactive closed complex via DNA loop either with the help of integration host factor (IHF) or without it depending upon the bacterium being examined. When cells detect C_4-dicarboxylates, DctB autophosphorylates the histidine residue in its conserved C-terminal cytoplasmic domain. This phosphate is then transferred to aspartate at position 55 in the N-terminal receiver domain of DctD. Once phosphorylated, DctD undergoes conformational changes, involving the formation of an oligomer, probably a hexamer (Porter et al., 1993; Wyman et al., 1997; Neuwald et al., 1999). This causes the ATPase domain to be released from inhibition by the N-terminal domains. The ATPase then takes part in converting the closed complex of the $E\sigma^{54}$-promoter complex to open one so that cells can make more DctA protein, a transmembrane permease, to facilitate uptake of C_4-dicarboxylates. Although DctB is actually responsible for sensing the presence of C_4-dicarboxylates, studies of *dct*A::Tn*pho*A mutants (Yarosh et al., 1989) as well as *dct*A::*lac*Z fusions (Ronson et al., 1985) show that the *dct*A gene is expressed constitutively in $DctA^-$ cells in the absence of C_4-dicarboxylates.

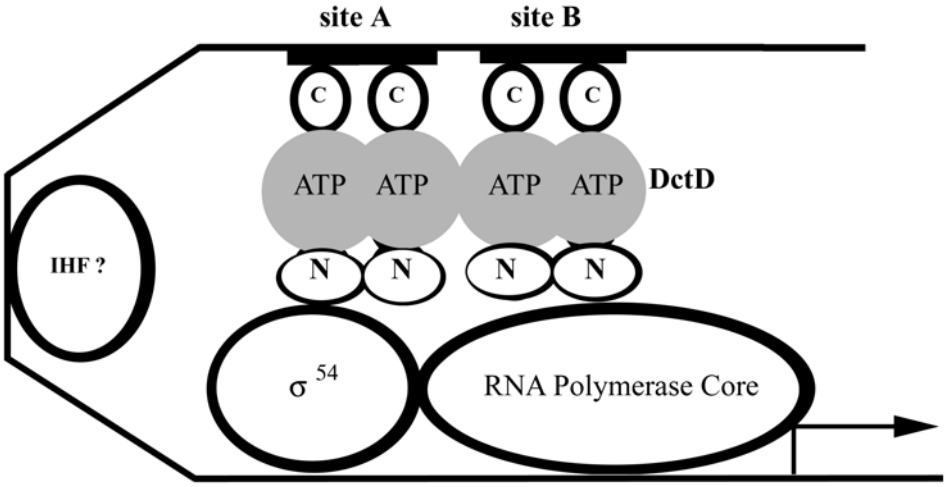
Figure 3. A) Schematic diagram for the C_4-dicarboxylate transport system. The dicarboxylic acid transport system consists of three proteins. DctB senses C_4-dicarboxylic acids and turns on the activity of its C-terminal histidine kinase domain. The phosphoryl group is then transferred to aspartyl residue 55 in the N-terminal receiver domain of DctD. Upon phosphorylation, DctD is converted to a transcriptionally active form, most likely a hexamer. This activated form causes the complex of $E\sigma^{54}$ and promoter to open and begin the transcription of *dctA*. The gene product of *dctA* is then used to transport C_4-dicarboxylic acids into the cell. This model is not intended to suggest specific regions of protein-protein interactions.

B) Model for transcriptional activation of *dctA*. DctD dimers bind to the two UAS sites located upstream of the *dctA* promoter. Each subunit has three domains. The amino-terminal, central ATPase and carboxyl-terminal domains of each subunit are labeled "N", "ATP" and "C", respectively. These proteins are shown to interact with $E\sigma^{54}$. IHF (integration host factor) may be required for this interaction, depending upon bacterial species.

A.


C4
DctB
C4
C4
P
ATP
ADP
DctD
P P
mRNA
dctA
Eσ54
P P P P
hexamer?
C4
DctA
C4

B.


site A
site B
C
C
C
C
ATP
ATP
ATP
ATP
DctD
IHF ?
N
N
N
N
σ 54
RNA Polymerase Core

This suggests that the DctA protein is also involved in controlling its own synthesis. More recently, Reid et al. showed that although DctB alone can detect C_4-dicarboxylates in the absence of DctA and initiate transcription via DctD, it had broader substrate specificity and cross talk with other operons. Therefore it appears that while DctB is a direct sensor for C_4-dicarboxylates, it is regulated by DctA for proper signal detection and signaling specificity. Though communications within and between proteins are involved in this signal transduction, relatively little is known about the detailed molecular mechanisms of how the cells receive environmental stimuli and then initiate signal transduction into activation of transcription at the *dctA* promoter.

4. Literature reviews of structural studies on two component receiver domains

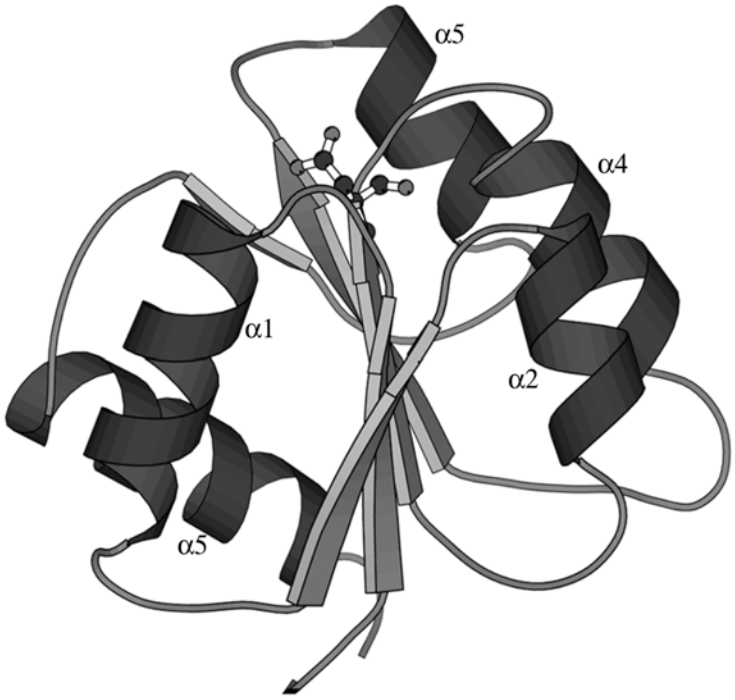
Structures of 'inactive forms' of two component receiver domains

Until recently, receiver domain structures were limited to CheY (Halkides et al., 2000; Stock et al., 1989; Volz, 1991; and Moy et al., 1994) from both *S. typhimurium* and *E. coli*, CheB (Djordjevic et al., 1998) from *S. typhimurium*, Spo0F (Madhusudan et al., 1997) from *B. subtilis*, the receiver domains of *Salmonella* NtrC (Volkman et al., 1995) and of *E. coli* PhoB (Sol et al., 1999), and NarL (Baikalov et al., 1996; Baikalov et al., 1998) from *E. coli*, which have been determined by X-ray or NMR techniques.

In CheY, the conserved aspartic acid at phosphorylation site is on top of a barrel-like structure that consists of a five-stranded parallel β sheet core surrounded by five α-helices (Fig. 4).

Figure 4. Cartoon diagram of *E. coli* CheY.

The central β-sheet is surrounded by α-helices. The site of phosphorylation, Asp 57, is shown on the carboxyl-terminal end of the central β-strand.


α5
α4
α1
α2
α5

As seen in the figure 4, helix 1 and 5 are on one side of the β sheet, and helix 2, 3, and 4 are on the other side. The aspartate is in a loop connecting β-strand 3 and α-helix 3, located between a lysine residue and an acidic pocket that has been identified as a binding site for a magnesium ion.

Although all of these receiver domains have the same secondary structure and have very similar regions near the phosphorylation site, there are some differences between them. The most conspicuous difference between the receiver domain of PhoB and that of other proteins in the same category is a displacement of helix α4 towards its amino terminus. Comparison of CheY and NarL structures reveals different placements of helix α5 relative to the rest of the structures. NarL contains a second small helix-turn-helix DNA binding domain at the carboxy terminus of the receiver domain. The carboxy terminal DNA binding domain of NarL is connected to the receiver domain by residues 126 to 159. Residues 143 to 154 of this linker region are disordered in the crystal lattice (Baikalov et al., 1996; Baikalov et al., 1998). Helices 1 and 5 are most displaced in an overlay of CheY and Spo0F. One of the main differences in the structures between CheY and NtrC receiver domain involves α–helix 4. In the NMR structure of NtrC, the helix is poorly defined, but it appears to be significantly displaced relative to its position in CheY.

Recently, a his-tagged fragment of DctD has been crystallized in the Nixon's lab. The fragment contains the protein's two-component receiver module and adjacent linker, which in the native protein joins the receiver domain to a σ^{54}-dependent ATPase domain. Its X-ray crystal structure at 1.7 Å reveals a novel dimerization interface for a two-component receiver domain. The dimer interface consists of helix α4, strand β5, helix α5 of the DctD receiver domain, and the adjacent linker that extends helix α5.

The unphosphorylated DctD receiver domain and linker form a dimer in which the α4-β5-α5 face of the receiver domain forms the subunit interface, with an extension of helix 5 through the adjoining linker forming a coiled-coil like structure that supports the dimer (Fig. 5).

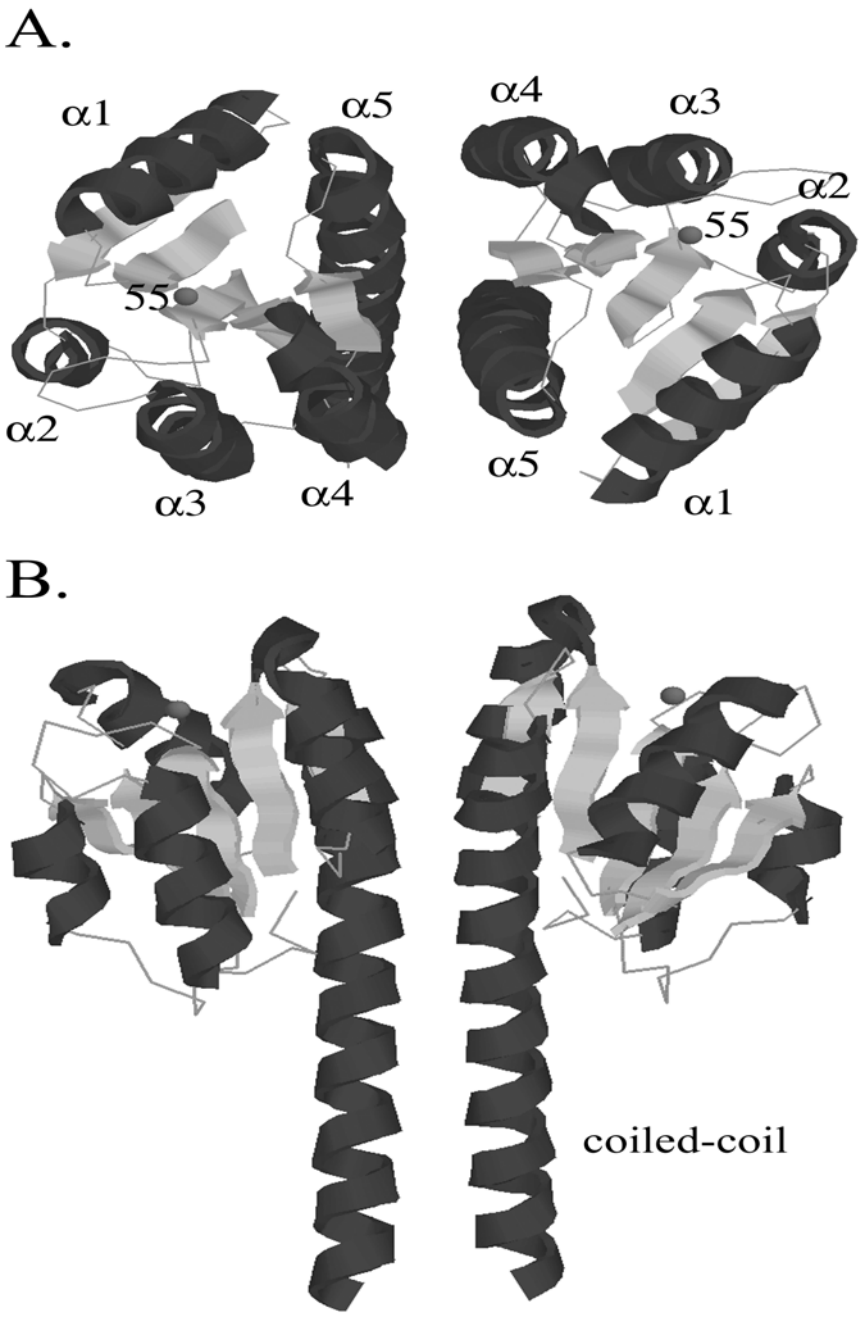
Based on sequence analysis, Nixon et al. proposed that a similar extended, coiled-coil helix 5 is involved in signal transduction in 54 of the 1148 known two-component receiver domains (Nixon, B. T., 2000). The predicted coiled-coil linkers do not appear in all members of a given output class of TCST proteins, or even in genetic homologs. For example, NtrC is not predicted to be like DctD, and the coiled-coil linker predicted for FixJ of *S. meliloti* is not predicted for FixJ proteins of other species. If these predictions prove to be correct, they will dramatically underscore diversities among structural mechanisms that link common changes in phosphorylated receiver domains with changes in the functional state of response regulator output domains.

Structures of 'activated forms' of two-component receiver domains

In bacteria, protein phosphotransfer reactions play a central role in the adaptive responses to changing environmental conditions. While the transient nature of phosphotransfer reactions effectively resets the signaling system in cells, it has also hindered the structural studies of the phosphorylated active forms. But recently there have been shown that high concentrations of organic phosphates such as carbamyl phosphate or acetyl phosphate can maintain the phosphorylated form of some receiver domains for a few days (Lukat et al., 1992; Feng et al., 1992; Kern et al., 1999), and BeF_3^{-1} can be used to mimic phosphorylation of NtrC, CheY, Spo0F, NarL,

Figure 5. A) Top view of cartoon diagram of DctDNL crystal structure. Like CheY, the central β-sheet is surrounded by peripheral α-helices. The phosphorylation site is located at the center of barrel-like structure.

B) Side view of cartoon diagram of DctDNL crystal structure. DctD monomer is associated with another monomer via the dimer interface that consists of the α4-β5-α5 face of each monomer. An extension of helix 5 through the adjoining linker forms a coiled-coil like structure that supports the dimer.


A.
α1
α5
55
α2
α3
α4
α4
α3
α2
55
α5
α1
B.
coiled-coil

OmpR and DctD (Yan et al., 1999). In the recent years, structures were reported for the phosphorylated or BeF_x bound forms of some of these proteins, or in most cases for fragments bearing their receiver domains (Halkides et al., 2000; Pelton et al., 1999; Cho et al., 2000; Lewis et al., 1999; Kern et al., 1999; Birck et al., 1999; Zhu et al., 1997). Phosphonomethylation of D57C in CheY mimics phosphorylation of D57, but supports only some of the structural changes seen upon binding BeF_x (Halkides et al., 2000). Comparing these 'activated' structures with the 'inactive' ones gives us some insights on the structural bases of two-component signal transduction system.

NMR structure of beryllofluoride-CheY showed T87 has moved to permit hydrogen bonding between its hydroxyl group and the phosphate analog, creating a hydrophobic cavity that is filled by inward rotation of the Y106 side chain (Cho et al., 2000). This motion was anticipated by structural and phenotypic studies of amino acid substitution variants of CheY (Zhu et al., 1997). Similar motion was just reported for BeF_3-CheY bound to a fragment of FliM, the target of activated CheY (McEvoy et al., 1999; Lee et al., 2000). The FliM peptide bound to helix α4 / strand β5 interface, requiring Y106 to move out of the way, affecting the Cα positions of loop β5-α5. The two residues T87 and Y106 are highly conserved in two-component receiver domains, and the same reorganization of them occurs upon phosphorylation of FixJN and Spo0AN (Lewis et al., 1999; Birck et al., 1999). The movement permits a dimerization interface to develop between helix α4 and strand β5 in FixJN (Birck et al., 1999), and dimerization of FixJ results in increased DNA binding and transcription activation by its associated output function (Da Re et al., 1999). Only the aromatic residue moves in phosphonomethylated D57C CheY, which nonetheless shows evidence of being

activated, so movement of the conserved threonine may not occur in all two component signal transduction systems (Halkides et al., 2000).

Phosphorylation of NtrCN causes other, more dramatic changes (Kern et al., 1999; Nohaile et al., 1997; Hwang et al., 1999). It displaces strands β4 and β5, and helices α3 and α4 away from the active site. It also causes a register shift and an axial rotation in helix α4. These changes expose a hydrophobic patch at the N-terminal end of helix α4, which is thought to activate the central ATPase domain by acting as a contact surface for another part of the protein. Phosphorylation dependent cleavage from Fe-BABE ((S)-1-[*p*-(bromoacetamido)benzyl]-EDTA-Fe) conjugated to the N-terminal portion of helix α4 took place at the beginning of the central domain in the partner subunit of NtrC (Lee et al., 2000). This evidence supports the hypothesis that helix α4 interacts with the beginning of the central domain for signal propagation in NtrC (Fig. 6).

While activation in DctD could occur by a similar mechanism, there are a number of differences between DctD and NtrC. While DctDNL is a stable dimer and followed by a α-helical linker, NtrCN is a stable monomer in the absence of the linker region between it and the ATPase domain (Volkman et al., 1995) and is predicted to be followed by a random-coiled linker (Bantscheff et al., 1999). The removal of N-terminal domain causes NtrC to lose the ability to hydrolyze ATP or activate transcription, while it does not affect the ability of NtrC to bind to DNA. But the effect of N-terminal domain deletion can be overcome by substitutions in the central domain (Drummond et al., 1990; Flashner et al., 1995).

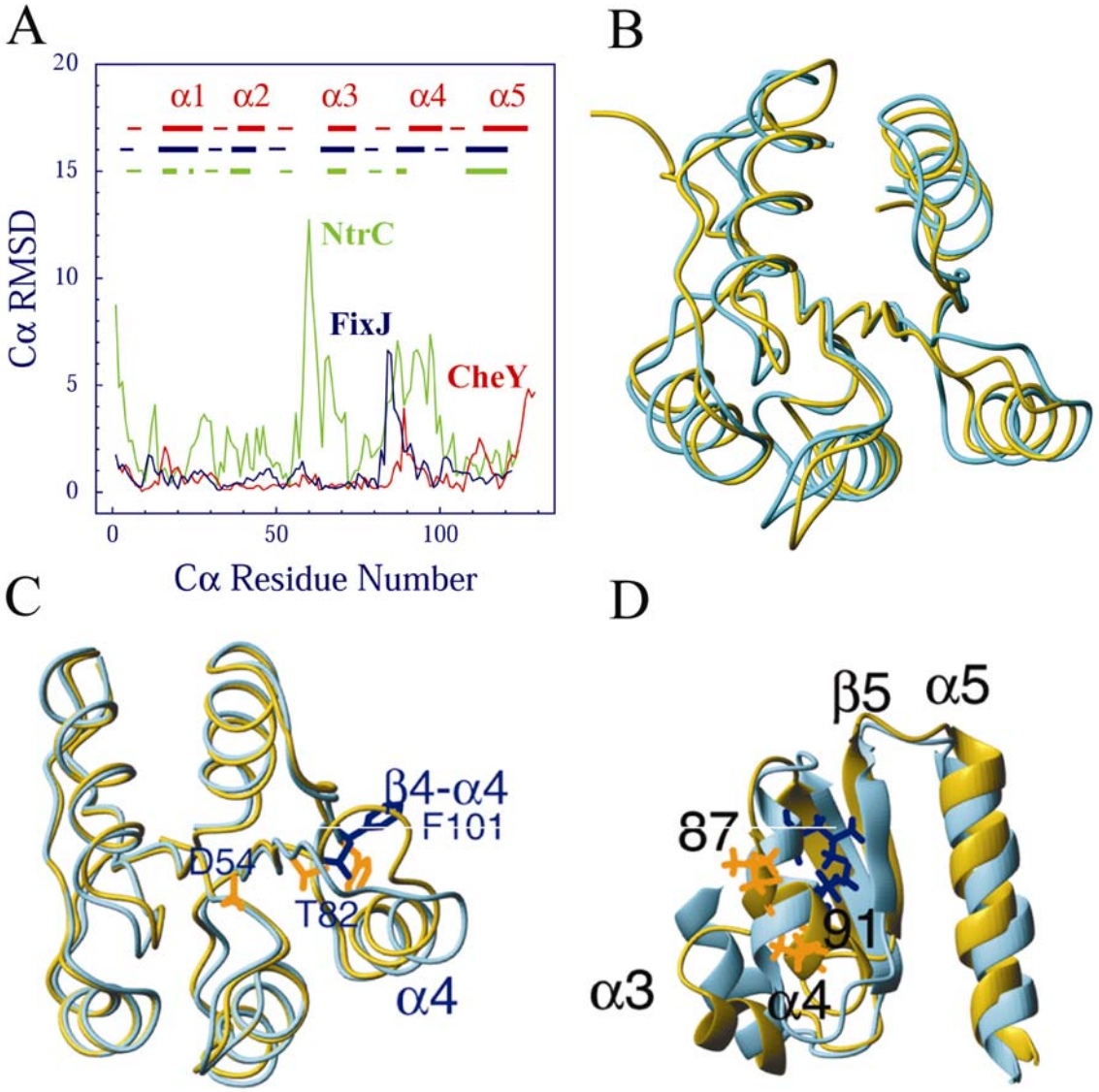
Figure 6. A) Plot of the RMSDs of C_α positions observed in the inactive and the active receiver domains of CheY, FixJ and NtrC. In the plot, CheY, FixJ and NtrC are represented by red, blue, and green color, respectively. Secondary structures are also included within the diagram: α-helix (rectangles); and β-strands (lines).

B) Strand-type diagram of an overlay of CheY and CheY~P. Only noticeable changes are observed in the loop β4/α4. Upon activation, this region forms part of the surface for binding with FliM.

C) Strand-type diagram of an overlay of FixJN and FixJN~P. Note that T82 is relocated toward D54~P, and correspondingly, F101 is displaced to support changes in loop β4/α4 and helix α4.

D) Ribbon diagram of an overlay of NtrCN and NtrCN~P. Major changes were found again in the β4/α4.

In all of the above crystal structures, an unphosphorylated form and a phosphorylated form of each protein are colored in cyan and gold, respectively.


A
20
15
10
5
0
α1 α2 α3 α4 α5
NtrC
FixJ
CheY
Cα RMSD
0
50
100
Cα Residue Number
B
C
β4-α4
F101
D54
T82
α4
D
β5 α5
87
91
α3
α4

This contrasts strikingly with the constitutive activity found in similarly truncated DctD proteins (Huala et al., 1992; Lee et al., 1994). Substitutions that give rise to constitutively active NtrC proteins (D54E, D86N and or A89T) map in the site of phosphorylation or in the N-terminal portion of helix α4 (Nohaile et al., 1997); none of the substitutions in DctD maps to the corresponding residues. The D55E substitution in DctD, analogous to D54E in NtrC, failed to activate DctD *in vivo* while it does activate NtrC (Klose et al., 1993). Contrasting results were also seen for substitution T83I: in DctD it activated the ATPase, while the corresponding substitution has a dominant negative affect on NtrC (and CheY) (Zhu et al., 1997). The major dimerization element of NtrC is provided by an extra helix preceding those of the helix-turn-helix DNA binding domain (Pelton et al., 1999; Klose et al., 1994). This extra helix is absent in DctD, which has dimerization determinants in its NT- and central domains. These observations suggest that structural details of the subunit-subunit and inter-domain interactions in DctD are different from those found in NtrC.

Consistent with this prediction, the X-ray crystal structure of DctDNL revealed a novel interface for dimerization that is not found in NtrC. This might be part of the way DctD maintains the 'off' state of the central ATPase domain in DctD. This could explain several observations about DctD: random substitutions that cause the central ATPase domain to partially escape repression by the receiver domain cluster in the dimer interface; at least two of these, K122ENL and E121KNL, reduce dimer stability 10-20 fold *in vitro* (Meyer et al., 2001); and phosphorylation of the full-length E121K variant further stimulates its ATPase activity.

Accordingly, one can think of a simple model for activation of DctD in which

phosphorylation destabilizes the receiver domain dimer to form monomer. Preliminary ^{1}H-^{15}N HSQC spectra confirm that phosphorylation or binding BeF_x alters the backbone environment of DctD residue Y100, which is clearly involved in the dimer interface (Hong Zhang, Masters Thesis, The Pennsylvania State University). However, analytical ultracentrifugation studies show that phosphorylating the DctDNL fragment does not reduce dimer stability *in vitro* (Meyer et al., 2001). The simple model is thus incorrect; rather, activation involves changing the equilibrium and/or kinetics of exchange between inactive and active conformation of the dimer. Phosphorylation of DctDNL not only decreases its sedimentation velocity but also increases its fluorescence anisotropy, confirming the occurrence of a conformational change (Nixon et al., unpublished). My work presented in this thesis will help clarify this conformational change.

5. Structural models for the ATPase domain of σ54-dependent activators.

The multiple sequence alignment of the EBPs (bacterial enhancer-binding proteins) revealed that the central domain is common to all member of the EBP family and is highly conserved along its entire length. Those conserved residues are arranged in the seven regions from C1 to C7 (Morett & Segovia, 1993). Sequence analysis of *dctD* from *R. leguminosarum* has shown that the central domain of DctD has the Walker A and B motifs which are shared by all of the mononucleotide-binding proteins (Ronson et al., 1987).

Based upon secondary structure prediction and fold pattern search, Osuna et al. proposed a fold architecture for the central domain of EBPs using the EF-Tu as working model. Since mutations in the region of NtrC and DmpR, which is equivalent to helix αN in the EF-Tu, caused the mutant proteins to activate transcription in a constitutive manner,

it has been speculated that helix αN forms part of the interface between the central domain and the receiver domain. Such an interface would be important for the receiver domains to regulate the central domains of σ^{54}-dependent activators. However, as they pointed out, their analysis is limited to the prediction of a likely fold and is not intended to be more than a working model (Flashner et al., 1995; Shingler, 1995; Osuna et al., 1997).

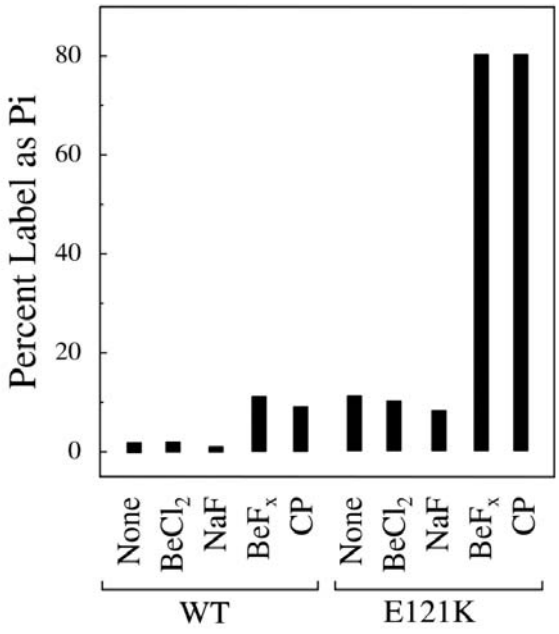
In recent years, it has been reported that prokaryotic NtrC-related transcription regulators belong to the AAA^+ class (Neuwald et al., 1999), whose protein sequences are related to the AAA family (ATPases associated with a variety of cellular activities). The AAA^+ module is spread out in higher cells as well as in bacteria. It consists of a P-loop ATPase motor domain upon which an AAA^+-specific component is mounted, and plays a role in a number of cellular activities. Many of the ATPases in the AAA^+ class serve as molecular matchmakers in the formation or activation of DNA-protein complexes, and are predicted to form a single ringed hexamer or two-ringed dodecamer with a hole in its center. It appears that bacteria and higher cells handle the same kinds of tasks with the same molecular strategies, even though the mechanistic details often differ. This suggests that crystallographic studies on $DctD_{2\text{-}384}$ containing an AAA^+ module will not only give some insight on other σ^{54}-dependent two-component response regulators, but also have a far-reaching impact on the understanding of numerous AAA^+ module related activities.

Here I report the results of structural studies on DctD and its variants in related to ATP binding and activation. To achieve this goal, firstly, I used trypsin to detect ATP-induced conformational changes in DctD, since site-specific proteases such as trypsin or chymotrypsin are known to be useful as probes for conformational changes. Secondly, I tried to examine the binding kinetics of DctD on mant-ATP. Thirdly, I have attempted to

crystallize both E121KNL, a constitutively active DctD mutant, and DctDNL complexed with BeF_3^{-1} to gain some insight on structural changes related to activation. This was based upon the following experimental results: 1) phosphorylation via the kinase (a truncated form of DctB), which maintains the phosphorylated state for only a few percent of the protein (Meyer et al., 2001); 2) phosphorylation using carbamyl phosphate, which by mass spectroscopy is seen to be efficient and quantitative but after several hours carbamylation is also evident (Meyer et al., 2001); 3) phosphonomethylation of substitution variant D55C, which is neither efficient nor necessarily specific for C55; 4) introduction of amino acid substitutions that bypass the need for phosphorylation; 5) incubation with BeF_3^{-1}, which clearly mimics phosphorylation (Yan et al., 1999); 6) the crystal structure of E121KNL turned out to be the same as that of DctDNL (Meyer et al., 2001); and 7) E121KNL is more readily phosphorylated and its phosphorylated form has a longer half-life compared to DctDNL (Meyer et al., 2001); and 8) In the presence of BeF_3^{-1}, the ATPase activity of full-length E121K is turned on to a greater extent than wild type full-length DctD as shown in Figure 7 (Meyer et al., 2001). Fourthly, I sought to crystallize $DctD_{2\text{-}162}$, which is basically same as DctDNL except that it is extended to helix αN adjoining to the receiver domain. Lastly, I attempted to crystallize $DctD_{2\text{-}384}$, which contains the entire ATPase domain except the DNA binding domain.

Figure 7. ATPase assay for full-length DctD and E121K

All assays were performed with DctD at final concentrations of 1 μM. Final concentrations of carbamyl phosphate, $BeCl_2$ and NaF are 20 mM, 200 μM and 10 mM, respectively. Each reaction mixture was incubated for 10 min. prior to the addition of ATP. Data were plotted as % label as Pi. The condition for ATPase assay was not optimized (Meyer et al., 2001).


Percent Label as Pi
0
20
40
60
80
None
BeCl2
NaF
BeFx
CP
None
BeCl2
NaF
BeFx
CP
WT
E121K

MATERIALS AND METHODS

Materials

Materials were purchased from the following vendors: trypsin and phosphocellelose from Sigma Chemical Co.; nucleotides from Sigma Chemical Co. or Amersham / Phamarcia; chromatographic columns and accessories from Amersham / Phamarcia; N-methylisatoic anhydride from Molecular Probes (Eugene, OR), oligonucleotides for PCRs from Integrated DNA Technologies, Inc., and FOMBLIN and $BeCl_2$ from Aldrich Chemical Company, Inc.

Gene clonings

The *Eco*N1 to *Nru*I fragment of pET11a (Novagen) containing the LacI repressor expressed by the $lacI^q$ promoter was cloned into the *Eco*R1 site located downstream of *dct*D in pTRH1 (Lee et al., 1994), making pBTN1 for expression of full-length DctD.

For wild type $dctD_{1\text{-}143}$, PCR primers were designed to amplify codons 1 to 143 of *Sinorhizobium meliloti dct*D. The primers introduced an *Nde*I restriction site at codon 1 and a *Hin*dIII restriction site immediately following codon 143. The *Nde*I/*Hin*dIII fragment was cloned into pET21a (Novagen) to create pT143WT, and was verified by DNA sequencing (Staley et al., 1998). Insertion into the pET21a vector resulted in the addition of the following sequence to the end of residue 143: KLAAALEHHHHHH. In intact DctD, this sequence would be PLIGQTPVMERLR. To construct the E121K substitution, I used the QuickChange™ site-directed mutagenesis kit available from Strategene and verified the product using DNA sequencing by the Nucleic Acid facility.

For *dctD*$_{1\text{-}162}$, a *Hind*III site was inserted right after residue 162 in pQE30DCTD (OriB), a expression vector. An *Xho*I site in the codons for 6 histidine residues in pT143WT vector was removed. Then, the *XhoI-Hind*III fragment from pQE30DCTD (OriB) was cloned into pT143WT to make pT162WT.

For *dctD*$_{1\text{-}384}$, a *Hind*III site was inserted right after residue 384 in pQE30DCTD (OriB). Then, the *XhoI-Hind*III fragment from pQE30DCTD (OriB) was cloned into pT162WT to make pT384WT.

Expression and purification of proteins

Plasmid pBTN1 was used to express wild type DctD in JM109 cells. This allows us to increase the expression of DctD up to 18% of total protein under the induction of isopropyl-β-D-thiogalactopyranoside (IPTG). The cells frozen after harvest were resuspended in ice-cold lysis buffer (20 mM Tris, 300 mM KCl, 5% glycerol, pH 7.9). Then the cells were broken for 10 cycles of 30 s sonication followed by 30 s of cooling using a Fisher model 300 Sonic Dismembrator. The solution was cleared by centrifugation at 36K rpm in a Beckman Ti 70 rotor for 45 min. at 4 ^{o}C. Dry ammonium sulfate was added to the supernatant with constant, gentle mixing to a final concentration of 35% saturation, which was incubated for 30 min. at 0 ^{o}C, and then centrifuged at 12K rpm at 4 ^{o}C using a Sorvall SS-34 rotor. The pellet was redissolved in buffer A (50 mM Tris, 50 mM KCL, 5% glycerol, pH 7.5), and dialyzed overnight against buffer A. The dialyzed solution was loaded onto a 10 ml phosphocellulose column pre-equilibrated with buffer A. DctD was included in fractions eluted between 150 mM and 250 mM KCl in a linear gradient from 50 mM to 300 mM KCl over a volume of 60 ml. The pooled

fractions were dialyzed overnight against buffer A at 4 ^{o}C, passed through a 1.0 ml Mono Q column, and then loaded onto a 1.0 ml Mono S column. DctD protein was eluted in a single peak at around 150 mM KCl in a salt gradient from 50 mM to 300 mM over a volume of 30 ml. Fractions were pooled and concentrated at 4 ^{o}C against buffer B (50 mM Tris, 150 mM KCl, 5% glycerol, pH 7.9) using Collodion membranes (Schleicher and Schuell). DctD was estimated to be more than 95% pure, as measured by laser densitometry of silver-stained SDS-PAGE gels.

His-tagged DctDNL from pT143WT was expressed under the control of an IPTG inducible T7 phage promoter (Dubendorff & Studier, 1991) in BL21(DE3) *E. coli* cells (Novagen) containing pLysS. Cells were grown in a 20 L Superboth media (400 g tryptone, 800 g yeast extract, 480 g glycerol, 140 g K_2HPO_4, 160 g KH_2PO_4, 2 g ampicillin and 0.68 g chloramphenicol) using a 30 liter Micros microbial fermenter at 32 ^{o}C. The cells were induced at an $O.D._{600\ nm}$ of 10 with 1 mM IPTG. After 4 h, the cells were harvested by centrifugation and placed at –70 ^{o}C freezer. Typical yields were 1 to 1.5 kg of cell paste. 10 g of cells were thawed and homogenized in a 20 ml of ice-cold lysis buffer (20 mM Tris pH 8.0, 0.5 M NaCl) by 10 cycles of 30 s sonications followed by 30 s coolings. The lysate was centrifuged at 100,000*g* for 45 min at 4 ^{o}C. The supernatant was loaded onto a 1 ml cobalt column made by replacing nickel on superflow NTA (Qiagen). After charging with cobalt, the column was pre-equilibrated with 20 ml of lysis buffer at a rate of 0.5 ml/min before loading sample at 0.5 ml/min at 5 ^{o}C. After washing with 30 ml of lysis buffer (0.5 ml/min) to achieve a stable baseline, the protein was eluted with 200 mM imidazole, 50 mM Na-phosphate buffer, pH 8.0. Imidazole and phosphate in the eluted solution were removed by a 5 ml HiTrap desalting column

(Pharmacia) using 20 mM Tris-HCl, pH 8.0. The desalted fractions were passed through a Mono S column and then loaded onto a Mono Q column. 6xHis-tagged DctDNL protein was eluted at around 400 mM KCl in a linear gradient from 200 mM to 500 mM and then concentrated in collodion membranes (Schleicher & Schuell). The Braford assay was used to quantify protein concentration, which was sometimes confirmed by optical density in 6M Guanidine HCl. The protein was immediately used for crystallization. Each time protein was made freshly before use, and typically found to be >99% pure by SDS-PAGE and laser densitometry.

His-tagged E121KNL protein, one of substitutive mutant proteins of DctDNL, was also purified as described above. Typically about 30% as much substituted protein was obtained from a given weight of cell paste compared to the wild type protein.

Tryptic digestion

Final concentrations were 2 μM for DctD, 5 μM for trypsin, and from 0 to 25 mM for ATP. Buffer C (20 mM Tris-HCl, 50 mM KCl, pH 7.9) was used to dissolve lyophilized trypsin, and buffer C plus 100 mM $MgCl_2$ was used to dilute DctD, dissolve the sodium salt of ATP, and to bring mixtures to required volumes. DctD, with or without ATP, was preincubated at 15 ^{o}C for 20 minutes prior to placing it and a trypsin stock at 25 ^{o}C. After 1.5 min, a control aliquot of the DctD sample was mixed with 2x loading buffer (100 mM Tris, 200 mM dithiothreitol, 4% sodium dodecyl sulfate, 0.2% bromophenol blue, 20% glycerol, ph 6.8) and placed in a dry ice/ethanol bath. After 2.5 minutes, portions of the warmed DctD and trypsin solutions were mixed and aliquots of the mixture were processed with loading buffer and frozen as above. The volume of the

aliquot taken for the control sample was less than those taken for the experimental samples, so that equal amounts of DctD-derived protein were present in each aliquot. Once all samples were collected, they were placed in boiling water for 3 min. and applied to a 10% Tricine-SDS polyacrylamide gel electrophoresis (Schägger, 1987). The gels were stained with Coomassie Blue, destained over several hours with 3 changes of destaining solution, and then equilibrated overnight with excess water. After destaining, gels were scanned and quantified using a laser densitometer and ImageQuaNT software (Molecular Dynamics). Images were processed for figures using Photoshop 6.0 (Adobe). To examine the effect of ionic strength on trypsin activity that might result from the high concentration of ATP, BSA was treated in the same way as DctD with or without ATP as a control reaction.

Measurements of equilibrium dissociation constant of the 44 kD tryptic fragment.

To estimate an equilibrium dissociation constant for ATP binding to the 44 kD tryptic fragment, we considered a two species decay process in which k_{NO} is the decaying with no ATP bound, and k_{ATP} is the decaying rate with ATP bound. This treatment assumes a single cleavage site between the receiver domain and the ATPase domain. By arbitrarily choosing t_o where no more 50 kD or 47 kD species exist to generate new 44 kD fragment, any subsequent data was considered to be the addition of the two decay processes:

$$\frac{[44kD]_t}{[44kD]_{t_0}} = F_U e^{-k_{NO}t} + F_B e^{-k_{ATP}t}$$

Where F_U and F_B refer to the unbound and bound fractions of the 44 kD fragment. By casting F_U and F_B in terms of the equilibrium dissociation constant, we obtain a formula that can be used with a version of the program NONLIN to estimate K_d by a global, nonlinear regression analysis of the decay data for digestions at different concentrations of ATP:

$$\frac{[44kD]_t}{[44kD]_{t_0}} = \left(\frac{K_d}{K_d + [ATP]}\right)e^{-k_{NO}t} + \left(1 - \left(\frac{K_d}{K_d + [ATP]}\right)\right)e^{-k_{ATP}t}$$

NONLIN estimates error by searching two coordinate systems of the variance space for an F-statistic corresponding to a 67% confidence probability (Johnson et al., 1992). For presentation of the decay data, the determined parameter values were used with the above equation to calculate decay curves for each concentration of ATP, and these were plotted with the measured data points.

Identification of tryptic fragments

Protein treated with trypsin was separated by 10% SDS-PAGE. To sequence a major species in each protein band, regions with the desired protein bands were blotted to PVDF membranes (BioRad) using 10 mM CAPS (3-[cyclohexylamino]-1-propanesulfonic acid; pH 11) in 10% methanol. The PVDF membrane with bound protein was rinsed with distilled water for 5 min., placed in methanol for 2 sec., and stained for 50 sec. with 0.1% Coomassie blue-R250, 1% acetic acid, and 40% methanol. Stained blots were destained with 50% methanol for 5 min., washed extensively with water, and the bands were excised and sequenced by Edman degradation analysis in Macromolecular Core Facility, Hershey Medical School.

To assess the homogeneity of each band, and to determine more precisely the molecular mass of the associated protein(s), 5 μM DctD treated with 0.1 μM trypsin for 2 hours was applied to a C_4-reverse phase chromatography column using 0.1% trifluoroacetic acid as buffer A and 100% isopropanol as buffer B. Each peak was then subject to electrospray ionization mass spectrometry (the Mass Spectrometry Facility of the Chemistry Department at Pennsylvania State University).

Synthesis and purification of mant-ATP

The N-methylanthraniloyl derivatives of nucleotides were synthesized with minor modifications of the published protocol (Hiratsuka 1983a), and the products characterized essentially as described by Woodward et al. Specifically, ATP or ADP (1mmol) was dissolved in 25 ml distilled water at room temperature with adjustment of pH to higher than 9.6 with 2 M NaOH. To this solution, a dry powder of methylisatoic anhydride (0.6 mmole) was added with continuous stirring, and incubated at 37 ^{o}C for 2 hours with a vigorous shaking while keeping pH above 9.6. To minimize formation of doubly substituted derivatives, the reaction was stopped by dropping pH down to 7.5 with 1 N HCl. The reaction mixture was filtered, diluted by 2 fold with distilled water. One-third of diluted reaction mixture was then placed onto a DEAE-cellulose column, and eluted with a linear gradient of 20 mM to 1 M triethylammonium bicarbonate. The monosubstituted derivative was eluted between the unreacted nucleotide and the bisubstituted derivative and well separated from the others. The pooled fractions were dried at –20 ^{o}C under vacuum with a removal of remaining triethylamine by evaporation. The dried mant-nucleotides were redissolved in distilled water. The A_{256}/A_{356} ratio for the

solution was 4 rather than 2.7, indicating that the fluorescent derivative was monosubstituted on either the 2′-position or the 3′-position rather than a bisubstituted derivative.

Binding of DctD to mant-ATP

Kinetics of mantATP binding was measured using a KinTek StopFlow System (Model SF-2001, KinTek Corp., State College, PA) fitted with a 100 W Hg arc lamp in Dr. K. A. Johnson's lab. The dead time of a KinTek StopFlow System was about 2 msec. An excitation wavelength of 280 nm was used. Since the binding of ATP to DctD was not found to induce a change in intrinsic fluorescence of the protein, the mant fluorophore was excited by way of energy transfer from tryptophan and/or tyrosine residues present in DctD protein. Fluorescence emission was measured at 450 nm using s25 450 bandpass filter. All reactions were performed at 20 ^{o}C, and the reported concentrations were the final values after mixing. The reaction was initiated by rapidly mixing DctD in 20 mM Tris-HCl, 150 mM KCl, 5% glycerol, pH 7.5 and varying amounts of mant-ATP, which were pre-equilibrated with the same buffer plus $MgCl_2$ to create Mg^{2+}-mant-ATP. The volume of the sample cell in a KinTek StopFlow System was 20 µl. The volume of 80 µl per shot was used to flush out the cell with fresh reactants prior to collecting data. The path length of the quartz observation cell was 0.5 cm. An average of three runs was used for data analysis and a minimum of a 10-fold molar excess of mant-ATP over DctD was used to allow analysis as a pseudo-first-order rate constant. The rate equation was simplified by assuming that the rates of steps following mant-ATP binding are sufficiently slow so that they can be neglected. The data were fit

to a single exponential by an iterative, nonlinear least-squares program available in Dr. K. A. Johnson's lab (KinTekSim). The rate of the exponential is plotted as a function of mant-ATP concentration, and fit to a straight line by a linear regression analysis. The slope and the intercept of the line were used to estimate an equilibrium dissociation constant.

CD (circular dichroism) spectra

Far UV-circular dichroism (CD) spectra were recorded at 25 ^{o}C using a 1 mm path length cell and an AVIV 62DS spectrophotometer (kindly made available by Dr. Matthews of the Chemistry Department). The molar circular dichroism ($\Delta\varepsilon$ in liter/mol.cm) is defined as described (Schmid, 1989): $\Delta\varepsilon = \theta_{obs} / (33 \times cl)$, where θ_{obs} is the observed ellipticity in degrees, c is the concentration in moles per liter, and l is the path length in centimeters. Data was collected between 195 nm to 300 nm with a time constant of 1 sec, every 1 nm and a 1.5 nm constant spectral bandwidth. DctD, E121K, and K122E were subject to CD analysis in 50 mM sodium phosphate, 1 mM β-mercaptoethanol, 5% glycerol, pH 7.2. Each protein was present at 2 μM as a final concentration.

Crystallization

In initial crystallization trials with the sparse matrix screen (Jancarik, 1991) using the hanging drop method, 18 conditions yielded crystals (Staley et al., 1998). Crystal growth conditions for wild type DctDNL and its constitutive active mutant protein, E121KNL, were further refined and subsequent crystals were grown using a well solution

of 50 mM succinate (pH 5.6) and 75 mM $NH_4H_2PO_4$. The protein was then filtered through a 0.1 μM filter prior to setting up crystallizations to remove most of nuclei. After filtration, only 1-4 crystals appear in each drop and grow to the size of $600 \times 400 \times 300$ μm. For E121KNL protein, the mixture of 90% filtered protein and 10 % filtered one was used to get large crystals like those of wild type.

To grow crystals of BeFx-bound proteins, BeFx was added in the condition in which both DctDNL and E121KNL did crystallize, and other screens using varied conditions were examined. Crystals of the BeF_3^{-1}-E121KNL variant were obtained as follows. The E121KNL protein solution was mixed with 5 mM $MgCl_2$, 14.3 mM $BeCl_2$ and 59 mM NaF dissolved in the same buffer as the protein, filtered using a 0.1 μm filter and set up for crystallization using 0.1 M HEPES (pH 7.0), 0.5 M $(Li)_2SO_4$ and 19% PEG 8K as a precipitant solution. Crystals grown in the presence of BeF_3 were cryoprotected in the buffer (5 mM HEPES pH 7.0, 50 mM $MgCl_2$, 7.15 mM $BeCl_2$ and 29.5 mM NaF) supplemented with 50 % glycerol. To obtain control crystals, varying conditions were explored. The mother liquor 0.1 M HEPES (pH 7.5), 0.5 M $(Li)_2SO_4$ and 5% PEG 8K was found to give such crystals. These crystals were cryoprotected in half the concentration of mother liquor supplemented with 50% glycerol. Both E121KNL crystals were flash-frozen in liquid nitrogen cooled air. As will be explained in Results section, no crystals of the wild type were obtained in the presence of BeFx.

For crystallization of $DctD_{1\text{-}162}$, it was dialyzed against 20 mM Tris-HCl (pH 8.0) overnight, and concentrated to 4 mg/ml with a change of buffer. Then 5 μl of protein was mixed with the same volume of mother liquor (50 mM succinate-acetate (pH 5.6), 13% PEG 4000) on a glass cover slip and equilibrated with 1 ml of mother liquor. The small

crystals formed were subsequently seeded into a fresh mixture of protein and mother liquor. Crystals were then mounted in thin-walled quartz capillaries. Later, a crystal grown in the same condition was flash-frozen in FOMBLIN (perfluoropolyether vacuum pump oil, average M.W. 1800, Aldrich #31,792-6) for searching systematic absences for a right space group.

Data collection

Each data set was collected using either a Rigaku R-AXIS IV or an R-AXIS IV^{++} imaging plate detector with a Rigaku RU200 rotating anode x-ray generator operating at 50 kV and 100 mA. To enrich for CuKα radiation, the X-ray beam was passed through a graphite monochromator for use with the R-AXIS IV detector, or an Osmic Mirror set for use with the R-AXIS IV^{++} detector. In the latter case, the intensity of the beam is about 8-times that of the former, necessitating the use of cooled crystals to avoid radiation damage. This was achieved using an X-stream low temperature system from Rigaku (Molecular Structure Corporation). Data at –160 ^{o}C were collected with exposure of 2 minutes and 0.5º oscillations using either all E121KNL crystals or DctD$_{1\text{-}162}$ crystal on the R-AXIS IV^{++} detector. Data at room temperature were collected with exposure of 90 minutes and 1.5^{o} using DctD$_{1\text{-}162}$ crystal on the R-AXIS IV detector. The diffraction images for each data set were indexed, integrated, scaled and merged using the programs Crystal Clear for all the data collected on the R-AXIS IV^{++}, or DENZO and SCALEPACK for data collected on the R-AXIS IV.

Determination of E121KNL crystal structure

The structure of the E121KNL fragment at -160 ^{o}C was determined by molecular replacement using the program AMoRe, which is part of the CCP4 package (Collaborative Computational Project 4). The room temperature E121KNL structure was used as an initial model for molecular replacement. In the rotation function, the model is rotated and its corresponding Patterson map is compared to that calculated from all reflections between a chosen upper and lower resolution (in this case, between 20.0 and 3.0 Å). Using the top 7 solutions from the rotation function, the translation function was then used to determine the translation parameters to position the model most correctly relative to the orientation of the unit cell. The final solution was then chosen as the one that gave the lowest 'R factor' between the observed structure factor amplitudes for the crystal and the calculated structure factor amplitudes for the best model of the translation search ($R = \sum | | F_o| - | F_c| | / \sum | F_o|$, where F_o is the observed amplitude and F_c is the calculated one). Using these rotation and translation parameters, a rigid body refinement at 3.0 Å was done using the FITFUN program of the CCP4 package. Before further refinement using CNS_Solve1.0, 10% of all measured reflections were randomly chosen and set apart to cross-validate the model throughout refinement (Brunger, 1992). Subsequent simulated annealing refinement was done to allow the model to escape from possible local energy minimum before further refinement. The resulting model was refined by conventional positional and B-factor refinement at 2.0 Å. The electron density peaks in the solvent region were interpreted as water molecules if their peak heights were greater than 3.0σ in $F_c - F_o$ (or 1.5σ in $2F_c - F_c$) maps. After each refinement, water molecules were removed if their temperature factors were greater than 50 Å^2. The model

was then subjected to visual inspection and model building using the program O. The alternating cycles of refinement and model building were continued until there was no further drop in the free R factor. Residues Lys144 and Leu145 were modeled as Ala because there was no density for the side chain. The quality of the final model was assessed by PROCHECK (Laskowski et al., 1993). The final coordinates of the structure have been submitted to The Protein Data Bank as 1L5Z.

Determination of E121KNL-Mg^{2+}-BeF_3^{-1} crystal structure

$E121K_{5\text{-}135}$ was used as a search model in rotation and translation functions for molecular replacement using AmoRe. The 21 top solutions from the rotation function were used for the translation function. Since the asymmetric unit for the space group $P2_12_12_1$ contains two molecules, the best solution from the translation function was fixed and a second translation function was solved to identify the second molecule. The orientations of the two molecules in the asymmetric unit were refined by treating each molecule as a single rigid body at 3.0 Å resolution. Following this, 10% of data were set aside for cross-validation analysis. The model was then refined by simulated annealing, positional, and B factor refinement with or without NCS (Non-crystallographic symmetry) constraints or restraints using CNS_Solve1.0. Once $2F_o$-F_c map was obtained, the model was modified and built up based on visual inspection by the program O. To improve the phase information in the model, residues not fitting the density map well were checked by simulated anneal omit maps, and adjusted to fit the density. New residues were gradually added as density maps became clearer until there was no more improvement in the $2F_o$-F_c, sigmaa map. In each step of the refinement following a model

building, the validity for the changes was assessed by monitoring the crystallographic R factor and the free R factor. Water molecules were initially placed in the solvent region if their peak heights were greater than 2.0σ in a $2F_o$ - F_c map. As the model improved, additional waters were added by slowly decreasing the cutoff to 1.5σ. There were no waters with B factors > 50 $Å^2$, so all were kept. At this stage, the optimum weight between the crystallographic and geometric terms ('wa' in CNS_Solve) was determined and used in the later refinements. The optimum weight value of the restraint for individual B-factor refinement ('rweight' in CNS Solve) was also determined and used in all the succeeding refinements. As refinement got close to the end, several cycles of the $F_o - F_c$ map were used to detect and correct the errors in the structure that remained undetected by $2F_c - F_o$ map. The validity of model rebuilds by the $F_o - F_c$ map was determined by monitoring the R factor and the free R factor during refinement following each correction. Residue Glu137 of chain A was modeled as alanine due to no density for its side chain. The stereochemistry and statistics of the structure were examined by PROCHECK and WHATIF. The final structure has been deposited in the Protein Data Bank as 1L5Y.

Determination of $DctD_{2-162}$ crystal structure

Data analysis for $DctD_{2-162}$ using the program Denzo was started with space group P6. But a lot of data were not included in the scaled and averaged data set. Additional data containing the 00l reflections were collected to clarify a space group by systematic absences. The presence of *l=3n* reflections clearly eliminated P6 space groups 168-170, leaving groups 151-154 and 171-172 as possible candidates. So the data were reduced

again using $P3_221$ as the space group, and molecular replacement solutions were sought for the trigonal space groups 151-154. Molecular replacement using the model $DctD_{5\text{-}128}$ as well as a larger model $DctD_{5\text{-}144,}$ gave reasonable solutions only for $P3_221$ (number 154; correlation coefficient = 0.687 and the R factor = 0.364). Since there is a possibility that the 2-fold axis on the b axis may not be perfect, space group $P3_2$ was also tested using $DctD_{5\text{-}122}$, which did not have the linker and beyond, as the search model for molecular replacement. It did give a solution that had correlation coefficient of 0.404 and the R factor of 0.489. Refinements for both space groups were therefore conducted in parallel until the proper space group stood out. The model was built up to residue 133 by alternating cycles of refinements and model rebuilds. Beyond this point, the $2F_o - F_c$ map was not well defined. Since each monomer in the asymmetric unit gave the electron density map of different resolution, one monomer was traced to residue number 143 while the other one was extended to residue number 156.

***In vivo* beta-galactosidase assays**

Plasimd pBTN1 was used to add a termination codon right after codon 384 of dctD, making pMAS1. An additional variant (pMAS1-E121K) containing E121K substitution was made by replacing the *Hind*III to *Sac*I fragment of pMAS1 with the corresponding fragment of pBTN1-E121K that has e121k substitution. These plasmids were then used to transform *E. coli* JM109 containing a previously made ϕ(*dctA´-lacZ´*)(Hyb) reporter genes carried on pRK290 (Ledebur et al., 1992). The activities of $DctD_{2\text{-}384}$ and $E121K_{2\text{-}384}$ were measured by the expression of the *lacZ* gene under the control of the *dctA* promoter. Control experiments were conducted with pBTN1 for full-

length DctD, and pBTN1-E121K for full-length E121K. The β-galactosidase assay was performed as described (Gu et al., 1994). Transformed cells were monitored for growth, and sampled in triplicate at the various time points. The supernatant was removed and stored on ice until all sampling process was done. Then cells were resuspended, sonicated, and clarified in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$). Beta-galactosidase activity, reported as nmoles of ONPG hydrolysed per minute per mg of supernatant protein, was measured by monitoring absorbance at 420 nm for the appearance of cleavage product of o-nitrophenyl-β-D-galactopyranoside (ONPG) at room temperature using a Beckman DU50 spectrophotometer. To normalize observed β-galactosidase activites to total 'soluble' protein, the protein content of the resulting supernatants was determined using the Bradford Protein assay (BioRad) using bovine serum albumin as a standard.

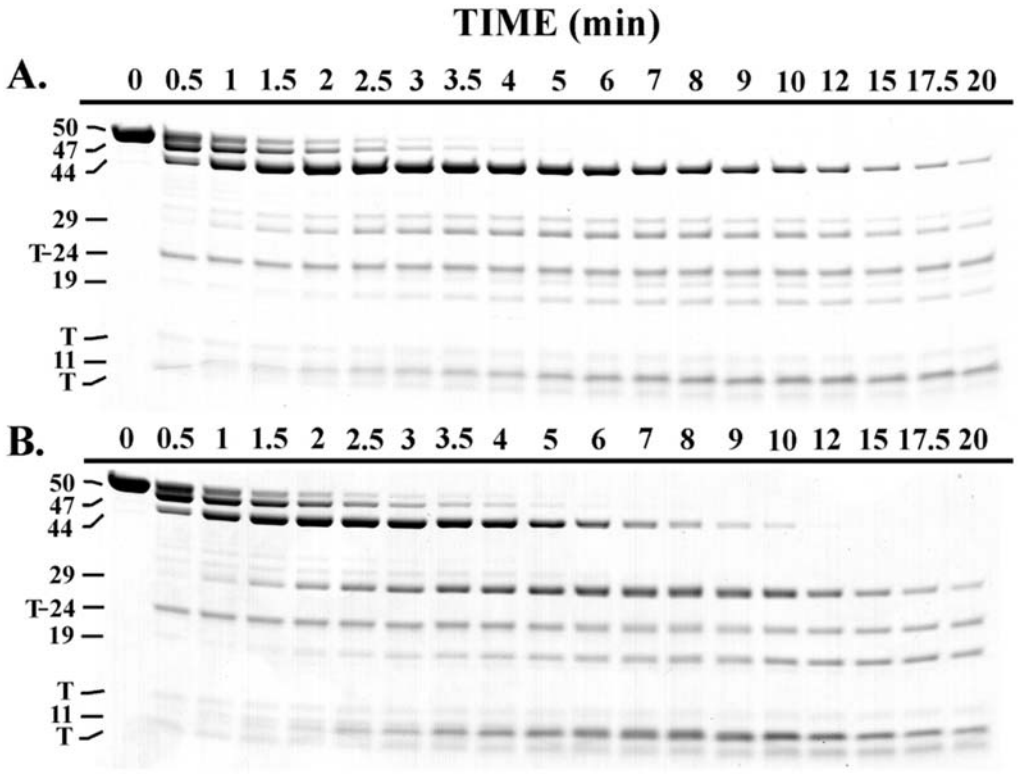
RESULTS

Tryptic digestion

It is known that phosphorylated DctD hydrolyzes ATP to convert σ^{54}-RNA polymerase complex into a transcriptionally competent form, while unphosphorylated DctD does not. But the role of phosphorylation in the activation process is still unknown. It could be involved in any of ATP binding, ATP hydrolysis, ADP release, protein assembly, etc. There have been no reports that unphosphorylated DctD binds ATP. So it is quite possible that phosphorylation might regulate the ability of DctD to bind ATP.

To test this, I used trypsin as a probe of conformational changes of DctD upon ATP binding, because proteases are more likely to gain access to exposed loops within domains or the linker region(s) connecting domains, which could either be present in native structure or arise from conformational changes in the protein. DctD was found to be rather resistant to trypsin. It took a 2.5:1 molar ratio of trypsin to DctD to give a complete digestion in 20 minutes. Preliminary experiments showed that ATP did change the digestion profile of DctD while equal molar amount of it did not affect the digestion profile of BSA (bovine serum albumin) in the control reaction. The change was saturated at about 10 mM ATP. Tryptic digestion was performed either with or without 12 mM ATP. It turned out that the ATP hydrolysis by one or more of the tryptic fragments during the whole course of tryptic digestion was negligible, indicating that ADP or phosphate was not responsible for any changes in the tryptic digestion pattern (data not shown). Figure 8 shows fragments of DctD separated by SDS-PAGE that were present in samples taken at various time points during tryptic digestion.

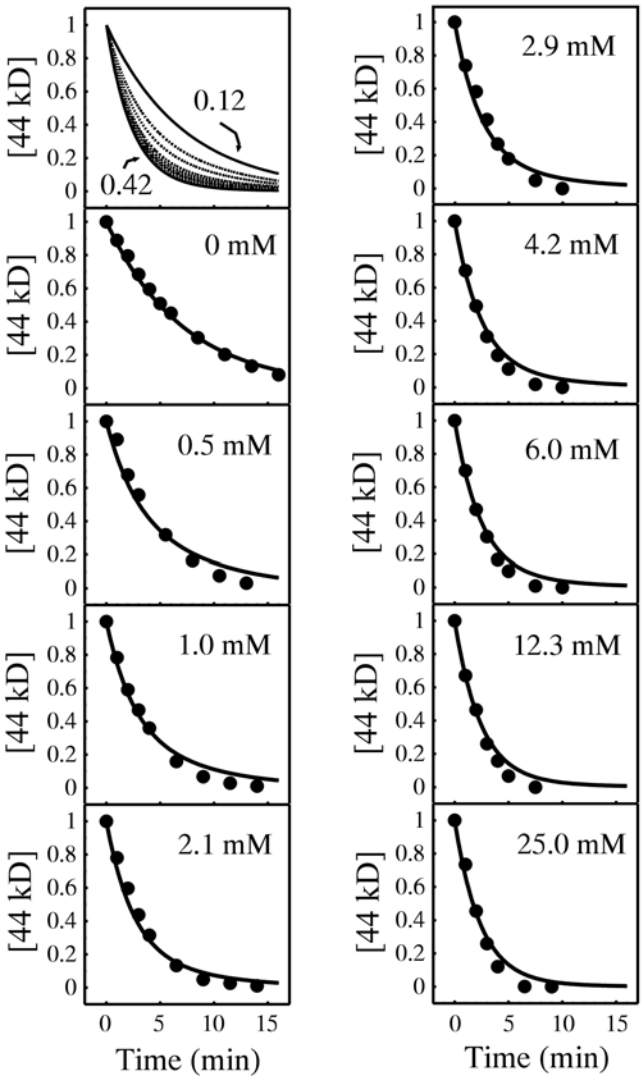
Figure 8. A) Tryptic digestion pattern of DctD in the absence of ATP. B) Tryptic digestion pattern of DctD in the presence of ATP. As described in Materials and Methods, 5 μM trypsin was used to digest 2 μM DctD without or with ATP of 12 mM being present. Samples from the indicated times were collected in loading buffer and then frozen in dry ice/ethanol bath. Once all were collected, they were subjected to SDS-PAGE. T-24 represents trypsin and its molecular weight.


TIME (min)
A.
0 0.5 1 1.5 2 2.5 3 3.5 4 5 6 7 8 9 10 12 15 17.5 20
50
47
44
29
T-24
19
T
11
T
B.
0 0.5 1 1.5 2 2.5 3 3.5 4 5 6 7 8 9 10 12 15 17.5 20
50
47
44
29
T-24
19
T
11
T

In addition to trypsin, which migrated as a major band of 24 kD, bands were seen at the apparent molecular masses 50 kD, 47 kD, 44 kD, 29 kD, 19 kD and 14 kD. Appropriate digests were conducted to allow unstained material corresponding to bands labeled 50 kD, 44 kD, 29 KD and 14 kD to be blotted to PVDF membranes. Then they were subject to several cycles of Edman degradation to sequence the major species in each band. The 50 kD and 44 kD bands both begin with the sequence SAAPS which starts at residue 2 predicted from the DNA sequence of *dctD*. The 29 kD band begins with the sequence AAEAA which starts with residue 135 of DctD. No useful information was obtained for the 14 kD. In the presence of ATP, the 44 kD fragment appeared to degrade more quickly, giving more intense bands for 29 kD and 14 kD while the 29 kD disappears more slowly. These results clearly indicated that unphosphorylated DctD binds ATP, and ATP binding induces conformational changes in DctD. Since the decay of 44 kD seems to be dependent on ATP binding and exists for quite a long period of time, we tried to estimate the K_d for ATP binding to 44 kD, which lost part of the DNA binding domain. We treated data taken in the presence of ATP as a sum of two first order decay rates: one for 44 kD unbound to ATP, an the other for 44 kD bound to ATP. The observed decay profiles during the same period of time, with ATP concentrations ranging from 0 to 25 mM, were related to the equilibrium constant as described in Materials and Methods. The estimated rate at 67 % confidence interval was 0.14 ± 0.02 min^{-1} in the absence ATP and 0.42 ± 0.03 min^{-1} in the presence of ATP, with an equilibrium constant of 0.7 (0.5, 1.2) mM ATP (Fig. 9).

Figure 9. Titration of ATP and decay of the 44 kD tryptic fragment of DctD.

Decay of the 44 kD fragment was monitored in the presence of varied amounts of ATP. Only data from time points after which all of the 44 kD precursors had decayed were considered, and they were treated as a two-species decay process in global, nonlinear regression analysis as described in the Materials and Methods. The estimated rates (0.12 and 0.42 min^{-1}) and equilibrium dissociation constant (0.7 mM) were then used to plot the curve for each of the tested concentrations of ATP with its corresponding decay data.


[44 kD]
0.12
0.42
0 mM
0.5 mM
1.0 mM
2.1 mM
2.9 mM
4.2 mM
6.0 mM
12.3 mM
25.0 mM
0 5 10 15
Time (min)

Subsequent measurements of the 44 kD fragment by mass spectroscopy revealed that it was a mixture of heterogeneous fragments, mainly differing in their C-termini. This result was consistent with that of protein sequencing showing that the major species in the band corresponding to 44 kD started with SAAP (data not shown). Since previous deletion studies indicated that the 44 kD is still under the regulation of the receiver domain (Gu et al., 1994), one might think that the experimental results on the 44 kD can be extended to full-length DctD.

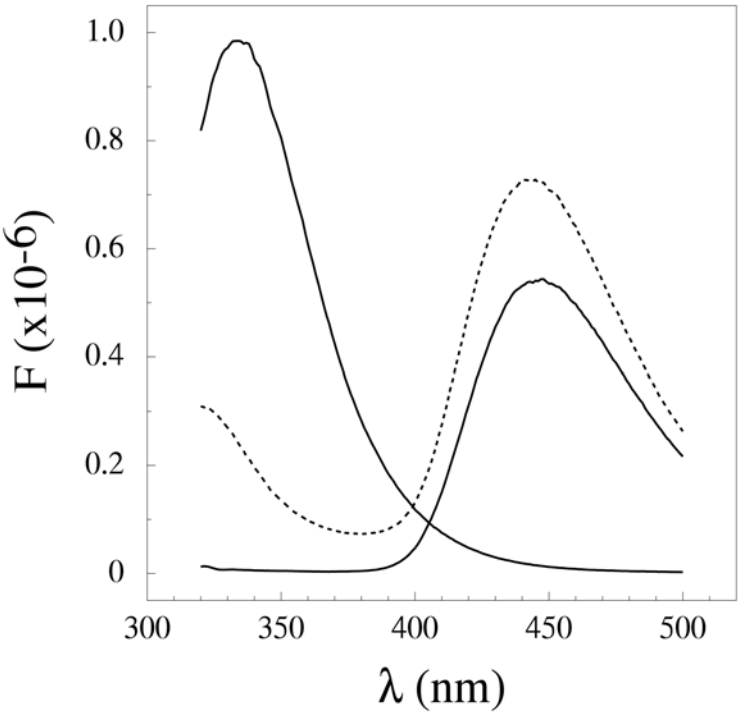
Pre-steady state kinetics of mant-ATP binding to DctD

To test this, I tried to examine ATP binding of DctD more carefully using a homogeneous, intact DctD rather than heterogeneous DctD proteolytic variants. Since ATP binding did not bring any change in intrinsic fluorescence of DctD, mant-ATP was synthesized as the source of extrinsic fluorescence to study ATP binding of DctD. Fluorescence energy transfer was observed between DctD and mant-ATP (Fig. 10). Since the efficiency of energy transfer between an intrinsic fluorophore and an extrinsic fluorophore is dependent not only upon the distance between them but also upon the conformation of a protein, mant-ATP was used to study ATP binding of DctD and its effect on the conformation of DctD.

To study ATP binding, I made an assumption that ATP binding is relatively fast, so that it is not affected by the subsequent steps during the period of time being used. To reduce a second-order kinetic equation to a pseudo-first-order kinetic equation, I used mant-ATP in excess over DctD.

Figure 10. Energy transfer from DctD to mant-ATP.

Fluorescence emission was scanned from 320 to 500 nm for DctD (peak emission at 340 nm), mant-ATP (peak emission at 444 nm), and a mixture of the two (dashed line). Excitation was at 290 nm. Energy transfer is evident from the loss of emission from DctD and gain of emission from mant-ATP.


1.0
0.8
0.6
0.4
0.2
0
F (x10-6)
300
350
400
450
500
λ (nm)

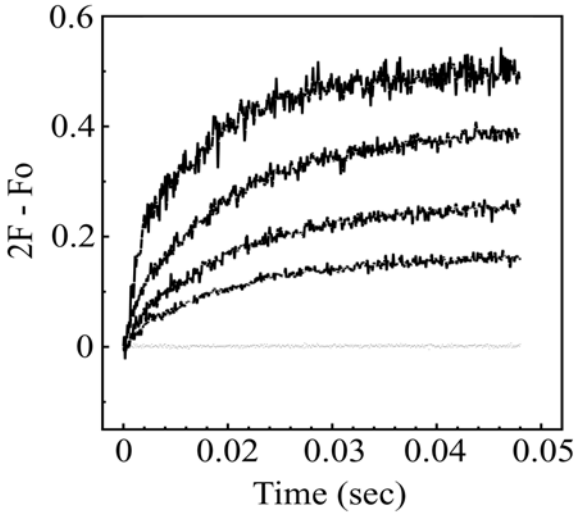
Under this condition, if $k_1[S]_o + k_{-1}$ is determined at several concentrations of mant-ATP, a plot of k_{obs} (= observed rate) versus $[S]_o$ (= total mant-ATP concentration) is linear with slope of k_1 and an intercept of k_{-1}. The ratio of rate constants is equal to the dissociation constant. Figure 11 shows the result of pre-steady-state kinetics of mant-ATP binding to DctD. It shows the rate (= observed rate) of the exponential term in rate equation as a function of the mant-ATP concentration.

The simplifying assumption used in this experiment seemed to be justified by the linear relationship between the observed rate and mant-ATP concentration. The second-order rate constant for mant-ATP binding to DctD estimated from the slope of the line was 0.087 $\mu M^{-1}sec^{-1}$. A dissociation rate constant estimated from the intercept was 76 sec^{-1}. The ratio of two rate constants gives the dissociation constant of 874 μM. Under these conditions, photobleaching, a potential cause for degradation of mant-ATP, did not contribute to the change in fluorescence (data not shown). The estimated dissociation constant is close to that estimated from tryptic digestion experiment, though mant-ATP and intact DctD instead of normal ATP and heterogeneous DctD proteolytic variants were used in this experiment. These results indicate that unphosphorylated DctD binds mant-ATP, and binding of ATP causes conformational changes in DctD. So the receiver domain is not likely to regulate the ATPase domain simply by preventing the ATPase domain from binding ATP, while the possibility that the receiver domain regulates the affinity of the ATPase domain for ATP still remains. But it should be noted that this dissociation constant is only a coarse estimation because a small error in data fitting process could result in a large error in the estimation of dissociation rate constant from y-axis intercept, and there were also inner filter effects by mant-ATP.

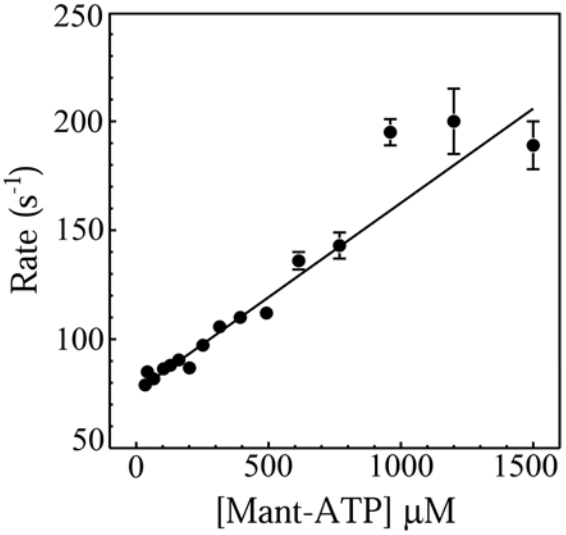
Figure 11. A) Kinetics of Mant-ATP binding to DctD. Increased fluorescence by mant-ATP via energy transfer from DctD irradiated with light of 280 nm. At 2 msec, 1) protein buffer or 2 to 5) increasing concentration of mant-ATP (45 μM, 90 μM, 180 μM, and 910 μM, respectively) was mixed with DctD of 1.5 μM. All concentrations reported represent the final concentrations after mixing 1:1 in stopped-flow instrument. All traces shown represent the average of three stopped-flow traces.

B) Dependence of the observed rate on mant-ATP concentration. The rate of exponential is plotted as a function of mant-ATP concentration. These data were fitted to a straight line by NONLIN. The slope of the line predicts an apparent second-order rate constant (on rate) for mant-ATP binding of 0.087 (0.069, 0.104) $\mu M^{-1}s^{-1}$ and a dissociation rate (off rate) of 76 (72, 81) s^{-1} is obtained from the intercept. A steady-state-dissociation constant of 870 (690, 1020) μM is obtained from the ratio of the off rate divided by the on rate.

A.


0.6
0.4
0.2
0
2F - Fo
0
0.02
0.03
0.04
0.05
Time (sec)

B.


250
200
150
100
50
Rate (s-1)
0
500
1000
1500
[Mant-ATP] µM

In addition, transient kinetic data over longer period of time revealed complex changes in fluorescence upon mixing of DctD and mant-ATP, further complicating the interpretation of the kinetic data (Fig. 12). As shown in Figure 12, the data fit poorly to a single exponential, better to a sum of two exponentials, and best to a sum of three exponentials. This brings three potential problems to the conventional data fitting methods. First, fitting data to multiple exponential functions leads to large errors in the rates and the observed rate may not be directly related to an analytical solution involving multiple exponentials. Second, the simplifying assumptions needed to achieve an analytical solution to the rate equations may not be valid and for any complex pathway, secondary effects due to other kinetic steps can alter the fit to the step that is being measured. Third, in a situation where the complete kinetic mechanism is not available, it is not possible to monitor conventional data analysis with a global fit of the data to the complete mechanism by nonlinear regression using computer simulation programs such as the KINSIM and FITSIM (Moyer et al., 1998). All these barriers, our lack of expertise in kinetic studies, and the development of subsequent interests did not allow us to pursue this line of investigation any further.

CD spectra – DctDNL, E121KNL, DctD, E121K and K122E

As the studies of binding of mant-ATP to DctD turned into a conundrum, our focus was moved to the structural studies on activation of DctD by phosphorylation or substitution. As part of these efforts, we measured CD spectra of DctDNL, E121KNL, full-length DctD, and full-length E121K, and full-length K122E. For comparison, difference CD spectra were plotted (Fig.13).

Figure 12. Time course of changes in fluorescence upon mixing mant-ATP with DctD. Fluorescence from the mant-ATP (443-456 nm) was monitored in two time domains (22 data points from 0 to 200 msec, and 300 data points from then to 20 sec). Heavy line indicates the best fitting model. Thin ones indicate less complex models, but do not fit as well as heavy one.

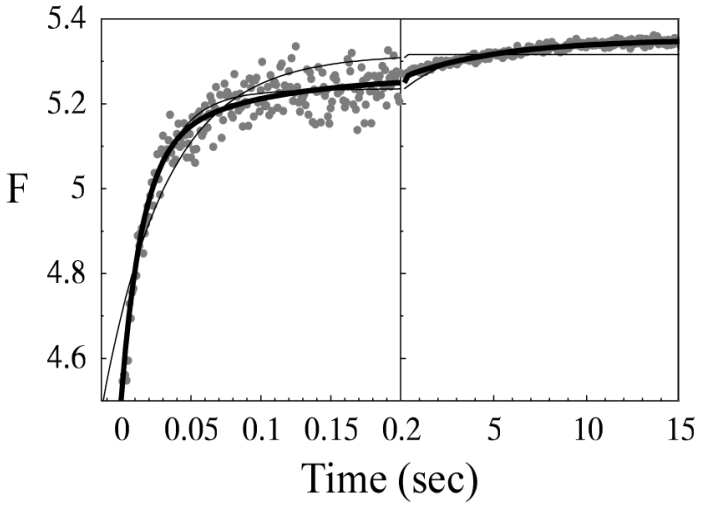
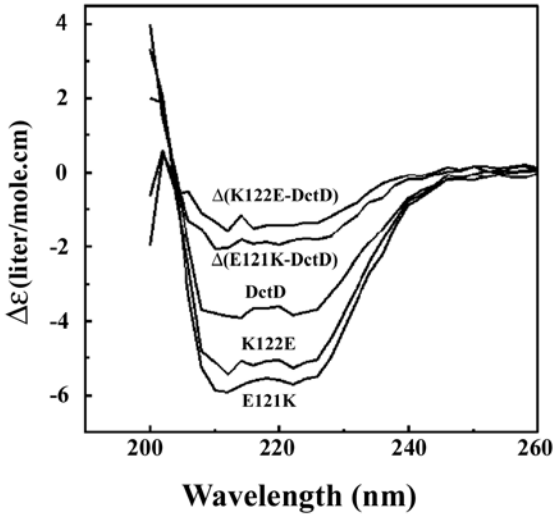

5.4
5.2
F
5
4.8
4.6
0
0.05
0.1
0.15
0.2
5
10
15
Time (sec)

Figure 13. CD (circular dichroism) spectra and differential CD spectra of full-length DctD, E121K, and K122E. The molar circular dichroism, $\Delta\varepsilon$, which is defined as the difference between the extinction coefficients for the two types of circularly polarized light, was plotted against the wavelength of 200 - 260 nm.


4
2
0
-2
-4
-6
Δε(liter/mole.cm)
Δ(K122E-DctD)
Δ(E121K-DctD)
DctD
K122E
E121K
200
220
240
260
Wavelength (nm)

The results showed that E121K substitution in the full-length protein caused an increase in optical ellipticity to an extent that is unusual for point mutations (Park et al., submitted). The difference CD spectra between full-length DctD and E121K or K122E suggested that there was an increase in α-helix content. But without any structural information available for the full-length protein, it was not possible to interpret the CD curve further. For the receiver domain, there was much less change in the CD signals between DctDNL and E121KNL, indicating that there were no significant changes in conformation (data not shown). Meanwhile, the crystal structures for DctDNL and E121kNL have been revealed. They showed that there were no structural differences between those two, though they represented the structures in crystal lattice at much higher concentration.

The crystal structure of E121KNL

Crystals of E121KNL are in space group I222 with unit cell dimensions of a=57.1, b=58.1, c=168.8, and $\alpha = \beta = \gamma = 90^{o}$. The asymmetric unit contains one monomer. The best solution of the rotation function was found at the second highest peak 12.1σ ($\alpha = 98.95^{o}$, $\beta = 77.28^{o}$, $\gamma = 353.24^{o}$). The best translation solution was obtained for the top rotation solution, giving a correlation coefficient value of 0.56 and the lowest reliability index of 0.44. This solution was from a 15.6σ density peak with Tx = 0.0459, Ty = 0.3085, and Tz = 0.1917. A rigid body refinement at 3.0 Å using these rotation and translation parameters (the CCP4 package) gave an R-factor of 0.36 and a correlation coefficient of 0.77. The final model contains 1,105 non-hydrogen protein atoms, 144 water molecules, one sulfate ion, and 4 glycerol molecules. The crystallographic R factor

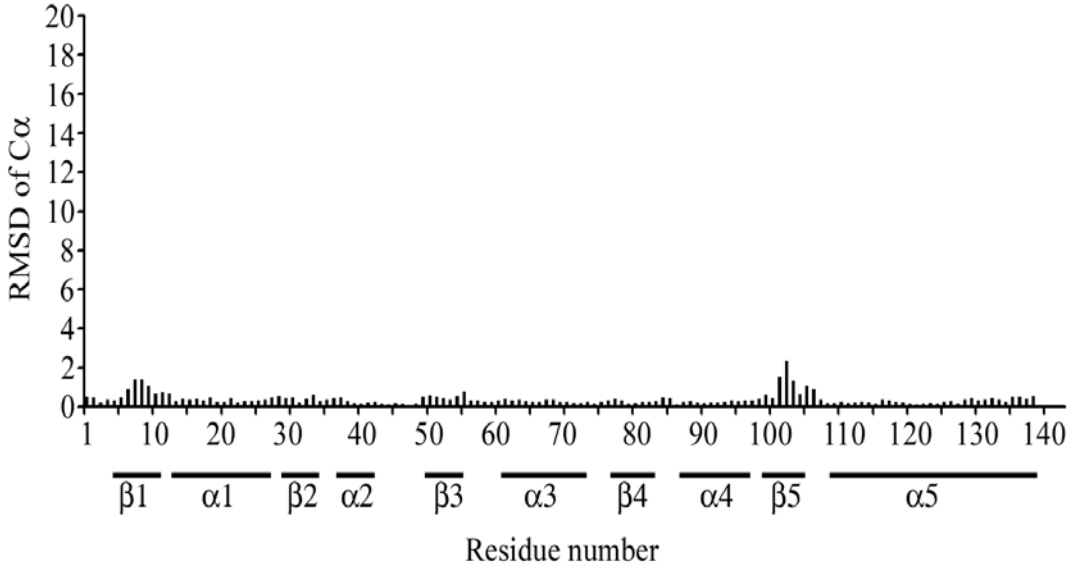
is 19.6 % with the free R factor of 23.2 % (Table 2).

Despite the fact that the E121K substitution removes one of cross chain hydrogen bonds that is present in the wild-type DctDNL, thereby weakening the dimer by ~20-fold (Meyer et al., 2001), it still crystallized into the I222 space group like the wild type DctDNL. An overlay of the two structures (C_α atoms of residues 5-83 and 90-134, RMSD 0.47 Å) shows them to be largely same except for loop β5-α5 (Fig. 14). The difference may be due to the use of different cryoprotectants, as there is a glycerol bound near this loop in E121KNL that is not present in the wild type structure. This loop is seen to adopt yet a third conformation in response to temperature, salt and amino acid substitution (Meyer et al., unpublished). So most structural features of wild type DctDNL were also found in the E121KNL structure. Like wild type DctDNL (Meyer et al., 2001), E121KNL has five parallel β-strands forming a single sheet that is surrounded by five α-helices, which is typical of receiver domains such as FixJN and CheY. The amphipathic helix 5 consisting of residues 108-143 is over 50 Å long and joins the receiver domain to the rest of the DctD protein. The calculations of surface area using CNS_Solve1.0 showed that roughly 1,805 $Å^2$ of solvent accessible surface area, which is exposed to solvent in the free monomers, is buried in the dimer. These buried regions in the dimer consist of loop β4-α4, helix 4, loop α4-β5, strand 5 and helix 5, including the linker region. The cross-chain hydrogen bonds formed by Val92 O – Gln114 Nε2, Tyr100 O – Arg118 NH1, Tyr100 H – Tyr100 H, and Glu128 Oε2 – Asn129 Oδ1 are the same as those of wild-type DctDNL, except that E121KNL does not have cross-chain hydrogen bond between E121 Oε2 and K122 Nζ.

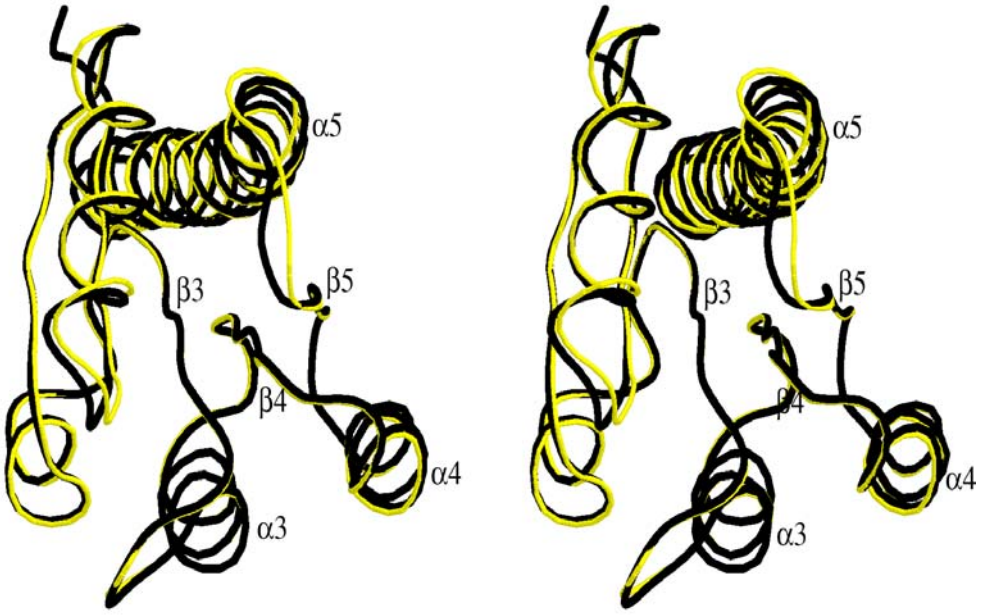
Figure 14. A) Plot of the RMSD in C_α positions of off-state DctDNL and off-state E121KNL. Secondary structures are also included within the diagram.

B) Stereo view of the overlay of off-state DctDNL and off-state E121KNL. DctDNL and E121KNL are colored in yellow and black, respectively. The regions covering 5-83 and 90-134 were used for superimposition of the two structures with an RMSD of 0.47 Å.

A.


RMSD of Cα
20
18
16
14
12
10
8
6
4
2
0
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
β1
α1
β2
α2
β3
α3
β4
α4
β5
α5
Residue number

B.


α5
β3
β5
β4
α4
α3
α5
β3
β5
β4
α4
α3

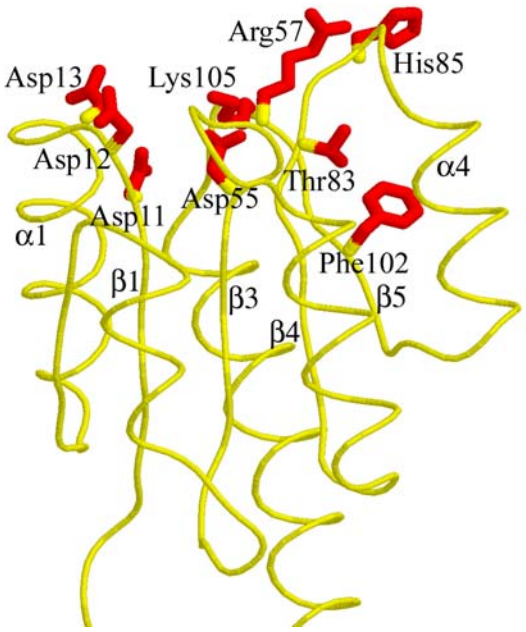
Like wild-type DctDNL dimer, helices α5 stack upon and cross each other, making a pseudo coiled-coil. In a tail-to-tail orientation, two dimers form a tetramer in an interlocking fashion, which is supported by two cross-dimer hydrogen bonds: Ser140 Oγ - Ser140 Oγ (2.96 Å) and Ser140 Oγ – Ser140 O (3.35 Å) and by two cross-dimer van der Waals contacts: Ala139 Cβ – Leu143 Cδ1 (3.61 Å) and Ala139 Cβ - Leu143 Cδ2 (3.68 Å). Crystal packing probably causes this contact between the two dimers. The Asp55 Oδ2 of the active site in strand β3 is hydrogen-bonded to the Lys 105 Nζ in strand β5, which seem to make E121KNL protein in crystals refractory to phosphorylation like DctDNL protein. The side chain of Phe102 in strand β5 assumed an external position, which is typical of other two-component receiver domains in their inactive states. The side chain of conserved Thr83 in the bottom of strand β4 is 6.5 Å apart from Asp55 Oδ1 and occupies the space between strand β4 and helix α4. Such positioning prevents the Phe102 side chain from flipping into an inward position (Fig. 15).

The crystal structure of E121KNL-Mg^{2+}-BeF_3^{-1}

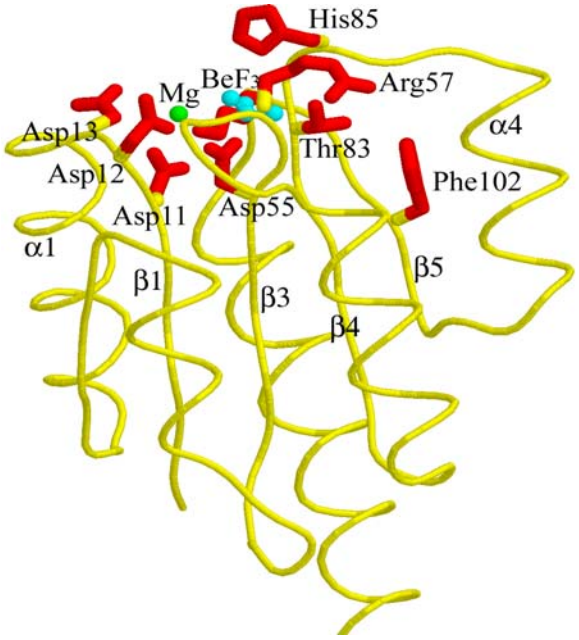
The crystals of E121KNL-Mg^{2+}-BeF_3^{-1} are in space group $P2_12_12_1$ with unit cell dimensions of a=52.0, b=76.8, and c= 77.8, and $\alpha = \beta = \gamma = 90^o$. The asymmetric unit contains two molecules. The solutions of the rotation function to give the best translation solutions for the first molecule (Tx = 0.0494, Ty = 0.4425, Tz = 0.7409, R = 51.8, and CC_F = 30.8) and the second molecule (Tx = 0.0782, Ty = 0.2935, Tz = 0.3390, R = 55.3, and CC_F = 17.7) were found at the highest peak 11σ ($\alpha = 69.91^o$, $\beta = 50.16^o$, $\gamma = 265.35^o$) and the 9th highest peak 9.2σ ($\alpha = 103.11^o$, $\beta = 90.00^o$, $\gamma = 70.97^o$), respectively.

Figure 15. A) Strand type view of the active site of off-state E121KNL. B) Strand type view of the active site of E121KNL-Mg^{2+}-BeF_3^{-1}. The Mg^{2+} and BeF_3^- in the active site were colored in green and cyan, respectively. C) Stereo view of concerted movements in the active site upon Mg^{2+}-BeF_3 binding. The off-state monomer (yellow) and the on-state monomer (black) of E121KNL were superimposed. The Mg^{2+} and BeF_3^- in the active site were indicated by green and cyan color, respectively.

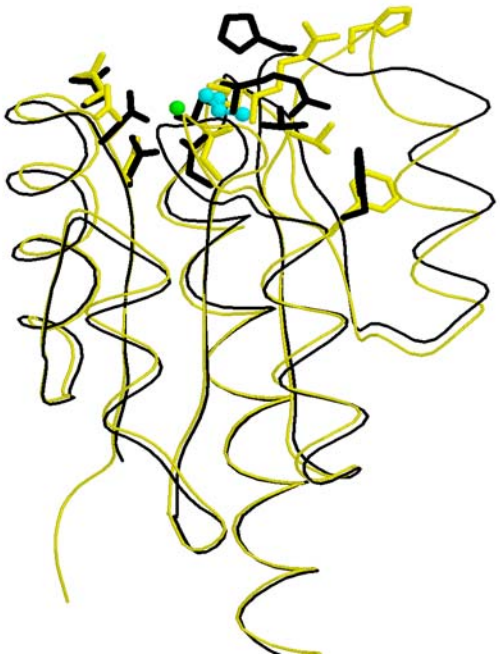
A.


Arg57
His85
Asp13
Lys105
Asp12
Thr83
α4
Asp11
Asp55
α1
Phe102
β1
β3
β5
β4

B.


His85
Mg
BeF3
Arg57
Asp13
α4
Thr83
Asp12
Phe102
Asp11
Asp55
α1
β5
β1
β3
β4

C.



Subsequent rigid body refinement at 3.5 Å gave a R-factor of 46.8 % and a correlation coefficient of 57.9 % for both molecules. During subsequent refinements, NCS restraint or constraint was ignored because the electron density map clearly showed that there was a difference between the two monomers in the asymmetric unit. The comparison of the main or side chain torsion angles of the corresponding residues between NCS-related molecules also justified ignoring of NCS constraints or restraints (Fig.16). The final model obtained after several cycles of refinements and model rebuilds contains 2,283 non-hydrogen protein atoms, 224 water molecules, 2 Mg^{2+}, 4 SO_4^{-2}, 10 BeF_2, 1 BeF_4^{-2}, and 3 BeF_3^{-1}. The crystallographic R factor and the free R factor were 18.40 % and 22.9 %, respectively (Table 2).

Density for Mg^{2+} and BeF_3^- is clearly seen in a simulated_annealed omit map calculated after removing a 5 Å sphere surrounding the ligands from the model (Fig. 17). The metal ion is coordinated to Oδ1 of Asp12, O of Arg57, Oδ2 of Asp55, F of BeF_3^-, and two water molecules. The Nζ atom of the highly conserved K105 residue is potentially hydrogen bonded to one of these water molecules (2.93 Å), Oδ1 of Asp55 (3.07 Å), Oδ2 of Asp11 (2.51 Å) and F of BeF_3^- (3.04 Å). The third fluorine atom of the bound BeF_3^- is hydrogen bonded to Oγ1 of Thr83 (2.74 Å). This pulls strand β4 toward the active site, by which loop β4-α4 is forced to move to a new position. This creates a space between strand β4 and helix α4 that is occupied by an inwardly oriented side chain of residue Phe102. The side chain of Phe102 is stabilized by several van der Waals contacts with Phe102 Cδ1 – I88 Cγ2 (3.71 Å), Phe102 Cδ2 – Ala94 Cβ (3.71 Å), Phe102 Cζ – Met90 Sδ (4.15 Å), Phe102 Cε1 – Ala91 Cβ (3.39 Å), Phe102 Cζ – Thr83 Cγ2 (3.56 Å), Phe102 Cε1 – Thr83 Cγ2 (3.74 Å), and Phe102 Cε2 – Ile56 Cδ.

Figure 16. $\Delta\phi/\Delta\psi$ for the two NCS-related molecules of Mg^{2+}-BeF_3^{-1}-E121KNL at 2.1 Å resolution. The solid black line shows the difference between the main-chain torsion angles of NCS-related residues (= $\Delta\phi$). The solid green line shows the difference between the side-chain torsion angles of corresponding residues (= $\Delta\psi$).

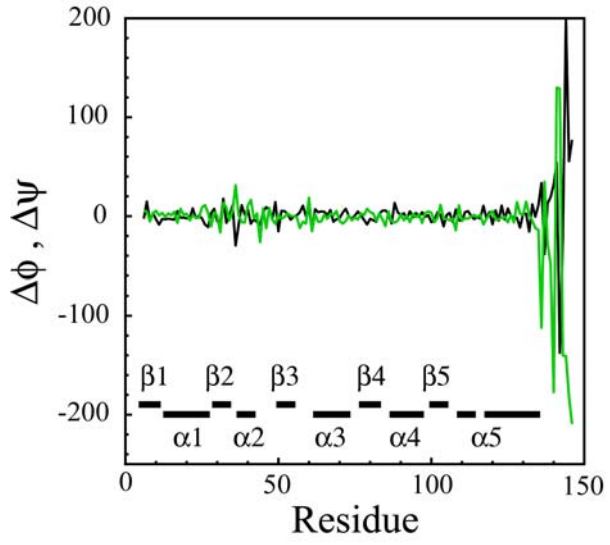
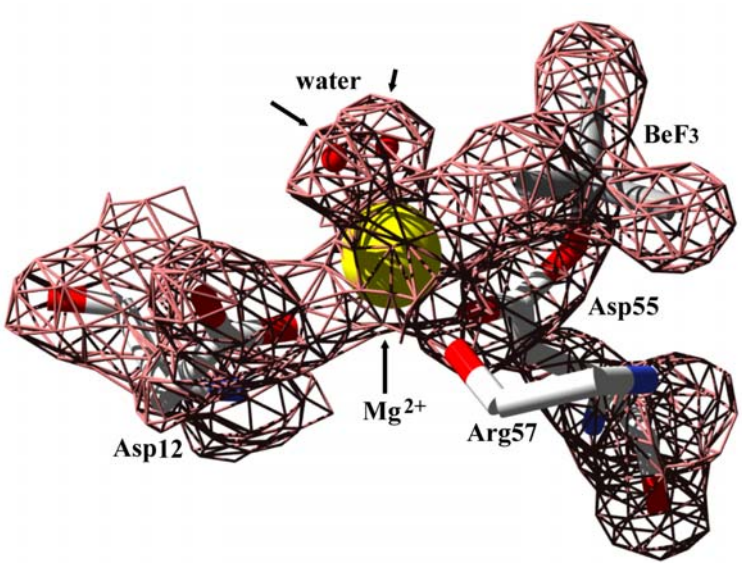

200
100
0
-100
-200
Δφ , Δψ
β1
β2
β3
β4
β5
α1
α2
α3
α4
α5
0
50
100
150
Residue

Figure 17. $2F_o - F_c$ SA (simulated annealed) omit map for the active site of Mg^{2+}-BeF_3^{-1}-E121KNL. The map was contoured at 1.3σ above the mean density. This map shows clearly that BeF_3 and Mg^{2+} are present in the active site.


water
BeF3
Asp55
Mg2+
Arg57
Asp12

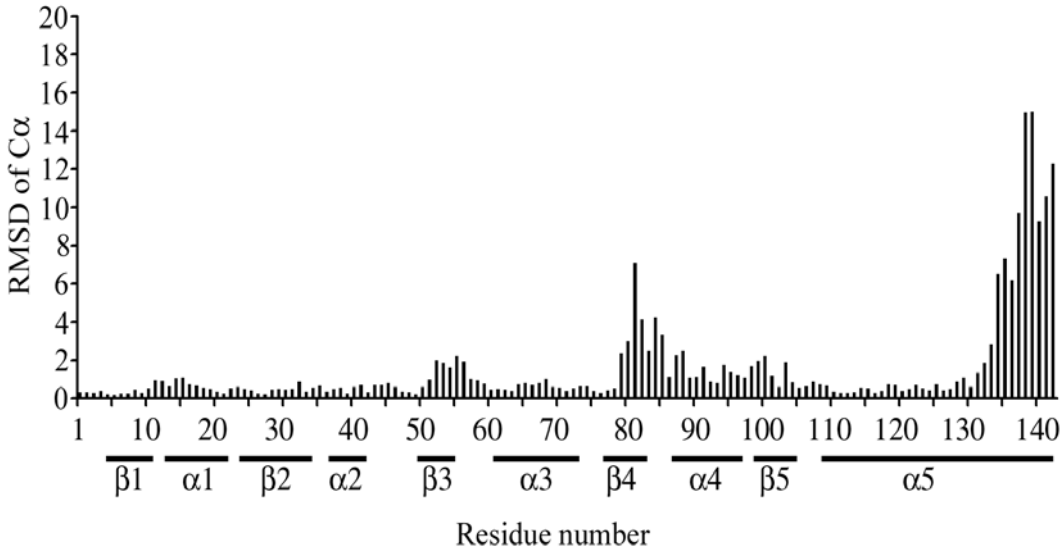
The active site just described is significantly different from that of the inactive state (Fig. 15). The terminal oxygen atoms of the Asp55 side chain in the inactive state are held in an orientation rotated 90º from that seen in the active state, supported in the latter by potential hydrogen bonds Asp55 Oδ1 – Arg57 N (3.01 Å), Asp55 Oδ1 – Arg57 O (3.24 Å), Asp55 Oδ2 – Lys105 Nζ (3.23 Å), Asp55 Oδ2 – Asp11 Oδ1 (2.54 Å), Asp11 Oδ2 – Asp12 N (2.97 Å) and Asp11 Oδ2 – Asp13 N (2.79 Å). Different rotamers are favored for Asp13, with Mg^{2+} stabilizing the one seen in the active state. The highly conserved Thr83 Oγ1 is 4.8 Å from Asp55 Oδ1 instead of 6.7 Å in the inactive state, reflecting the inward relocation of the upper portion of β4 and the side chain of Phe102 in β5. The off-state monomer and each of the two on-state monomers are superimposed using residues 5-83, and 90-134 with RMSDs of 0.87 Å (Fig. 18A and B) and 0.95 Å (Fig. 19A and B), respectively. It shows that there are major displacements in the β4-α4-β5 region, with most significant changes in the loop β4-α4, which caused the reorientation of two thirds of α4 to be tilted away from the active site. These regions are involved in protein-protein contacts in the crystal lattice. Over the same region, the two on-state monomers are superimposed with an RMSD of 0.55 Å (Fig. 20A and B). It shows that there is a little change in the loop β4-α4, with significant changes in the helix α5. But the latter changes seem to be caused by crystal packing.

The off-state dimer is not present in the crystal lattice of the BeF_3^--bound protein. In its place are four possible alternatives (Fig. 21). The first dimer is made by symmetric contact between helix α4 and strand β5 with a buried surface area of 931 Å. The second is maintained by symmetric contact between helix α1 and α5, burying surface area of 562 Å.

Figure 18. A) Plot of the RMSD in C_α positions of off-state E121KNL and the monomer1 of on-state E121KNL. Secondary structures are also included within the diagram.

B) Stereo view of the overlay of off-state E121KNL and the monomer1 of on-state E121KNL. The off-state E121KNL and the on-state E121KNL are colored in yellow and black, respectively. The regions covering 5-83 and 90-134 were used for superimposition of the two structures with an RMSD of 0.87 Å. Mg (blue) and BeF_3 (red) in the on-state monomer1 were represented by ball-and-stick model.

A.


RMSD of Cα
20
18
16
14
12
10
8
6
4
2
0
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
β1
α1
β2
α2
β3
α3
β4
α4
β5
α5
Residue number

B.

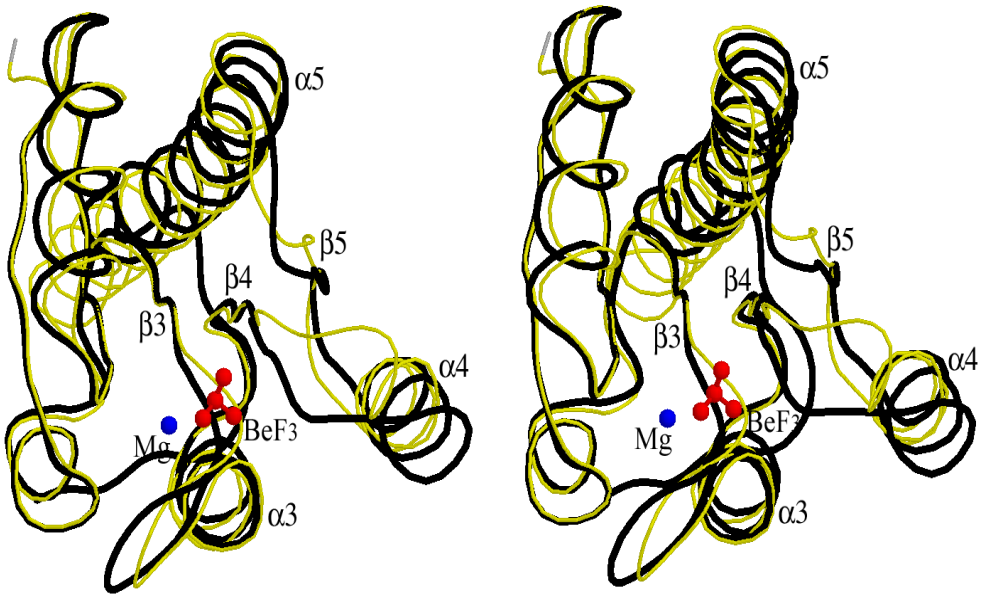
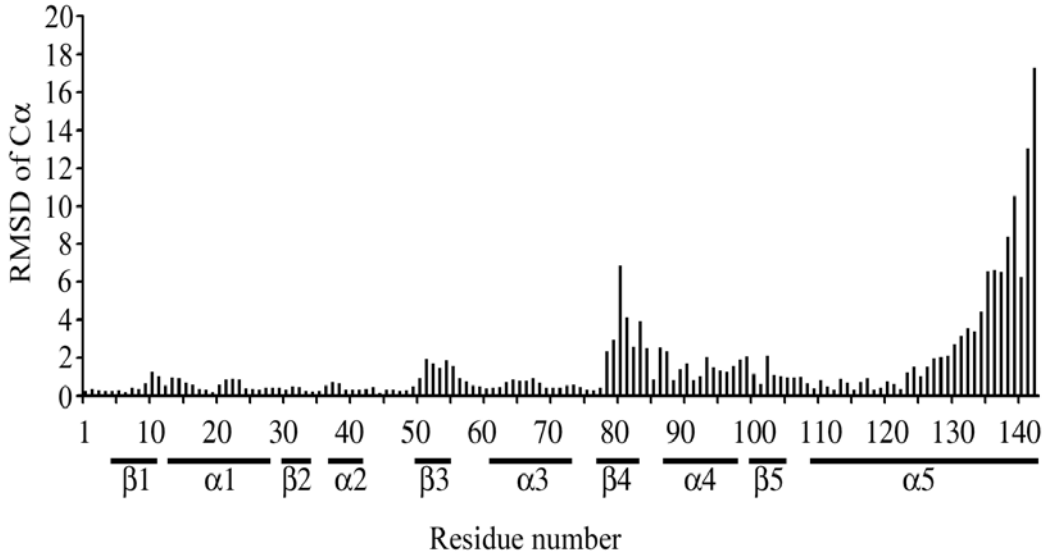

α5
β5
β4
β3
α4
Mg
BeF3
α3
α5
β5
β4
β3
α4
Mg
BeF3
α3

Figure 19. A) Plot of the RMSD in C_α positions of off-state E121KNL and the monomer2 of on-state E121KNL. Secondary structures are also included within the diagram.

B) Stereo view of the overlay of off-state E121KNL and the monomer2 of on-state E121KNL. The off-state E121KNL and the on-state E121KNL are colored in yellow and black, respectively. The regions covering 5-83 and 90-134 were used for superimposition of the two structures with an RMSD of 0.95 Å. Mg (blue) and BeF_3 (red) in the on-state monomer were represented by ball-and-stick model.

A.


RMSD of Cα
0
2
4
6
8
10
12
14
16
18
20
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
β1
α1
β2
α2
β3
α3
β4
α4
β5
α5
Residue number

B.

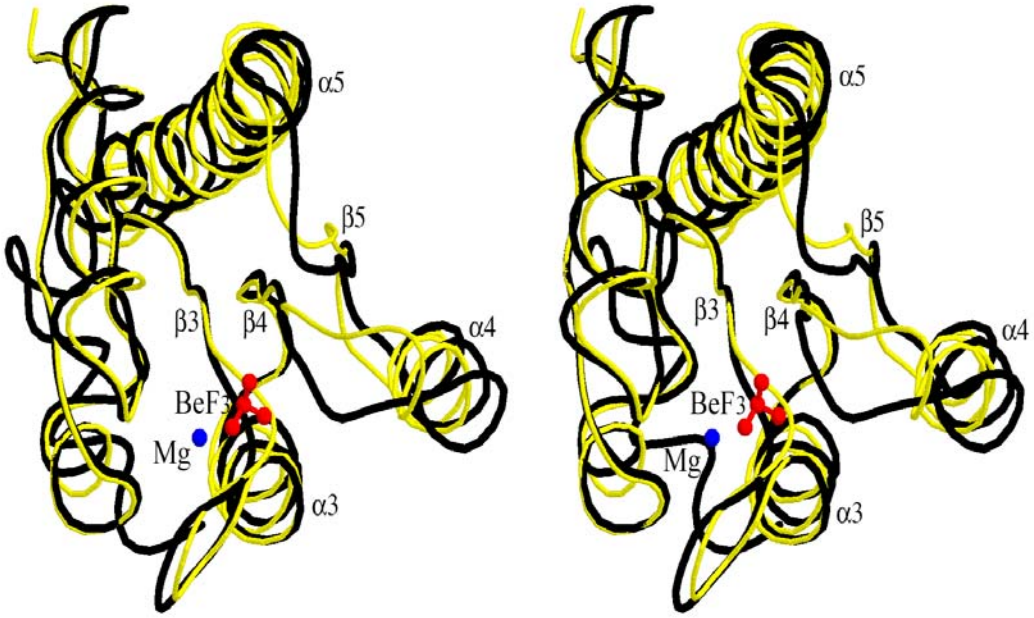
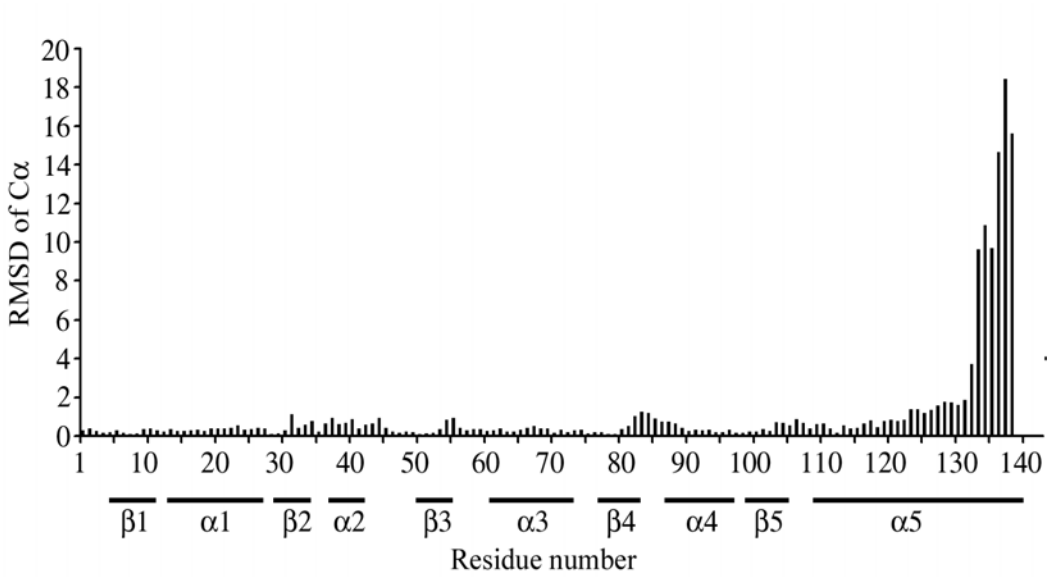

α5
β5
β3
β4
α4
BeF3
Mg
α3
α5
β5
β3
β4
α4
BeF3
Mg
α3

Figure 20. A) Plot of the RMSD in C_α positions of the two monomers of on-state E121KNL. Secondary structures are also included within the diagram.

B) Stereo view of the main-chain traces of the two monomers of on-state E121KNL. The monomer1 and the monomer2 of on-state E121KNL are colored in yellow and black, respectively. The regions covering 5-83 and 90-134 were used for superimposition of the two structures with a RMSD of 0.55 Å. Mg and BeF_3 in the monomer1 and 2 were represented by yellow and red color, respectively.

A.


RMSD of Cα
20
18
16
14
12
10
8
6
4
2
0
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
β1
α1
β2
α2
β3
α3
β4
α4
β5
α5
Residue number

B.

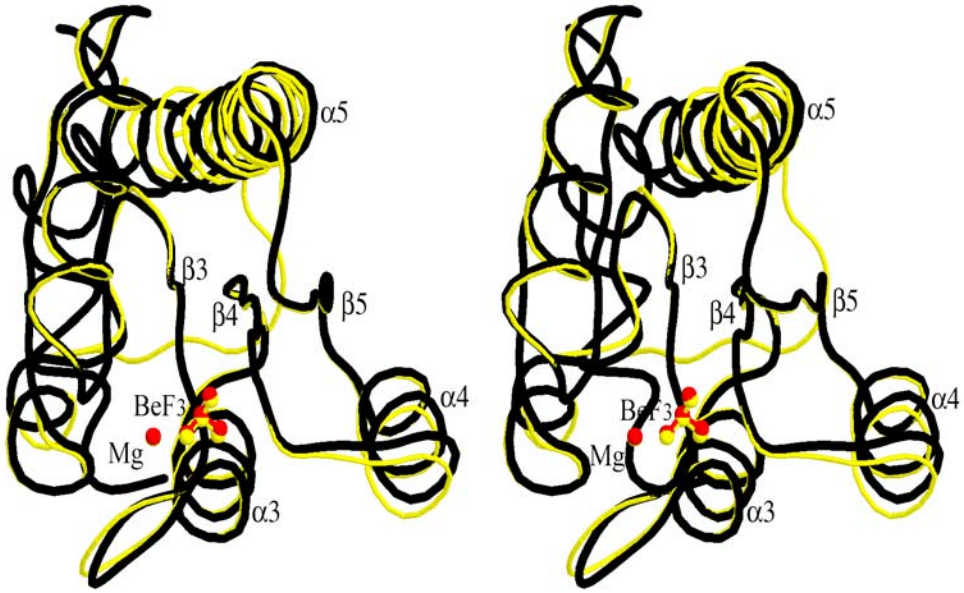

α5
β3
β4
β5
α4
BeF3
Mg
α3
α5
β3
β4
β5
α4
BeF3
Mg
α3

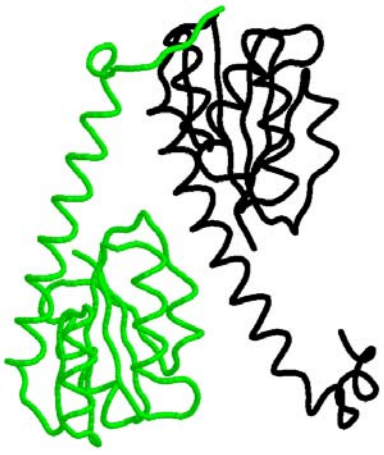
Figure 21. A) Dimer1 in the crystal lattice of E121KNL-Mg^{2+}-BeF_3^{-1}. B) Dimer2 in the crystal lattice of E121KNL-Mg^{2+}-BeF_3^{-1}. C) Dimer3 in the crystal lattice of E121KNL-Mg^{2+}-BeF_3^{-1}. D) Dimer4 in the crystal lattice of E121KNL-Mg^{2+}-BeF_3^{-1}.

A.

B.

C.

D.



The third involves the asymmetric interactions between strand α1-β2-α2 surface from one receiver domain and loop β3-α3 and top part of helix α3 from the other, burying surface area of 647 Å. The fourth is mediated with asymmetric contact between loop β5-α5 and helix α5 from one receiver domain and helix α2, loop α2-β3, and helix α5 from its partner, burying surface area of 1,760 Å. Among these, the first dimer is believed to be biological relevant one based upon the following reasons: first, it has symmetric contact; second, it has a relatively large surface area involved in the formation of dimer; third, it is the only dimer whose dimer interface is made by the signaling surface common in other two-component receiver domains, while other three dimers are using the region outside the signaling surface in forming dimer; and fourth, it has almost same dimer interface as that of FixJN~P (Fig. 22), which was proposed to be biologically relevant one (Birck et al., 1999).

Under the assumption that the first dimer is biologically relevant, the details of its interface are compared with those of the off-state's interface (Fig. 23A & B). The superimposition of one monomer of the first dimer with one monomer of the off-state dimer shows that, in the unaligned partner molecules, the C_α atoms of the active site residue Asp55 are displaced by 21.7 Å, along with about 50^o rotation from each other (Fig. 23C). This interface is maintained mainly by van der Waals interactions provided by Ile88 Cγ2 – Phe102 Cδ1 (3.69 Å), Pro89 Cβ – Arg111 NH1 (3.79 Å), Pro89 Cβ – Arg111 NH2 (3.63 Å), Ile95 Cδ1 – Val92 Cδ1 (3.98 Å), Gln Nε2 – Tyr100 O (3.28 Å), Phe102 Cδ1 – Ile 88 Cδ1 (3.80 Å), and Phe102 Cδ1 – Ile 88 Cδ1 (3.80 Å). The only hydrogen bond is formed between O of Tyr100 from loop α4-β5 and Nε2 of Gln96 from strand β5.

Figure 22. Overlay of the dimer interfaces of E121KNL-Mg^{2+}-BeF_3^{-1} and FixJN~P. Residues 87-89, 91, 92, 95, 96, 101, and 102 of E121KNL-Mg^{2+}-BeF_3^{-1} (red) were superimposed with the corresponding residues 86-88, 90, 91, 94, 95, 100, and 101 of FixJN~P (from PDB ID 1D5W, cyan) with a RMSD of 0.56 Å.

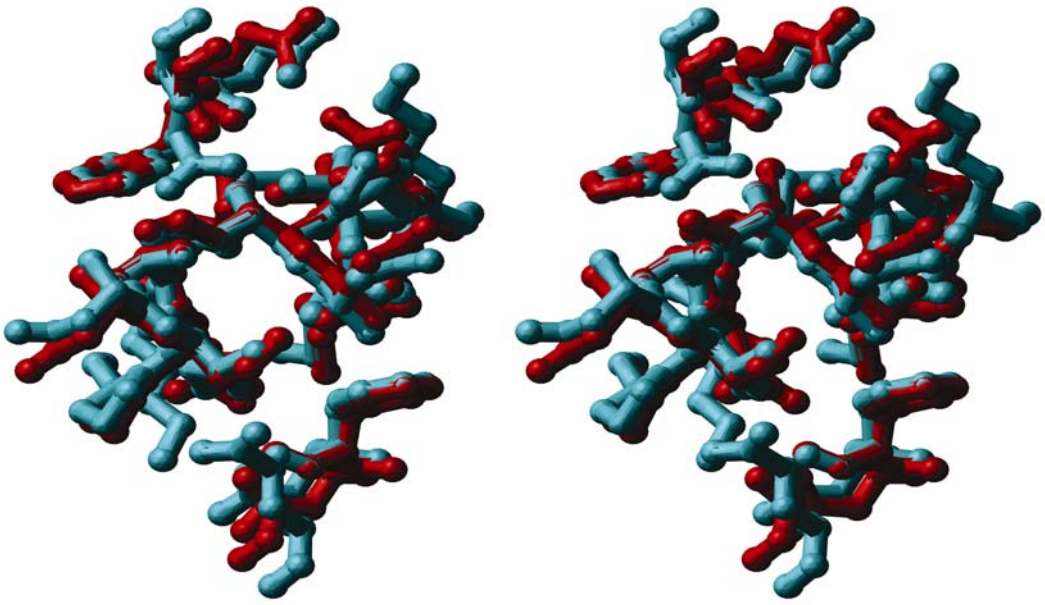
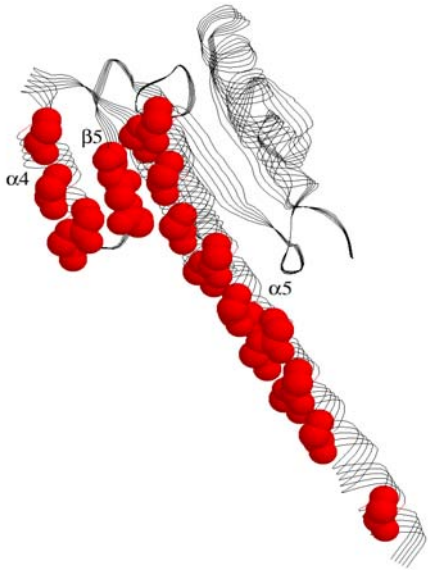
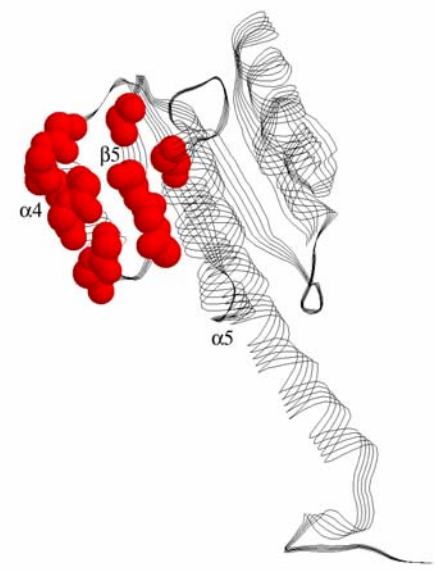


Figure 23. A) Dimer interface of off-state E121KNL. B) Dimer interface of on-state E121KNL. The main chain atoms (red color) involved in dimer interface were indicated by space-fill model. C) Switching between alternative dimers of the receiver domain upon activation. One monomer of the off-state dimer (red) has been superimposed with the on-state dimer1 (green). The regions covering 5-93 and 100-137 were used for superimposition of the two dimers with a RMSD of 1.08 Å. Side chains for residues Asp55 (or BeF_3^{-1}-Asp55), Thr83 and Phe102 are represented by wireframe, and identified by number in the partner monomers. The C_α atoms of the active site residues Asp55 are displaced by 21.7 Å in the unaligned partners.

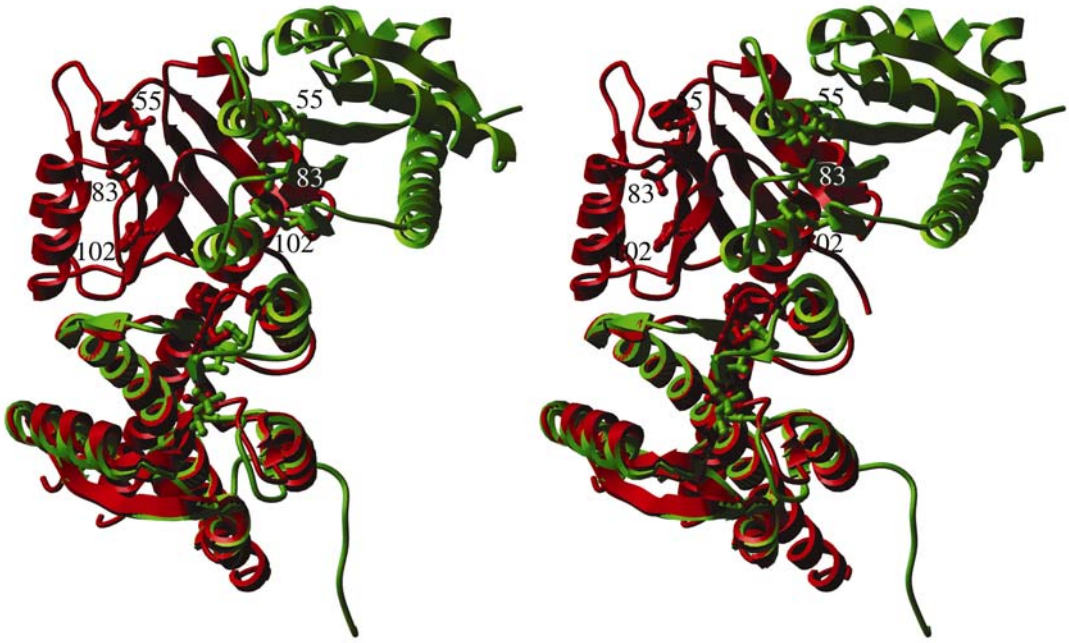
A.



B.



C.

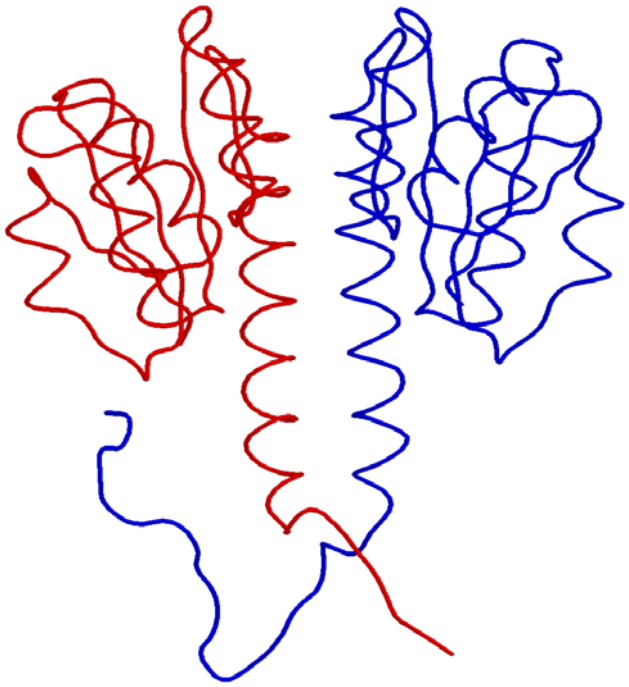


But this is significantly less than the one of the off-state dimer. The activation by BeF_3^- binding brought an increase in the radius of gyration from 19.6 Å to 22.6 Å (when calculated using residues 5-143 from the X-ray structure) as well as the retardation of movement due to increased frictional drag in sedimentation velocity experiments (Meyer et al., 2001).

The crystal structure of $DctD_{2-162}$

The crystals of $DctD_{2-162}$ belong to the trigonal space group $P3_2$ with the unit cell dimensions of a=54.4 Å, b=54.4 Å, c=103.4 Å, and $\beta=90^o$. The asymmetric unit contains two monomers. In the molecular replacement using $DctD_{5-125}$ as a search model, the best solutions for a rotational function for each of two molecules in the asymmetric unit were found at ($\alpha = 46.37^o$, $\beta = 131.73^o$, $\gamma = 25.63^o$) and at ($\alpha = 79.05^o$, $\beta = 49.10^o$, $\gamma = 198.20^o$), respectively. The best solutions of a translational function were found at (Tx = 0.7941, Ty = 0.2941, Tz = 0.0000, R = 50, and CC_F = 35.3) and at (Tx = 0.2998, Ty = 0.7940, Tz = 0.1691, R = 44.3, and CC_F = 50.8), respectively. The electron density map was not clear beyond residue 134, suggesting the presence of a mobile random coil instead of a well-defined secondary structure. As expected, the sequential addition of alanine after residue 134 showed that the helical structure seen in the DctDNL was turned into a random coil after residue 132 in $DctD_{2-162}$ (Fig. 24). In model builds and refinements using reflection data greater than 2σ, one finds no novel interactions between the N-terminal region of the central domain and the receiver domain, along with much higher B-factors. In addition to this, as the fact that the $DctD_{2-162}$ crystal was suffering from the packing artifacts became clear, I decided not to try model builds and refinements using the whole reflection data.

Figure 24. Strand-type representation of the dimeric structure of $DctD_{2\text{-}162}$



At this point, the one chain was extended to residue 143, while the other was to residue 156. The incomplete final structure of $DctD_{2\text{-}162}$ contains 2,162 non-hydrogen protein atoms. At this point, the crystallographic R factor and the free R factor were 20.86 % and 25.78 %, respectively (Table 2).

In vivo beta-galactosidase assays of $DctD_{2\text{-}384}$ and $E121K_{2\text{-}384}$

Due to the difficulties to crystallize the full-length DctD, we decided to remove the flexible DNA binding domain. Previous *in vivo* function assay of deletion mutants showed that the C-terminal boundary of the ATPase domain was between residue 381 and 384 (Gu et al., unpublished). So we constructed a recombinant gene encoding from residue 1 to 384 and tested the activity of the resulting product using the ϕ(*dctA´-lacZ´*)(Hyb) reporter gene. Like the full-length DctD, $DctD_{2\text{-}384}$ was expressed well and was soluble, but it was unable to activate transcription of *lacZ* under control of the *dctA* promoter. However, when the E121K substitution was introduced into $dctD_{1\text{-}384}$, the new recombinant protein stimulated the expression of β-galactosidase. The activity of $E121K_{2\text{-}384}$ was about 2.5 % of that of the full-length E121K, which has an intact DNA binding domain (Fig. 25A and B). This means that $DctD_{2\text{-}384}$ is still regulated by the two-component receiver domain. But several attempts to purify $DctD_{2\text{-}384}$ had not been successful, so I introduced a $(His)_6$ tag into its C-terminus to make purification much easier. Removal of the DNA binding domain made the protein more soluble than the full-length protein. Dynamic light scattering measurements (Dynapro) clearly indicated that the protein exists as a stable, homogeneous tetramer at the concentration greater than 0.3 mg/ml. But it has not been crystallized in any of the conditions I have tried so far.

Table 2. Crystallographic Data Statistics

Crystallographic data				
Data Set	**DCTDNL-SeMet[1]**	**E121KNL**	**Mg^{2+}-BeF_3^{-1}-E121KNL**	**$DctD_{1-162}$**
Data collection statistics				
Space group	I222	I222	$P2_12_12_1$	$P3_2$
Unit cell				
a (Å)	58.65	57.1	52	54.4
b (Å)	58.77	58.1	76.8	54.4
c (Å)	167.89	168.8	77.8	103.4
β (°)	90	90	90	90
Data collection				
Temperature (K)	100	111	111	295 or 111
Completeness /last shell (%)	99.6	98.9/98.1	99.3/100	93.7
Unique reflections	61,207	19,264	18,672	11,121
Measured reflections	428,225	59,135	69,850	20,738
Redundancy	7	3.07	3.74	1.86
Signal (=I/σ(I))	8.3	7.1	4.9	7.4
R_{merge}[2]	0.059	0.061	0.105	0.059
Resolution range (Å)	1.7	2.0	2.1	2.5
Asymmetric unit content	1	1	2	2
Solvent content (%)	71	70	46	46.8
V_m	4.31	4.17	2.32	2.315
Refinement statistics				
Model				
σ cut-off (=F/σ(F))	0	0	0	0
R-factor	0.218	0.196	0.184	0.209
R-free	0.241	0.232	0.229	0.258
Test-set size (%)	9.9	9.8	9.9	
No. of protein atoms	1,065	1,105	2,284	2,162
No. of waters	212	144	224	n.a.
Stereochemistry[3]				
RMS deviation from ideality				
bond length (A)	0.0218	0.015	0.014	n.a.
bond angles (°)	1.8	1.7	1.7	n.a.
Ramachandran[2] (%)				
most favorable	99.2	96.9	94.6	n.a.
additionally allowed	0.8	3.1	5.4	n.a.
generously allowed	0.0	0.0	0.0	n.a.
disallowed	0.0	0.0	0.0	n.a.
Average B-factors				
main chain monomer 1	23.69	19.6	15.5	n.a.
side chain monomer 1	29.22	22.03	18.08	n.a.
main chain monomer 2	n.a.	n.a.	16.6	n.a.
side chain monomer 2	n.a.	n.a.	19.4.	n.a.

[1] quoted from Nixon's NIH grant proposal.

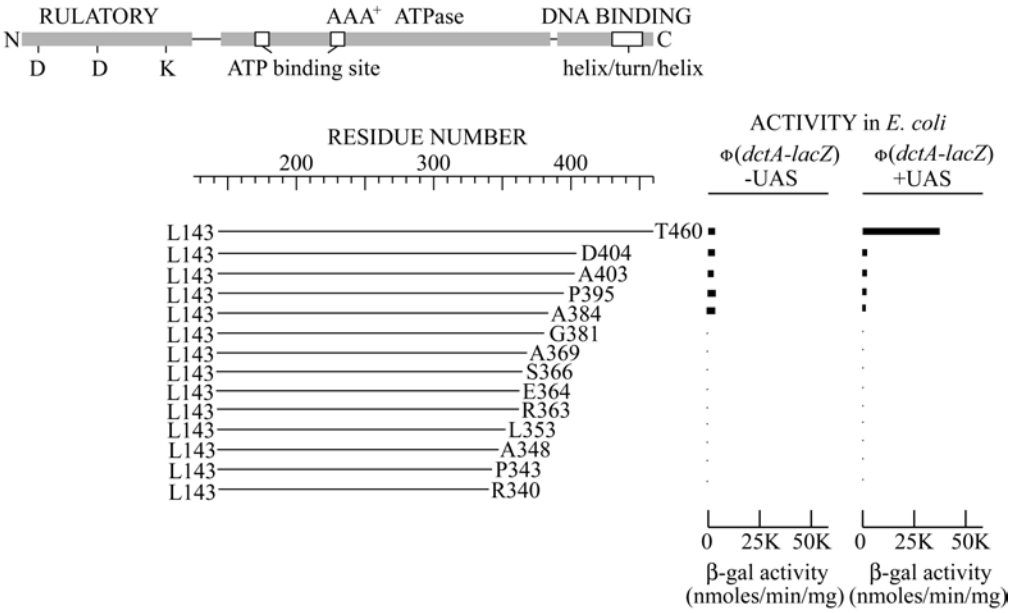
[2] $R_{merge} = \Sigma \|I_i - \langle I \rangle\| / \Sigma \langle I \rangle$, where I_i is the measured intensity of the reflection i and <I> is the averaged intensity of the multiple measurements of the same reflection.

[3] stereochemical check by PROCHECK

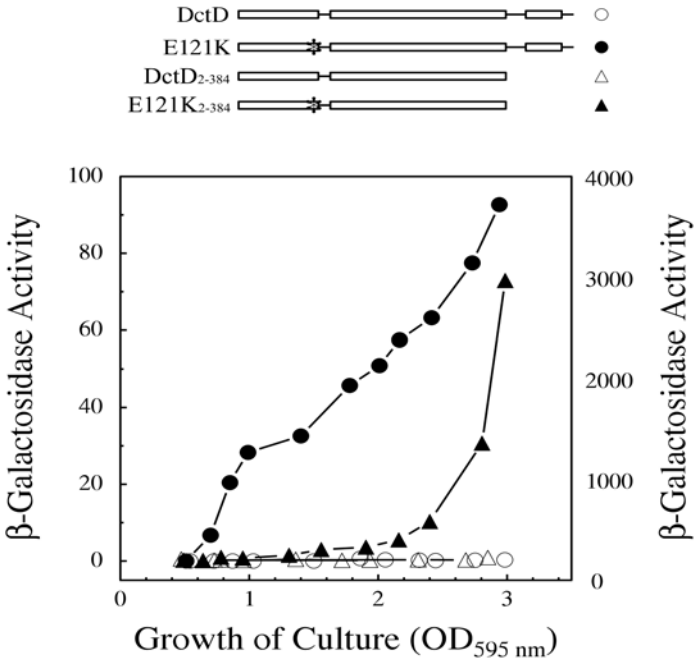
Figure 25. A) UAS-binding and transcriptional activities of truncated DctD proteins. The recombinant gene *dctD*Δ(1-142) was previously shown to encode a constitutively active deletion mutant, $DctD_{143-460}$. A series of deletions in the 3′ coding region of *dctD*Δ(1-142) were generated. The resulting genes were expressed in overnight cultures of *E. coli* strain JM109 to determine their ability to activate transcription of Φ(*dctA-lacZ*) reporter genes with or without the *dctA* UAS being present. Background β-galactosidase activity supported by the reporter gene alone was less than 10 units (Gu et al., 1994).

B) Transcription activation activities of DctD, $DctD_{2-384}$, E121K and $E121K_{2-384}$ in *E. coli*. The E. coli strain containing the *dctA-lacZ* reporter gene was transformed with pBTN1, encoding DctD (open circles), or its recombinant variants encoding E121K (closed circles), $DctD_{2-384}$ (open triangles) or $E121K_{2-384}$ (closed triangles). Cells grown at 32 ^{o}C were harvested for β-galactosidase activity assays. The values for E121K are plotted versus the right axis, while all others were plotted versus the left axis (Park et al., unpublished).

A.


RULATORY
AAA+ ATPase
DNA BINDING
N
C
D
D
K
ATP binding site
helix/turn/helix
RESIDUE NUMBER
200
300
400
ACTIVITY in E. coli
Φ(dctA-lacZ) -UAS
Φ(dctA-lacZ) +UAS
L143 T460
L143 D404
L143 A403
L143 P395
L143 A384
L143 G381
L143 A369
L143 S366
L143 E364
L143 R363
L143 L353
L143 A348
L143 P343
L143 R340
0 25K 50K
0 25K 50K
β-gal activity (nmoles/min/mg)
β-gal activity (nmoles/min/mg)

B.


DctD
E121K
DctD2-384
E121K2-384
β-Galactosidase Activity
β-Galactosidase Activity
Growth of Culture (OD595 nm)
0
20
40
60
80
100
0
1000
2000
3000
4000
0
1
2
3

DISCUSSION

The pattern of tryptic digestion suggests that there are conformational changes in the 44 kD tryptic fragment upon ATP binding, which increases sensitivity of the linker region between the N-terminus and central domain of DctD while making the putative ATP binding site in the central domain less susceptible to trypsin. The C-terminal DNA binding domain seems to have a loosened structure because it disappeared most quickly during tryptic digestions whether ATP was present or not.

Evidence for conformational changes in an unphosphorylated, intact DctD upon ATP binding was provided by a preliminary transient kinetic experiment using mant-ATP, a fluorophore-containing ATP analog. The result of this experiment clearly indicated that the binding of DctD to ATP was a muti-step process, each step causing a difference in the intensity of fluorescence emission. The structural bases for such changes still remain unknown. These observations were consistent with the results of the tryptic digestion experiments on the 44 kD fragment. I did not pursue kinetic studies on DctD any further because of the several problems mentioned earlier.

Substitution E121K or substitution K122E caused a significant increase in CD signal in the full-length protein but not in its N-terminal fragment. Such an increase, which is a strong indicator for a conformational change, might be induced either by a change in tertiary interactions or by a change in protein-protein associations or both. The nature of conformational changes has to be further clarified by other studies such as analytical ultracentrifugation, light scattering, X-ray crystallography, etc.

In the meantime, the crystal of E121KNL in complex with Mg^{2+} and BeF_3^{-1} was obtained and its crystal structure was determined. This made it possible to compare the

crystal structure of the off-state E121KNL dimer with that of its active form, Mg^{2+}-BeF_3^{-1}-E121KNL dimer. Although there have been a few reports of metal-bound structures of two-component receiver domains in their active states such as Ca^{2+}-Spo0AN-P (Lewis et al., 1999) and Mn^{2+}-CheY-BeF_3^{-1} (Lee et al., 2001), this is only the second receiver domain whose structure has been determined for BeF_3^{-1} and Mg^{2+} bound to active site; CheY-FliM was the first (Lee et al., 2001). The active site features seen in prior studies on receiver domains were clearly present in the structure of Mg^{2+}-BeF_3^{-1}-E121KNL. This study confirms a common switching mechanism prevalent among two-component receiver domains, along with two new insights: that is, how a specific dimeric structure can maintain the highly conserved lysine and aspartate in a state that is refractory to phosphorylation; and how the active site, when activated, can be stabilized in a dramatically different dimeric state to regulate the function of an associated output domain. The orientations of the conserved residues D55, T83, and F102 in E121KNL are very similar to those found in the FixJN-P receiver domain, and similar but less so to those present in the activated forms of the CheY and NtrC receiver domains.

One of the dimers in the Mg^{2+}-BeF_3^{-1}-E121KNL crystal lattice is almost the same as the FixJN-P dimer that is believed to be biologically relevant. In FixJN-P, amino acid substitutions mapping in the α4 portion of the dimer interface prevent activation *in vivo* and dimerization in solution, both suggesting biological relevance of that particular dimer (Da Re S et al., 1999). In addition, since the α4-β5 region involved in the formation of that dimer is believed to be part of the signaling surface common among two-component systems, it is reasonable to assume that the FixJN-P-like dimer in the E121KNL crystal has biological relevance. However, it is still possible that the activated E121KNL dimer

described above is an artifact of having isolated the receiver domain and linker from the ATPase and DNA binding domains of DctD. It is also premature to dismiss the other two dimeric forms seen in the crystal lattice of Mn^{2+}-BeF_3^{-1}-DctDNL, as genetic and biochemical studies have not been performed to test the presence of these structures in solution.

There is evidence of multiple conformations in both states of the DctDNL fragment (Park et al, submitted). The relevance of each of these conformations to signal transduction remains to be determined. Recent NMR studies on NtrCN and NtrCN-P emphasize the presence of a preexisting equilibrium between fairly homogeneous inactive and active conformations that is pushed toward the active state by phosphorylation (Volkman et al., 2001). Similarly, comparison of CheY and Mn^{2+}-BeF_3^--CheY revealed multiple conformations for β4/α4 loop (Lee et al., 2001). A 1.08 Å crystal structure for CheY revealed a mixture of two conformations, with one mimicking the active state in many but not all features. It was suggested that this structure might be responsible for basal signaling in the absence of phosphorylation of CheY (Simonovic et al., 2001).

Even though the structural change in the hydrophobic α4-β5 surface appears to form a switch alternating between repressing and stimulating the ATPase domain, how it performs that role in the context of the intact protein remains to be answered. The switch may involve direct contact between the receiver and ATPase domains within the same or another protein. Or it may play that role in a passive manner, simply by freeing the ATPase from repressive contact. In FixJ, alteration of the signaling surface by phosphorylation brings a transition from a monomeric receiver domain to a dimeric one

(Da Re S et al., 1999). In NtrC, a σ54-dependent activator like DctD, it is suggested that the α4 region of the receiver domain of one monomer interacts with an AAA^+ ATPase domain of another (Lee et al., 2000). So signaling seems to involve the alteration of a hydrophobic α4-β5 surface in such a way that new types of protein-protein interactions or domain-domain interactions might be stabilized.

As part of the effort to understand signal propagation from the activated receiver domain to the ATPase domain in the same or different subunit, I made an extended version of DctD receiver domain called $DctD_{2\text{-}162}$, which contained the N-terminal portion of the ATPase domain. It was crystallized, but there were no novel interactions between the receiver domain and the αN region of the ATPase domain. I then tried to crystallize $DctD_{2\text{-}384}$ that contains the receiver domain and the ATPase domain. But I have not been successful in crystallizing $DctD_{2\text{-}384}$ so far.

Here I hypothesize a 'working' model of how DctD is activated by its phosphorylation. In this model, the receiver domain of DctD is maintained in equilibrium between distinct dimeric forms: unphosphorylatable form and phosphorylatable form. The unphosphorylatable form consists of a stable dimer in which the receiver domain is refractory to phosphorylation and inhibits oligomerization of the ATPase, which keeps the ATPase from being active. The change into a phosphorylatable, dimeric form involves dissociation of an unphosphorylatable, dimeric form to its monomeric form. So another equilibrium between an unphosphorylatable, monomeric form and a phosphorylatable, monomeric form of the receiver domain may be included in this model, even though it has a very transient nature. Stabilization of a phosphorylatable, dimeric form of the receiver domain by phosphorylation or BeF_3^- binding drives the equilibrium

to favor the phosphorylatable, dimeric form of the receiver domain, which fails to inhibit or even facilitates oligomerization of the ATPase, probably hexamerization. According to this model, one can expect that either destabilizing an unphosphorylatable, dimeric form of the receiver domain or stabilizing a phosphorylatable, dimeric form of the receiver domain will increase the activity of DctD. The overall process is briefly described in the following equation.

$$D \rightleftharpoons (M \rightleftharpoons M^*) \rightleftharpoons D^* \rightleftharpoons H\text{-}P^{**}$$

where D is an unphosphorylatable, dimeric form of the receiver domain, M is an unphosphorylatable, monomeric form of the receiver domain, M* is a phosphorylatable, monomeric form of the receiver domain and D* is a phosphorylatable, dimeric form of the receiver domain, respectively. H-P** denotes an active, hexameric form of the full-length DctD. Monomeric species are included in parenthesis to denote their transient nature.

Several lines of evidence support or at least are consistent with this hypothesis. Firstly, both active forms and inactive forms of DctDNL exist as a stable dimer in solution, suggesting that monomeric forms, if present, should be short-lived. Secondly, two mutants that have amino acid substitutions in the dimer interface of the receiver domain showed an increase in activity as well as in dissociation constant. This indicates that dissociation into monomers is closely related to being activated. The fact that the structure of activated E121KNL has a completely different dimer interface compared to E121KNL also supports the presence of monomeric intermediates of the receiver domain. Thirdly, DctDNL and E121KNL in the I222 crystal lattice cannot be phosphorylated by organic phospho-donors such as carbamyl-phosphate, even though both proteins are

phosphorylatable in solution. The lattice forces of DctDNL or E121KNL crystals seem to stabilize an inactive form that is refractory to phosphorylation. This is supported by the fact that the structures of both proteins in the crystals turned out to be basically the same. Fourthly, CD experiments under denaturing conditions using heat or GdmCl indicated that there is more than one conformation in both states of DctDNL (Park et al., submitted). Recent X-ray crystallographic studies of CheY revealed that both an inactive form and a meta-active form are present in the apo-CheY population (Simonovic et al., 2001). These observations are consistent with this model in that the receiver domain assumes multiple conformations in both states. Lastly, the class serving as molecular mechanoenzymes among AAA^+ superfamily often assembles into a hexameric structure to remodel DNA-protein complexes (Neuwald et al., 1999). Since DctD belongs to the same class, it is likely to form a hexamer. However, it should be noted that this model exploits some experimental results of other related proteins rather than DctD. So this model should be assessed by further studies on DctD and modified accordingly.

REFERENCES

Baikalov, I., I. Schroeder, M. Kacor-Grzeskowiak, K. Grzeskowiak, R.P. Gunsalus, and Dickerson, R.E. (1996). Structure of the *Escherichia coli* Response Regulator NarL. *Biochemistry* 35: 11053-11061.

Baikalov, I., Schroder, I., Kaczor-Grzeskowiak, M., Cascio, D., Gunsalus, R.P., Dickerson, R.E. (1998). NarL dimerization? Suggestive evidence from a new crystal form. *Biochemistry* 37: 3665-3676.

Bantscheff, M., Weiss, V., and 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.

Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, G., Rousseau, P., Kahn, D., and Samama, J.-P. (1999). Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure* 7: 1505–1515.

Bolton, E., Higgisson, B., Harrington, A., and O'Gara, F. (1986). Dicarboxylic acid transport in *Rhizobium meliloti*: isolation of mutants and cloning of dicarboxylic acid transport genes. *Arch. Microbiol*. 144: 142-146.

Brunger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355: 472-475.

Buck, M., and Cannon, W. (1992). Specific binding of the transcription factor sigma-54 to promoter DNA. *Nature (London)* 358: 422-424.

Cho, H.S., Lee, S.-Y., Yan, D., Pan, X., Parkinson, J.S., Kustu, S., Wemmer, D.E., and Pelton, J.G. (2000). NMR structure of activated CheY. *J. Mol. Biol*. 297: 543-551.

Da Re, S., Schumacher J, Rousseau P, Fourment J, Ebel C, Kahn D (1999). Phosphorylation-induced dimerization of the FixJ receiver domain. *Mol. Microbiol*. 34 (3): 504-11.

Djordjevic, S., Goudreau, P.N., Xu, Q., Stock, A.M., West, A.H. (1998). Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. *Proc. Natl. Acad. Sci. USA* 95: 1381-1386.

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

Dubendorff, J. W. and Studier, F. W. (1991). Creation of a T7 autogene: Cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promoter. *J. Mol. Biol.* 219: 45-59.

Engelke, T., Jagadish, M. N., and Puhler, A. (1987). Biochemical and genetic analysis of *Rhizobium meliloti* mutants defective in C_4-dicarboxylate transport. *J. Gen. Microbiol*. 133: 3019-3029.

Feng, J., M.A. Atkinson, W. McCleary, J.B. Stock, B.L. Wanner, and A.J. Ninfa (1992). Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in *Escherichia coli. J. Bacteriol.* 174: 6061-6070.

Finan, T., Wood, J. M., and Jordan, D. C. (1981). Succinate transport in *Rhizobium Leguminosarum*. *J. Bacteriol*. 148: 193-202.

Fischer, H.-M. (1994). Genetic regulation of nitrogen fixation in *Rhizobia. Microbiological Reviews.* 58: 352-386.

Fisher, R.F., and S.R. Long (1989). DNA footprinting analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible nod gene promoters. *J. Bacteriol*. 171: 5492-5502.

Flashner, Y., D.S. Weiss, J. Keener, and S. Kustu (1995). Constitutive forms of the enhancer-binding protein NtrC: evidence that essential oligomerization determinants lie in the central activation domain. *J. Mol. Biol.* 249: 700-713.

Gardiol, A., Arias, A., Cervenansky, C., and Martinez-Drets, G. (1982). Succinate dehydrogenase mutants of *Rhizobium meliloti. J. Bacteriol*. 151:1621-1623.

Giblin L., B. Boesten, S. Turk, P. Hooykaas, and F. O'Gara (1995). Signal transduction in the *Rhizobium meliloti* dicarboxylic acid transport system. *FEMS Microbiol. Let.* 126: 25-30.

Gu, B., J. H. Lee, T. R. Hoover, D. Scholl and B. T. Nixon (1994). *Rhizobium meliloti* DctD, a σ^{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.

Halkides, C.J., McEvoy, M.M., Casper, E., Matsumura, P., Volz, K., and Dahlquist, F.W. (2000). The 1.9 Å resolution crystal structure of phosphono-CheY, an analogue of the active form of the response regulator, CheY. *Biochemistry* 39: 5280-5286.

Hiatsuka, T. (1983). New ribose-modified fluorescent analogs of adenine and guanine nucleotides available as substrates for various enzymes. *Biochim. Biophys. Acta* 742: 196-508.

Hill, S. (1992). Physiology of nitrogen fixation in free-living heterotrophs, p. 87-134. in G. Stacey, R. H. Burris, and H.J. Evans (ed.), *Biological nitrogen fixation*. Chapman & Hall, New York.

Honeycutt, R. J., M. McClelland, and B. Sobral (1993). Physical map of the genome of *Rhizobium meliloti* 1021. *J. Bacteriol*. 175: 6945-6952.

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

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.

Jancarik, J., and Kim, S. H. (1991). Sparse matrix sampling: A screening method for crystallization of proteins. *J. Appl. Cryst*. 24: 409-411.

Jiang, J., Gu, B, Albright , L.M., and Nixon, B.T. (1989). Conservation between coding and regulatory elements of *Rhizobuim meliloti* and *Rhizobium leguminosarum dct* genes. *J. Bacterio.* 171: 5244-5253.

Johnson, M.L., and Faunt, L.M. (1992). *Meth*. *Enzymol*. 210: 1-37.

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

Kern, D., Wemmer, D., Kustu, S. (1999). Presented at the FASEB meeting on transcription initiation in prokaryotes, July 1999.

Klose, K.E., Weiss, D.S., and Kustu, S. (1993). Glutamate at the site of phosphorylation of nitrogen-regulatory protein NTRC mimics aspartyl-phosphate and activates the protein. *J Mol Biol.* 232: 67-78.

Kutsu, S., A. K. North, and D. S. Weiss (1991). Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends biochem. Sci.* 16: 397-402.

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. Cryst.* 26, 283-291.

Ledebur, H., Gu, B., Sodja III, J., and Nixon, B. T. (1990). *Rhizobium meliloti* and *Rhizobium leguminosarum dctD* gene products bind to tandem sites in an activation sequence located upstream of σ^{54}-dependent *dctA* promoter. *J. Bacteriol*. 172: 3888-3897.

Lee, J. G., J. T. Owens, I. G. Hwang, C. Meares, and S. Kustu (2000). Phosphorylation-induced signal propagation in the response regulator NtrC. *J. Bacteriology* 182: 5188-5195.

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

Lee, J., Owens, J.T., Hwang, I., Meares, C., and Kustu, S. (2000). Phosphorylation-induced signal propagation in the response regulator NtrC. *J. Bacteriol.* 182: 5188-5195.

Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L. S., Kustu, S., Berry, E. A., Wemmer, D. E. (2001). Crystal structure of an activated response regulator bound to its target *Nat. Struct. Biol.* 8: 52-56.

Lehninger, A. L., D. L. Nelson, and M. M, Cox (1993). *Principles of Biochemistry 2nd edition*. p 690.

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.

Lukat, G. S., McCleary, W.R., Stock, A.M., Stock, J.B. (1992). Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. USA* 89: 718-722.

Madhusudan, M., J. Zapf, J.A. Hoch, J.M. Whiteley, N.H. Xuong and K.I. Varughese (1997). A response regulatory protein with the site of phosphorylation blocked by an arginine interaction: crystal structure of Spo0F from *Bacillus subtilis. Biochemistry* 36: 12739-12745.

McEvoy, M.M., Anat Bren, A., Eisenbach, M., Dahlquist, F.W. (1999). Identification of the binding interfaces on CheY for two of its targets, the phosphatase CheZ and the flagellar switch protein FliM. *J. Mol. Biol.* 289: 1423-1433.

Meyer, M., S. D. Park, L. Zeingue, M. Staley, M. McKinstry, R. I. Kaufman, H. Zhang, D. Yan, N. Yennawar, H. Yennawar, G. K. Farber, and B. T. Nixon (2001). A dimeric two-component receiver domain inhibits the σ^{54}-dependent ATPase in DctD. *The FASEB Journal Express Article* 10.1096/fj.00-0516fje.

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

Moy, F. J., D. F. Lowry, P. Matsumura, F. W. Dahlquist, J. E. Krywko, and P. J. Domaille (1994). Assignments, secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (^{13}C, ^{15}N) NMR spectroscopy. *Biochemistry* 33: 10731-10742.

Moyer, M. L., S P. Gilbert, and K. A. Johnson (1998). Pathway of ATP hydrolysis by monomeric and dimeric kinesin. *Biochemistry* 37: 800-813.

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

Nixon, B. (2000). Coiled-coil linkers and two component response regulators.

http://www.bmb.psu.edu/nixon/coils.

Nohaile, M., Kern. D., Wemmer. D., Stedman. K., Kustu. S. (1997). Structural and functional analyses of activating amino acid substitutions in the receiver domain of NtrC: evidence for an activating surface. *J Mol. Biol.* 273: 299-316.

Osuna, J., Xavier S., and Enrique M. (1997). A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. *Protein Science* 6: 543-555.

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.

Pieternel van Rhijn and Jos Vanderleyden (1995). The Rhizobium-Plant symbiosis. *Microbiological Reviews*. 59: 124-142.

Porter, S. C., A. K. North, A. B. Wedel, and S. Kustu (1993). Oligomerization of NtrC at the *glnA* enhancer is required for transcriptional activation. *Genes Dev*. 7: 2258-2273.

Reid, C. J., and P. S. Poole (1999). Roles of DctA and DctB in signal detection by the dicarboxylic acid transport system of *Rhizobium leguminosarum*. *J. Bacteriol.* 180: 2660-2669.

Ronson, C. W., and Astwood, P. M. (1985). In *Nitrogen Fixation Research Progress*. Evans, Bottomly, and Newton (eds.): 201-207, Martinus Nijhoff, Dordrecht.

Ronson, C. W., Astwood, P. M., and Downie, A. (1984). Molecular cloning and genetic organization of C_4-dicarboxylate transport genes from *Rhizobium leguminosarum*. *J. Bacteriol*. 160: 903-909.

Ronson, C. W., Astwood, P. M., Nixon, B. T., and Ausubel, F. W. (1987). Deduced products of C_4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen gene products. *Nucleic Acids Res*. 15: 7921-7934.

Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel (1987b). *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. *J. Bacteriol*. 169: 2424-2431.

Ronson, C. W., Lyttleton, P., and Robertson, J. G. (1981). C_4-dicarboxylate transport mutants of *Rhizobium trifoli* form ineffective nodules on *Trifolium repens*. *Proc. Natl. Acad. Sci. USA* 78: 4284-4288.

Ronson, C. W., P. Astwood, B. T. Nixon, and F. Ausubel (1987a). Deduced products of C_4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products. *Nucleic Acids Res*. 15: 7921-7934.

Sasse-Dwight, S., and J. D. Gralla (1990). Role of eukaryotic-type domains found in the prokaryotic enhancer receptor factor σ^{54}. *Cell* 62: 945-950.

Schägger, H., and g. v. Jagow (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem*. 166: 368-379.

Schmid, F. X. (1989). *Protein structure : A Practical Approach* p. 283, IRL, Oxford.

Scholl, D., and B. T. Nixon (1996). Cooperative biding of DctD to the dctA upstream activation sequence of *Rhizobium meliloti* is enhanced in a constitutively active truncated mutant. *J. Biol. Chem*. 271: 26435-26442.

Shingler V, Pavel H. (1995). Direct regulation of the ATPase activity of the transcriptional activator *DmpR* by aromatic compound. *Mol. Microbiol*. 17: 505-513.

Siminovic, M. and K. Volz (2001). A distinct meta-active conformation in the 1.1 Å resolution structure of wild-type apoCheY. *J.Biol.Chem.* 276: 28637-28640.

Sol, M., F.X. Gomis, L. Serrano, A. Gonzalez and M. Coll (1999). Three-dimensional crystal structure of the transcription factor PhoB receiver domain. *J Mol Biol* 285: 675-687.

Staley, M., Zeringue, L.C., Kidd, R.D., Farber, G.K., Nixon, B.T. (1998). Crystallization and characterization of the *Rhizobium meliloti* DctD two-component receiver domain. *Acta Crystallographica, D* 54(2 (Pt 6)): 1416-8.

Stock, A. M., J. M. Muttonnen, J. B. Stock, and C. E. Schutt (1989). Three dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature* 337: 745-749.

Stock, J. B., A. J. Ninfa, and A. M. Stock (1989). Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiological Reviews.* 53: 450-490.

Volkman BF, Lipson D, Wemmer DE, Kern D (2001). Two-state allosteric behavior in a single-domain signaling protein. *Science* 23: 2429-33.

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

Volz, K., and P. Matsumura (1991). Crystal structure of *Escherichia coli* CheY. *J. Biol. Chem.* 266: 15511-15519.

Watson, R. J., Y.-K. Chan., R. Wheatcroft, A.-F. Yang, and S. Han (1988). *Rhizobium meliloti* genes required for C_4-dicarboxylate transport and symbiotic nitrogen fixation are located on a megaplasmid. *J. bacterial*. 170: 7486-7489.

Woodward, S. K. A., J. F. Eccleston, and M. A. Geeves (1991). Kinetics of the interaction of 2′(3′)-O-(N-methylanthraniloyl)-ATP with myosin subfragment 1 and actomyosin subfragment 1: Characterization of two acto.s1.ADP complexes. *Biochemistry* 30: 422-430.

Wyman, C., Rombel, I., North, A. K., Bustamante, C., and S. Kutsu (1997). Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* 275: 1658-1661.

Yan, D., Cho, H.S., Hastings, C.A., Igo, M.M., Lee, S.Y., Pelton, J.G., Stewart, V., Wemmer, D.E., and Kustu, S. (1999). Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators. *Proc. Natl. Acad. Sci. U S A* 96: 14789-14794.

Yarosh, O.K., Charles, T. C., and Finan, T. M. (1989). Analysis of C_4-dicarboxylate genes in *Rhizobium meliloti. Mol. Micro*. 3(6): 813-823.

Zhu, X., Rebello, J., Matsumura, P., and Volz, K. (1997). Crystal structures of CheY mutants Y106W and T87I/Y106W. CheY activation correlates with movement of residue 106. *J Biol Chem* 272: 5000-5006.

Vita

PERSONAL INFORMATION

- Name: Park, Sungdae
- E-mail: sxp27@psu.edu

EDUCATION

- August, 1994 – present: The Pennsylvania State University
  Ph.D. in Biochemistry, Microbiology, and Molecular Biology
- March, 1989 – August, 1991: Kyungpook National University, Korea
  M.S. in Microbiology
- March, 1985 – February, 1989: Kyungpook National University, Korea
  B.S. in Microbiology

TEACHING EXPERIENCE

- Biochemistry and Molecular Biology 102 (A. M. Daniels, Instructor)
- Biochemistry and Molecular Biology 443W (A. M. Daniels, Instructor)

DISSERTATION

- Title: Structural studies on *Sinorhizobium meliloti* DctD related to ATP binding and activation (Advisor: Dr. B. Tracy Nixon)

PUBLICATIONS

- Matthew Meyer, **Sungdae Park**, Lori Zeringue, Mark Staley, Mike McKinstry, R Ilene Kaufman, Hong Zhang, Dalai Yan, Neela Yennawar, Hemant Yennawar, Gregory K. Farber, and B. Tracy Nixon. A dimeric two-component receiver domain inhibits the σ54-dependent ATPase in DctD (2001). *FASEB J*. 15(7): 1326-8.

- **Sungdae Park**, Matthew Meyer, Daniel A. Jones, Hemant P. Yennawar, Neela H. Yennawar and B. Tracy Nixon. Two-component signaling in the AAA+ ATPase DctD: binding Mg^{2+} and BeF_3^- selects between alternate dimeric states of the receiver domain. Submitted to *FASEB J*.

- **Sungdae Park**, Hong Zhang, Daniel A. Jones and B. Tracy Nixon. Biochemical evidence for multiple dimeric states of the *Sinorhizobium meliloti* DctD receiver domain. Submitted to *Biochemistry*