Structural studies of conformational changes of proteins upon phosphorylation: Structures of activated CheY, CheY-N16-FliM complex, and AAA^+ ATPase domain of NtrC1 in both inactive and active states

by

Seok-Yong Lee

B.S. (Yonsei University, Korea) 1998

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biophysics

In the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor David E.Wemmer, Chair

Professor Robert M. Glaeser

Professor James M. Berger

Professor Sydney Kustu

Spring 2003

Abstract

Structural studies of conformational changes of proteins upon phosphorylation: Structures of activated CheY, CheY-N16-FliM complex, and AAA^+ ATPase domain of NtrC1 in both inactive and active states

by

Seok-Yong Lee

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor David E. Wemmer, Chair

Protein phosphorylation is a general mechanism for signal transduction as well as regulation of cellular function. Unlike phosphorylation in eukaryotic systems that uses Ser/Thr for the sites of modification, two-component signal transduction systems, which are prevalent in bacteria, archea, and lower eukaryotes, use an aspartate as the site of phosphorylation. Two-component systems comprise a histidine kinase and a receiver domain. The conformational change of the receiver domain upon phosphorylation leads to signal transfer to the downstream target, a process that had not been understood well at the molecular level. The transient nature of the phospho-Asp bond had made structural studies difficult. The discovery of an excellent analogue for acylphosphate, BeF_3^-, enabled structural study of activated receiver domains. The structure of activated Chemotaxis protein Y (CheY) was determined both by NMR spectroscopy and X-ray crystallography.

These structures revealed the molecular basis of the conformational change that is coupled to phosphorylation. Phosphorylation of the conserved Asp residue in the active site allows hydrogen bonding of the T87 Oγ to phospho-aspartate, which in turn leads to the rotation of Y106 into the "in" position (termed Y-T coupling). The structure of activated CheY complexed with the 16 N-terminal residues of FliM (N16-FliM), its target, was also determined by X-ray crystallography and confirmed the proposed mechanism of activation (Y-T coupling). First, N16-FliM binds to the region on CheY that undergoes a significant conformational change. Second, the "in" position of Y106 presents a better binding surface for FliM because the sidechain of Y106 in the inactive form of CheY ("out" position) sterically interferes with binding of N16-FliM. In addition to confirmation of Y-T coupling, the structure of the activated CheY-N16-FliM complex suggested that the N16-FliM might be sandwiched between CheY and the remainder of FliM to change the direction of flagellar rotation.

Nitrogen Regulatory protein C (NtrC) is the response regulator that is involved in σ^{54} dependent transcription upon nitrogen depletion in enteric bacteria. It comprises a receiver, a central AAA^+ ATPase, and a DNA binding domains. An NMR study of the BeF_3^- - activated receiver domain of NtrC ($NtrC^R$) showed that there was a large conformational change including secondary structure rearrangement in H4 and the loop between H4 and α4 unlike that seen in activated CheY. The high-resolution solution structure of BeF_3^- - activated $NtrC^R$ showed that Y-T coupling may also be the key to trigger the large conformational change, extending the notion that basic Y-T coupling can cause different structural changes in different receiver domains, allowing fidelity in signaling pathways.

NtrC is an activator of σ^{54}-RNA polymerase (σ^{54} activator). σ^{54}-RNA polymerase is distinct from σ^{70}-RNA polymerase in that it cannot isomerize unless σ^{54} activators contact and remodel the polymerase using ATP. This ATP-dependent action is carried out by the central AAA^+ ATPase domain in the activator. Regulated oligomerization is required for the ATP hydrolysis and the remodeling the polymerase. Two main questions remained to be answered at the start of this work: 1) What are the molecular bases for regulation of the central ATPase domain, which appears to be coupled to the oligomerization? 2) How does energy coupling between ATP hydrolysis and RNA polymerase remodeling occur? To provide the framework for answering these questions, structures of the combined regulatory and central ATPase domains ($NtrC1^{RC}$) and the central ATPase domain ($NtrC1^{C}$) alone of NtrC1 from *Aquifex aeolicus* were determined by X-ray crystallography. *In vivo* and *in vitro* studies showed that unmodified $NtrC1^{RC}$ is inactive whereas $NtrC1^{C}$ is active. $NtrC1^{RC}$ crystallized as a dimer whereas $NtrC1^{C}$ crystallized as a heptamer. Surprisingly, the sequence that is known to be critical for energy coupling formed an unusual structural motif: a loop, connected with two strands, protruding from the middle of a helix. These motifs from multiple subunits form the central pore in the heptamer that can contact σ^{54} for energy coupling. Comparison with other AAA^+ ATPases gives an outline for understanding energy coupling.

Comparison of the inactive dimer ($NtrC1^{RC}$) and the active heptamer ($NtrC1^{C}$) provides insight into the molecular basis of regulation and oligomerization. The inactive dimer ($NtrC1^{RC}$) has extensive contact between receiver and central domains, preventing ATPase domain oligomerization. Activation by phosphorylation similarly disrupts these interactions between receiver and central domains. In the absence of the regulatory

domain, the ATPase ($NtrC1^{C}$) forms an active, heptameric ring using surfaces that were buried in the dimer. This very simple way of regulation seems to have many advantages. By using same surfaces for dimer and heptamer, the free energy cost for switching between states is less and should prevent promiscuous interactions.

Approved: _____

Date

TABLE OF CONTENTS

List of Figures

List of Tables

ACKNOWLEDGEMENTS

Finally, time has come for this "acknowledging" section, a hallmark of the end of graduate school. Believe it or not, I've been curious about how I would feel at this moment ever since I came to the United States in the fall of 1998 (I was then so naïve, full of hope, and optimistic). A quick 5 years later, as I write the acknolwedgements to my thesis, I am filled with mixed emotions. In retrospect, the last 5 years in my life have been satisfactory in that I learned a lot and accomplished more than I expected thanks to the help of many. On the contrary, it wasn't very satisfactory in that I could have managed better my life; I am still horrible separating "work" from "leisure" because of my horrible time managing skills and loss of sound sleep at night due to the incredible stress I have had.

I'd like to thank sincerely Prof. Wemmer for all his support, encouragement, and insightful advice. Every time I had some "dumb" idea and dropped in his office without any notice, he kindly explained why this idea would not likely work but nevertheless encouraged me. I found that nourishing in my growth as a scientist.

I'd like to thank four persons who trained me with infinite patience: Ho S. Cho, Dong-Hae Shin, Edward Berry, and Jeff Pelton. Ho and Jeff were my big brothers in my beginning years at Berkeley and forced me to grow up (I deeply appreciate all the hard training), Ed and Dong-Hae from the Kim group awakened me to the world of "X"tallography. I also thank Prof. Kustu, Prof. Nixon, and Dalai Yan for their great support and helpful discussion. With their important findings and advice I was able to complete the CheY and NtrC project, my first and last projects in grad school.

I am highly indebted to David King who solved all the problems that I had with my projects, enabling me to graduate on-time with his great expertise in mass-spectroscopy and peptide synthesis (FliM peptide, M-28, oxidation of Met…).

Spending the last four and half years with "chemists" in the Wemmer group was so much fun and intellectually stimulating (?) to me. I thank Armando De la Torre for his numerous nights setting up xtal trays, Seth "Fungus" Rubin for providing gum for the last four and half years, Steve "Smoking" Damo for being my smoking buddy, Annie "Nice" Roberts for playing tennis and the San Diego trip, Samson Jarso and Priya Rai for their encouragement, and Tuan "Linguist" Nguyen, Esq. for help with proofreading many manuscripts of mine including this thesis. I also thank the past and present members of the Wemmer group and Lishar Huang for their helpful advice.

There are many friends back in Korea who I really appreciate. I thank Jong-Ho, Ji-Min, Heegon, Manhee, Sangjoon, and Heejung for our 20 years of friendship and for the splendid welcoming parties (!) every time I visited Korea. I have no doubt that our friendship will last forever even though I am far away here in the U. S.

I also thank Jesoo, Eosoo, Kyun, Kunwook, Baeksoo, Eunchan, Jinwon, and Miyoung, who are my college friends who have convinced me to pursue a Ph.D. (Guys, look what you've done). Hux, Ilmo, Yongseok, Sungwook, Yontae and Youngim I thank for hanging out with me on my visit to Korea, making my vacations exciting and highly anticipated (So many nights we were drunk together!).

Last, but not the least, I do want to share this moment with my family. I dedicate this thesis to my dad and mom whose boundless love has motivated me to continue on my work despite all the frustrations of graduate life. I thank my sisters, Si-Eun and her

husband, Yu-Seok, and their lovely kids, Joseph and Ester; Ji-Eun and her husband, Heon-Jo, and their lively daughter, Yu-seon; Ji-Soo and my younger brother, Sang-Yong. Without all the support of these people, I would not be where I am today.

CHAPTER 1

Introduction

This chapter provides a basic introduction to two-component signal transduction, bacterial chemotaxis, σ^{54}-dependent transcription, and the AAA^{+} ATPases superfamily.

1.1 Two-component signal transduction

Two-component signal transduction systems are involved in a variety of cellular processes including chemotaxis and gene expression in bacteria, archea, and lower eukaryotes [1-4]. Signal transduction is mediated by phosphotransfer from a histidine residue of a histidine kinase to an aspartate residue of a response regulator. In general, histidine kinases are transmembrane proteins having various extracellular domains sensing environmental signals. Histidine kinases autophosphorylate a conserved histidine residue using ATP, with one histidine kinase monomer phosphorylating the other histidine kinase in the dimer. The phosphoryl group in the histidine kinase is then transferred to the cognate response regulator. Once phosphorylated, response regulators undergo a conformational change that leads to an output signal. In general, response regulators are composed of a receiver domain and an effector domain. To date, more than 300 response regulators have been identified based on sequence homology within a domain of approximately 120 residues, commonly referred to as a receiver domain [5]. The structures of several receiver domains have been solved [6]. They all have a similar $(\alpha/\beta)_5$ fold with an active site comprised of five highly conserved residues including three aspartates, a lysine, and either threonine or serine [6]. The conformational change that occurs upon phosphorylation of the aspartyl residue in the active site in a response regulator remained to be characterized at the start of this work.

1.1.1 Bacterial Chemotaxis.

One of the best-studied two component signal transduction systems is used in bacterial chemotaxis [6,7]. Bacteria move toward favorable chemical environments by regulation of periods of smooth swimming and tumbling in a process known as

chemotaxis [7]. The regulation of tumbling frequency is key to chemotaxis since the default direction of flagellar rotation is counterclockwise, which leads to smooth swimming [7]. Bacteria decide whether not to tumble by sensing temporal concentrations of attractants and repellents [8]. Repellent binding to a transmembrane chemoreceptor induces a signal that is transferred to the histidine kinase CheA, which enhances the probability of phosphorylation of a conserved histidine residue in CheA. CheA in turn transfers a phosphoryl group to the conserved aspartate of the response regulator CheY, which induces a conformational change that increases its affinity for FliM, a component of the flagellar switch. Interaction between activated CheY and FliM reverses the direction of flagellar rotation from counterclockwise (smooth swimming) to clockwise (tumbling motion) [6,7]. The phosphorylation state of CheY is controlled by the opposing activities of the sensor kinase CheA and the phosphatase CheZ [6].

CheA and CheY comprise a two-component signaling system, a predominant signal transduction system in bacteria [1-4]. CheY has been studied extensively as a model for understanding the conformational change upon phosphorylation in the receiver domain. Phosphorylation of D57 [9] induces a structural change [10,11] that enhances the interaction of CheY with the flagellar basal body protein FliM [12] and with CheZ [13], but reduces its affinity for CheA [14]. Mutational studies suggested that the structural changes caused by phosphorylation, but not the phosphoryl group itself, are important for activation [6]. Six highly conserved residues are very important in the function of the receiver domain. D57 was established as the site of phosphorylation [9], and along with D12 and D13, is required for magnesium binding [15-17]. While T87 and K109 were implicated in post-phosphorylation events [16,18-20], their roles had not been defined.

Finally, the rotameric position of Y106, another conserved residue adjacent to the active site, was thought to be correlated with the signaling state of CheY [21]. Although the importance of these residues has been noted in numerous studies of inactive and mutant forms of CheY as well as other response regulators [22,23], a detailed understanding of the mechanism of the phosphorylation-induced transformation from an inactive to an active conformation could not be reached because of the short half-life of the aspartyl-phosphate linkage (half-life of a few seconds to hours).

The nature of the phospho-CheY-FliM interaction and the mechanism by which phospho-CheY (P-CheY) induces a change in the direction of flagellar rotation had been poorly understood, as were interactions between activated receiver domains and their targets generally at the time that I started my work. Genetic and physiological studies suggests that the H4-β5-H5 face of CheY interacts with FliM [24-26]. Deletion analysis shows that binding of P-CheY occurs largely to the first third of FliM [27] and experiments employing specific point mutations indicates that most binding determinants are contained within the 16 N-terminal residues [28]. A peptide consisting of these residues (N16-FliM) binds 20-fold more tightly to phosphorylated than unphosphorylated CheY [28], as is true for full length FliM [29]. In agreement with earlier findings, fluorescence and NMR chemical shift mapping studies [30] confirm that N16-FliM binds to the H4-β5-H5 face of CheY. In Chapter 2 and 3 structural studies providing insights into details of activation of CheY and its interaction with N16-FliM are presented.

1.2. AAA$^+$ ATPase superfamily

ATPases Associated with various cellular Activities (AAA) is a class of ATPases involved in diverse and key cellular actions. AAA ATPases are found in virtually all organisms including bacteria, archea, and eukaryotes. Using better database searches and multiple sequence alignments, Neuwald et al. found that there are far more AAA ATPases than reported previously [31]. Many classes of proteins that are involved in proteolysis, DNA replication, recombination, membrane fusion, transcription, and restriction are now newly classified as AAA$^+$ ATPase superfamily proteins. Since there are extremely diverse functions of this superfamily, it is hard to narrow down to a common mechanism for the action of this protein superfamily. However, there are three types of evidence that show members of this ATPase superfamily share some similarity in their function. First, recent structural studies of AAA$^+$ ATPases show that the fold for this superfamily is conserved even though there are some variations and insertions [32]. AAA$^+$ ATPases have conserved ~230 residues whose structural fold has a α/β folded domain typical of P-loop type NTPases and α-helical domain that is a distinctive feature of AAA$^+$ ATPases [31]. Comparison of structures of AAA$^+$ ATPases shows that structural variations are mostly localized in the α-helical domain [31,32]. Second, almost all AAA$^+$ ATPase superfamily proteins share a common feature in their functions: substrate remodeling. In the case of the heat shock locus protein U (HslU), HslU binds partially unfolded proteins and fully unfolds them for proteolysis by the protease HslV, which forms a complex with HslU. In the case of σ^{54}-dependent transcription, σ^{54} activators bind and remodel σ^{54} in order to make the σ^{54}-RNA polymerase capable of transcription. Third, AAA$^+$ ATPases function as oligomeric rings, in most cases forming hexameric

rings, although the stoichiometry of the oligomeric ring is variable among AAA^+ ATPases. In addition to homohexamers such as NSF and p97, heteropentamers such as DNA clamp loader γ-complex, homoheptamers such as ClpB, a mixture of homohexamer and homoheptamer such as RuvB and HslU have also been observed [33-39].

Many questions about AAA^+ ATPases remain to be answered. First of all, what conformational change governs the substrate remodeling and what is the nature of this conformational change? Little is known about how AAA^+ ATPases change conformations during the ATP hydrolysis cycle and how that conformational change is coupled to substrate remodeling. Second, how do AAA^+ ATPases create rings at the molecular level? Nature has provided three ways for AAA^+ ATPases to create rings: 1) formation of regulated assembly of oligomeric rings from a monomer or dimer: 2) stable two-tiered rings stacked together, which remain oligomerized; and 3) formation of a ring from one polypeptide chain containing six AAA^+ ATPase domains [40]. Since AAA^+ ATPases function as rings, oligomerization is critical for function. Variations of ring formation suggest that there is diversity in enzymatic mechanisms among AAA^+ ATPases.

1.2.1 σ^{54} dependent bacterial transcription.

Several different σ factors mediate transcriptional initiation in bacteria. Transcription by σ^{54} RNA polymerase depends on activators that belong to the AAA^+ class of ATPases. A significant difference between σ^{54} and σ^{70} dependent transcription is that σ^{54} -RNA polymerase cannot isomerize to form open complexes at the promoter without activators. Activators of the σ^{54}-RNA polymerase catalyze an ATP-dependent

conformational change in the polymerase that allows it to form open complexes that are transcriptionally productive at promoters [41-44]. Remodeling depends on previous binding of the polymerase to a promoter in a closed complex. The activators, which are usually composed of regulatory, ATPase (AAA^+), and DNA binding domains, respond to different environmental signals and are unusual in binding to DNA enhancer sequences between 100 and 150 bp upstream of promoters. In many cases, their activity is controlled by phosphorylation of an aspartate residue in the receiver domain, which induces a conformational change that allows the central ATPase domain to oligomerize. Oligomerization is essential for ATP hydrolysis and contact with σ^{54}-holoenzyme [45,46].

A well-studied σ^{54} activator is the nitrogen regulatory protein C (NtrC) in *Salmonella typhimurium* that plays a central role in control of genes involved in nitrogen metabolism. It is phosphorylated by the kinase NtrB, whose activity is ultimately regulated by the ratio of α-ketoglutarate to glutamine in the cell. There are two enhancers for NtrC upstream (-108 and –140) to the major promotor for *glnA* gene, which encodes glutamine synthetase. Two NtrC dimers bind to the enhancers. Phophosphorylation of the NtrC receiver domain ($P\text{-}NtrC^R$) stimulates oligomerization and contact to the σ^{54}-polymerase by way of DNA looping (Fig. 1.2) [45,47,48]. Using ATP hydrolysis, P-NtrC oligomers remodel σ^{54} in the holoenzyme thus allowing σ^{54}-RNA polymerase to melt the DNA at the transcription start site (Fig. 1.2) [45,47,49]. Although the structures of isolated regulatory and DNA binding domains of NtrC have been solved [50-52], there is not yet a structure for the oligomerization/ATPase domain of an activator of σ^{54}-holoenzyme and hence many questions remain unanswered: What are the molecular bases for regulation of the central ATPase domain by the regulatory domain? How do

the activators, which are dimers in their inactive state, form active oligomers? How is ATP hydrolysis by the activators coupled to the conformational change in σ^{54}-holoenzyme that allows DNA melting at promoters? Structural studies of a closely related central ATPase domain described in Chapter 5 provide a framework to answer these questions.

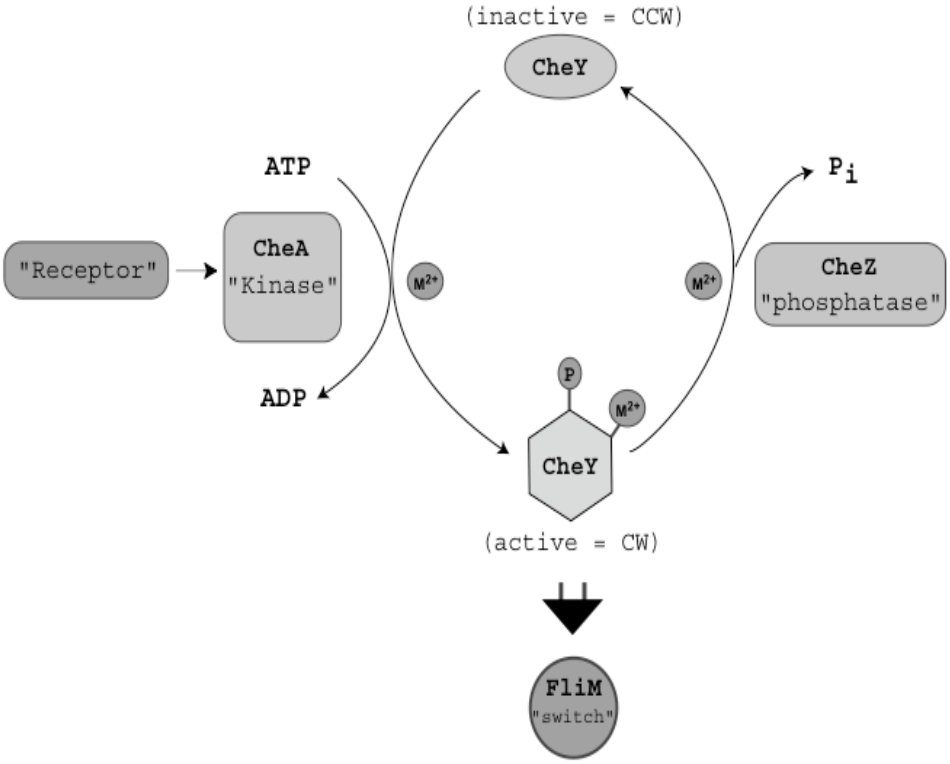


Figure 1.1 Bacterial chemotaxis diagram. CCW and CW denotes the directions of the flagellar rotation counterclockwise and clockwise, respectively. M^{2+} denotes Mg^{2+} and other divalent cations. P indicates a phosphate.

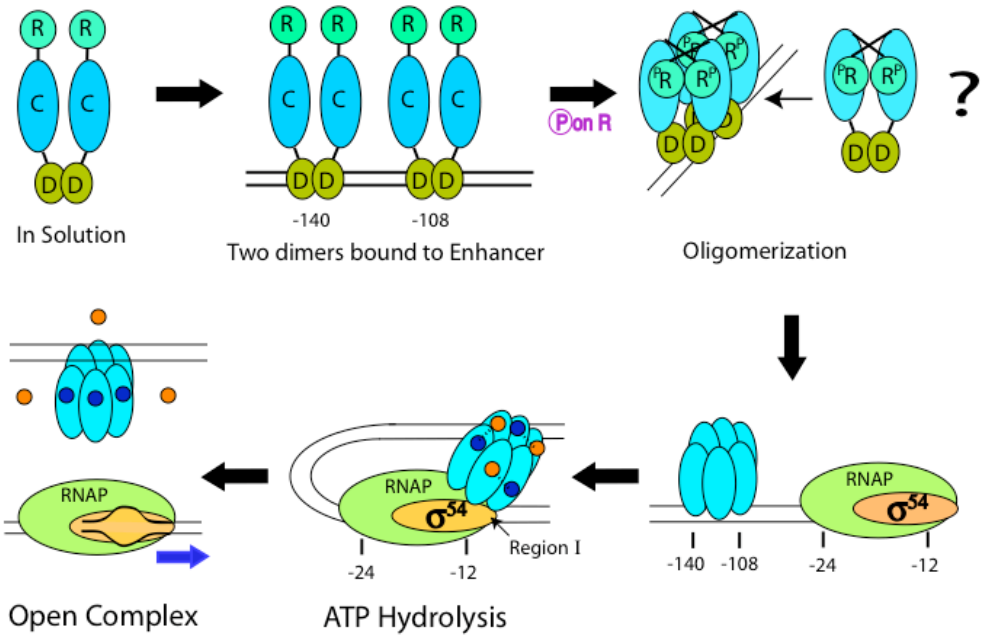


Figure 1.2 σ^{54} dependent bacterial transcription. R, C, and D denote the receiver, the central ATPase, and the DNA binding domains in the σ^{54} activator, respectively. The enhancer sites are for NtrC from *Salmonella typhimurium*. P denotes phosphorylation (i.e. R^P for the phosphorylated receiver domain). RNAP is the abbreviation for a core RNA polymerase. Region I is the N-terminal region of σ^{54} that interacts with σ^{54} activators. Blue and orange dots are for an ADP and a γ-phosphate, respectively. The dotted lines between the blue and orange dots show the transition state of ATP. Note that only three ATPs were shown for simplicity. The blue arrow shows the direction for the transcription. Note that the activators were drawn as a single ellipsoid after the oligomerization for simplicity.

1.3 References

1. Nixon, B. T., Ronson, C. W. & Ausubel, F. M. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. *Proc. Natl. Acad. Sci. U.S.A.* 83, 7850-4 (1986).
2. Ota, I. M. & Varshavsky, A. A yeast protein similar to bacterial two-component regulators. *Science* 262, 566-569 (1993).
3. Parkinson, J. S. & Kofoid, E. C. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 26, 71-112 (1992).
4. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183-215 (2000).
5. Grebe, T. W. & Stock, J. B. The histidine protein kinase superfamily. *Adv. Microb. Physiol.* 41, 139-227 (1999).
6. Djordjevic, S. & Stock, A. M. Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system. *J. Struct. Biol.* 124, 189-200 (1998).
7. Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. The two-component signaling pathway of bacterial chemotaxis. *Annu. Rev. Cell. Dev. Biol.* 13, 457-512 (1997).
8. Macnab, R. M. & Koshland, D. E., Jr. The gradient-sensing mechanism in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U.S.A.* 69, 2509-12 (1972).

9. Sanders, D. A., Gillece-Castro, B. L., Stock, A. M., Burlingame, A. L. & Koshland Jr., D. E. Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. *J. Biol. Chem.* 264, 21770-21778 (1989).

10. Lowry, D. F. et al. Signal transduction in chemotaxis. A propagating conformation change upon phosphorylation of CheY. *J. Biol. Chem.* 269, 26358-62 (1994).

11. Drake, S. K., Bourret, R. B., Luck, L. A., Simon, M. I. & Flake, J. J. Activation of the phosphosignaling protein CheY. *J. Biol. Chem.* 268, 13081 (1993).

12. Welch, M., Oosawa, K., Aizawa, S.-I. & Eisenbach, M. Phosphorylation dependent binding of a signal molecule to the flagellar switch of bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 90, 8787-8791 (1993).

13. Blat, Y. & Eisenbach, M. Phosphorylation-dependent binding of the chemotaxis signal molecule CheY to its phosphatase CheZ. *Biochemistry* 33, 902-906 (1994).

14. Li, J., Swanson, R. V., Simon, M. I. & Weis, R. M. The response regulators CheB and CheY exhibit competitive binding to the kinase CheA. *Biochemistry* 34, 14626 (1995).

15. Needham, J. V., Chen, T. Y. & Falke, J. J. Novel ion specificity of a carboxylate cluster Mg(II) binding site: strong charge selectivity and weak size selectivity. *Biochemistry* 32, 3363-7 (1993).

16. Lukat, G. S., Lee, B. H., Mottonen, J. M., Stock, A. M. & Stock, J. B. Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. *J. Biol. Chem.* 266, 8348-8354 (1991).

17. Lukat, G. S., Stock, A. M. & Stock, J. B. Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. *Biochemistry* 29, 5436-42 (1990).

18. Appleby, J. L. & Bourret, R. B. Proposed signal transduction role for conserved CheY residue Thr87, a member of the response regulator active-site quintet. *J. Bacteriol.* 180, 3563 (1998).

19. Volz, K. & Matsumura, P. Crystal structure of *Escherichia coli* CheY refined at 1.7 Å resolution. *J. Biol. Chem.* 266, 15511-15519 (1991).

20. Zhu, X., Rebello, J., Matsumura, P. & Volz, K. Crystal structures of CheY mutants Y106W and T87I/Y106W. *J. Biol. Chem.* 272, 5000-5006 (1997).

21. Zhu, X., Amsler, C. D., Volz, K. & Matsumura, P. Tyrosine 106 of CheY plays an important role in chemotaxis signal transduction in *Escherichia coli*. *J. Bacteriol.* 178, 4208-4215 (1996).

22. Weinstein, M., Lois, A. F., Monson, E. K., Ditta, G. S. & Helinski, D. R. Isolation of phosphorylation-deficient mutants of the *Rhizobium meliloti* two-component regulatory protein, FixJ. *Mol. Microbiol.* 6, 2041-9 (1992).

23. Feher, V. A. & Cavanagh, J. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature* 400, 289-93 (1999).

24. Sockett, H., Yamaguchi, S., Kihara, M., Irikura, V. M. & Macnab, R. M. Molecular analysis of the flagellar switch protein FliM of *Salmonella typhimurium*. *J. Bacteriol.* 174, 793-806 (1992).

25. Roman, S. J., Meyers, M., Volz, K. & Matsumura, P. A chemotactic signaling surface on CheY defined by suppressors of flagellar switch mutations. *J. Bacteriol.* 174, 6247-55 (1992).

26. Barak, R. & Eisenbach, M. Regulation of interaction between signaling protein CheY and flagellar motor during bacterial chemotaxis. *Curr. Top. Cell. Regul.* 34, 137-58 (1996).

27. Toker, A. S. & Macnab, R. M. Distinct regions of bacterial flagellar switch protein FliM interact with FliG, FliN and CheY. *J. Mol. Biol.* 273, 623-34 (1997).

28. Bren, A. & Eisenbach, M. The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. *J. Mol. Biol.* 278, 507-14 (1998).

29. Welch, M., Oosawa, K., Shin-Ichi, A. & Eisenbach, M. Effects of phosphorylation, Mg^{2+} and conformation of the chemotaxis protein CheY on its binding to FliM. *Biochemistry* 33, 10470 (1994).

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

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

32. Ogura, T. & Wilkinson, A. J. AAA^+ superfamily ATPases: common structure--diverse function. *Genes Cells.* 6, 575-97 (2001).

33. Rohrwild, M. et al. The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nat. Struct. Biol.* 4, 133-9 (1997).

34. Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell* 106, 429-41 (2001).

35. Kim, K. I. et al. Heptameric ring structure of the heat-shock protein ClpB, a protein-activated ATPase in *Escherichia coli*. *J. Mol. Biol.* 303, 655-66 (2000).

36. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* 94, 525-36 (1998).

37. Miyata, T. et al. Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J. Struct. Biol.* 131, 83-9 (2000).

38. Yu, R. C., Hanson, P. I., Jahn, R. & Brunger, A. T. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat. Struct. Biol.* 5, 803-11 (1998).

39. Zhang, X. et al. Structure of the AAA ATPase p97. *Mol. Cell* 6, 1473-84 (2000).

40. Vale, R. D. AAA proteins. Lords of the ring. *J. Cell. Biol.* 150, F13-9 (2000).

41. Xu, H. & Hoover, T. R. Transcriptional regulation at a distance in bacteria. *Curr. Opin. Microbiol.* 4, 138-44 (2001).

42. Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. L. & Gralla, J. D. The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *Journal of Bacteriology* 182, 4129-4136 (2000).

43. Rombel, I., North, A., Hwang, I., Wyman, C. & Kustu, S. The bacterial enhancer-binding protein NtrC as a molecular machine. *Cold Spring Harb. Symp. Quant. Biol.* 63, 157-66 (1998).

44. Zhang, X. et al. Mechanochemical ATPases and transcriptional activation. *Mol. Microbiol.* 45, 895-903 (2002).

45. Weiss, D. S., Batut, J., Klose, K. E., Keener, J. & Kustu, S. The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell* 67, 155-67 (1991).

46. Wyman, C., Rombel, I., North, A. K., Bustamante, C. & Kustu, S. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* 275, 1658-61 (1997).

47. Weiss, V., Claverie-Martin, F. & Magasanik, B. Phosphorylation of nitrogen regulator I of *Escherichia coli* induces strong cooperative binding to DNA essential for activation of transcription. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5088-92 (1992).

48. Weiss, V. & Magasanik, B. Phosphorylation of nitrogen regulator I (NRI) of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 85, 8919-23 (1988).

49. Porter, S. C., North, A. K., Wedel, A. B. & Kustu, S. Oligomerization of NTRC at the glnA enhancer is required for transcriptional activation. *Genes Dev.* 7, 2258-73 (1993).

50. Pelton, J. G., Kustu, S. & Wemmer, D. E. Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. *J. Mol. Biol.* 292, 1095-110 (1999).

51. Volkman, B. F., Nohaile, M. J., Amy, N. K., Kustu, S. & Wemmer, D. E. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. *Biochemistry* 34, 1413-24 (1995).

52. Kern, D. et al. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* 402, 894-8 (1999).

CHAPTER 2

Structure of Activated CheY

Chemotaxis protein Y (CheY) is a well-studied system for protein conformational changes induced by phosphorylation. The transient nature of phosphoryl group in CheY (and receiver domains in general) had hampered the structural study of activated CheY. Using an acylphosphate anaogue, BeF_3^-, the structure of activated CheY was solved both by NMR and X-ray crystallography. Structures of activated CheY allow us to understand the molecular basis of the conformational change that is coupled to phosphorylation.

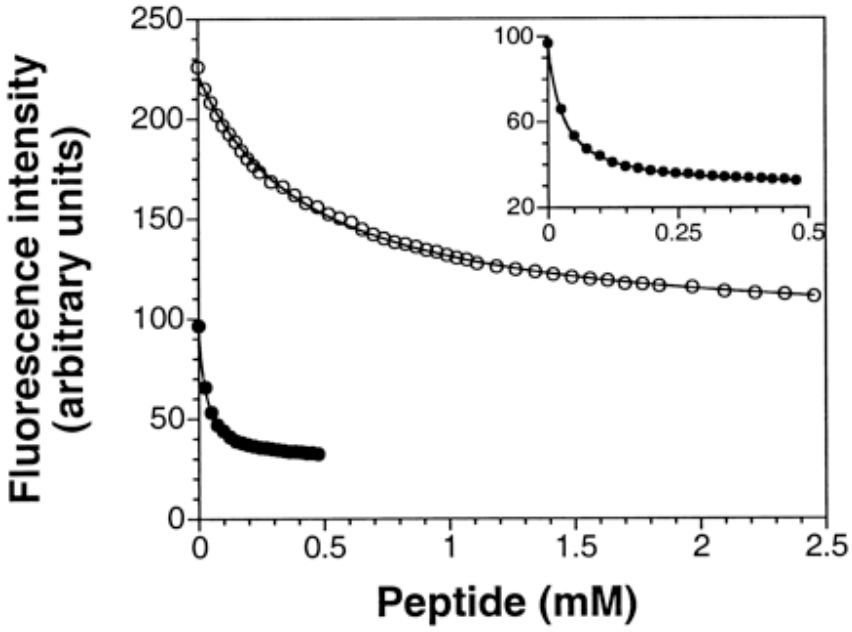
2.1 Beryllofluoride: An excellent analogue for the phospho-Asp bond

Two-component systems are a dominant mode for bacterial signal transduction. These systems are composed of a histidine kinase and a response regulator. Signal is transferred by way of phosphorylation of an aspartate residue in a receiver domain in the response regulator. The half-life of the phospho-aspartate bond ranges from several seconds to hours among various response regulators. The short half-life of the phospho-aspartate bond has hindered the structural study of the conformational change of the receiver domain upon phosphorylation. Recently Yan and Kustu [1] discovered that beryllofluoride mimic the chemistry of aspartyl phosphate in the response regulator NtrC, a σ^{54} dependent transcriptional activator, and in other response regulators [1]. To test the ability of BeF_X to activate other receiver domains that contain the highly conserved active site residues, several receiver domain proteins (CheY, Spo0F, OmpR, NarL) were tested for BeFx binding. Among these receiver domains, my contribution was BeFx-binding assays for both CheY and Spo0F under the guidance of Ho S. Cho [1]. Since BeF_X is toxic to cells, there is no good *in vivo* assay for the activity of BeFx-CheY in changing the direction of flagellar rotation. As an alternative assay, two methods were used to characterize BeF_X-CheY in comparison with phosphorylated CheY (P-CheY). The first method was to monitor the binding of BeF_X-CheY to the 16 N-terminal residues of FliM protein (N16-FliM), a protein that is the target protein for P-CheY. The binding affinity was measured using the intrinsic tryptophan fluorescence quenching of CheY. It was known that the fluorescence of the W58, which is located close to the active site, decreased ~60% upon phosphorylation and decreased further upon binding N16-FliM [2] (Fig. 2.1a). The dissociation constants for binding of BeFx·CheY and apo-CheY to the

N16-FliM peptide, 26 and 500 M, respectively were the same as those reported for P-CheY and apo-CheY (Fig. 2.1a). In order to see whether BeF_X-CheY undergoes a similar conformational change to P-CheY, NMR spectroscopy was used to compare the chemical shift changes between apo- and BeF_X-CheY and those between apo- and P-CheY (Fig. 2.1b). Chemical shifts for isotope-labeled proteins have been reported for the inactive state (apo) of CheY and P-CheY [3-5]. By adjusting the concentrations of Be^{2+}, F^-, and Mg^{2+}, it was possible to convert the protein quantitatively to the BeF_X complex. The chemical shift differences between BeF_X·CheY and CheY were very similar to those between P-CheY and CheY (Fig. 2.1b), particularly for the residues that showed the largest changes. The correlation in the patterns argues strongly that very similar structural changes are likely to occur after activation by either mechanism.

Both BeF_4^{2-} and BeF_3^- exist in solution as well as BeF_2. Whether BeF_4^{2-} or BeF_3^- generated active complexes with NtrC was a question. BeF_X dependent quenching of CheY fluorescence was monitored at different fluoride concentrations and provided evidence that only BeF_3^- can generate an active species [6].

a)


Fluorescence intensity
(arbitrary units)
250
200
150
100
50
0
0
0.5
1
1.5
2
2.5
Peptide (mM)
100
60
20
0
0.25
0.5

b)

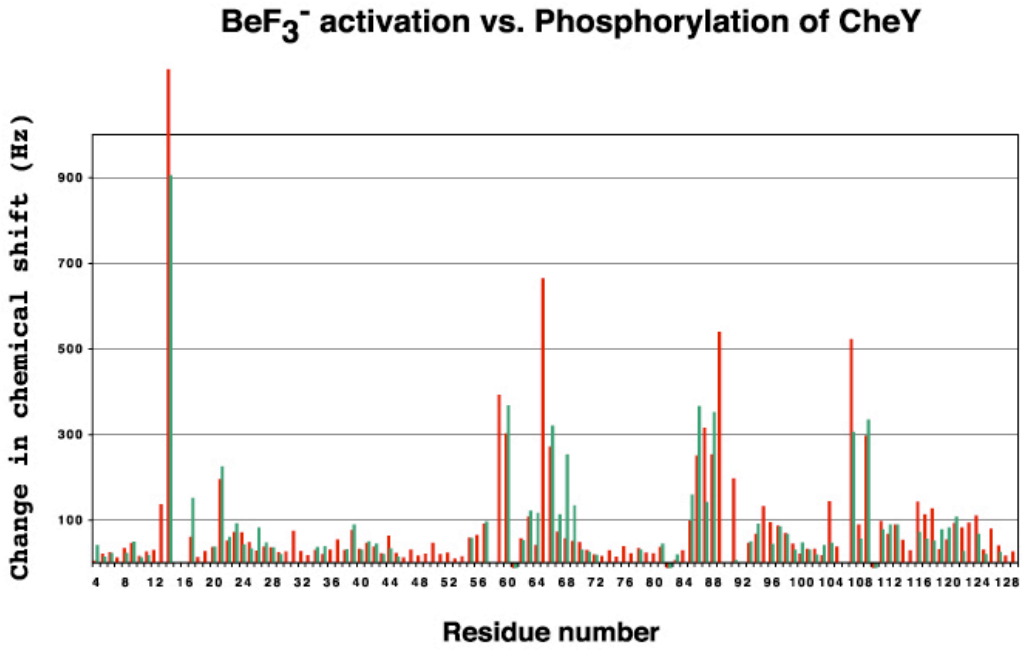

BeF3- activation vs. Phosphorylation of CheY
Change in chemical shift (Hz)
900
700
500
300
100
4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96 100 104 108 112 116 120 124 128
Residue number

Figure 2.1 **a)** BeF_X induced binding of CheY to N16-FliM peptide (residues 1-16). Fluorescence intensity of CheY (10□M) at 340nm as as a function of added FliM peptide in the absence (open circles) or presence (closed circles and inset) of 2□mM $BeCl_2$/20 mM NaF/20 mM $MgCl_2$. **b)** Chemical shift changes of amide nitrogens and protons, $(\Delta\delta_{N2}+\Delta\delta_{H2})^{1/2}$, in Hz, after activation of CheY. Red bars indicate the difference between the BeFx-activated and inactive states of CheY, whereas green bars represent the phosphorylated and inactive states of CheY. Because of the instability of the phosphorylated CheY (P-CheY), there are more peaks unassigned for P-CheY than those for BeF_3^--CheY, leading to less green bars than red bars. Small negative bars designate prolines, which have no NH.

2.2 Methods and materials

2.2.1. Sample preparation and solution structure determination

The gene for E. coli CheY was cloned into a pET21a vector (Novagen, Madison, WI) using PCR. The sequence was confirmed by DNA sequencing (UC BERKELEY DNA sequencing facility). The protein was over-expressed in *E. coli* strain BL21(DE3)/pACYC and was purified as described previously [3]. Uniformly ^{15}N and $^{15}N/^{13}C$ labeled samples were prepared by growth in M9 minimal medium supplemented with ^{15}N ammonium chloride or ^{15}N ammonium chloride and ^{13}C glucose. The BeF_3^--activated sample conditions were 4 mM CheY, 16 mM $BeCl_2$, 100 mM NaF, 20 mM $MgCl_2$, at pH 6.7, and 10% D_2O. Data collection and structure calculation were done together with Ho S. Cho under the direction of Jeff Pelton. NMR spectra were recorded on AMX 600 and DRX 500 NMR spectrometers at 298K. Backbone resonances were assigned with 3D ^{15}N NOESY-FHSQC [7], HNCACB [8], CBCA(CO)NH [9], and HNCA[10] spectra. Side-chain aliphatic $^{13}C/^{1}H$ pairs were assigned with 3D ^{15}N TOCSY-HSQC [11], HCCH-TOCSY [12] and CBCA(CO)NH spectra. In each of the experiments listed above, purge-type pulsed field gradients were used to suppress artifacts and the solvent signal [13]. Aromatic assignments were obtained from DQF-COSY [14] and $^{13}C/^{1}H$ HMQC spectra [15]. The assignment process was also aided by making reference to published chemical shifts for CheY [3,5]. Phi torsion angle restraints were obtained from a ^{15}N HMQC-J spectrum [16]. Stereospecific assignments for Val and Leu methyl groups were obtained by comparison of ct-HSQC spectra of uniformly ^{13}C-labeled and 10% uniformly ^{13}C-labeled samples [17,18]. NOEs identified in 3D NOESY-FHSQC, 4D $^{13}C/^{15}N$ HMQC-NOESY-FHSQC and 4D $^{13}C/^{13}C$ HMQC-NOESY-HMQC (all recorded with a 100 ms mixing time) [19] spectra were

classified as strong (2.9Å upper distance limit), medium (3.3Å), or weak (5.0Å). A total of 972 non-trivial NOE restraints (213 intra-residue, 271 sequential, 238 medium-range, and 250 long-range) were used as input to DYANA [20], along with 78 phi torsion angle restraints and 17 chi1 restraints for the Val, Ile, and Thr residues. Once sets of 20 (out of 60) structures reached a backbone r.m.s.d. of 1Å, 47 hydrogen bonds (94 upper and 94 lower distance restraints (H-O distance restraint 1.8-2.0Å; N-O 2.7-3.0Å)), identified on the basis of slow amide proton exchange rates (protection factors greater than 75) and short donor/acceptor distances, were included in the calculations. Structures resulting from DYANA calculations with a pseudoatom (van der Waals radius 2.5Å) corresponding to BeF_3^- attached to the side-chain of D57 resulted in a backbone r.m.s.d. of only 0.4 Å when compared to structures without the additional pseudoatom. The 27 of 60 structures (BeF_3^- pseudoatom not included) with residual target function values less than 1.0 $Å^2$ (Table 1; target function before energy minimization was 0.3 ± 0.2 $Å^2$) were subjected to restrained energy minimization using the AMBER94 forcefield [21] implemented in the program OPAL [22]. Conjugate gradient minimization (1500 steps) included bond, angle, dihedral, improper dihedral, van der Waals, electrostatic, NMR distance, and NMR torsion angle terms. The minimization was performed in a shell of water at least 6 Å thick, with the dielectric constant set to 1, and with no cutoff for nonbonded interactions. Analysis [23] of the structures by Procheck revealed that 99% of the residues fall within the allowed or generously allowed regions of the Ramachandran map. The 27 energy-minimized structures are used to represent the solution structure of CheY complexed with beryllofluoride and magnesium. The coordinates are deposited in the Protein Data Bank under pdb accession code 1DJM.

2.2.2. Crystallization and crystal structure determination

Crytallization, data collection and phasing were done under the guidance of Li-shar Huang and Edward Berry. *Escherichia coli* CheY was prepared as a solution containing 2 mM CheY, 8 mM $BeCl_2$, 50 mM NaF, and 4 mM $MnCl_2$ at pH 8.4. Crystals of BeF_3^--CheY were obtained at room temperature using the hanging drop vapor diffusion method using a well solution containing 1.8 M ammonium sulfate, 5-10% glycerol, and 100 mM Tris (pH 8.4). The crystallization droplets contained the CheY solution mixed with an equal volume of the well solution. Crystals appeared after 1 day and grew to approximately 0.7 x 0.4 x 0.4 mm after 3 days. The concentration of glycerol in the well solution was increased (5% at each step) every two days to a final concentration of 25 % (v/v). The glycerol was added as a cryoprotectant to allow freezing for data acquisition. The protein crystallized in space group $P2_12_12_1$ with unit cell dimensions a=53.5 Å, b=53.8 Å and c=161.3 Å with two molecules in the asymmetric unit.

The diffraction data were collected at 100 K on the Mar345 detector at the Stanford Synchrotron Radiation Laboratory (SSRL) using beamline 7 - 1 (wavelength of 1.08 Å). A crystal-detector distance of 180 mm was used to collect data to 2.37Å resolution. The data set was integrated and scaled to 2.37 Å resolution, using DENZO and SCALEPACK [24].

Initial phases were determined by molecular replacement (MR) using AMoRe [25]. The structure of BeF_3^--CheY from a BeF_3^--CheY-N16-FliM complex (PDB accession code 1F4V) was used as the search model with Y106 replaced by alanine. Two solutions were easily found with a correlation coefficient of 67% and an R factor of 34% (20-3.5 Å). The MR model was refined by several rounds of simulated annealing and B-group

refinements using CNS [26] to an R factor and R_{free} of 29.5%, and 31.4%. Structure factors were calculated from this partially refined MR model using SFALL of the CCP4 package [27]. These phases were improved by 20 rounds of 2-fold noncrystallographic symmetry (NCS) averaging and solvent flattening at 2.4 Å using the RAVE package [28], with a mask made from the MR solution and operators from the MR solution refined by the IMP program using the $2F_O$-F_C MR map. An unbiased ($2F_O$-F_C) map calculated with phases and Fc from the final symmetry-averaged, solvent-flattened map was displayed using the graphics program O [29] and used as a guide in modeling Y106, positioning Mn^{2+} ions, and manually modifying the model in places where it did not fit the electron density.

Refinement was performed using CNS [26]. Anisotropic B factor and bulk solvent corrections as well as the cross validation method [30] were applied throughout the refinement. 15.0 to 2.37 Å data were included in the refinement with tight NCS restraints (300 kcal mol^{-1} $Å^2$). Water picking was performed after the R factor/R_{free} dropped to 24%/27% using CNS.

The electron density for beryllofluoride on D57 was clearly seen (10□) in the resulting CNS F_o-F_c SIGMAA weighted map. This moiety was modeled on both protomers and refinement continued giving a final R factor and R-free of 21.0 / 24.0%. Geometric parameters for the structure were monitored using PROCHECK [23] and WHAT_CHECK [31]. The coordinates are deposited in the Protein Data Bank under pdb accession code 1FQW.

2.3 Solution structure of activated CheY

All of the backbone amide ^{1}H and ^{15}N resonances except for A2, S15, and T16 (Fig. 2.3a), and 89% of the nonexchangeable side-chain resonances were assigned. The ensemble of structures (Fig. 2.3b) reveals the $(\alpha/\beta)_5$ fold of receiver domains. The five helices [H1 (15-26); H2 (39-46); H3 (65-73); H4 (92-101); and H5 (113 -126)] and the five-stranded parallel β-sheet [β1 (7-11); β2 (33-35); β3 (53-57); β4 (83-87); and β5 (105 -108)] are well defined by the data, yielding a backbone r.m.s.d. of 0.40 ± 0.06 Å (Table 2.1).

Comparison of backbone coordinates for BeF_3^--CheY with previously determined crystal structures of apo [32], magnesium-bound [33], CheA-complexed [34], hyperactive mutant form Y106W [35], as well as with a magnesium-bound NMR structures [5] revealed low r.m.s.d.s (range 0.9 to 1.4 Å). In addition, $^{13}C\alpha$ shifts, which are sensitive to changes in electronic environment and structure, are similar (difference less than 1 ppm) for most, but not all residues of BeF_3^--activated and the magnesium-bound forms of the protein. The largest backbone structural change is movement of the C-terminal portion of β4, which includes T87, toward D57 on β3. This is likely due to interaction of T87 with the BeF_3^- -D57 entity (see below).

Small differences observed in the backbone conformation of inactive CheY, especially in H4 and the loop between β4 and H4, led to the idea that the protein is structurally malleable and that this particular region can adopt multiple conformations [36]. For example, in the NMR structure of magnesium-bound CheY [5] and the crystal structures of apo CheY [32] and magnesium-bound (10 mM Mg^{2+}) CheY [37], H4 begins at K92, but in another crystal structure of magnesium-bound CheY that was crystallized in

the presence of 100 mM Mg^{2+} [33], H4 begins at residue N94 (Fig. 2.2c). The conformational flexibility of the β4-H4 loop and marginal stability of H4 are also supported by ^{15}N NMR relaxation and hydrogen-deuterium exchange studies of the inactive magnesium-bound protein [5].

In BeF_3-activated CheY, H4 begins at residue K92, which is similar to the apo and Mg^{2+}-CheY structures. However, significant differences in $^{13}C\alpha$ shifts were observed between the BeF_3-activated and inactive magnesium-bound [5] forms for V86 in β4 as well as N94 and K91, which are located in H4 and the loop between β4 and H4, respectively. These data point to possible structural changes resulting from activation. Comparison of BeF_3-activated CheY with the inactive structures shown in Fig. 2.2c reveals movement of H4 up and toward the active site. In addition, comparison of solution structures of Mg^{2+} - and BeF_3^--CheY shows that activation by BeF_3^- causes the stabilization of this helix based on the evidence of substantial protection for the amide of residues I95 through A99. Interestingly, phosphorylation of the receiver domain of NtrC appears to induce large conformational changes as well as secondary structure rearrangements in the H4 region [38]. In contrast, our data show that activation of CheY induces small but significant structural rearrangements, but no large conformational changes. This is not surprising since the activation of CheY increases its affinity to FliM only 20-fold based on the previous biochemical data [39].

In almost all known receiver domains the residue equivalent to position 87 in CheY is conserved as either threonine or serine and position 106 is conserved as either tyrosine or phenylalanine [40]. It was suggested that the side-chain of Y106 [41] and the hydroxyl group of T87 [42] participate in conformational changes upon phosphorylation

based on previous biochemical and genetic data. In addition, structural studies of inactive wildtype and mutant forms of CheY indicated that there is steric competition between these two side-chains for a hydrophobic pocket created by β4, H4, and the intervening loop (β4-H4-pocket) [35,41,43] (Fig. 2.3). Except for Y106W, which is hyperactive upon phosphorylation, other mutational substitutions at this position caused either a major decrease (Phe) or complete loss of activity [41]. In the crystal structure of the unphosphorylated mutant form Y106W, the Trp side-chain was found in the β4-H4-pocket, whereas in the mutant forms T87I and T87I/Y106W, which are inactive *in vivo*, the aromatic ring of 106 was solvent exposed and the pocket occupied by the side-chain of I87 [35] (not shown). Analogous to the case for I87 in T87I and T87I/Y106W, in inactive magnesium-bound CheY Y106 was exposed and a significant portion of the β4-H4 pocket was occupied by the methyl group of T87 [5,33]. Our structure of BeF_3^--activated CheY is more similar to Y106W, in that the side-chain of Y106 is buried in the β4-H4-pocket, making contacts with W58, T87, N94, I95 and A98 of H4, whereas the methyl group of T87 is displaced from the pocket (see below).

We have identified the resonance for the hydroxyl proton of T87 in BeF_3^--activated CheY as an exchangeable, non nitrogen-bound proton with NOEs to the amide protons of W58, A88, and E89 (Fig. 2.3). These data strongly suggest that a hydrogen bond is formed between T87 and an active site acceptor, presumably BeF_3^--D57. In the inactive state, the methyl group of T87 is in the β4-H4-pocket, with Y106 in the solvent exposed position (Fig. 2.3). Activation moves the hydroxyl group of T87 (and β4) toward the active site promoting formation of a hydrogen bond. This results in removal of the T87 methyl group from the β4-H4-pocket, which is then filled by Y106 (Y-T

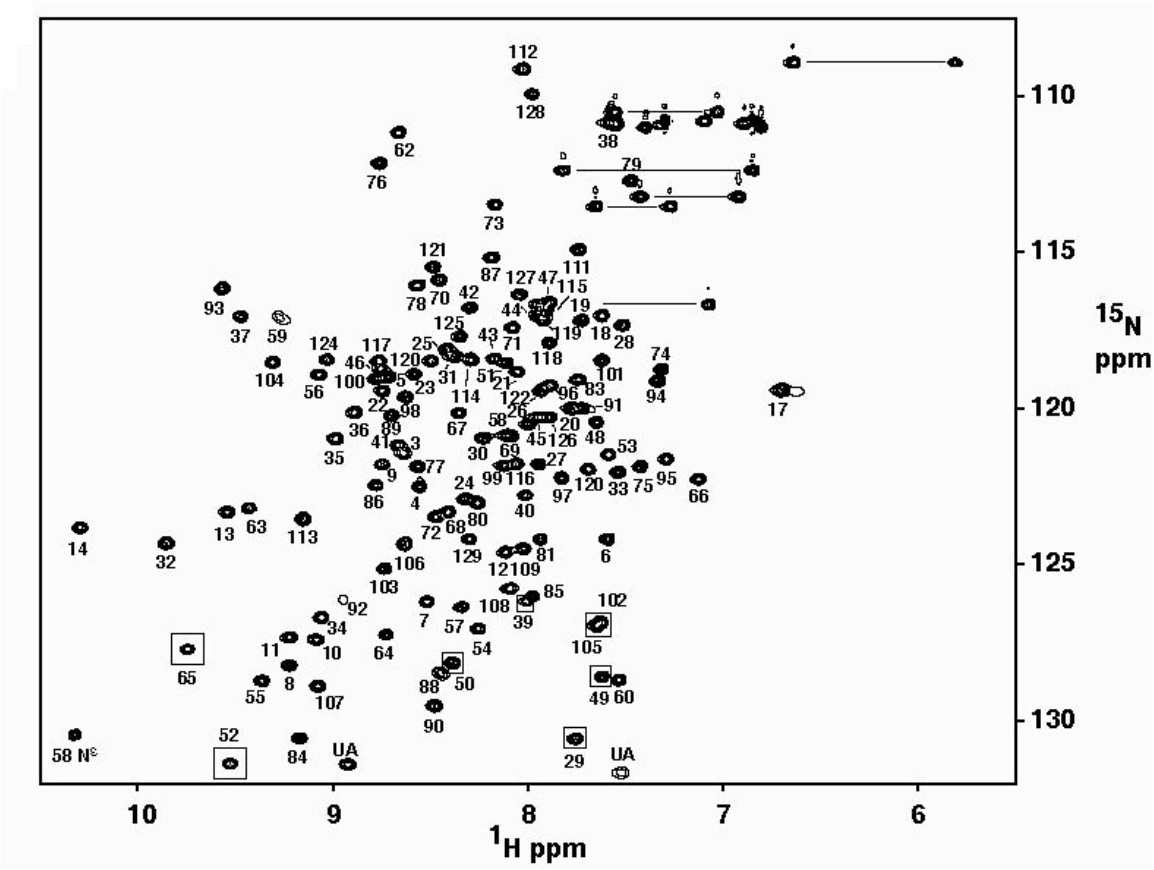
coupling) (Fig. 2.3). Thus, the buried-exposed conformational change of Y106 is coupled via T87 to phosphorylation of D57 [44]. A similar arrangement of the equivalent threonine and tyrosine in the phosphorylated response regulators Spo0A [45] and FixJ [46] was recently reported. The high conservation of these residues combined with the structural data suggest that this may be a general mechanism for inducing structural changes leading to the activation of receiver domains. Although resolution in the active site structure of the phosphorylated receiver domain of NtrC was not sufficient to define the mechanism of activation, both structural [38] and genetic/biochemical studies [47,48] (Yan & Kustu, unpublished data) indicate that there are significant differences between the activation of NtrC and the activation of both CheY and the receiver domain of Spo0A [45]. The high-resolution structure of BeF_3^--$NtrC^R$ is discussed in Chapter 4.

The H4-β5-H5 face of CheY is an important interaction surface for FliM binding (discussed in Chapter 3) [36,49]. Similar to results for apo-CheY [2], titrations of BeF_3^--activated CheY with the 16 N-terminal residues of the FliM protein resulted in large shifts for amide resonances of many residues in this face, including K92, Y106, and V107 (data not shown). In our activated structures, Y106 (in β5) is exclusively in the buried conformation. It is therefore plausible that Y106 in the exposed position (magnesium-bound inactive state) contributes to a decrease in the affinity of FliM for CheY by sterically blocking interaction of FliM with the H4-β5-H5 face. However, failure of other residues to substitute for Y106 indicates that other subtle structural changes which occur upon re-orientation of the tyrosine side-chain are likely to be important for normal function [41]. For example, upon activation there is stabilization of H4, presumably due to hydrophobic interactions between Y106 and the β4-H4-pocket,

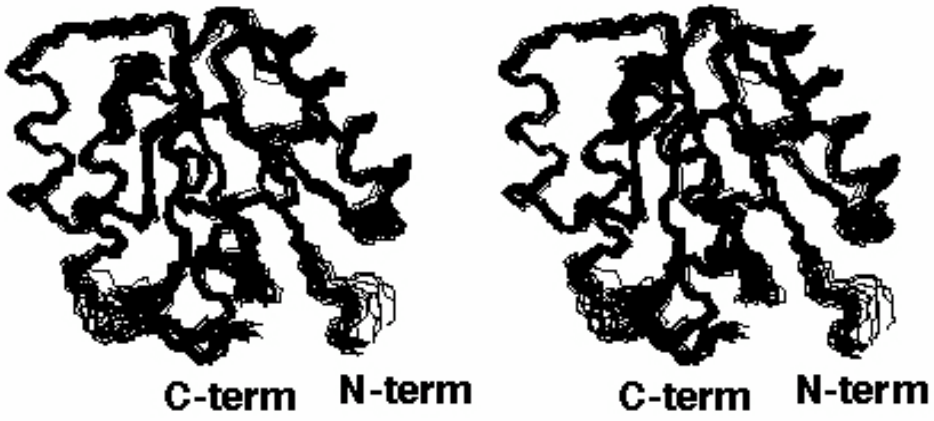
and there is a displacement of H4 in conjunction with α4. There is also a significant $^{13}C\alpha$ chemical shift change for highly conserved K109 and a slight distortion of the backbone in the region near it in activated CheY compared to inactive CheY [5,33]. Unfortunately, the sidechain resonances for K109 are weak and fall in a very crowded region, and could not be assigned. Consequently, we were unable to define its position in the active state, however this was adressed through crystallographic studies of BeF_3^-- activated CheY, section 2.4.

From the NMR structural studies of BeF_3^--activated CheY, we proposed that the rotameric position of the Y106 side-chain is important in changing the affinity to two opposing targets, FliM and CheA, which is highly correlated with the signaling state of CheY.

a)


110
115
120
125
130
15N ppm
10
9
8
7
6
1H ppm

b)


C-term
N-term
C-term
N-term

c)

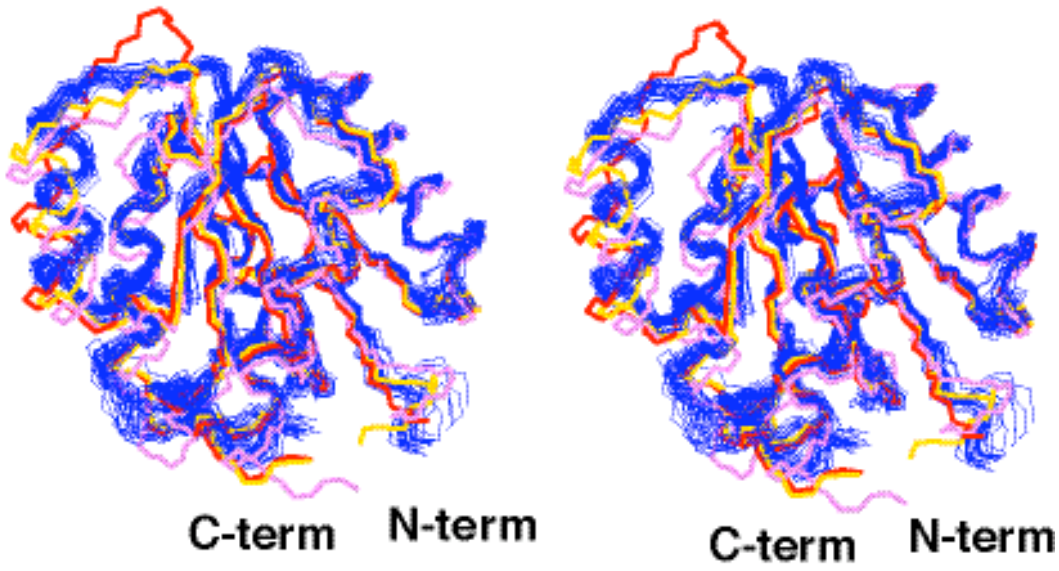


Figure 2.2 **a)** ^{15}N-^{1}H FHSQC spectrum of BeF_3-CheY. Peaks are labeled with residue numbers. Unassigned backbone resonances are labeled UA. Pairs of sidechain NH_2 resonances are connected by horizontal lines. Signals enclosed in boxes are folded in the ^{15}N dimension. **b)** The 27 structures of BeF_3-activated CheY. Backbone coordinates for residues in the five helices and 5-stranded □-sheet were superimposed. **c)** Comparison of BeF_3-activated and inactive CheY structures. Superposition of the 27 structures of BeF_3-activated CheY (blue), with apo X-ray [32] (gold), Mg^{2+}-bound X-ray [33] (red), and mean Mg^{2+}-bound NMR [5] (magenta) structures. Superposition included backbone coordinates for residues in H1, H2, □1, □2, and □3. The figure was produced with the program MOLMOL [50].

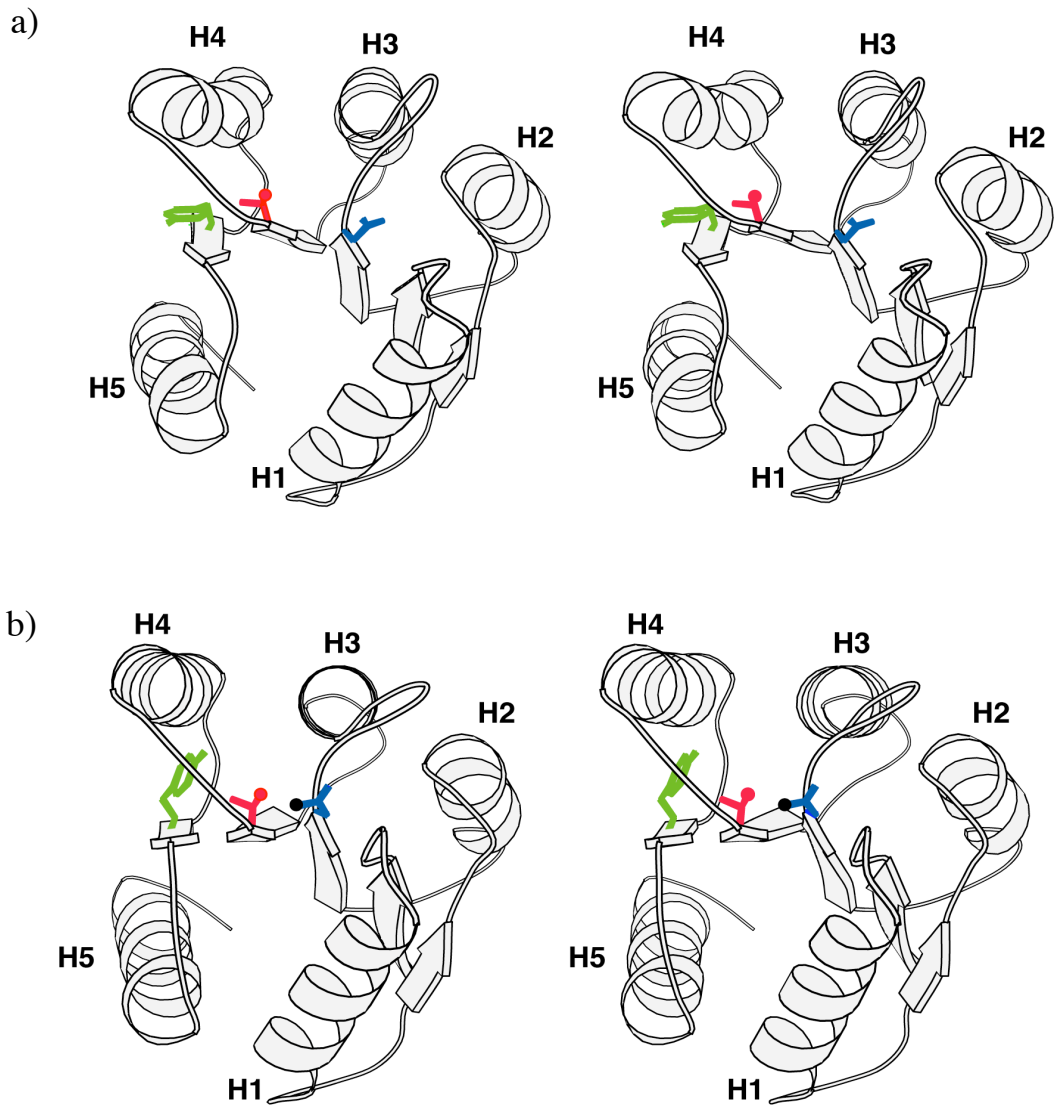


Figure 2.3 Ribbon diagrams of CheY in stereo showing movement of side-chains T87 and Y106 upon activation. Relative to Figure 2.2b, the structures are rotated 90° about a horizontal axis in the page, affording a view (top) of the active site. **a)** CheY taken from the inactive Mg^{2+}-bound NMR structure [5] and **b)** representative NMR structure of BeF_3-activated CheY. D57 (blue) is the site of phosphorylation. Highly conserved Y106 (green) and T87 (red) are also shown. The hydroxyl group of T87 is represented by a small ball. BeF_3^- is modeled as a black ball attached to D57. The figure was created with the program MOLSCRIPT [51].

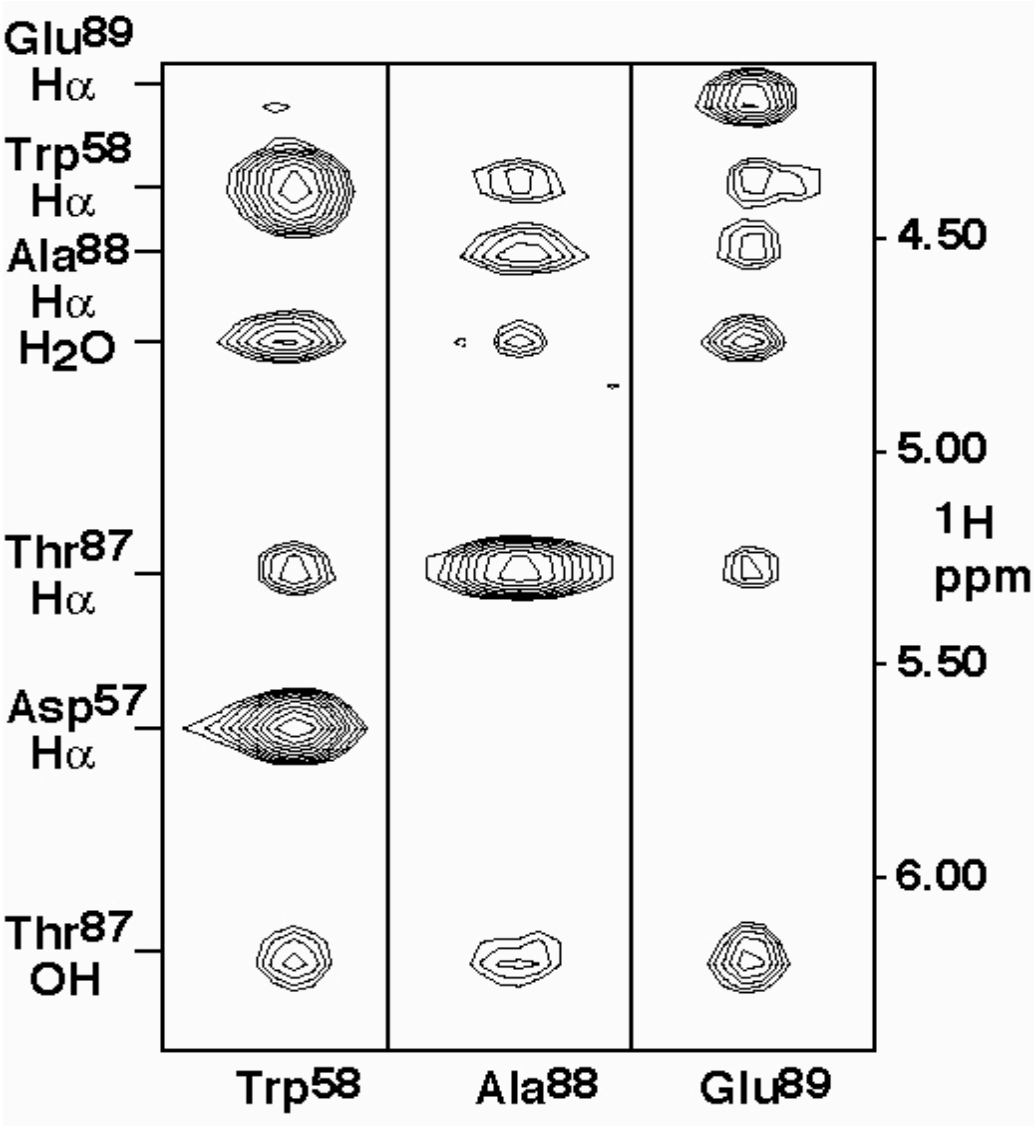


Figure 2.4 Strip-plot of ^{15}N-edited 3D NOESY spectra. Vertical and horizontal axes correspond to ^{1}H and ^{1}H-^{15}N. Three residues (W58, A88, and E89) that are located close to T87 are shown. Note that the peaks at ~6.2 ppm on spectra correspond to hydroxy group of T87. Normally a sidechain hydroxyl group is not detectable due to its exchange with solvent unless forming H-bond, this OH has typical chemical shift ~6.2 ppm.

Table 2.1. Statistics of the energy-minimized NMR structures of BeF_3^- -CheY*.

Parameter	27 Conformers†
Residual distance restraint violations	
Number ≥ 0.1 Å	0.2 ± 0.4
Maximum (Å)	0.1 ± 0.01
Residual dihedral angle violations	
Number ≥ 2.5°	0.1 ± 0.3
Maximum (°)	2.1 ± 0.3
AMBER energies (kcal mol^{-1})	
Total	-5184 ± 95
van der Waals	-467 ± 24
Electrostatic	-5609 ± 87
R.m.s. deviation from ideal geometry	
Bond lengths (Å)	0.0070 ± 0.0001
Bond angles (°)	1.87 ± 0.04
Peptide bonds (°)	0.16 ± 0.04
Atomic r.m.s. differences‡ (Å)	
Backbone residues 6 - 129	0.58 ± 0.08
Backbone for helices and □-sheet	0.40 ± 0.06
Backbone-sidechain < 40% solvent exposed	1.0 ± 0.1

*The average residual target function value for the 27 best DYANA conformers before energy minimization was 0.3 ± 0.2 $Å^2$. †For each entry the average for the 27 conformers and the standard deviation among the 27 conformers is given. ‡Backbone atoms include N, C□, and C'.

2.4 Crystal structure of activated CheY

The crystal structure of BeF_3^--activated CheY, with manganese in the magnesium binding site, was determined at 2.4 Å resolution. The overall crystal structure of BeF_3^--CheY retains the $(\alpha/\beta)_5$ fold of receiver domains [36] (Fig. 2.5a) and is very similar to the NMR structure of BeF_3^--CheY [6] as well as the crystal structure of inactive Mg^{2+}-bound CheY [37]. Superposition of Cα coordinates (residues 6-125) of the BeF_3^--CheY X-ray structure with the mean BeF_3^--CheY NMR structure and the Mg^{2+}-bound CheY crystal structure yielded r.m.s.d.s of only 1.2 Å and 0.8 Å, respectively. In the active site, BeF_3^- bonds to D57, the site of phosphorylation, forming a hydrogen bond and salt bridge with T87 and K109, respectively (Fig. 2.5). The six coordination sites for manganese are satisfied by a fluorine of BeF_3^-, the sidechain oxygens of D13 and D57, the carbonyl oxygen of N59, and two water molecules (Fig. 2.5b). All of the active site interactions seen for BeF_3^--CheY are also observed in P-Spo0A^R. Thus, BeF_3^- activates CheY as well as other receiver domains by mimicking both the tetrahedral geometry and electrostatic potential of a phosphoryl group. The aromatic ring of Y106 is found buried within a hydrophobic pocket formed by β-strand β4 and helix 4 (H4) (Fig. 2.5b). The tyrosine sidechain is stabilized in this conformation by a hydrogen bond between the hydroxyl group and the backbone carbonyl oxygen of E89. This hydrogen bond appears to stabilize the active conformation of the β4/H4 loop. Comparison of the backbone coordinates for the active and inactive states of CheY reveals that only modest changes occur upon activation, except in the loops, with the largest changes occuring in the β4/H4 loop. This region is known to be conformationally flexible in inactive CheY and part of the surface used by activated CheY for binding its target, FliM. The pattern of

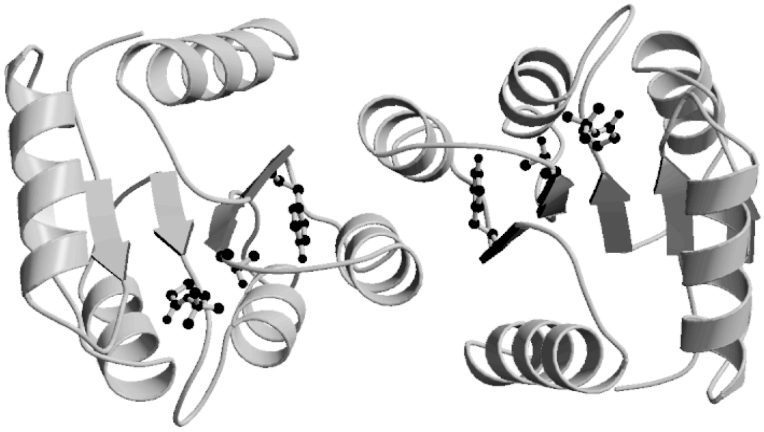
activation-induced backbone coordinate changes is similar to that seen in $FixJ^R$ (R indicates receiver domain throughout this chapter). A common feature in the active sites of BeF_3^--CheY, P-$Spo0A^R$, P-$FixJ^R$, and phosphono-CheY is a salt bridge between K109 Nζ and the phosphate or its equivalent, beryllofluoride. This suggests that, in addition to the concerted movements of T87 and Y106 (Y-T coupling), formation of the K109-PO_3^- salt bridge is directly involved in the activation of receiver domains generally.

Differences in Cα coordinates between BeF_3^--CheY and the crystal structure of inactive Mg^{2+}-bound CheY -- based on a superposition of the residues least affected by activation -- are shown in Figure 2.6a. The biggest changes are observed in the β4/H4 loop, the β5/H5 loop, and the N-terminus of helix H5. The significance of the changes in the β4/H4 loop is particularly hard to interpret because it adopts different conformations in the various crystal structures of inactive CheY [36]. Indeed, dynamics studies of Mg^{2+}-bound CheY showed that this region is flexible in solution [5], and a superposition of the crystal structure of BeF_3^--CheY with the solution structures of Mg^{2+}-bound CheY shows that the β4/H4 loop of the active (crystal) structure falls on the edge of the bundle formed by the inactive (solution, Mg^{2+}-bound) structures. Rather than a conformational change, we prefer to view the activation-induced changes in the β4/H4 loop as a stabilization of the active conformation that may be sampled by the inactive protein. Unfortunately, it is hard to make similar conclusions for the β5/H5 loop, because the relaxation data for residues in this loop could not be reliably interpreted due to complications caused by chemical exchange of Mg^{2+} in the active site [5].

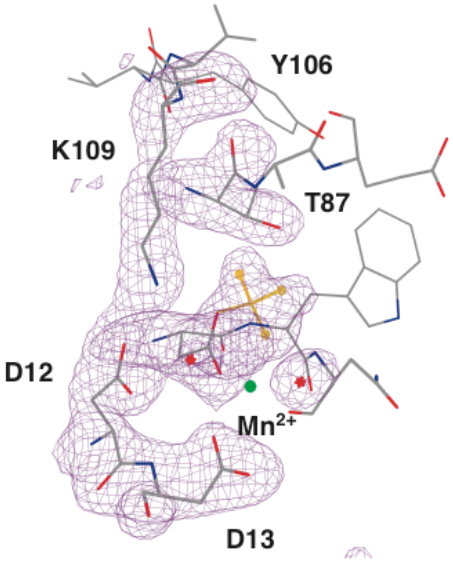
From the NMR structure of BeF_3^--CheY we determined that the switch from an inactive to an active conformation involves hydrogen bond formation between the

hydroxyl group of T87 and an active site residue, presumably a fluorine atom of BeF_3^-. As a consequence of, or in conjunction with, formation of this hydrogen bond, β-strand β4 (along with T87) is displaced toward the active site and the aromatic ring of Y106 becomes buried in a hydrophobic pocket between helix H4 and β5. However, the NMR data did not define the positions of either the BeF_3^- moiety or the divalent cation. In addition, the NMR data for K109, a residue known to be critical for switching to the active conformation [52], were insufficient to define the position of the sidechain in the active site accurately. The BeF_3^--CheY crystal structure verifies the previous conclusions, and extends the detail in the active site (Fig. 2.5). The hydroxyl group of T87 does hydrogen bond with one of the fluorine atoms of the BeF_3^- moiety that is bonded to D57 Oδ (Oδ-Be distance 1.5 Å) in a tetrahedral configuration. The aromatic ring of Y106 is seen exclusively in the buried position – stabilized in this rotameric conformation by a hydrogen bond that was not previously identified in the NMR studies between the tyrosine hydroxyl group and the backbone carbonyl oxygen of E89 as well as hydrophobic interactions. The divalent cation (Mn^{2+}) is located adjacent to D57- BeF_3^- in the crystal structure and is coordinated by D13 Oδ, D57 Oδ, the backbone carbonyl oxygen of N59, a fluorine atom, and two water molecules. Finally, the sidechain of K109 forms a salt bridge with BeF_3^- and D12 Oδ.

a)



b)



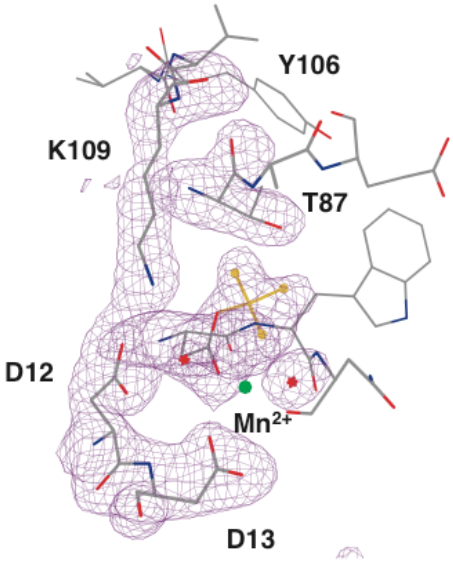


Figure 2.5 **a)** Ribbon diagram of the two BeF_3^--activated CheY molecules in the asymmetric unit. The active sites are directed toward the reader. Side chains are shown for BeF_3^- -D57, T87, and Y106. **b)** Stereo view of the active site of BeF_3^- -CheY. Carbon, nitrogen, oxygen beryllofluoride, and manganese atoms are colored gray, dark blue, red, yellow, and green, respectively. OMIT map contoured at 3.0 covering D12, D13, BeF_3^--D57, T87, K109, and two water molecules. This map was calculated with the occupancies for these residues set to zero.

Table 2.2 Data collection and refinement statistics

Crystal:	
Space Group	$P2_12_12_1$
Cell dimensions a, b, c	53.3, 53.8, 161.2 Å
Data collection (cryogenic):	
Resolution limit (Å)	29.2- 2.37
Measured reflections	89,177
Unique reflections	17,703
R_{sym} [1] (%) (overall / last shell)	7.5 / 23.0
Completeness (%) / last shell	89.1 / 52.2
Ave. B-factor (Wilson 2.4 -- 3.0Å)	43.4
Refinement:	
No. of molecules in AU	2
No. of amino acids	256
No. of solvent molecules	173
Resolution used (Å)	15.0 - 2.37
Sigma cutoff	0.0
No. of reflections work / test	16748 / 846
Final R factor / R_{free} (| F | > 0 s)	21.0 / 24.0
Average B factor ($Å^2$)	44.8
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.011
Bond angles (deg)	1.5

[1] The resolution range in the highest bin is 2.41 - 2.37 Å.

2.5 Comparison with other receiver domains

2.5.1 Active site comparison with other receiver domains

Comparison of the BeF_3^--activated CheY active-site with those of phosphorylated receiver domains determined to high resolution, including P-Spo0A^R [45], P-FixJ^R [46], and phosphono-CheY [53], provides a structural basis by which BeF_3^- mimics the phosphoryl group (Fig. 2.7). Of these, BeF_3^--CheY is best compared to P-Spo0A^R because both contain a divalent cation, hexavalent Mn^{2+} and heptavalent Ca^{2+}, respectively, in the active site. Similar to the phospho-aspartate in P-Spo0A^R, BeF_3^--aspartate acts as a ligand to the divalent metal atom, forms a salt bridge with K109 Nζ, and hydrogen bonds with T87 Oγ and the backbone amides of W58, N59, and A88. The measured distances for the common interactions in the structural models of BeF_3^--CheY and P-Spo0A^R are within coordinate uncertainties (~0.3Å) (Table 2.3). It appears that although calcium has an extra coordination site relative to manganese (and magnesium), which is occupied by a water molecule in P-Spo0A^R, the extra ligand is accommodated without requiring significantly different active site geometry.

Although lacking a divalent cation, the distances for the analogous interactions in P-FixJ^R are also similar to those seen for BeF_3^--CheY and P-Spo0A^R (Table 2.3). The only exception is the large distance (4.1 Å) between K109 Nζ and D12 Oδ in P-FixJ^R, indicating that this salt bridge is broken in the absence of metal. It is interesting to note that, although a divalent cation is necessary for the chemistry of phosphorylation and dephosphorylation of CheY, removal of the metal after phosphorylation apparently does not alter the affinity of P-CheY for FliM [54]. Similarly, the fact that P-FixJ^R purifies as a dimer in the absence of metal, consistent with its activated state, suggests the metal is not

required for inducing the active conformation of $FixJ^R$. This may be a general feature of receiver domains.

In phosphono-CheY, except for the salt bridge between K109 Nζ and a phosphonate oxygen, the distances measured for the hydrogen bonds analogous to those in other activated receiver domains are outside of the acceptable range (2.5 – 3.1Å) (Table 2.3). The absence of these interactions leads to much smaller changes in the β4/H4 and β5/H5 loops (Fig. 2.6c). The modest structural differences relative to inactive CheY appear to be consistent with the partial activity of phosphono-CheY, which shows an 8-fold increase in affinity for N16-FliM [53], whereas BeF_3^-- and phosphorylation-activated CheY show a 25-fold increase in affinity [1,2]. Considering that the Sβ-Cβ bond in phosphono-cysteine is only 0.5 Å longer than the Cβ-Oγ bond in phospho-aspartate, it is surprising that the phosphonate analog does not better activate CheY. Since the salt bridge formed by K109 Nζ and an active site partner (BeF_3^-, PO_3^-, phosphonate) is the only common interaction in P-Spo0A^R, P-$FixJ^R$, BeF_3^--CheY, and phosphono-CheY, it appears to be an important part of the active site interactions that together induce the fully active conformation. Based on crystal structures of mutant forms of CheY, it was previously suggested that K109 plays a role in positioning the β5/H5 loop [33,55].

2.5.2 Activation induced conformation change

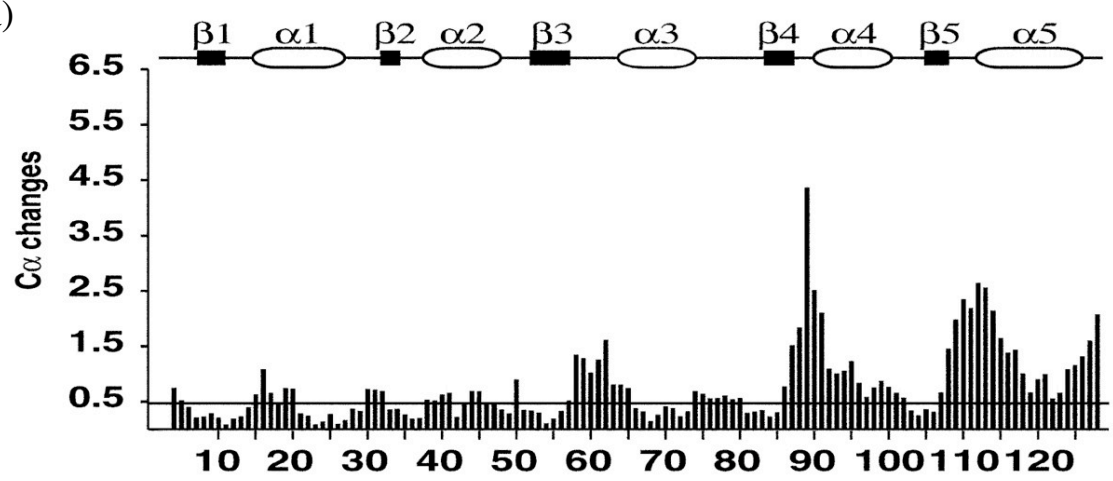
CheY and $FixJ^R$ were the only receiver domains that had been solved with sufficient resolution in both active [46] and inactive [56,57] states to allow a detailed comparison of activation-induced structural changes at the time this work was done. The largest activation-induced Cα coordinate changes for both proteins occur in loop regions, particularly the β4/H4 loop. In addition, the β5/H5 loop shows significant displacement

in CheY, potentially due to the K109 Nζ – BeF_3^- salt bridge, but the analogous conformational change is not seen in $FixJ^R$. As stated previously, for inactive CheY the β4/H4 loop is conformationally flexible according to solution NMR studies [5]. Activation results in the formation of a new hydrogen bond between the hydroxyl of Y106 and the backbone carbonyl of E89. This likely helps to stabilize the active conformation of this loop. Thus, a comparison of just the active and inactive CheY crystal structures could lead one to conclude that activation induced dramatic conformational changes in the β4/H4 loop. However, in light of the NMR data, the exact magnitude of this change is hard to quantify. In $FixJ^R$ the residue homologous to Y106 is a phenylalanine that cannot stabilize the β4/H4 loop through a sidechain-backbone hydrogen bond. It would be interesting to determine whether this loop in $FixJ^R$ is also conformationally flexible in the inactive state and becomes stabilized upon activation.

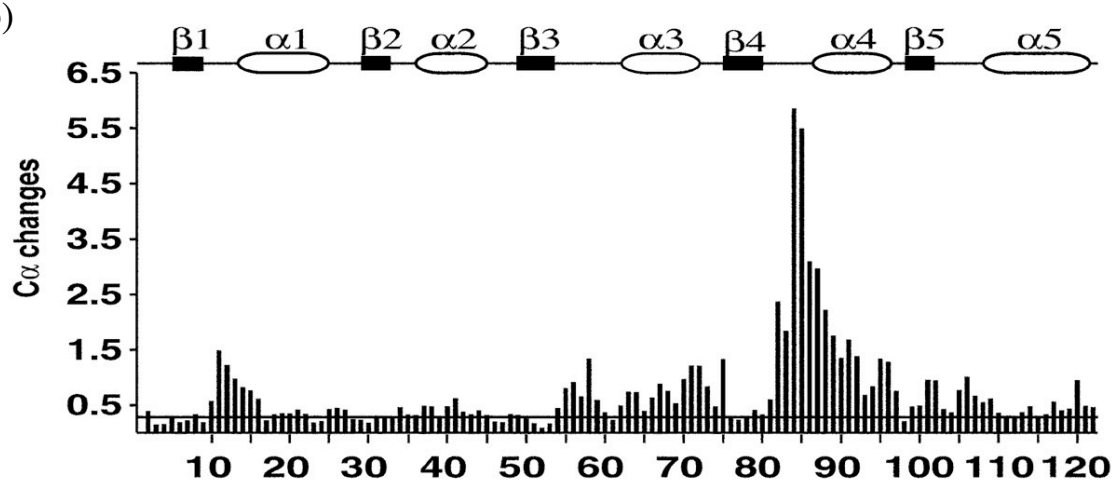
Even though the loops show significant activation-induced changes, activation of CheY and $FixJ^R$ does not result in any major structural rearrangements. Whereas some β-strands and α-helices are slightly displaced, the actual residues that define these elements of secondary structure remain unchanged in both proteins. In both BeF_3^--CheY and P-$FixJ^R$ the N-terminus of H4 moves slightly upon activation, and in CheY there is also a small displacement of the N-terminus of H5. Indeed, even when compared as a group (Fig. 2.8), including P-$Spo0A^R$, there are no dramatic structural differences among either the active or inactive forms of the receiver domains. Although there are small differences in the tilt and inclination of the helices, these differences do not give rise to changes in atomic coordinates of more than a few Ångstroms.

The structures of Spo0A^R and NtrC^R have also been determined in both the active [38,45] and inactive states [58,59]. It was difficult to analyze the activation-induced structural changes for Spo0A^R because the inactive form crystallized as a domain-swapped dimer, the biological relevance of which is unclear. In contrast to CheY and FixJ^R, the low-resolution NMR structures of active and inactive NtrC^R show major structural differences, especially for residues that define helix H4. A higher resolution structure of BeF_3^--activated NtrC^R clearly defined these changes, and is discussed in Chapter 4.

a)


β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110 120

b)


β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110 120

c)

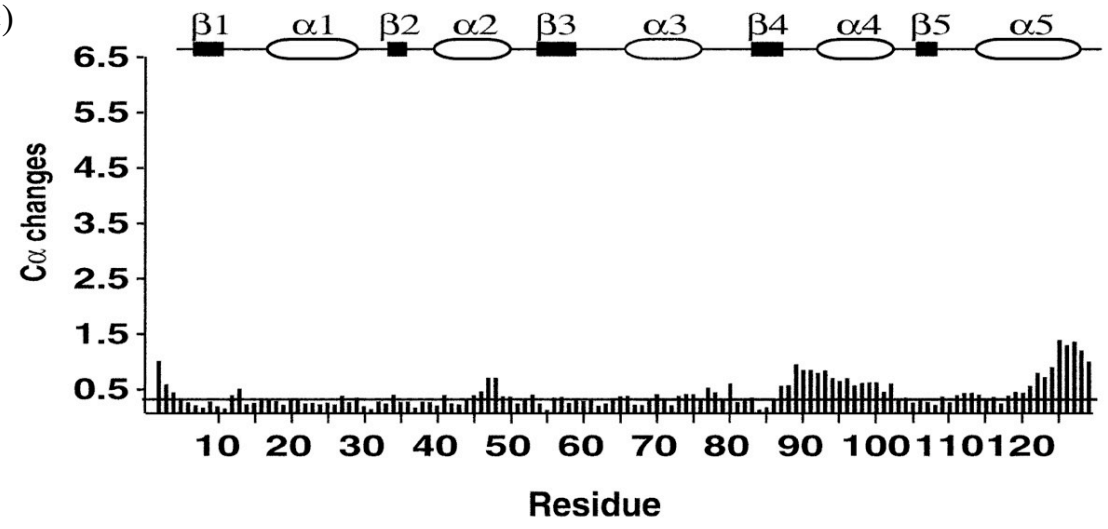
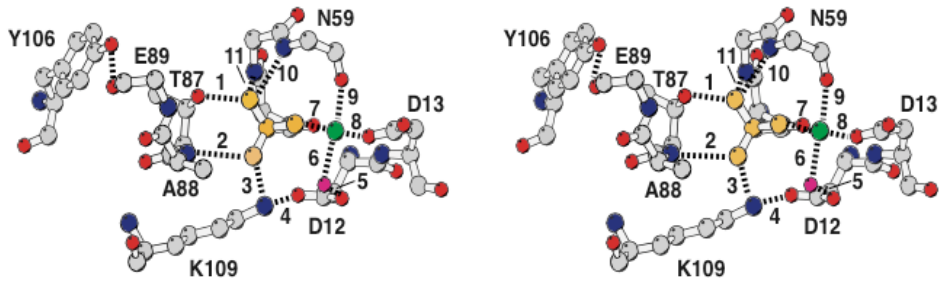

β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110 120
Residue

Figure 2.6 Activation-induced Cα coordinate changes (active-inactive) for CheY. **a)** CheY. **b)** $FixJ^R$. **c)** Phosphono-CheY. For CheY, α-distance plot comparing crystal structures of Mg^{2+} and BeF_3^--CheY showed that residues 5-55 and 65-84 were the least affected by activation. These residues were used to superimpose the Mg^{2+} and BeF_3^--CheY structures from which changes in Cα positions were calculated. For $FixJ^R$, residues least influenced by phosphorylation (residues 1-8, 16-52, and 106-122) were used to superimpose Mn^{2+}-bound and phosphorylated (no metal) structures from which changes in Cα positions were calculated. For phosphono-CheY, residues least influenced by the phosphono group (residues 5-50) were used to superimpose Mg^{2+}-CheY and phosphono-CheY (no metal) structures. The horizontal line in each plot denotes the overall backbone root mean square deviation for each pair of structures.

a)



b)

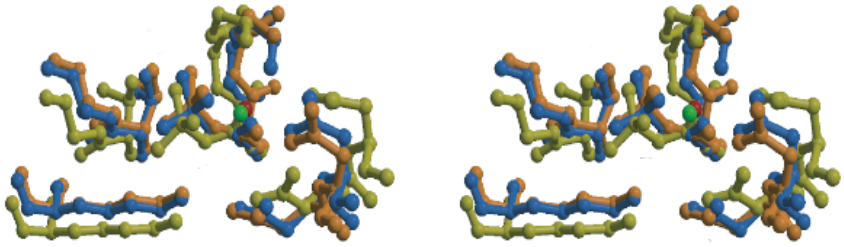


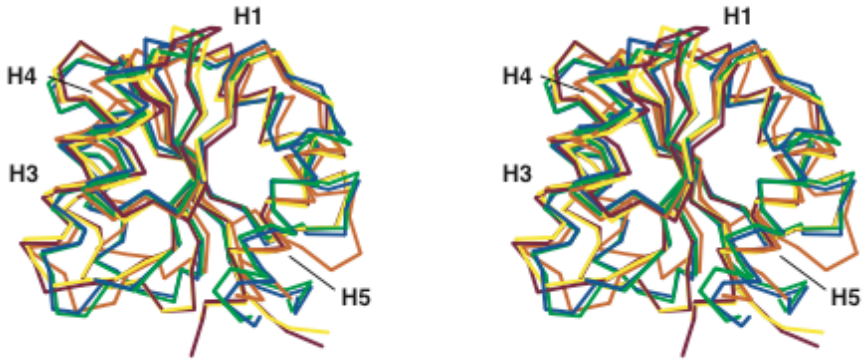
Figure 2.7 **a)** Ball-and-stick diagram of the BeF_3^--activated CheY active site. Dashed lines and numbers denote active site interactions defined in Table 2.3. **b)** Stereo view of active site residues for BeF_3^--activated CheY (Mn^{2+}) (blue), phosphorylated $FixJ^R$ (no metal) (lime), and phosphorylated $Spo0A^R$ (Ca^{2+}) (copper). Mn^{2+} and Ca^{2+} are shown as red and green *balls*, respectively. Residue numbers are based on *E.□oli* CheY. For clarity, phosphono-CheY was not included.

Table 2.3 Active sites distances[1,2] (Å)

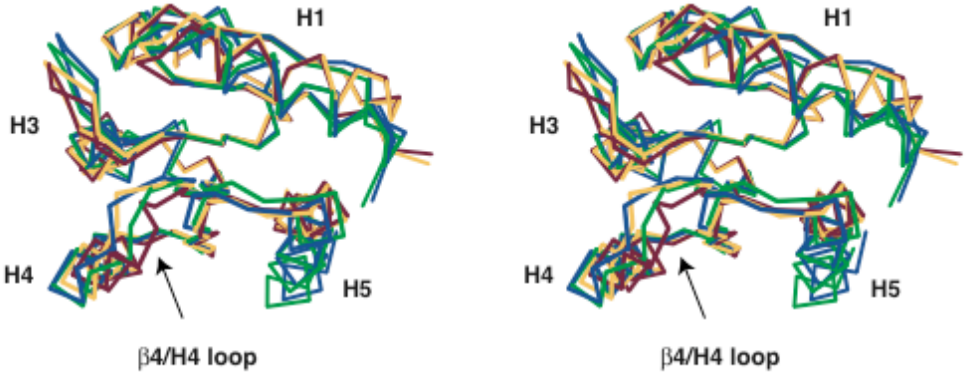
	BeF_3^--CheY	P-Spo0A^R	P-FixJ^R	Phospono-CheY
1. PO_3 – T87O□	2.5	2.5	2.7	6.4
2. PO_3 – A88 NH	2.9	2.8	3.0	4.3
3. PO_3 – K109 NH	2.9	3.0	3.1	3.0
4. K109 NH - D12 O□	2.8	2.6	4.1	5.8
5. D12 O□ - H_2O	2.6	2.5	N/A	N/A
6. M^{2+} – H_2O	2.4	2.6	N/A	N/A
7. M^{2+} – PO_3	2.2	2.4	N/A	N/A
8. M^{2+} – D13 O□	2.3	2.6	N/A	N/A
9. M^{2+} – N59 CO	2.3	2.5	N/A	N.A
10. PO_3 – N59 NH	3.0	2.9	2.9	5.5
11. PO_3 – W58 NH	2.9	2.9	3.4	5.9
12. M^{2+} – D57 O□	2.2	2.6	N/A	N/A

[1] M^{2+} denotes divalent cation; N/A denotes not applicable.
[2] Residue numbers correspond to those of *E. coli* CheY.

a)


H1
H4
H3
H5
H1
H4
H3
H5

b)


H1
H3
H4
H5
β4/H4 loop
H1
H3
H4
H5
β4/H4 loop

c)

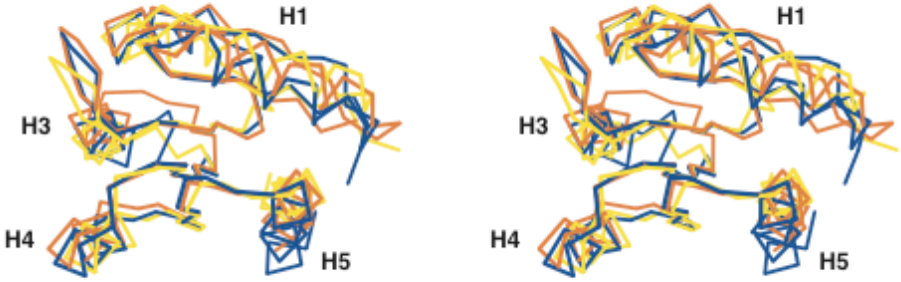

H1
H3
H4
H5
H1
H3
H4
H5

Figure 2.8 □Stereoview of C□ overlays of inactive and active response regulators. Inactive CheY (Mg^{2+}) (green) and BeF_3^--activated CheY (Mn^{2+}) (blue), inactive $FixJ^R$ (Mn^{2+}) (burgundy) and phospho-activated $FixJ^R$ (no metal) (yellow), and phospho-activated $Spo0A^R$ (Ca^{2+}) (orange). The superposition was generated by first superimposing residues in the five □-strands of the three active structures and then superimposing residues 5-55 and 65-84 for inactive CheY (Mg^{2+}) and residues 1-8, 16-52, and 106-122 for inactive $FixJ^R$ (Mn^{2+}) onto their respective active structures. **a)** Molecules are oriented with the H4, □5, H5 face directed toward the reader. **b)** Inactive and active CheY and $FixJ^R$ are shown oriented with the active site directed toward the reader. Compared with a) the structures are rotated 45° along the z-axis and 90° along the x-axis. **c)** Only the three active structures are shown with the same orientation as in b). Note H2 is not labeled.

2.6 References

1. Yan, D. et al. Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14789-94 (1999).
2. McEvoy, M. M., Bren, A., Eisenbach, M. & Dahlquist, F. W. Identification of the binding interfaces on CheY for two of its targets, CheZ and FliM. *J. Mol. Biol.* 289, 1423-1433 (1999).
3. Bruix, M. et al. ^{1}H and ^{15}N NMR assignment and solution structure of the chemotactic *E. coli* CheY protein. *Eur. J. Biochem.* 215, 573-585 (1993).
4. Lowry, D. F. et al. Signal transduction in chemotaxis. A propagating conformation change upon phosphorylation of CheY. *J. Biol. Chem.* 269, 26358-62. (1994).
5. Moy, F. J. et al. Assignments, secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (13C,15N) NMR spectroscopy. *Biochemistry* 33, 10731-42 (1994).
6. Cho, H. S. et al. NMR structure of activated CheY. *J. Mol. Biol.* 297, 543-51 (2000).
7. Talluri, S. & Wagner, G. An optimized 3D NOESY-HSQC. *J. Magn. Reson.* 112, 200-205 (1996).
8. Wittekind, M. & Mueller, L. HNCACB, a high sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the alpha-carbon and beta-carbon resonances in proteins. *J. Magn. Reson.* 101 B, 201-205 (1993).

9. Grzesiek, S. & Bax, A. Correlating backbone amide and sidechain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc* 114, 6291-6293 (1992).

10. Grzesiek, S. & Bax, A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* 99, 201-207 (1992).

11. Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T. & Gronenborn, A. M. Complete resonance assignment for the polypeptide backbone of interleukin 1 beta using three-dimensional heteronuclear NMR spectroscopy. *Biochemistry* 29, 3542-56 (1990).

12. Kay, L. E., Xu, G.-Y., Singer, A., Muhandram, D. & Forman-Kay, J. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson.* 101 B, 333-337 (1993).

13. Bax, A. & Pochapsky, S. Optimized recording of heteronuclear multidimensional NMR spectra using pulsed field gradients. *J. Magn. Reson.* 99, 638-643 (1992).

14. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G. & Ernst, R. R. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* 117, 479-485 (1983).

15. Bax, A., Ikura, M., Kay, L. E., Torchia, D. A. & Tschudin, R. Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Reson.* 86, 304-318 (1990).

16. Kay, L. E. & Bax, A. New methods for the measurement of NH-CαH coupling constants in ^{15}N-labeled proteins. *J. Magn. Reson.* 86, 110-126 (1990).

17. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. Stereospecific NMR assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetic directed fractional ^{13}C labeling. *Biochemistry* 28, 7510-7516 (1989).

18. Szyperski, T., Neri, D., Leiting, B., Otting, G. & Wüthrich, K. Support of ^{1}H NMR assignments in proteins by biosynthetic directed fractional ^{13}C-labeling. *J. Biomol. NMR* 2, 323-334 (1992).

19. Vuister, G. W. et al. Increased resolution and improved spectral quality in 4D 13C/13C-seperated HMQC-NOESY-HMQC spectra using pulsed field gradients. *J. Magn. Reson.* 101, 210-213 (1993).

20. Guntert, P., Mumenthaler, C. & Wuthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273, 283-98 (1997).

21. Cornell, W. et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117, 5179-5197 (1995).

22. Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR* 8, 136-146 (1996).

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

24. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* 276, 307-326 (1997).

25. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* A50, 157-163 (1994).

26. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D-Biol. Cryst.* 54, 905-21 (1998).

27. Bailey, S. The Ccp4 Suite - Programs For Protein Crystallography. *Acta Crystallogr. D-Biol. Cryst.* 50, 760-763 (1994).

28. Kleywegt, G. J. & Jones, T. A. Software for handling macromolecular envelopes. *Acta Crystallogr. D-Biol. Cryst.* 55 (Pt 4), 941-4 (1999).

29. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47 (Pt 2), 110-9 (1991).

30. Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5018-23 (1997).

31. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* 381, 272 (1996).

32. Volz, K. & Matsumura, P. Crystal structure of *Escherichia coli* CheY refined at 1.7 Å resolution. *J. Biol. Chem.* 266, 15511-15519 (1991).

33. Bellsolell, L., Prieto, J., Serrano, L. & Coll, M. Mg^{2+} binding to the bacterial chemotaxis protein CheY results in large conformational changes involving its functional surface. *J. Mol. Biol.* 238, 489-495 (1994).

34. Welch, M., Chinardet, N., Mourey, L., Birck, C. & Samama, J.-P. Structure of the CheY binding domain of histidine kinase CheA in complex with CheY. *Nat. Struct. Biol.* 5, 25-29 (1997).

35. Zhu, X., Rebello, J., Matsumura, P. & Volz, K. Crystal structures of CheY mutants Y106W and T87I/Y106W. *J. Biol. Chem.* 272, 5000-5006 (1997).

36. Djordjevic, S. & Stock, A. M. Structural analysis of bacterial chemotaxis proteins: components of a dynamic signalling system. *J. Struct. Biol.* 124, 189-200 (1998).

37. Stock, A. M. et al. Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry* 32, 13375-80 (1993).

38. Kern, D. et al. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* 402, 894-8 (1999).

39. Welch, M., Oosawa, K., Aizawa, S.-I. & Eisenbach, M. Phosphorylation dependent binding of a signal molecule to the flagellar switch of bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 90, 8787-8791 (1993).

40. Volz, K. Structural conservation in the CheY superfamily. *Biochemistry* 32, 11741-11753 (1993).

41. Zhu, X., Amsler, C. D., Volz, K. & Matsumura, P. Tyrosine 106 of CheY plays an important role in chemotaxis signal transduction in *Escherichia coli*. *J. Bacteriol.* 178, 4208-4215 (1996).

42. Appleby, J. L. & Bourret, R. B. Proposed signal transduction role for conserved CheY residue Thr87, a member of the response regulator active-site quintet. *J. Bacteriol.* 180, 3563-9 (1998).

43. Ganguli, S., Wang, H., Matsumura, P. & Volz, K. Uncoupled phosphorylation and activation in bacterial chemotaxis. The 2.1 Å structure of a threonine to isoleucine mutant at position 87 of CheY. *J. Biol. Chem.* 270, 17386-17393 (1995).

44. Sanders, D. A., Gillece-Castro, B. L., Stock, A. M., Burlingame, A. L. & Koshland Jr., D. E. Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. *J. Biol. Chem.* 264, 21770-21778 (1989).

45. Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I. & Wilkinson, A. J. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol.* 294, 9-15 (1999).

46. Birck, C. et al. Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure Fold. Des.* 7, 1505-15 (1999).

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

48. Moore, J. B., Shiau, S. P. & Reitzer, L. J. Alterations of highly conserved residues in the regulatory domain of nitrogen regulator I (NtrC) of *E. coli*. *J. Bacteriol.* 175, 2692-2701 (1993).

49. Shukla, D., Zhu, X. Y. & Matsumura, P. Flagellar motor switch binding face of CheY and the biochemical basis of suppression by CheY mutants that compensate for motor switch defects in *E.coli*. *J. Biol. Chem.* 273, 23993-23999 (1998).

50. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55 (1996).

51. Kraulis, P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* 24, 946-950 (1991).

52. Lukat, G. S., Lee, B. H., Mottonen, J. M., Stock, A. M. & Stock, J. B. Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. *J. Biol. Chem.* 266, 8348-8354 (1991).

53. Halkides, C. J. et al. The 1.9 A resolution crystal structure of phosphono-CheY, an analogue of the active form of the response regulator, CheY. *Biochemistry* 39, 5280-6 (2000).

54. Welch, M., Oosawa, K., Aizawa, S. I. & Eisenbach, M. Effects of phosphorylation, Mg2+, and conformation of the chemotaxis protein CheY on its binding to the flagellar switch protein FliM. *Biochemistry* 33, 10470-6 (1994).

55. Sola, M. et al. Towards understanding a molecular switch mechanism: thermodynamic and crystallographic studies of the signal transduction protein CheY. *J. Mol. Biol.* 303, 213-25 (2000).

56. Gouet, P. et al. Structural transitions in the FixJ receiver domain. *Structure Fold. Des.* 7, 1517-26 (1999).

57. Stock, A. M. et al. Structural analysis of bacterial chemotaxis proteins - Review. *Biochemistry* 32, 13375 (1993).

58. Lewis, R. J. et al. Domain swapping in the sporulation response regulator Spo0A. *J. Mol. Biol.* 297, 757-70 (2000).

59. Volkman, B. F., Nohaile, M. J., Amy, N. K., Kustu, S. & Wemmer, D. E. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. *Biochemistry* 34, 1413-24 (1995).

CHAPTER 3

Structure of the Activated CheY-N16-FliM complex

The crystal structure of BeF_3^--activated CheY complexed with the N-terminal 16 residues of FliM (N16-FliM), its target, showed that the first 7 residues from the peptide pack against the β4-H4 loop of CheY in an extended conformation (two H-bonds typical of antiparallel β strands), whereas residues 8 through 15 form two turns of helix and pack against the H4-β5-H5 face. Interestingly, the region where N16-FliM binds to CheY coincides with the region on CheY that undergoes noticeable conformational change upon activation. Analysis of the detailed interactions between CheY and N16-FliM suggest that N16-FliM is sandwiched between activated CheY and the rest of FliM to reverse the direction of flagellar rotation.

3.1 Data collection and phasing

Escherichia coli CheY was overexpressed and purified as described [1,2]. A peptide containing the first 16 N-terminal residues of the FliM protein from *E. coli* (MGDSILSQAEIDALLN) was synthesized by FMOC/HOBT-DCC chemistry and purified by reversed-phase HPLC. Purity of both the protein and the petide was confirmed by ESI-mass spectrometry**.** The peptide was added to a solution containing 1.3 mM CheY, 5.3 mM $BeCl_2$, 33.3 mM NaF, and 6.7 mM $MgCl_2$ (pH 8.4) to a final concentration of 2 mM. The selenomethionine derivative of CheY was expressed using *E. coli* B834 (DE3) in minimal medium as described [3]. Crystals of the complex grew at room temperature with the hanging-drop vapor diffusion method using a well solution containing 1.6 M ammonium sulfate, 10% glycerol, and 100 mM Tris (pH 8.4). The crystallization droplets contained the CheY-peptide solution described above mixed with an equal volume of the well solution. Crystals appeared after 3 days and grew to approximately 0.2 mm in each dimension after one week (hexagonal plate shape). Crystals were transferred to well solution plus 20% glycerol for 5 min then flash frozen and stored in liquid nitrogen until data collection. The complex crystallized in space group $P3_221$ with unit cell dimensions a=b=54.2 Å and c=347.4 Å with three complexes in the asymmetric unit. Native data were collected to 2.2 Å Bragg spacing at 100 K on beamline 5.0.2 of the Advanced Light Source (ALS, Berkeley, CA), and a four-wavelength, 3.0 Å MAD dataset of a selenomethionine derivative of the complex was collected at 100 K at beamline 1-5 of the Stanford Synchrotron Radiation Laboratory (SSRL). The initial phases were determined from the native dataset by molecular replacement (MR) using the structure of apo-CheY (3CHY) [4] as the search model with

Y106 replaced by alanine. A three-molecule solution was found giving a correlation coefficient of 49.3% and an R factor of 46.4 % (20-3.5 Å) with AMoRe [5]. Using this partial phase from molecular replacement, the difference-Fourier map was calculated to find Se sites. Eleven Se sites were found by this way and confirmed with partial molecular replacement solutions. Eleven methionine S atoms in the three molecules were then taken as selenium sites for refinement and SIR/MAD phasing to 3.0 Å in the MLPHARE program of the CCP4 suite [6], using the ALS native dataset and the four wavelengths of the MAD dataset from SSRL. The resulting MAD/SIR phases gave an interpretable map that was improved by 15 rounds of 3-fold NCS-averaging and solvent flattening at 3.0 Å using the RAVE package [7]. The phases were then extended to 2.2 Å in steps of 0.1 Å by molecular averaging and solvent flattening. Two independent unbiased ($2F_O$-F_C) maps were calculated from (A) the partially refined molecular replacement model and (B) the symmetry-averaged, solvent-flattened MAD map. These were used as a guide in modeling Y106, positioning three Mg^{2+} ions, and manually modifying the model in places where it did not fit the electron density.

Refinement was performed against the ALS native data using CNS [8], with anisotropic B factors and bulk solvent corrections. Successive cycles of rebuilding and refinement against data from 20.0 to 2.8 Å with tight noncrystallographic symmetry restraints reduced the R and R_{free} factors to 26.6% and 29.2% respectively. The NCS restraints were then released and resolution was extended in small steps to 2.2 Å. At this stage beryllofluoride on D57 in all three monomers, and the first two residues of the FliM peptide on chain F were modeled and refinement continued giving a final R and R_{free} factors of 21.7% / 25.8%. Geometric parameters of the structure were monitored using

PROCHECK [9] and WHAT_CHECK [10]. The crystallographic data and statistics are detailed in Table 3.1. The coordinates are deposited in the Protein Data Bank under accession code 1F4V.

Table 3.1 Data collection, phasing, and refinement statistics

Crystal	Space Group	$P3_221$			
	Cell dimensions a, b, c	54.2, 54.2, 347.4Å			
Data collection		Native	Edge	Peak	Remote
	Resoluton limit (Å)	24.7-2.2	29.9-3.0	29.9-3.0	29.9-3.0
	Wavelength (Å)	1.2	0.9799	0.9796	0.9253
	Measured reflections	121,390	50,774	59,218	56,611
	Unique reflections	28,982	12,209	12,161	12,163
	Completeness/last shell	92.9/84.7	94.0/80.4	94.2/80.3	94.2/73
	R_{merge}[1](%) (overall/last shell)	8.0 / 31.8	8.5 / 26.4	8.7 / 26.7	10.2 / 31.8
	R_{cullis}[2](%) (iso/ano)		0.69/0.96	0.68/0.93	0.67/0.96
	I/δI	24.9	26.4	25.4	16.1
	B-factor[3] (wilson)	29.3	51.1	52.2	53.9
Phasing					
	Mean figure of merit			0.67	
Refinement					
	No. of molecules in AU	3	No. of reflections (work/test)		27181/1442
	Average B factor (Å)	55.1	Resolution used (Å)		20.0-2.22
	No. of amino acids	429	Effective resolution (Å)		2.27
	No. of solvent molecules	282	Sigma cutoff		0.0
	Final R factor / R_{free} ($\|F\| > 0\ \sigma$)	21.7/25.8			
R.m.s.deviations from ideal geometry					
	Bond lengths (Å)	0.010	Bond angles (deg)		1.50

[1]The resolution range in the highest bin is 2.24 - 2.20 Å for the native dataset and 3.05 - 3.00 Å for the MAD dataset, respectively

[2]$Rcullis=\sum||F_{PH}-F_P|-|F_{HCALC}||/\sum|F_{PH}-F_P|$

[3]B-factor was estimated for the native dataset from a Wilson plot over the resolution range 3.0-2.2Å. For the lower-resolution MAD datasets, a relative wilson plot was made against the native dataset, and the resulting relative B-factor was added to that of the native dataset to give the value shown.

3.2 Crystal structure of activated CheY bound to N16-FliM peptide

An introduction of the interactions between activated CheY and FliM was described in Chap.1.

Crystals of activated CheY complexed with N16-FliM peptide form in the $P3_221$ space group with three molecules (complexes) in the asymmeteric unit. The electron density is well defined for all residues for one complex, whereas several N and C-terminal residues of the protein and peptide were not observed for the other two complexes. The conformation of the peptide (residues 4 through 16) is similar in all three complexes even though crystallographic contacts are different in each of the complexes, especially around the FliM peptide, indicating that crystallographic packing does not significantly affect its conformation.

The overall fold of BeF_3^--CheY in complex with N termianl 16 residue of FliM (N16-FliM) (Fig. 3.1) is similar to that in both active and inactive CheY [1,11]. N16-FliM binds to the H4-β5-H5 face of BeF_3^--CheY, and is oriented parallel to H4 (Fig. 3.1). The r.m.s.d. for backbone atoms of secondary structures in the complex is 0.4 Å and 1 Å with respect to their positions of crystal and solution structures of BeF_3^--CheY, respectively. Thus the conformation of the protein is little altered by interaction with N16-FliM.

In the Chapter 2, Y-T coupling (once BeF_3^- addition or phosphorylation occurred on D57 in inactive CheY, the hydroxyl group of T87 becomes hydrogen bonded to BeF_3^-, β-strand 4 is displaced toward the active site with respect to that seen in the inactive CheY structure, and the sidechain of Y106 is buried) was proposed for the activation mechanism [1,12,13]. The high-resolution crystal structure of CheY-N16-FliM not only confirms the previously proposed activation mechanism but also provides information

that was not observed in the NMR structure. First, the positions of the BeF_3^- moiety and Mg^{2+} are clearly defined in the structure (Fig. 3.2). Beryllium is not observed because of its low scattering density, but electron density is clearly visible for each of the three fluorine atoms. Based on the short Be-O□1 distance (1.5 Å) in the deduced structure, BeF_3^- appears bound to D57, the site of phosphorylation. The six coordination sites for Mg^{2+} are satisfied by a fluorine from BeF_3^-, sidechain carbonyl oxygens of D13, D57, backbone carbonyl oxygen N59, and two water molecules (Fig. 3.2). Second, a previously unreported hydrogen bond forms between the hydroxyl group of Y106 and the backbone carbonyl oxygen of E89, which is located in the loop between □4 and H4 (Fig. 3.2). Reorientation of the sidechain of Y106 in activated CheY alters the surface of the H4-□5-H5 face (Fig. 3.2 and Fig. 3.3), presenting a better binding surface. The "out" position of Y106 should sterically hinder the N16-FliM binding. Y106 also stabilizes the loop between □4 and H4 through the hydrogen bond formed upon activation, and stabilizes H4 through hydrophobic interactions between residues of H4 and the aromatic ring of Y106 when it is in an "in" position. All of these effects are likely to be important for binding of CheY to the FliM peptide.

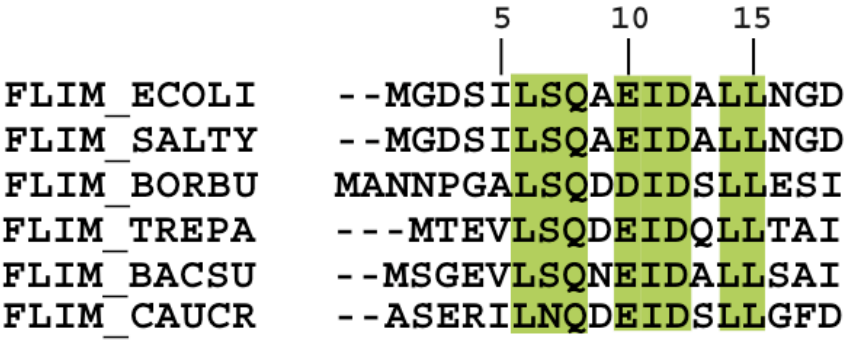
The CheY-N16-FliM complex is stabilized by a combination of hydrophobic interactions, hydrogen bonds, and salt bridges (Fig. 3.3). The total interface spans 1110 $Å^2$, which is slightly below the range (1150 $Å^2$ to 4660 $Å^2$) found in a survey of heteroprotein complexes [14]. Hydrophobic interactions account for 67% of the interface (Fig. 3.3), in which peptide residues L6, I11, and L15 interact with CheY residues I95 and A99 of H4, and Y106 and V108 of □5 to form a hydrophobic core. Peptide residues D3 to S7 adopt an extended conformation and are anchored to CheY by backbone

hydrogen bonds (characteristic of antiparallel β-sheet structure). These occur between S4-CO and L6-HN of the peptide and K92-HN and Ala 90-CO of CheY, respectively. Residues Q8 through L15 of the peptide form two helical turns that are anchored to CheY by hydrogen bonds between Q8-Oε and N16-Oδ of N16-FliM and V108-HN and Y106-HN of CheY, respectively, and salt bridges between D12-Oδ and N16-COO^- of N16-FliM and K119-Nζ and K122-Nζ of CheY, respectively. The N-terminus of the peptide helix is stabilized by a classic S(7) - X - X - E(10) capping box or type 1α motif [15], as evidenced by backbone-sidechain hydrogen bonds between S7 and E10, and hydrophobic interactions between the sidechains of L6 and I11. This helix is also stabilized by the sidechain-sidechain hydrogen bond between Q8 and D12.

Sequence alignment of the N-terminal residues for FliM from six bacterial species shows that residues L6, S7, Q8, E10, I11, D12, L14, and L15 are highly conserved (Fig. 3.1a), suggesting that they may play a role in binding activated CheY or otherwise controlling the direction of flagellar rotation. Indeed, mutational change of L6 to Ile, S7 to Tyr, or E10 to Gly in *E. coli* FliM (full-length) reduces its binding affinity for CheY by 7-, 11-, or 35-fold, respectively [16]. Conversely, mutational change of I95 of CheY to Val increases its affinity for N16-FliM by 8-fold [17]. Inspection of the CheY-N16-FliM structure suggests why these residues are important. L6, S7, E10, and I11 of N16-FliM participate in an N-terminal helix capping box motif; L6, I11, and L15 form hydrophobic interactions with I95 of CheY; and the sidechains of Q8 and D12 are involved in an intermolecular hydrogen bond and salt bridge, respectively. L14 is the only conserved residue that apparently does not stabilize the peptide conformation or participate in peptide-protein binding. It is located adjacent to H4 of CheY, but its sidechain is directed

toward solvent. It may bind to the remainder of FliM (see below). The hydrophobic component of the CheY – N16-FliM interface [14] is on the high end of the range seen in a survey of heteroprotein complexes. Similarly, the dimerization interface of phosphorylated FixJ^r, which includes H4 and α5, is composed mainly of hydrophobic residues [12]. Likewise, phosphorylation of NtrC^r induces a conformational change in helix H4 that results in an increased hydrophobic surface [18]; this surface appears to be important for binding of NtrC^r to its effector [19,20].

a)



b)

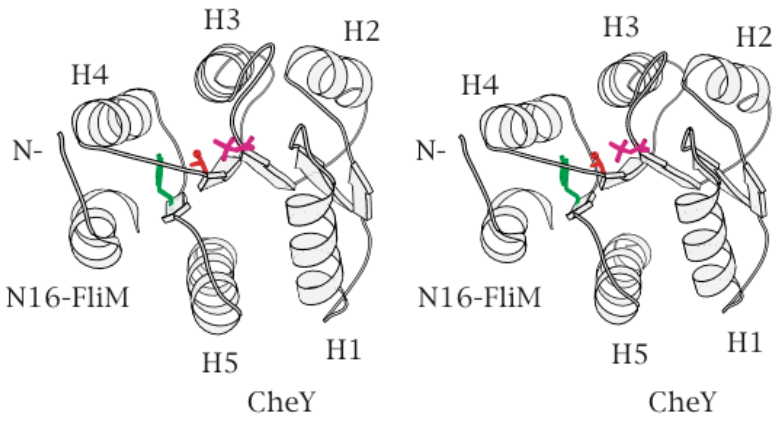


Figure 3.1 Sequence alignment and ribbon structure of the BeF_3^--CheY-N16-FliM complex. **a)** Sequence comparison of the N-terminal 16 residues of FliM from six bacterial species. Residues highlighted in green are highly conserved. Sequence numbering is for *E. coli*. Abbreviations are as follows: ECOLI, *E coli*; SALTY, *S. typhimurium;* BORBU, *B. burgdorferi*; TREPA, *T. pallidum;* BACSU, *B. subtilis*; CAUCR, *C. crescentus*. **b)** Ribbon diagram of the structure of BeF_3^--activated CheY complexed with a peptide corresponding to the first 16 residues of FliM. The complex is shown in stereo with the active site directed toward the reader. The active site residues BeF_3^--D57, T87, and Y106 are shown in magenta, red, and green respectively.

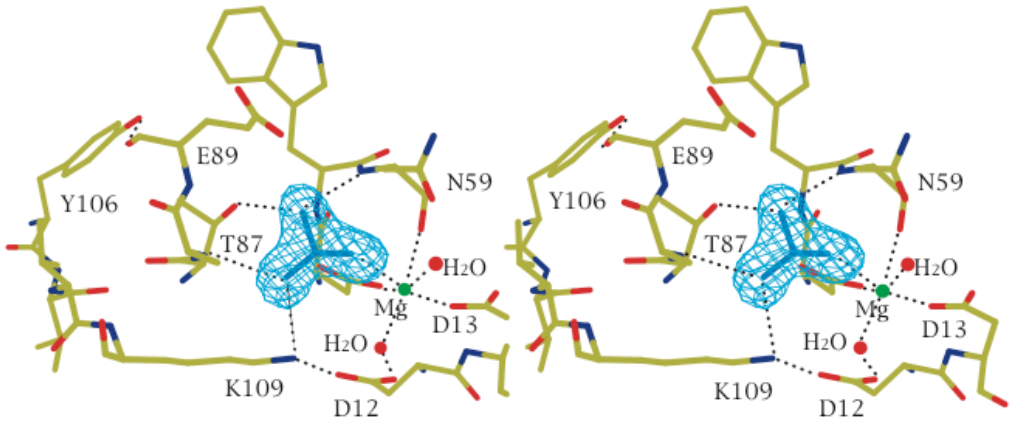
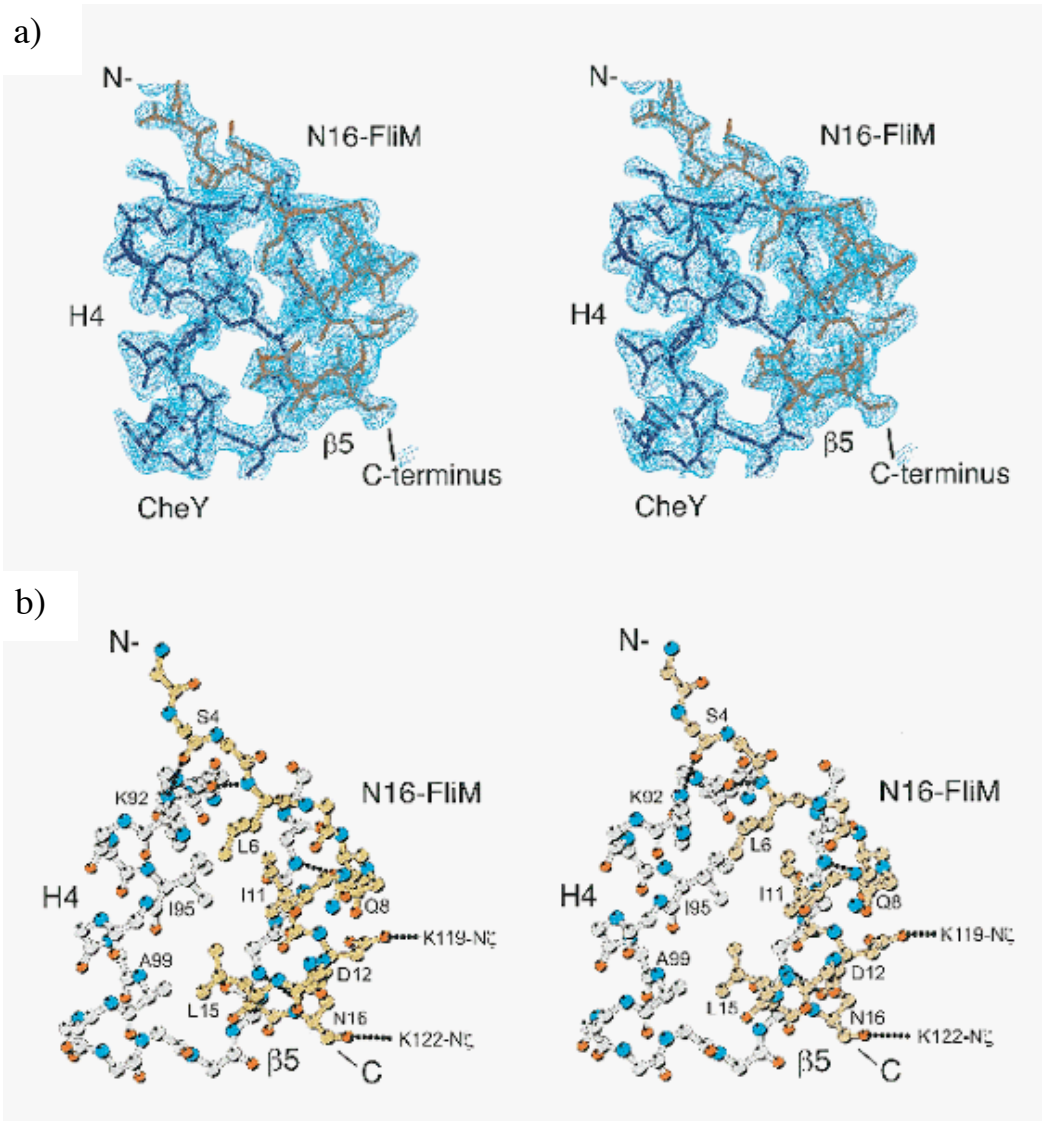


Figure 3.2 Stereoview of the active site of CheY showing the electron density for the BeF_3^- moiety contoured at 5.5□. This map was calculated with F_O-F_C Fourier coefficients with the occupancy of the BeF_3^- moiety set to zero (OMIT map). The three fluorine atoms, shown in light blue, adopt a tetrahedral geometry around Be and hydrogen bond with the hydroxyl group of T87, N□ of K109, and backbone amides of A88 and N59. BeF_3^- also serves as a ligand for Mg^{2+} (dark green). The two water molecules that coordinate with Mg^{2+} are shown in red. Carbon, nitrogen, and oxygen atoms are shown in green, dark blue, and red, respectively. For additional interactions see text.


a)
N-
N16-FliM
H4
β5
C-terminus
CheY
N-
N16-FliM
H4
β5
C-terminus
CheY
b)
N-
S4
K92
N16-FliM
L6
H4
I95
I11
Q8
K119-Nζ
A99
D12
L15
N16
K122-Nζ
β5
C
N-
S4
K92
N16-FliM
L6
H4
I95
I11
Q8
K119-Nζ
A99
D12
L15
N16
K122-Nζ
β5
C

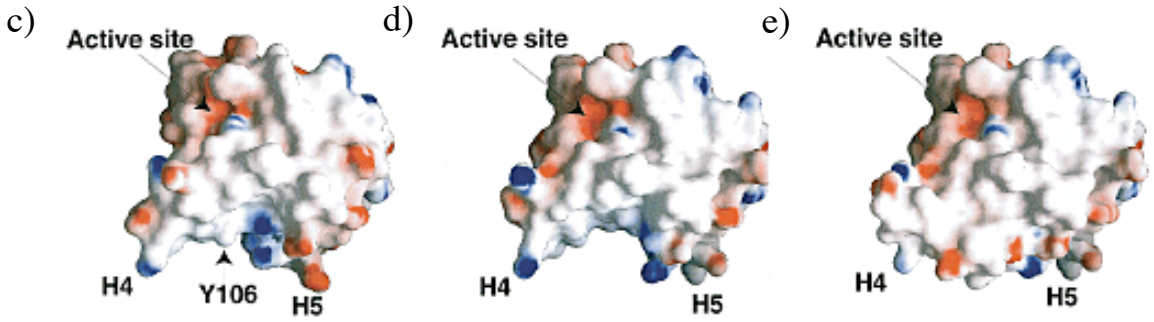

c)
Active site
H4
Y106
H5
d)
Active site
H4
H5
e)
Active site
H4
H5

Figure 3.3 Stereoview of the CheY- N16-FliM interface. **a)** CheY residues A88 to K109 (blue), including H4 (K92 - A101), β5 (G105 - V108), and the loop between β4 and H4 (A88 - K91), and N16-FliM residues M1-N16 (brown). The protein and peptide are covered by two separate OMIT maps (3.3σ) calculated with the occupancy of the protein residues (A88 - K109) or the peptide residues (M1 - N16) set to zero, respectively. **b)** Ball-and-stick representation showing the interactions between CheY (grey) and peptide (brown) residues. Sidechains are shown only for labeled residues. Inter-molecular hydrogen bonds and salt bridges are denoted by dashed lines. The sidechains of K119 and K122 from H5 of CheY, which form salt bridges with N16-FliM residues D12 and N16, respectively, are simply denoted as K119-Nζ and K122-Nζ.

Figure c) - e) Electrostatic surface representations of **c)** apo-CheY with Y106 in the solvent exposed position. **d)** CheY alone from the CheY-N16-FliM complex, and **e)** the CheY-N16-FliM complex. Comparison of c) and d) shows that Y106 in the solvent exposed position sterically interferes with peptide binding to the H4-β5-H5 face.

3.3 Conclusions and implications

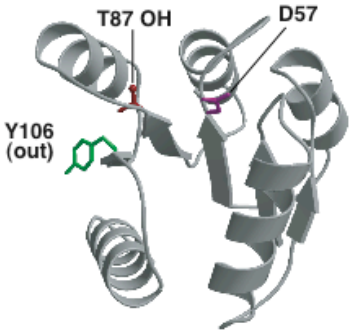
There are two ways that receiver domains regulate their effector domain or target protein: positively and negatively. The phosphorylated receiver domain interacts directly with its effector domain/target protein to convey the signal in positive regulation whereas phosphorylated receiver domain releases the inhibition of the effector domain/target protein in negative regulation [21-25]. There are several cases in which response regulators work in the positive sense. CheY is one of the response regulators that acts positively in that deletion of N16-FliM [16,26] results in a permanent counterclockwise rotation of flagellar thus smooth swimming [27,28]. Hence it is an interesting idea, although untested, that N16-FliM is sandwiched between activated CheY and the rest of FliM to alter the direction of flagellar rotation. Study of the complex of activated CheY and full-length FliM will answer this question but FliM in the flagellar rotor is membrane-associated making characterization difficult. Interestingly the activated receiver domain of $NtrC^{r}$ that is regulated positively, also binds to the N-terminal portion of its target, the central domain of NtrC [19,20] (Hastings and Kustu, personal communication).

3.3.1 Three static pictures of CheY

The structures of activated CheY and its complex with the FliM peptide, its target, allow us to view three static pictures of CheY in different functional stages —inacitve-CheY, activated CheY, and activated CheY bound to the target. We have shown that Y-T coupling is an important mechanism for the phosphorylation-induced conformational change. Phophorylation of the conserved aspartate residue in the active site allows the hydrogen bonding of T87 Oγ with phopho-aspartate that stabilizes the "in" position of Y106. Because the sidechain of Y106 in the inactive form of CheY ("out" position)

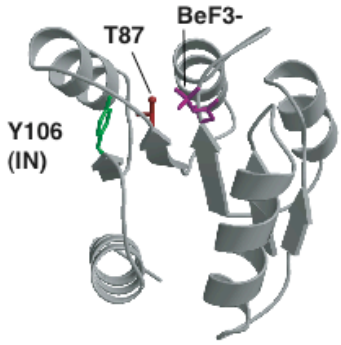
would sterically interfere with binding of N16-FliM, the "in" position of Y106 presents a better binding surface for FliM (Fig. 3.4). This is our understanding of the conformational change in CheY from three static pictures (Fig. 3.4). However, this rather simple mechansim of activation based on three static pictures needs to be supplemented with dynamics data. Based on the NMR studies, there is significant structural flexibility in CheY [11,29]. Most of regions that show flexibility overlap with the regions that show conformational changes, especially the β4–H4 loop and the N-terminus of H4. These regions are flexible in the inactive CheY but become stabilized in the active CheY. From our structure of the activated CheY, we can explain how this is achieved. Activation results in the formation of a new hydrogen bond between the hydroxyl of Y106 and the backbone carbonyl of E89 that is located in the middle of the β4–H4 loop. Additionally, in the "in" position the aromatic ring interacts with H4, which may provide the stabilization of H4. Taken together, these factors likely help stabilize the active conformation of CheY. We interpret the conformational change in CheY upon phosphorylation from the perspective of a dynamic equilibrium. In the inactive CheY, there is a dynamic equilibrium between the inactive and the active populations and what phosphorylation does is to shift the equilibrium toward the active state population by stabilizing active conformation. A recent high resolution crystal structure of apo-CheY confirmed our proposal that there are two distinct conformations in the apo-CheY–one corresponding to the inactive state the other to the meta-active or active state involving the β4/H4 loop and Y106 rotameric position [30].

a)



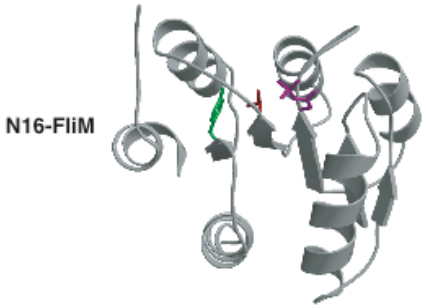
CheY - Inactive conformation

b)



CheY - Active conformation

c)



CheY - bound to its target domain (N16-FliM)

Figure 3.4 Three static pictures of CheY in different functional stages. **a)** Crystal structure of inactive CheY (PDB accession code: 2CHF). **b)** Crystal structure of activated CheY (PDB accession code: 1FQW). **c)** Crystal structure of activated CheY-N16-FliM complex (PDB accession code: 1F4V). Key residues that are involved in the activation are colored magenta (D57 and BeF_3^-), red (T87), green (Y106), respectively.

3.4 References

1. Cho, H. S. et al. NMR structure of activated CheY. *J. Mol. Biol.* 297, 543-51 (2000).
2. Bruix, M. et al. ^{1}H and ^{15}N NMR assignment and solution structure of the chemotactic *E. coli* CheY protein. *Eur. J. Biochem.* 215, 573-585 (1993).
3. Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P. & Hendrickson, W. A. Crystallization of a fragment of human fibronectin - introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. *Protein-Struct. Funct. Genet.* 19, 48-54 (1994).
4. Volz, K. & Matsumura, P. Crystal structure of *Escherichia coli* CheY refined at 1.7 Å resolution. *J. Biol. Chem.* 266, 15511-15519 (1991).
5. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* A50, 157-163 (1994).
6. Bailey, S. The CSP4 Suite - Programs For Protein Crystallography. *Acta Crystallogr. D-Biol. Cryst.* 50, 760-763 (1994).
7. Kleywegt, G. J. & Jones, T. A. Software for handling macromolecular envelopes. *Acta Crystallogr. D-Biol. Cryst.* 55, 941-4 (1999).
8. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D-Biol. Cryst.* 54, 905-21 (1998).
9. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the streochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283-291 (1993).

10. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* 381, 272 (1996).

11. Djordjevic, S. & Stock, A. M. Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system. *J. Struct. Biol.* 124, 189-200 (1998).

12. Birck, C. et al. Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure Fold. Des.* 7, 1505-15 (1999).

13. Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I. & Wilkinson, A. J. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol.* 294, 9-15 (1999).

14. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* 285, 2177-98 (1999).

15. Aurora, R. & Rose, G. D. Helix capping. *Protein Sci.* 7, 21-38 (1998).

16. Bren, A. & Eisenbach, M. The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. *J. Mol. Biol.* 278, 507-14 (1998).

17. Schuster, M., Zhao, R., Bourret, R. B. & Collins, E. J. Correlated switch binding and signaling in bacterial chemotaxis. *J. Biol. Chem.* 275, 19752-8 (2000).

18. Kern, D. et al. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* 402, 894-8 (1999).

19. Lee, J., Owens, J. T., Hwang, I., Meares, C. & Kustu, S. Phosphorylation-induced signal propagation in the response regulator NtrC. *J. Bacteriol.* 182, 5188-95 (2000).

20. Hwang, I., Thorgeirsson, T., Lee, J., Kustu, S. & Shin, Y. K. Physical evidence for a phosphorylation-dependent conformational change in the enhancer-binding protein NtrC. *Proc. Natl. Acad. Sci. U. S. A.* 96, 4880-5 (1999).

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

22. Weiss, D. S. et al. *Prokaryotic transcriptional enhancers.* (eds. McKnight, S. L. & Yamamoto, K. R.) (Cold Spring Harbor Laboratory, NY, 1992).

23. Simms, S. A., Keane, M. G. & Stock, J. Multiple forms of the CheB methylesterase in bacterial chemosensing. *J. Biol. Chem.* 260, 10161-8 (1985).

24. Huala, E., Stigter, J. & Ausubel, F. M. The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. *J. Bacteriol.* 174, 1428-31 (1992).

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

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

27. Toker, A. S., Kihara, M. & Macnab, R. M. Deletion analysis of the FliM flagellar switch protein of *Salmonella typhimurium*. *J. Bacteriol.* 178, 7069-79 (1996).

28. Mathews, M. A., Tang, H. L. & Blair, D. F. Domain analysis of the FliM protein of *Escherichia coli*. *J. Bacteriol.* 180, 5580-90 (1998).

29. Moy, F. J. et al. Assignments, secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (13C,15N) NMR spectroscopy. *Biochemistry* 33, 10731-42 (1994).

30. Simonovic, M. & Volz, K. A distinct meta-active conformation in the 1.1-A resolution structure of wild-type apo-CheY. *J. Biol. Chem.* 276, 28637-40 (2001).

CHAPTER 4

Solution Structure of BeF_3^--activated $NtrC^R$

Bacterial receiver domains mediate cellular responses to environmental changes through conformational changes associated with phosphorylation of a conserved aspartate residue. In most cases, structural characterization of these changes has been precluded by the labile aspartyl-phosphate bond. Recently, it was demonstrated that beryllofluoride (BeF_3^-) forms persistent complexes with receiver domains that give activities as high as the corresponding phosphorylated forms of the proteins. The structure of beryllofluoride-activated NtrC receiver domain (BeF_3^--$NtrC^R$) was determined using NMR data, including residual dipolar couplings, yielding a family of structures with a backbone r.m.s.d. of 0.57 +/- 0.08 Å. The solution structure of $BeF_3^-NtrC^R$ is compared with those of the transiently phosphorylated $NtrC^r$ (P-$NtrC^R$). Both phosphorylation and beryllofluoride addition induce a shift in register and an axial rotation of α-helix 4 (H4). Differences between the NMR structures of P-$NtrC^R$ and $BeF_3^-NtrC^R$ are confined to regions of the P-$NtrC^R$ structure that were poorly defined by the NMR data.

4.1 Data collection and structure calculation

This chapter describes data collection and structure calculation parts of the project. The protein preparation and structure calculation were done entirely by Curt Hastings. My contribution to this project was refinement of conformers, structure analysis, and understanding the mechanism of activation.

4.1.1 Sample preparation, data collection, and structure calculation

Uniformly ^{15}N-labeled, uniformly ^{15}N, ^{13}C-labeled, and 10% ^{13}C-labeled NtrC^{R} were prepared by growth on M9 minimal media. Recombinant protein was expressed in *E. coli* BL21(DE3)/pACYC cells and purified as described by Nohaile et al. [1]. Purified protein was concentrated to a final protein concentration of 1.1-1.5 mM in 50 mM sodium phosphate (pH 6.75) and 50 mM NaCl. NMR samples were made 5% $^{2}H_2O$, 4.4 mM $BeCl_2$, 7.2 mM $MgCl_2$, 29 mM NaF. Quantitative conversion to the beryllofluoride-bound state was verified by examination of the ^{15}N, ^{1}H-FHSQC spectra.

NMR data were acquired at 303K using a Bruker AMX-600 spectrometer equipped with a triple resonance probe with triple axis pulsed-field gradients, and a Bruker DRX-500 spectrometer equipped with a triple resonance probe with single axis pulsed-field gradients. Nearly complete backbone and sidechain resonance assignments were obtained through analysis of HNCA [2], CBCA(CO)NH [3], ^{15}N-edited TOCSY-HSQC [4] and NOESY-HSQC [5], HCCH-TOCSY [6], and ^{13}C-edited HSQC-NOESY [7] spectra. Distance constraints were obtained from assigned crosspeaks in 2D-NOESY, 3D ^{15}N-edited NOESY-HSQC, 3D ^{13}C-edited HSQC-NOESY, 4D ^{13}C, ^{15}N-edited HMQC-NOESY-HSQC and 4D ^{13}C, ^{13}C-edited HMQC-NOESY-HMQC [8] spectra. All NOESY spectra were collected with a mixing time of 100 ms. NOEs were categorized as strong,

medium, or weak according to the peak intensities. Corresponding initial upper distance limits for strong, medium, and weak NOEs were 2.8, 3.4, and 5.0 Å, respectively. Orientational constraints for ^{15}N-^{1}H bond vectors were obtained by comparison of IPAP-HSQC [9] spectra of an unaligned sample and a sample aligned using 30 mg/ml pf1 phage [10]. Amide protons that were resistant to deuterium exchange were identified by dissolving a lyophilized sample in 99.99% $^{2}H_2O$ (CIL) and measuring residual peak heights in FHSQC spectra as a function of time.

The program DYANA was used for initial rounds of iterative structure calculation and NOE assignment. Simulated annealing was carried out with all parameters set to their default values. Initially only NOE distance constraints, torsion angle constraints, and coupling constant constraints were included in the structure calculations. Once the backbone r.m.s.d. reached 1.1 Å, hydrogen bond constraints were added where unambiguous donor-acceptor pairs could be identified. Final rounds of structure calculations were performed using the simulated annealing protocol with the program CNS. A grid search was performed to identify values of the molecular alignment tensor leading to the lowest total energy upon inclusion of the orientational constraints. Multiple rounds of structure calculation and refinement of the distance constraint list were performed. The final scale factors for CNS energy terms were: NOE, 150; van der Waals, 4; dihedral angle, 200. The force constant for orientational constraints was 0.1. A final minimization stage consisting of 10 cycles of 200 steps of minimization was performed using CNS. Inclusion of the beryllofluoride group in the calculations has no measurable effect on the resulting structures [11], therefore, the beryllofluoride-aspartate group was represented by an aspartate. PROCHECK-NMR was used to analyze the

quality of the resulting structures, and SSIA was used to check consistency between the structure and the measured residual dipolar couplings.

A composite Ramachandran plot for the 26 structures in the final ensemble indicates that 97% of the backbone dihedral angles are in allowed and generously allowed regions as determined by the program PROCHECK-NMR [12]. Unstructured regions at the N terminus and C terminus and in the loops between β3 and β3 and β4 and β4 largely account for residues in disallowed regions. In regions of defined secondary structure (residues 5-9, 15-24, 39-47, 49-54, 63-72, 77-82, 89-95, 98-102, and 108-120) 95% of the residues lie in the most favored regions and 100% lie within all allowed regions. The structures were well defined by the NOE data, with an average of 9.1 meaningful internuclear distance restraints per residue used in the final structure calculations.

A second ensemble of structures was calculated using the final constraint set from the NMR data as well as additional distance constraints obtained by analysis of the crystal structure of beryllofluoride-activated CheY. The additional constraints were: (1) a hydrogen bond from T82 Hγ1 (OH) to phosphoryl-D54 OP*; (2) a salt bridge from K104 Hζ* to phosphoryl-D54 OP*; (3) a hydrogen bond from K104 Hζ* to D10 Oδ1. A phosphate group in the active site was used for these calculations. The annealing protocol was identical to that described for calculation of the NMR structure using CNS. PROCHECK-NMR was used to analyze the quality of the resulting structures [12]. The coordinates for 27 conformers, model structures incorporating active site contacts, average (energy minimized) structure with active site contacts are deposited at the Protein Data Bank under the accesion codes 1KRW, 1KRX, and 1J56, respectively.

4.2 Structure of BeF_3^-- $NtrC^R$

4.2.1 Overall architecture of $NtrC^R$.

The structures are well defined, with a backbone heavy-atom r.m.s.d. of 0.57 +/- 0.08 Å for residues 5-120 and an all heavy-atom r.m.s.d. of 1.03 +/- 0.08 Å for the same regions (Fig. 4.1). Importantly, α4 and the loop between β4 and α4, which undergo significant rearrangements upon activation in all receiver domains that have been structurally characterized to date, are well defined by the data. The packing of H3 and H4 against the β-sheet core of $NtrC^R$, which was not well defined in the NMR structure of P-$NtrC^r$, is defined by NOEs between α2 and α3 (a network of contacts between L40, L43, K68, and I69), α3 and α4 (NOEs between the methyl groups of L63 and the backbone amide of A89, between the methyl groups of A89 and A93 and HN and Hα of K67), α3 to the core (a network of contacts between I69 and L76 and V78) and α4 to the core (contacts between the side chain of Y101 and A90 and A93). The contacts between α3 and the core, primarily occurring at the C-terminus of α3, define an orientation of α3 in which the N-terminus of α3 is canted away from the core, and not the C-terminus as was observed in the structure of P-$NtrC^R$ though the definition of the structure in this region was poor.

4.2.2 Model structure incorporating active site contacts.

Sidechains of the residues that comprise the active site were not well defined by the NMR data. This is not surprising since many of the interactions that stabilize the local structure in this region are hydrogen bonds and salt bridges that are difficult to identify with NMR. To explore the possibility that the conserved residues D11, D12, and K104 adopt a hydrogen-bonding network analogous to the active site conformation observed in

the crystal structure of BeF_3^--CheY, an ensemble of structures was calculated using the final restraint set, and additional upper distance restraints consisting of a hydrogen bond between T82 OH and the active site phosphoryl group; the salt bridge between K104 and the phosphoryl group; and a hydrogen bond between K104 and the carboxylate of D10. The resulting structures remain well-defined, with a backbone r.m.s.d. of 0.56 +/- 0.08 Å and an all heavy-atom r.m.s.d. of 0.99 +/- 0.09 Å for 26 structures, and are within the structural envelope defined by the BeF_3^-NtrC^R ensemble, with the exception of the loop between □5 and □5, which moves slightly closer to the active site.

4.3 Conformational change upon activation

4.3.1 Similarity and differences of beryllofluoride-activated and phosphorylated NtrC receiver domains.

In the calculated structures P-NtrC^R and BeF_3^-NtrC^R show similar large-scale structural rearrangements relative to apo-NtrC^R, with the most pronounced changes occurring in α4. Most of structural differences between the P-NtrC^R and BeF_3^-NtrC^R structures are localized in α3 and loops. Although the detailed structures of α3 and α4 were not well defined by the NMR data for P-NtrC^R, the rotational orientation, overall position and change in residues involved in α4 in BeF_3^-NtrC^R and P-NtrC^R are quite similar [13]. In contrast, the position of α3 is different in the two active structures. In P-NtrC^R, the N-terminus of α3, which is marginally defined as a helix in this structure, is canted away from the central β-sheet. In BeF_3^-NtrC^R, the helix axis is roughly parallel to the sheet surface, a conformation more like those observed for CheY, Spo0A^R, and FixJ^R in their activated states [14-17]. To explore the source of this discrepancy, we measured ^{15}N-^{1}H residual dipolar coupling constants for P-NtrC^R in the presence of 10 mg/ml f1 phage (the maximum concentration at which a magnetically oriented state could be maintained in the presence of the phosphodonor and buffer salts). The coupling constants fit the BeF_3^-NtrC^R structure that is closest to the mean with a r.m.s.d. of 1.7 Hz (alignment tensor: D_a = 1.5 Hz, D_r/D_a = 0.49) while they fit the P-NtrC^R that is closest to the mean with an r.m.s.d. of 1.9 Hz (alignment tensor: D_a = 1.2 Hz, D_r/D_a = 0.38), supporting the hypothesis that apparent differences between the P-NtrC^R and BeF_3^-NtrC^R are largely a result of the regions of that are not well defined by the NMR data in P-NtrC^R, and that the BeF_3^-NtrC^R structure is in fact the same as the active phosphorylated form.

4.3.2 Overall structural change upon activation by BeF_3^-.

The overall structure adopts a $(\beta/\alpha)_5$ fold, which is found in all receiver domains. The family of twenty-six refined structures calculated is shown (Fig. 1). We generated an average minimized structure of BeF_3^-NtrC^R and compared it to the average minimized structure of apo-NtrC^R (PDB accession code 1DC7). As judged by a difference-distance plot, segments of the protein that are least affected by activation are 4-10, 13-34, 39-55, 63-68, 73-79, 101-105, and 111-122. We superimposed the active and inactive structures using these residues, and plotted the Cα coordinate differences between these two structures (Fig. 4.3). A sizeable structural change (6-9Å) was observed for residues 83-99, which are located in the β4 to α4 loop, α4 itself, and the loop between α4 and β5. The structural changes in NtrC^R upon activation are much larger than the changes observed in both CheY and FixJ^R (Fig. 4.3) [14,15]. Other regions that show significant changes are mostly localized in loops (β1-α1, β2-α2, β3-α3, β5-α5). The loop between β1 and α1 is known to be involved in Mg^{2+} binding (D10 and D11). Since the comparison is between the apo protein (no Mg^{+2}) and the active form, the structural change observed in this region is likely due to Mg^{2+} binding rather than phosphorylation or BeF_3^- addition. The remaining loops (α2-β2, α3-β3, β5-α5) are not well defined in apo-NtrC^R, making it difficult to assess the real structural change upon activation. In Fig. 4.2, we show ribbon structures of both apo-NtrC^R and BeF_3^-NtrC^R to visualize the structural changes localized in α4 and the loops before and after α4.

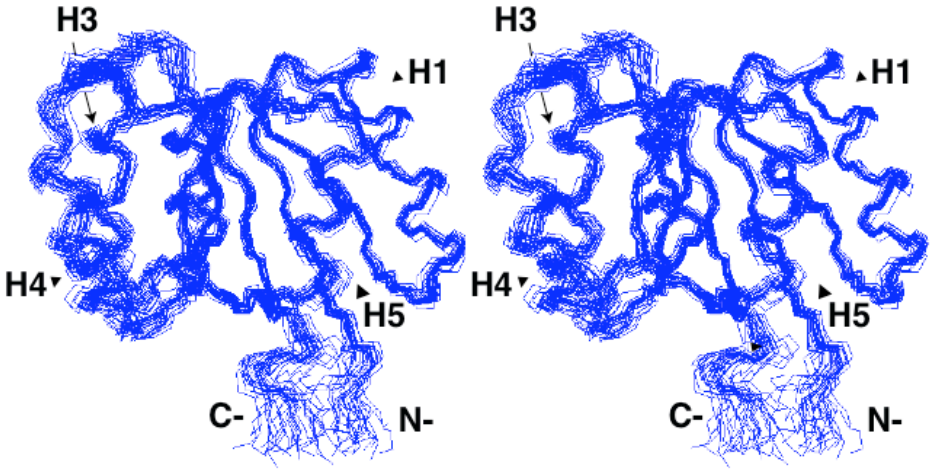


Figure 4.1 A stereoview of the 26 best NMR structures of BeF_3^-NtrC^R is shown with just backbone atoms. The N and C termini are indicated, as is the numbering of the helices.

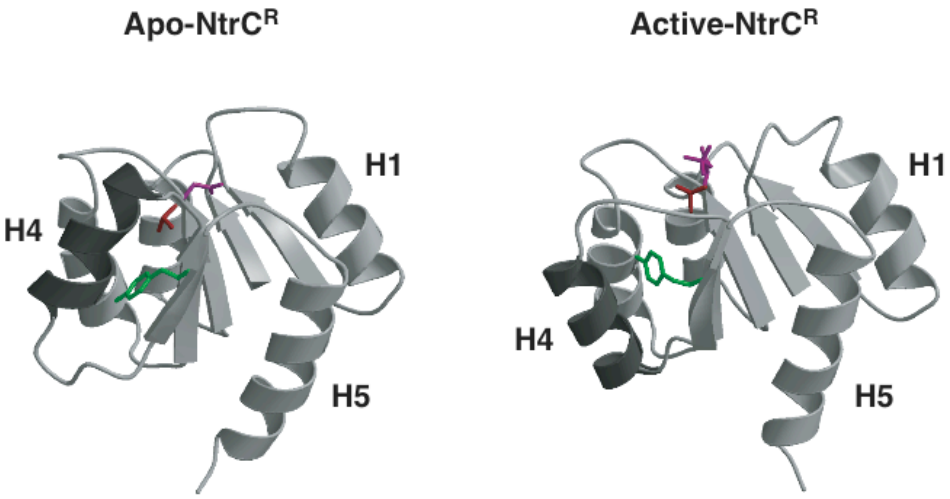
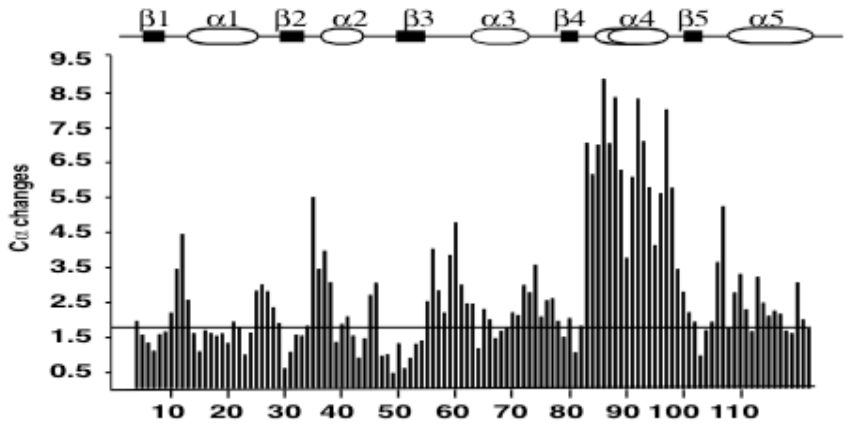
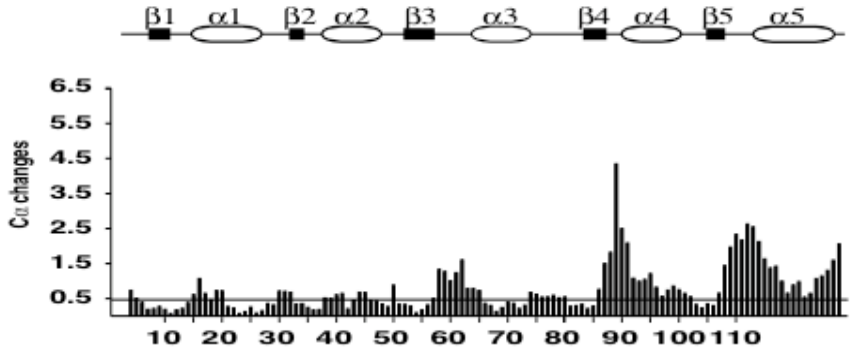


Figure 4.2 Ribbon representations of the apo-NtrC^R (left) and BeF_3^-NtrC^R (right) structures are shown. Side chains of resdues D54, T82, and Y101 are shown in magenta, red, and green, respectively. H4 in each structure is shaded dark to emphasize the structural changes that occur upon activation.

a)


β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110

b)


β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110

c)

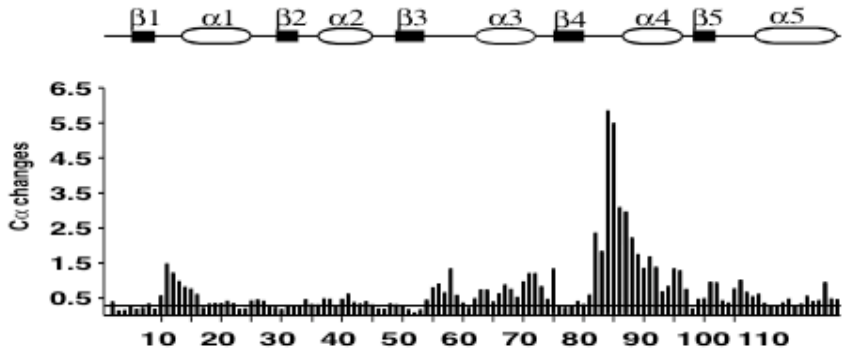

β1 α1 β2 α2 β3 α3 β4 α4 β5 α5
Cα changes
6.5
5.5
4.5
3.5
2.5
1.5
0.5
10 20 30 40 50 60 70 80 90 100 110

Figure 4.3 Changes in Cα coordinate positions upon activation are shown for $NtrC^R$ **(a)**, CheY **(b)** and $FixJ^R$ **(c)**. For $NtrC^R$, residues 4-10, 13-34, 39-55, 63-68, 73-79, 101-105, 111-122 were the least influenced by activation and were used to superimpose the apo and BeF_3^- activated forms. For CheY, residues 5-55 and 65-84 were the least influenced by activation and Cα positions were used to superimpose Mg^{2+} bound and BeF_3^- activated forms. For $FixJ^R$, residues least influenced by phosphorylation (residues 1-8, 16-52, and 106-122) were used to superimpose Mn^{2+} and phosphorylated (no metal) $FixJ^R$. The horizontal lines indicated r.m.s.d of superposition of inactive and active structures.

4.3.3 Detailed structural change of the β4-α4 loop and α4: hydrophobic core rearrangement.

The structural change in the region of α4 was previously described in P-NtrC^R and suggests that this structural change present a new binding surface for interaction with the central ATPase domain [13]. However, because of the low resolution of the P-NtrC^R structure, it was not possible to define details of sidechain interactions between α4 and the remainder of the protein. In the structure presented here, the contacts of α4 and the loop between α4 and β4, which account for most of the structural changes upon activation, can be characterized in detail.

The key interactions involving α4 are depicted in Fig. 4.4. Residues 84-86, which are part of the helix in apo-NtrC^R, adopt a defined, extended conformation that permits the N-terminus of α4 to move away from the central beta sheet; residues 95-97, defined as a loop in apo-NtrC^R become the C-terminus of α4 in BeF_3^-NtrC^R. There is some rotation of α4 about its long axis relative to apo-NtrC^R, leading to the exposure of the hydrophobic residue V91 to solvent. Interestingly, three aromatic rings showed significant movement upon activation: Y94, F99, and Y101 (Fig. 4.4). The rings of Y94 and F99 are buried in apo-NtrC^R, pointing into the pocket between α4 and β5, but are rotated in BeF_3^--NtrC^R. These changes arise from changes in backbone conformation, not just from side chain rotation. The ring of Y94 rotates out as a result of this residue's participation in helix formation upon activation. The movement of F99 is also due to backbone reorientation as the tilt of β5 is reversed upon activation (Fig. 4.4). Y101 adopts a buried conformation in BeF_3^-NtrC^R. Multiple long-range NOEs orient the aromatic side chain of Y101 in BeF_3^-NtrC^R, which adopts consistent χ^1 and χ^2 angles in

the ensemble of structures. Although this side chain is also buried in apo-$NtrC^R$, the orientation differs in the two forms. In $BeF_3^-NtrC^R$, the aromatic side chain is much closer to the loop between β4 and α4 and is canted differently with respect to the β-sheet core (Fig. 4.3 and Fig. 4.4). These ring positions are defined by multiple NOEs in both states. The change of Y101 is rather different from the equivalent residue in CheY, which moves from a solvent exposed state in the inactive conformation to a buried one in the active form. This change seems to be the key in modulating the affinity for the downstream target FliM [16]. The changes in $FixJ^R$ and $Spo0A^R$ upon activation are much closer to those in CheY than $NtrC^R$. It will be interesting to see whether there may be distinct classes of receiver domains according to the rearrangements upon activation.

4.3.4 Proposed mechanism of activation: A coupled rearrangement of T82 and Y101.

The resolution of the active $NtrC^R$ structure presented here made possible the identification of the sidechain interactions described above. In all structures of activated receiver domains which have been described to date, the highly conserved hydroxyl side chain of residue 82 (always Thr or Ser) forms a hydrogen bond with the phosphoryl (beryllofluoride) group at the active site [14-17]. As a result of the movement of this residue, a new pocket forms in the interior of the protein and the conserved aromatic side chain of residue 101 (Tyr or Phe) adopts a buried conformation. It is evident from the structure presented above that a similar correlated sidechain movement occurs upon beryllofluoride complexation by $NtrC^R$, and furthermore, that these side chain motions are related to the structural changes that occur in α4. In apo-$NtrC^R$, the side chain of T82 is close to H84, the first residue of α4 in apo-$NtrC^R$ (Fig. 4.4), and the loop between α4

and α4 (81-84) is compact. Once BeF_3^- (or a phosphoryl group) binds to D54, hydrogen bonding between the hydroxyl group of T82 and one fluorine (or oxygen in phosphate) stabilizes the geometry of the backbone with T82 flipped. With the backbone of T82 moved, the N-terminal residues of α4 in apo-$NtrC^R$ (84-86) unwind and take on an extended conformation, which also allows reorientation of the helix. This also makes it possible for two residues (95 and 96) to add to the C-terminus of α4. In conjunction with this rearrangement, Y101 fills the cavity generated by the unwinding of residues 84-86. Thus the stabilization of T82 in a new position leads to a large-scale structural rearrangement of α4 and α5. The reorganization of the sidechains of Y94 and F99 appears to be due to the helical rearrangement and change in position of the Y101 ring. F99 forms new interactions with α5 but the α5 conformation does not change significantly.

4.3.5 Implications for binding to the central ATPase domain in NtrC and comparison with other response regulators.

As with phosphorylation, beryllofluoride activation of NtrC results in a register shift, displacement relative to the active site, and rotation about the long axis of α4. The rotation of α4 alters the molecular surface by exposing V91. Helix 4's position is defined by an extensive network of NOE constraints to residues in both α3 and the β-sheet core (*vide infra*). In both P-NtrC and BeF_3^--NtrC, α4 lies at an angle with respect to the central beta sheet and is positioned toward the bottom (N-terminal end) of the sheet. The absence of D_{NN} NOEs and observed random-coil chemical shifts are consistent with the extended conformation calculated for residues 84-87, which are helical in apo-$NtrC^R$. Additionally, the observed ^{15}N-^1H residual dipolar coupling

constants imply that the HN-N bond vectors for these residues are not collinear with those in the rest of β4, and backbone dynamics data indicate that S85 is extremely mobile (order parameter 0.1). The structure of the complex of CheY and the N-terminal peptide from its downstream target FliM have shown that N-FliM binds in the pocket between β4 and β5, as well as to the loop between β4 and α4 [16]. N-terminal peptides derived from the ATPase domain of NtrC appear to bind in the same pocket as judged by NMR chemical shift changes induced by binding. These observations suggest that signaling in NtrC is mediated, at least in part, by conformational changes in β4 and β5, and confirms that the region showing the greatest structural change acts as a binding surface. I95 in CheY is involved in formation of hydrophobic core with FliM peptide; considering that V91 in BeF_3^-NtrC^R is equivalent to I95 in BeF_3^--CheY in terms of structure and sequence, and becomes exposed to the solvent upon activation, it is probable that V91 is involved in the interaction with the N-terminal region of central ATPase domain.

Both NtrC^R and CheY act positively, that is, they interact directly with a downstream partner to affect output. In NtrC the receiver domain has to bind to the central domain to stimulate assembly of the complex that contacts polymerase and hydrolyzes ATP. In CheY the rotameric state of Y106 seems to determine the signaling state. In the inactive state, Y106 is solvent exposed. When complexed with CheA (the histidine kinase in the chemotaxis system) it forms hydrogen bonds to two side chains. Activation of CheY causes Y106 to be buried and to form a hydrogen bond to the backbone of E89, stabilizing the active conformation of the β4-α4 loop and better defining the binding surface for FliM. In the Ntr system, NtrB (the histidine kinase)

seems to interact with only the receiver domain, and Y101 cannot participate directly in the interaction with NtrB since Y101 is buried in apo-NtrC^{r} as well. Instead, Y101 probably plays a key role in changing the binding surface to the central domain by filling the cavity that is created by reorientation of α4 and the β4-α4 loop. Thus, instead of a simple switch of the rotameric state of a single residue observed in CheY, a larger surface changes through a complicated reorganization by way of Y-T coupling in NtrC^{R}, and this surface change seems to determine the signaling state.

Several features of the active NtrC^{R} structure reported here, and the structure of apo-NtrC^{R} reported previously, suggest that while NtrC has many similarities with other members of the response regulator family, there are also significant differences. In particular, α4, which has been implicated in signal output in both NtrC and CheY, lies in a different position and at a different angle to the β-sheet core, and Y101 appears to change between buried configurations upon activation, rather than switching from exposed to buried. These features may be related to the intramolecular mode of NtrC regulation, and additional studies of the interaction between the regulatory and ATPase domains of NtrC will be required to further clarify differences.

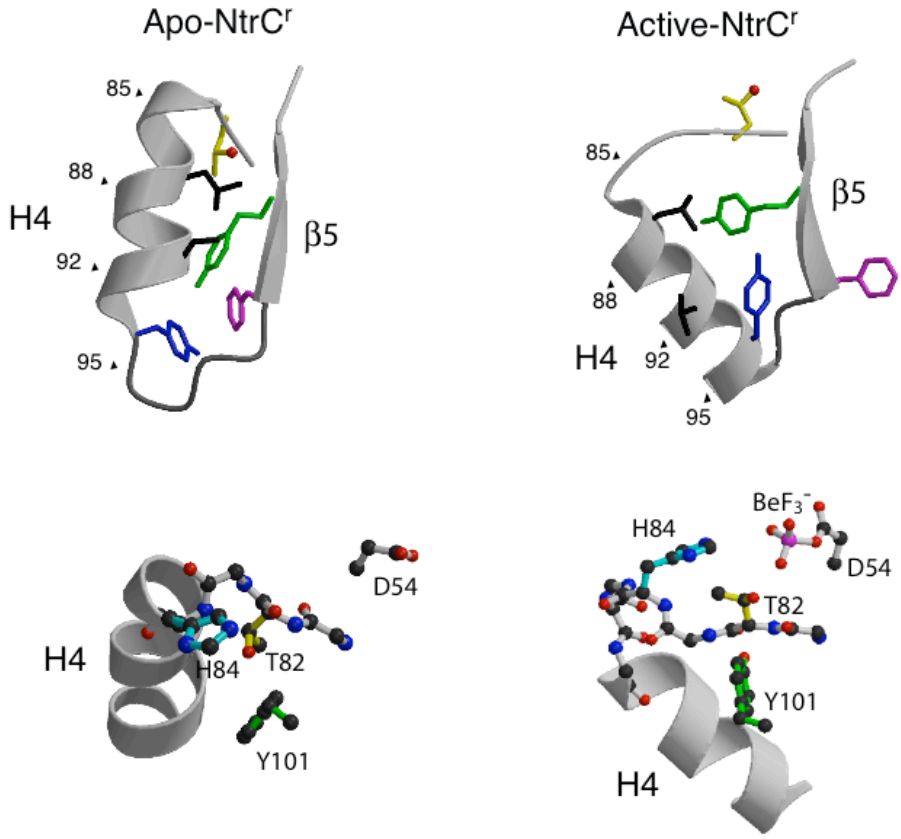


Figure 4.4 Details of the activation-induced conformational change in $NtrC^R$ are presented. **a)** The hydrophobic core region between □4 and □4 is shown with ribbon (backbone) and stick (sidechain) at key sites, T82 (yellow), L87 (black), V91 (black), Y94 (blue), F99 (blue) and Y101 (green). **b)** The loop between □4 and □4 along with □4 are shown in ball and stick and ribbon models respectively. Key side chains are also shown in ball and stick representations, including D54, T82 (yellow), H84 (cyan), and Y101 (green).

4.4 References

1. Nohaile, M., Kern, D., Wemmer, D., Stedman, K. & Kustu, S. 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 (1997).
2. Grzesiek, S. & Bax, A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201-207 (1992).
3. Grzesiek, S. & Bax, A. Correlating backbone amide and sidechain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc* **114**, 6291-6293 (1992).
4. Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T. & Gronenborn, A. M. Complete resonance assignment for the polypeptide backbone of interleukin 1 beta using three-dimensional heteronuclear NMR spectroscopy. *Biochemistry* **29**, 3542-56 (1990).
5. Talluri, S. & Wagner, G. An optimized 3D NOESY-HSQC. *J. Magn, Reson.* **112**, 200-205 (1996).
6. Kay, L. E., Xu, G.-Y., Singer, A., Muhandram, D. & Forman-Kay, J. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson.* **101 B**, 333-337 (1993).
7. Majumdar, A. & Zuiderweg, E. R. Improved ^{13}C-resolved HSQC-NOESY spectra in H2O, using pulsed field gradients. *J. Magn, Reson.* **102**, 242-244 (1993).

8. Vuister, G. W. et al. Increased resolution and improved spectral quality in 4D 13C/13C-seperated HMQC-NOESY-HMQC spectra using pulsed field gradients. *J. Magn. Reson.* **101**, 210-213 (1993).

9. Ottiger, M., Delaglio, F. & Bax, A. Measurement of *J* and Dipolar Couplings from Simplified Two-Dimensional NMR Spectra. *J. Magn. Reson.* **131**, 373-378 (1998).

10. Hansen, M. R., Mueller, L. & Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**, 1065-74 (1998).

11. Cho, H. S. et al. NMR structure of activated CheY. *J. Mol. Biol.* **297**, 543-51. (2000).

12. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* **8**, 477-86 (1996).

13. Kern, D. et al. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* **402**, 894-8 (1999).

14. Birck, C. et al. Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure Fold. Des.* **7**, 1505-15 (1999).

15. Lee, S. Y. et al. Crystal structure of activated CheY. Comparison with other activated receiver domains. *J. Biol. Chem.* **276**, 16425-31 (2001).

16. Lee, S. Y. et al. Crystal structure of an activated response regulator bound to its target. *Nat. Struct. Biol.* **8**, 52-6 (2001).

17. Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I. & Wilkinson, A. J. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol.* **294**, 9-15 (1999).

CHAPTER 5

Regulation of AAA$^+$ transcriptional activator NtrC1

NtrC (Nitrogen Regulatory protein C) family proteins are σ^{54} dependent bacterial transcription factors that are members of the AAA$^+$ ATPase family. NtrC family proteins interchange their oligomeric state from dimers, in the inactive state, to higher oligomers, in the activated state. Two crystal structures of NtrC1 from *Aquifex aeolicus*, one including the regulatory and central ATPase domains (inactive state), and one with the central ATPase domain alone (active state), were determined. The regulatory plus central domains crystallize as a dimer while the central ATPase domain crystallizes as a heptamer. Comparison of the inactive (dimer) and active (heptamer) structures explains how the regulation is achieved. In addition, a loop in the central domain, containing a conserved sequence that is involved in coupling ATP hydrolysis with RNA polymerase isomerization, forms a novel motif that is localized in a central pore within a heptamer ring and forms a surface to interact with σ^{54}.

5.1 Sample preparation and structure determination

5.1.1. Protein expression and purification

The *Aquifex aeolicus* genomic DNA was a gift from R.Huber. Both $NtrC1^{RC}$ (*Aquifex aeolicus*, GI # 2983588, residues 1-387) and $NtrC1^{C}$ (residues 121-387) were cloned from the genomic DNA into pET21a vectors (NOVAGEN). A selenomethionine derivative $NtrC1^{RC}$ protein was expressed using *E.coli* B834 (DE3) with SJS1244 in a M9 medium with selenomethionine [1]. $NtrC1^{C}$ protein was expressed using Rosetta (DE3) pLysS (to provide rare tRNAs). Purification of both $NtrC1^{RC}$ and $NtrC1^{C}$ is done in a very similar manner. Cells were harvested, and sonicated. Lysate was heated at 75-80°C for 25 min with 50 mM Tris buffer (pH 8.2), 1mM $MgCl_2$, 1 mM ATP, 1mM tris-2-carboxy-ethylphosphine (TCEP) and the supernatant was dialysed against 50 mM Tris buffer (pH 8.0) and 5 mM EDTA over night. The protein was further purified on a Q-sepharose column, and then a gel filtration column. The purified protein was extensively dialysed against 5 mM ammonium bicarbonate (2 liter chamber, 3 buffer changes for about 24hrs) and then lyophilized. Lyophilized $NtrC1^{RC}$ was redissolved in 50 mM Tris (pH 8.0), 20 mM ATP, 10 mM TCEP, 5 mM $MgCl_2$, 5% glycerol (v/v). $NtrC1^{C}$ was dissolved in the same buffer except without glycerol. The pH of TCEP and ATP solutions was adjusted to 7.0 before mixing with the protein solution and the final pH of the protein solution was 8.0. Protein solutions were filtered with a 0.2mm filter before use in crystallization trials. Gel filtration chromatography (Superdex200, 10/30) and dynamic light scattering confirmed the dimeric state of $NtrC^{RC}$ and the higher oligomeric state (M.W. ~250kDa) of $NtrC1^{C}$. The oligomeric state of $NtrC1^{C}$ is dependent on salt concentration with oligomer stabilized at low salt concentration (data not shown).

5.1.2. Crystallization and structure determination

Crystals of NtrC1^{RC} were grown at room temperature using the hanging drop vapor diffusion method. Six μl of protein solution (10 mg/ml) was mixed with an equal volume of well solution containing 0.6-0.7 M NaH_2PO_4/KH_2PO_4, 100mM citric acid (pH 6.5), 10mM imidazole, and 3% (v/v) methanol. Crystals appeared after 1 day and grew to approximately 0.2 x 0.2 x 1.0 mm (hexagonal rod shape) after one week. Crystals were transferred to well solution plus 10% glycerol for 1 min then transferred to well solution plus 20% glycerol for 5 min, after which they were flash frozen and stored in liquid nitrogen until used for data collection. These crystals have the space group $P3_221$, with unit cell dimensions a=b=94.8Å and c=195.0Å with two molecules in the asymmetric unit. SAD (Single wavelength Anomalous Diffraction) data were collected to 2.4 Å Bragg spacing at the beam line 5.0.2 of the Advanced Light Source (ALS) at Lawrence Berkeley National Lab. The data were processed and scaled with DENZO and SCALEPACK from the HKL program suite [2]. Heavy atoms sites (Se) were found with SOLVE at 3 Å [3]. Eleven out of 12 SeMet sites were found and 10 sites were related by two-fold rotations that are not related to crystallographic symmetry. Phasing was calculated with SHARP [4] up to 3Å resolution and was improved and extended by RESOLVE [5] using two-fold non-crystallographic symmetry (NCS) averaging and solvent flattening. Model building was done using O [6]. Using partial models, two separate NCS operators for both regulatory and central ATPase domains were calculated and two domain NCS-averaging was performed using DM [7]. This improved the map quality significantly and allowed us to complete the model building. The GAFTGA loop that is inserted into the H8 was not initially visible but the connecting strands were clear in the

very early stage of experimental map before NCS averaging. The loop was built after completion of the rest of chain. The structure was refined with CNS [8]. Anisotropic B factors and bulk solvent corrections as well as the cross validation methods were applied throughout the refinement. NCS restraints were applied only in the initial refinement stage and were released after the R-working and R-free factors dropped to 33%/35%. Water and solvent picking was performed after the $R_{working}$ and R_{free} dropped to 25%/28.5% using CNS. PROCHECK was used to monitor geometric parameters [9].

Crystals of $NtrC1^{C}$ were grown at room temperature using the hanging drop vapor diffusion method. 2 μl of protein solution (20mg/ml) was mixed with equal volume of well solution containing (pH 6.5), 50mM diammonium tartrate and 8% (v/v) PEG 3350. Crystals grew to approximately 0.2 x 0.2 x 0.2 mm (diamond shape) after one day. Crystals were transferred to well solution plus 10% glycerol for 2 min then transferred to well solution plus 25% glycerol for 5 min, and then flash frozen and stored in liquid nitrogen until data collection. These crystals have the space group P1, with unit cell dimensions a= 106.8Å, b=108.3Å, c=110.0Å, a=70.3 °, b=85.9 ° and g=73.3 °, with fourteen molecules in the unit cell. Data were collected at the beam line 8.3.1 of the Advanced Light Source (ALS). The data were processed and scaled with MOSFLM and SCALA using the Elves interface (J. Holton and T. Alber manuscript in preparation). Seven two-fold NCS peaks were found (at about 5σ) using the self-rotation function of Molrep [10]. Locked-cross-rotation was performed using the self–rotation solutions by Molrep at several different resolution ranges. Three sets of solutions (ten solutions each) were found at several different resolution ranges and sphere radius (25Å-40Å). Solutions that were not determined to form rings were discarded and the correct solutions were

confirmed by the decrease of R_{free} in the initial refinement with CNS. Eleven solutions were found in this way and the remaining three molecules were placed manually in the $2F_o$-F_c electron density map. After all fourteen solutions had been found, two-domain rigid body refinement (α/β domain and α-helical domain) improved $R_{working}$/ R_{free} factor significantly. Refinement was done with CNS. Coordinates are deposited in the Protein Data Bank under the accession codes 1NY5 ($NtrC1^{RC}$) and 1NY6 ($NtrC1^{C}$) respectivley.

Table 5.1. Data collection, Phasing, Refinement Statistics

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

[a] $R_{sym} = \sum|I_i - \langle I_i \rangle| / \sum I_i$, where $\langle I_i \rangle$ is the average intensity of symmetry equivalent reflections. Numbers in parenthesis are statistics for the highest resolution bin. Same for completeness.

[b] R_{cullis} (ano) = $\sum| \Delta_{ano}(obs) - \Delta_{ano}(cal) | / \Delta_{ano}(obs)$ for acentric reflections, where $\Delta_{ano}(obs)$ is the anomalous difference.

[C] Both figure of merit and R_{cullis} were calculated with reflections up to 3.0 Å

[d] R factor = $\sum|F(obs) - F(calc)| / \sum F(obs)$.

[e] R.m.s.d = root-mean-square-deviations.

5.2. Structure of the inactive NtrC1^{RC} dimer and active NtrC1^{C}

The structure of the joined N-terminal regulatory and central ATPase domains (NtrC1^{RC}) of NtrC1 from *Aquifex aeolicus* and the structure of the isolated central domain (NtrC1^{C}) were determined to resolutions of 2.4 Å and 3.1 Å, respectively by x-ray crystallography. Both were in their ADP-bound form. NtrC1^{RC} crystallized with a dimer in the asymmetric unit, whereas NtrC1^{C} crystallized with two heptamers. The dimeric form of the NtrC1^{RC} and the higher oligomeric form of NtrC1^{C} (~250kDa) in solution are supported by dynamic light scattering and gel filtration chromatography (data not shown). Data and refinement statistics are in table 5.1.

5.2.1. The inactive NtrC1^{RC} dimer and an unusual structural motif

NtrC1^{RC} is a dimer in which the regulatory domain of each monomer is adjacent to the central domain of its dimeric partner (Fig. 5.1a). The regulatory domain, often called a receiver domain that is a part of two-component signal transduction, has a typical $(\alpha/\beta)_5$ fold and contains the conserved active site residues in the same basic arrangements as they were in many previous structures of homologs [11,12]. However NtrC1^{R} has a substantial extension of C-terminal helix 5 (to residue 135, about 10 residues longer than the equivalent helix in *S. typhimurium* NtrC), and the extensions from the two monomers form a pair of crossed helices, creating a major dimerization interface of NtrC1^{RC}. The dimerization interface between the two receiver domains and the extended helices is very similar to the receiver domain of dicarboxylate transport regulator D (DctD) of *Sinorhizobium meliloti*, a homologous σ^{54} activator [13] (Fig. 5.3a).

The central ATPase domain (residues 143-387) of NtrC1 is composed of a α/β subdomain that is typical of P-loop NTPases and an α-helical subdomain that is a

distinctive feature of AAA$^+$ ATPases [14]. For comparison, five □-strands in the □/□ fold domain of our structure (central ATPase domain from NtrC1^{RC}) and that of N-ethylmaleimide sensitve factor domian 2 (NSF D2) were superimposed and aligned side by side. It is known that these five □□strands are structurally conserved among AAA$^+$ ATPases [15]. The r.m.s.d is 1.0 Å. Although the tilts and lengths of the □ helices in the □/□ subdomain are quite different, the overall structural folds are similar in the two structures (Fig. 5.2a).

There are two notable insertions protruding out from the central ATPase domain of NtrC1^{RC} that were not observed in NSF D2 (Fig. 5.2). These two ordered structures are within helix 8 (H8) and after helix 9 (H9). The □-hairpin loop after H9 contacts the insertion in H8. The □-hairpin loop after H9 was observed in two other AAA$^+$ ATPases, RuvB and HslU [15-17]. A closer view of the insertion in H8 shows an unusual structural fold, an ordered loop projecting out of the middle of this helix (Fig. 5.2b). The ordered loop is connected to H8 by way of two extended strands (containing two H-bonds typical of anti-parallel □ sheets). The segments of H8 before and after the insertion have a similar helical axis but are separated by one helical turn as a result of the insertion. Comparison to structures of the corresponding regions of other AAA$^+$ ATPases shows that the two helical segments of H8 in NtrC1 are likely to be one helix (*e.g.* H8 from NSF D2 and H4 from the □ subunit of the DNA clamp loader). Sequence comparison of these regions supports our observation (Fig. 5.2b). Based on sequence alignment of NSF D2, p97 and a subset of □^{54}-activators, the insertion begins after the second turn of H8 (AEL) and is followed by residues (FELA) that complete the helix. The portion of the inserted sequence corresponding to the ordered loop has been referred to as the GAFTGA motif in

NtrC1 and other σ^{54} activators (Fig. 5.2b). The GAFTGA loop itself is composed of two consecutive type-II β-turns. It is interesting to look into what interactions allow the unusual structural motif to exist. There are two conserved glycine (G210 and G225) residues before and after the inserted strands that allow for the distortion of the main chain configuration between helix and inserted strands. In addition, there are two conserved phenylalanine residues (F209 and F227) before and after the conserved glycine residues packed against L208 as well as neighboring hydrophobic residues holding the main helical configuration. These residues pack against L208 and other neighboring hydrophobic residues to support the structure of H8 (Fig. 5.2b and Fig. 5.2c).

A similar structural motif was observed in the actin structure [18] (1ijj.pdb) but in the case of the actin structure, the inserted loop is relatively short and not connected with strands. The helices before and after the loop are also along two different helical axes. This unusual structural motif has not been observed before to our best knowledge. Two proposed roles of this novel structural motif are discussed below: energy coupling and oligomerization.

5.2.2. The active NtrC1^{RC} heptamer

The central ATPase domain of NtrC1 alone (NtrC1^{C}) forms a ring-like heptamer (Fig. 5.1b). There are two such heptamers in the P1 unit cell, which are stacked on each other in a head-to-head fashion with an offset of ~30Å (not shown). The diameter and height of the rings are ~124Å and ~40 Å, respectively. The fold of the ATPase domain in the rings is identical to that in the dimer, except for a subdomain reorientation discussed in detail below. Calculation of rotation angles relating the seven protomers in the heptamer indicates that the heptamer rings are not perfectly symmetrical. Four of the protomers

(chains D, E, F, G in Fig. 1b) are related to each other by rotations of 51–55 º, and the other three protomers (A, B, C) are related to each other by rotations of 53-55 º, but the rotation angles between these two groups is 47 º. There are small gaps between the groups of three (A, B, C) and four (D, E, F, G) protomers. Whether this is due to crystal packing is unclear. The rotation angles and spacings between protomers are quite similar in the two rings.

5.2.3. Dimer and heptamer interfaces

In general, AAA^+ ATPases function as oligomeric rings [19]. Nature has used three ways to create such rings: 1) regulated formation of higher order oligomers from monomers or dimers; 2) formation of stable rings that remain oligomerized; or 3) formation of a ring from one polypeptide chain containing repeating AAA^+ ATPase domains [19]. Activators of σ^{54}-RNA polymerase fall in the first category. The formation of oligomers by these activators is often regulated by phosphorylation of receiver domains [20]. There are two ways that receiver domains regulate central AAA^+ ATPase domain oligomerization: positively and negatively [21-25]. The phosphorylated receiver domain interacts with central domain to activate the ATPase activity, and thus transcriptional activity, in positive regulation (example: NtrC from *Salmonella typhimurium*) [21,22] whereas the phosphorylated receiver domain releases the inhibition of the central ATPase domain in negative regulation (example: DctD from *Sinorhizobium meliloti*) [23-25]. The dimerization determinant of NtrC is localized in the DNA binding domain, but that of DctD is localized in the receiver domain. Both *in vivo* transcription and *in vitro* ATPase assays show that deletion of the receiver domain enhances both transcriptional and ATPase activity of NtrC1 from *Aquifex aeolicus* (B.T.Nixon, personal

communication). Thus, we can conclude that NtrC1 from *Aquifex aeolicus* is similar to DctD protein [13,22]. The extensive dimeric interface between receiver domains of $NtrC1^{RC}$ is very similar to that seen previously for the receiver domain of DctD (henceforth $DctD^{R}$) [26]. Recently, Nixon and colleagues determined the receiver domain structures of DctD revealing that long C-terminal helices (H5) from each receiver domain form a dimeric interface in a manner similar to that in the structure of NtrC1 and called the C-terminal extension of H5 "linker" [26] (Fig. 5.3a). Nixon and colleagues also demonstrated that mutations along the dimeric interface enhanced transcriptional activity and concluded that the dimerization is important for negative regulation of central ATPase activity in DctD [13]. Activated $DctD^{R}$ forms a very different dimer than the inactive protein, in which the dimerization is mediated by H4, rather than the crossed helices (H5) [26,27] (Fig. 5.3b).

Analysis of the dimerization interface in $NtrC1^{RC}$ shows that the dimerization interface between the receiver domains spans 2110 $Å^2$ (1055$Å^2$/monomer). The dimerization interfaces involving the interaction between the receiver and the central ATPase domains span 2840 $Å^2$ (1420$Å^2$) and that between the central domains spans 2100 $Å^2$ (1050 $Å^2$/ monomer). In $NtrC1^{RC}$ the interface between the receiver domains involves contacts between H4, β5, and H5 of the two monomers. The interface formed by the interaction between the receiver domains (including the linker) of opposite monomers in the dimer and the central domain involves strand β2, helix H8, and both the GAFTGA insertion in H8 and the β-hairpin after H9. Both β2 and helix H8 contact the C-terminal H5 extension from the receiver domain of the opposite monomer (Fig. 5.3a). Both the GAFTGA insertion in H8 and the β-hairpin after H9 contact the C-terminal

parts of H2 and H3 in this opposite receiver domain (Fig. 5.3a). The interface between the ATPase domains is primarily in the sensor II helix in the □-helical subdomain (residues 353-369). The total dimerization interface spans about 7100 Å^2 (3550 Å^2 per monomer).

Contacts mediated by the crossed H5 of NtrC1^{RC} appear to be major dimerization determinants by this analysis. Nixon tested whether amino acid substitutions predicted to cause destabilization of contacts between them would increase transcriptional activation by NtrC1^{RC} *in vivo* and confirmed that this was the case (B.T.Nixon, personal communication). Compared to the dimerization contacts discussed above, the amount of buried surface area between the ATPase domains in the heptamer of NtrC1^{C} is very similar to that on the ATPase domains in the dimer of NtrC1^{RC} and the surfaces involved are overlapping (Fig. 5.4). In the heptamer conformation, there are two surfaces in each ATPase domain participating in intermolecular interactions, which span ~2500 Å^2. The surface involved in the dimerization by interaction of the C-terminal extension of H5 (residue 125-135, termed "linker") and the central domain is overlapping with one of surfaces involved in heptamerization. The regions participating in these interactions are: □2, H8, and the two insertions (GAFTGA motif and □-hairpin) (Fig. 5.4). Once phosphorylation changes the way the receiver domains dimerize, the interactions between the receiver and the central domains are altered or are removed, enabling oligomerization. The regions on central ATPase domains (□2, H8, two insertions) then interact to form the heptamer formation (Fig. 5.3b, 5.4, and 5.7). The novel structural motif (GAFTGA motif) inserted into H8 and H8 itself may also participate in oligomerization. In the NtrC1^{RC} dimer, part of the □ hairpin after H9 and the insertion in H8 contact the receiver

domain and thus are buried, whereas in the $NtrC1^{C}$ heptamer, these regions are involved in contacts with other protomers (Fig. 5.3a and Fig. 5.4).

5.2.4. Subdomain rearrangements between dimer and heptamer

Comparison of the central ATPase domain in the inactive dimer ($NtrC1^{RC}$) to that in the active heptamer ($NtrC1^{C}$) shows a hinge rotation between the α/β subdomain and the α-helical subdomain, even though both structures have the same nucleotide, ADP, bound (Fig. 5.5a). A similar hinge rotations between α/β domain and α-helical domain were observed in Heat Shock Locus protein U (HslU) upon change in nucleotide [15,28-30]. The rotations between these subdomains in NtrC1 are about 10° counterclockwise going from the dimer to the heptamer (Fig. 5.5a), with the rotation axis at the midpoint between H7 and residue 308 and running parallel to H7. The α-helical domain along with H6 opens with respect to the α/β domain, the opened surface seeming more likely to adopt the heptameric configuration.

Closer examination of the ATP binding pocket shows that there are some rearrangements of ADP and residues that interact with ADP between dimer and heptamer, although the positions of residues comprising Walker A (also called P-loop; GXXXGKE motif) and Walker B (DEXX motif) do not change (Fig. 5.5b). These observations are consistent with the domain motions that are observed in our structures. Since the rotation axis is located near N-terminus of H7, Walker A, Walker B, and the two phosphates from ADP that are located close to the rotational axis show almost no change. In contrast, those residues that are located in the Sensor II (R357 and K360) region along with the ribose and the adenine base from ADP show significant changes.

The ADP in both of our structures is in the *anti* configuration that has been observed with p97, an NSF homolog [31].

5.2.5. Correlation of biochemical data with structural data

A wealth of genetic data is available for σ^{54} activators allowing us to map mutations that affect both ATP binding and ATP hydrolysis onto our structure (Fig. 5.5b). Most of the mutations that abolish ATP binding are localized in the Walker A (P-loop) and sensor II regions (G354 and R357: numbering from NtrC1 from *Aquifex aeolicus*) [32,33]. In our structures, the side-chain of R357 interacts with the γ-phosphate, implying that the interaction is important for holding the phosphate region of the nucleotide. R357 is located in the beginning of the helix that is categorized as the sensor II region in AAA^+ ATPases. G354 is located in the loop region before the sensor II helix. The glycine in this region is highly conserved throughout the σ^{54} activators; its effect may be to orient the sensor II helix so that R357 can interact with the γ-phosphate of nucleotide (Fig. 5.5). Mutations that abolish ATP hydrolysis are localized in the Walker B motif [34] (Fig.5.5 and Fig. 5.6). In our $NtrC1^{RC}$ structure, Mg^{2+} is observed in only one molecule of the dimer and interacts with D238. R293 (Box VII or AAA minimal consensus region), highly conserved among σ^{54} activators as well as AAA^+ ATPases, is involved in ATP hydrolysis rather than ATP binding [32,34]. The role of R293 has been proposed to be a γ-phosphate sensor. The corresponding residues (K631 for NSF D2, R359 for p97) from neighboring protomer make a contact with the γ-phosphate of ATP [35,36]. In our heptamer structure, R293 from the next protomer is located somewhat close to the ADP and will be able to reach γ-phosphate of ATP within the heptamer if ATP is bound (Fig. 5.5b). The GAFTGA portion of the insertion into H8 is highly conserved

among σ^{54} activators and has been shown to be involved in coupling the energy available from ATP hydrolysis to a conformational change of σ^{54}-holoenzyme [34,37-39]. Most mutations that abolish transcriptional activation but not ATP binding or hydrolysis are localized in the GAFTGA sequence (Fig. 5.6). Buck and colleagues recently showed that the GAFTGA loop is the region that binds to σ^{54} in phage shock protein F (PspF) with ADP-AlF_3 bound [40]. The GAFTGA loop is localized in the central pore region of the heptamer structure of $NtrC1^C$ (Fig. 5.1b). In this oligomer structure, the inserted GAFTGA loops form an extended binding surface for σ^{54} and may play a role in mediating the energy coupling that allows NtrC1 to change the conformation of σ^{54}-RNA polymerase at promoters. Taken together, mutations that abolish ATP binding, ATP hydrolysis, and transcriptional activity are mapped onto the structure of central ATPase domain (Fig. 5.6). It is interesting to see how residues that affect each functional stage (ATP binding, ATP hydrolysis, and transcription) are localized in different parts of the structure (Fig. 5.6).

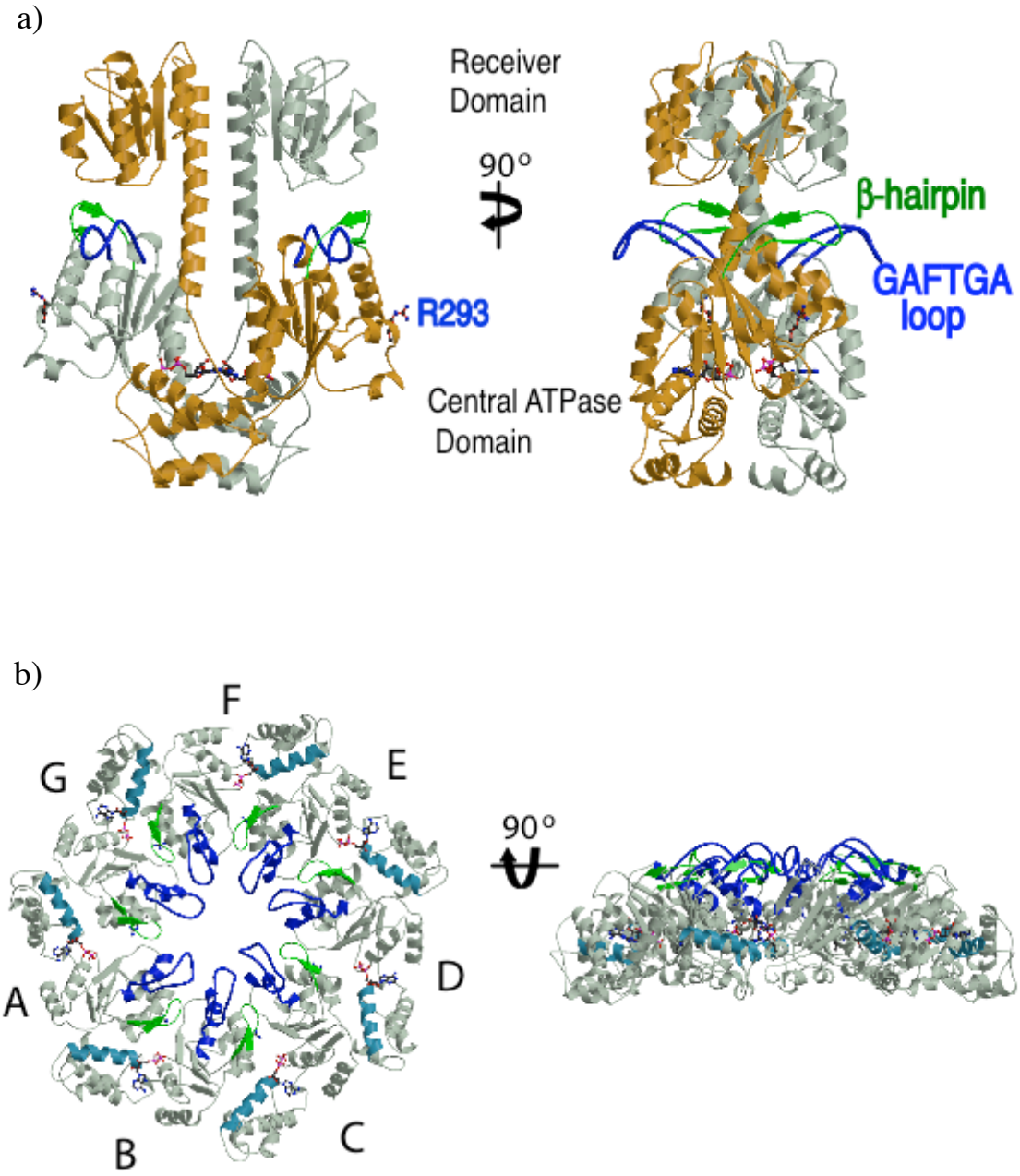
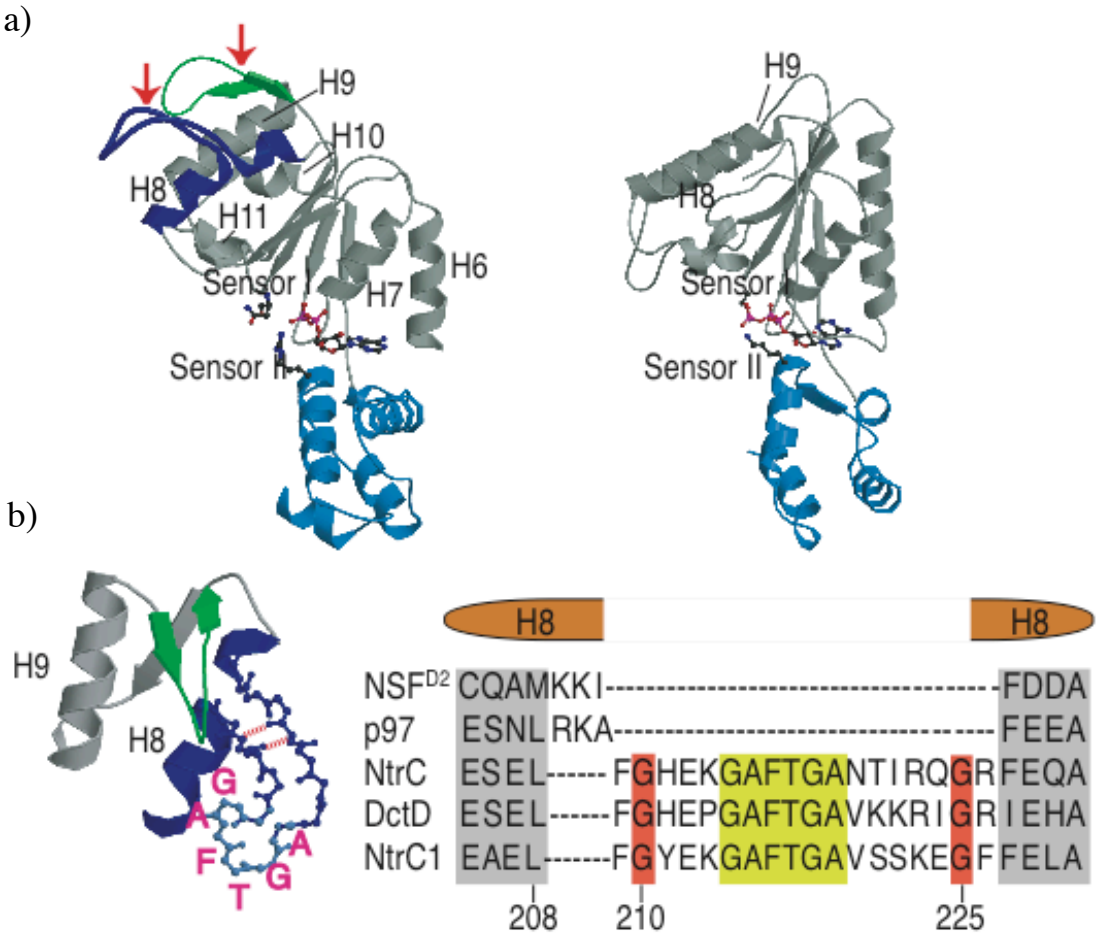

a)
Receiver
Domain
90°
β-hairpin
GAFTGA
loop
R293
Central ATPase
Domain
b)
F
G
E
A
D
B
C
90°

Figure 5.1 Overall structures of $NtrC1^{RC}$ and $NtrC1^{C}$. **a)** Structure of $NtrC1^{RC}$. $NtrC1^{RC}$ forms a dimer in the asymmetric unit. The monomers are colored gray and gold. The GAFTGA and β-hairpin insertions into the central ATPase domain are colored blue and green, respectively. ADP and the catalytic arginine residue (R293) are shown using a ball-and-stick representation. R293 cannot contact the nucleotide in the inactive dimer. **b)** Structure of the $NtrC1^{C}$ heptamer. The left side is a top view illustrating how the protomers pack in the context of the heptamer. Each protomer is labeled with a letter, A through G. The GAFTGA insertion (blue) and β-hairpin (green) comprise the pore in the heptamer. The Sensor II helix is colored cyan. On the right side is a side view.


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

c)

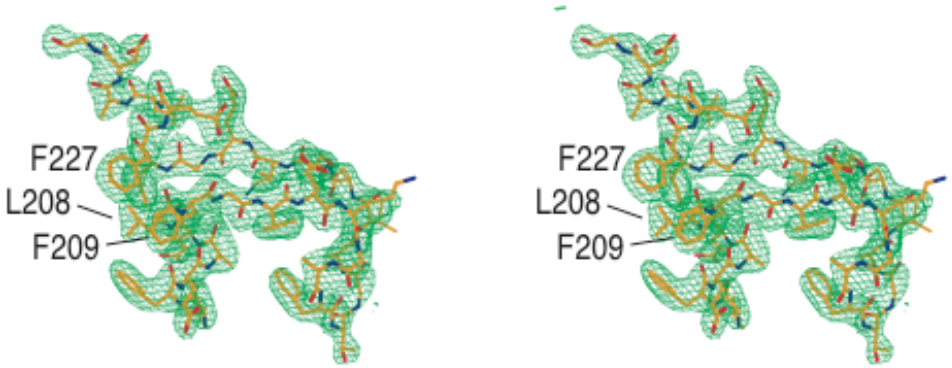

F227
L208
F209
F227
L208
F209

Figure 5.2 The GAFTGA insertion forms an unusual structural motif. **a)** Comparison of the central ATPase domain of $NtrC1^{RC}$ (left) and NSF D2 (right). The α/β and α-helical subdomains are colored gray and light blue, whereas the GAFTGA and β-hairpin insertions in NtrC1 are colored blue and green, respectively. The putative Sensor I residue (N280) and the Sensor II residue (R357) of NtrC1 are shown using a ball-and-stick representation. Note that ADP is bound in $NtrC1^{RC}$ but ATP is bound in NSF D2. **b)** A close view of the GAFTGA insertion. Two H-bonds, typical of anti-parallel β-strands, are shown as dashed lines. Sequence alignments among σ^{54} activators and other AAA^+ ATPases are shown with the conserved sequences highlighted. The sequences that form the last turn of H8 in other AAA^+ ATPases align well with those of σ^{54} activators. **c)** OMIT map of the GAFTGA region including H8 (residues 203-232) with density contoured at 3.5σ.

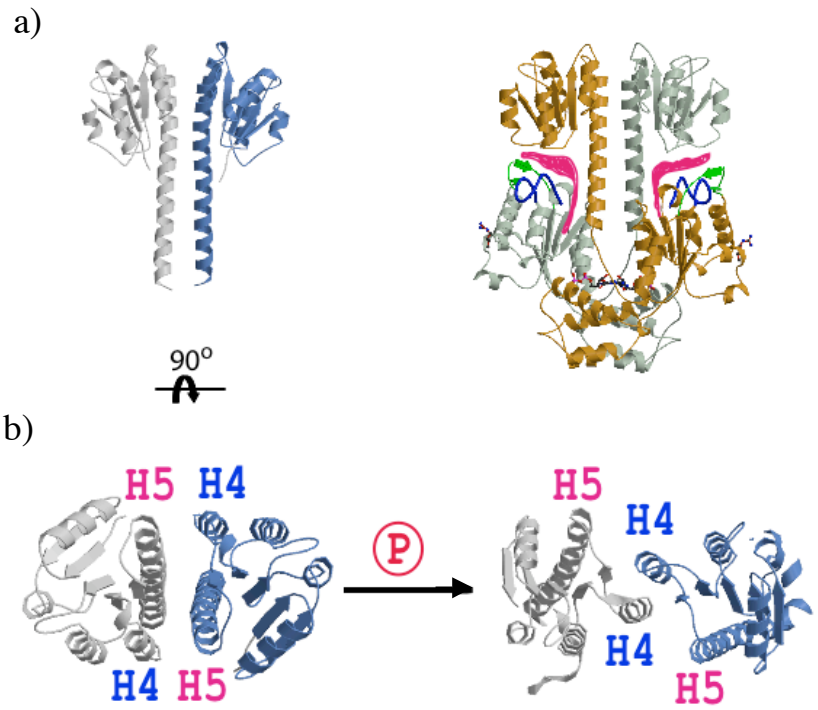


Figure 5.3 A different dimerization observed for DctD^{R} upon activation. **a)** The receiver domain structure of DctD (DctD^{R}) in the inactive state (left). Note the C-terminal extension of H5 that forms crossed helices. On the right side is NtrC1^{RC}. The regions that are colored pink are the dimer interface between the receiver and the central ATPase domains (see text). **b)** On the left is the inactive DctD dimer that is rotated 90° with respect to that in **a)** (PDB accession code 1QKK). P denotes phosphorylation. On the right side is the BeF_3^--activated DctD^{R} (PDB accession code 1L5Y). Note the change in dimerization surface between the inactive and active states.

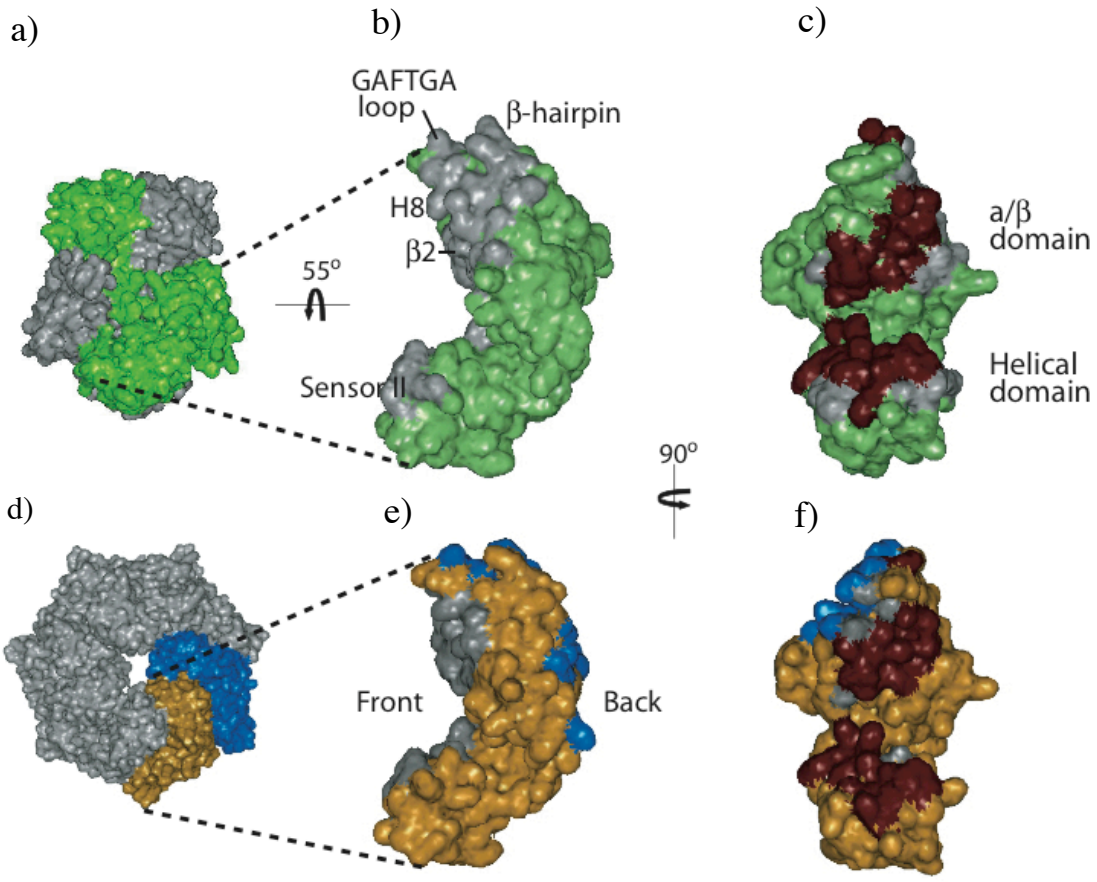
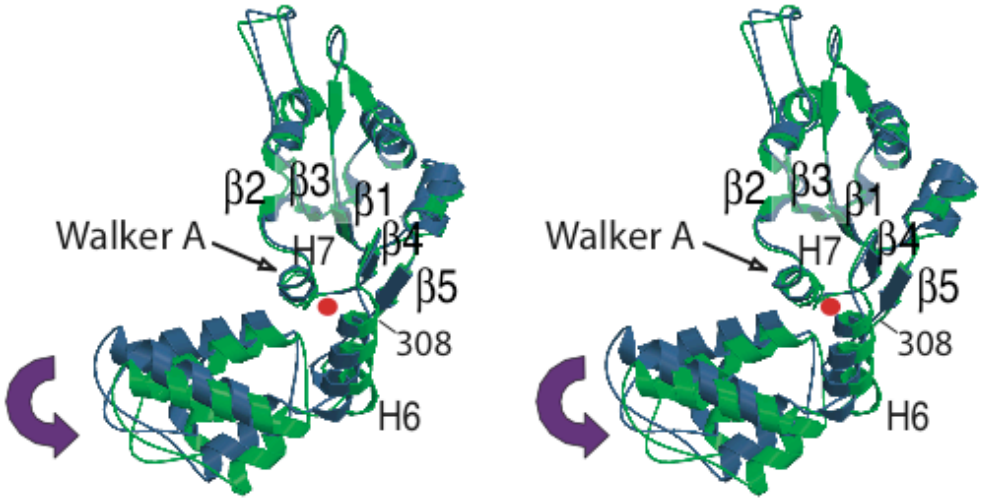


Figure 5.4 Surface representations of both the inactive NtrC1^{RC} dimer and the active NtrC1^{C} heptamer. **a)** Surface plot of NtrC1^{RC} dimer. The monomers are colored gray and green. **b)** The central domain of one monomer in the inactive dimer. Buried regions of contact with the other central domain in the dimer are shown in gray. **c)** Same as **b)** but rotated 90°. The common region that is buried in both the dimer and the heptamer structures is colored red. **d)** Surface plot of NtrC1^{C} heptamer. Two protomers from the heptamer are colored gold and blue. **e)** One protomer from the active heptamer. The contact regions with the preceding and following protomers in the heptamer are colored gray and blue, respectively. **f)** Same as **e)** but rotated 90°.

a)


β2
β3
β1
Walker A
H7
β4
β5
308
H6
β2
β3
β1
Walker A
H7
β4
β5
308
H6

b)

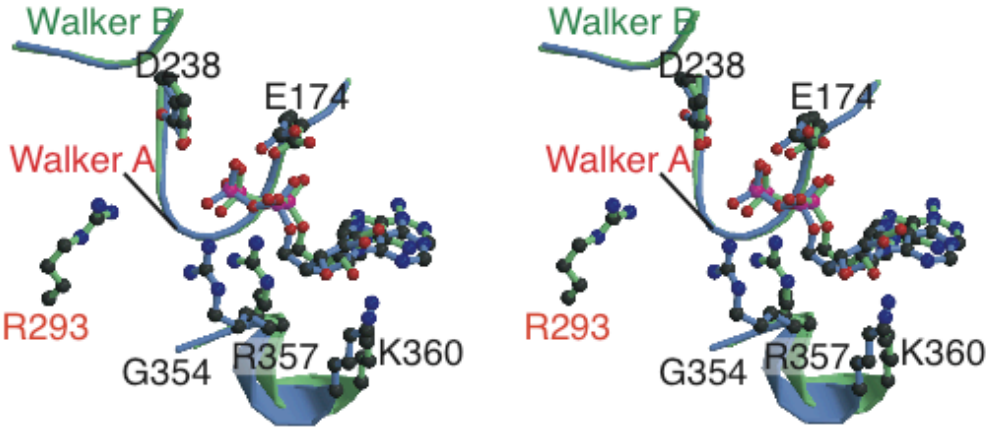

Walker B
D238
E174
Walker A
R293
G354
R357
K360
Walker B
D238
E174
Walker A
R293
G354
R357
K360

Figure 5.5 Subdomain motion in the central ATPase domain between inactive $NtrC1^{RC}$ and active $NtrC1^{C}$. **a)** The hinge motion between the □/□ subdomain and the □-helical subdomain. Residues 160 to 306 of inactive and active NtrC1 are superimposed. The central ATPase domain of the inactive $NtrC1^{RC}$ (dimer) is colored blue and that of the $NtrC1^{C}$ (heptamer) is colored green. The red dot indicates the rotation axis, which is perpendicular to the plane of the paper. **b)** The nucleotide binding sites for both $NtrC1^{RC}$ (blue) and $NtrC1^{C}$ (green) are shown. R293 (orange) from the adjacent protomer is located near the nucleotide-binding site in the active $NtrC1^{C}$ heptamer, whereas it is located on the opposite side in the inactive $NtrC1^{RC}$ dimer (see Fig. 5.1a).

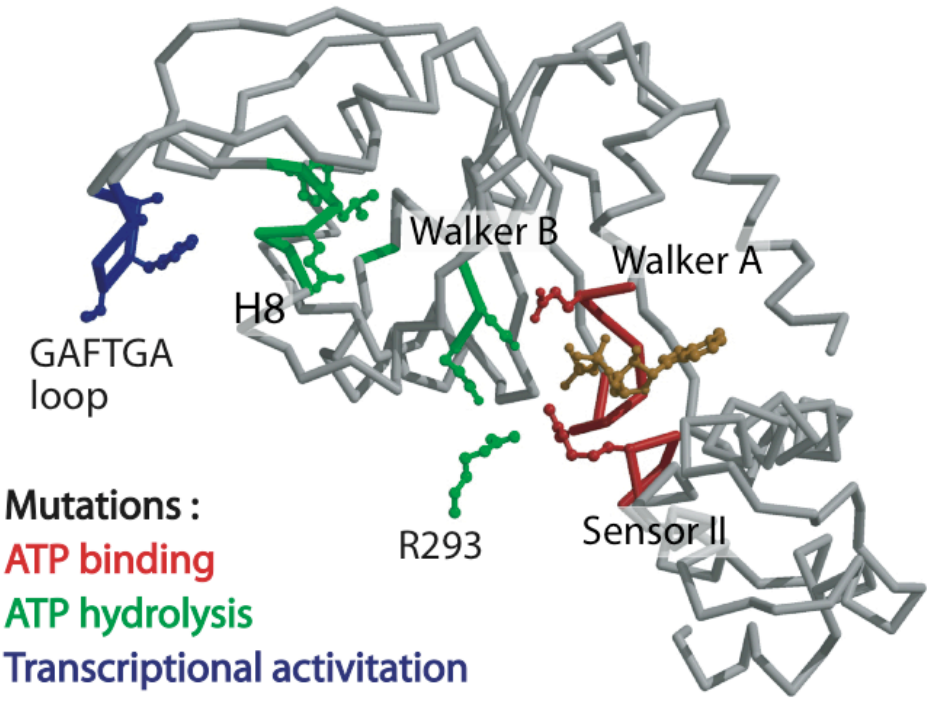


Figure 5.6 Correlation of structure with genetic and biochemical studies. Mutations that affect the function of σ^{54} activators are mapped onto the structure of the central ATPase domain. Mutations that affect ATP binding, ATP hydrolysis, and transcriptional activation are colored red, green, and blue, respectively. Residues shown in ball-and-stick representation are: E174 and R357 (red); E207, L208, F209, D237, D238 and R293 (different subunit) (green); and A215, F216, T217, and A219 (blue).

5.3. Oligomerization: from dimer to heptamer

The formation of the active oligomer is regulated by the receiver domain in σ^{54} activators. As noted earlier in this chapter, there are two ways that receiver domains can regulate central ATPase domain activity either positively or negatively. It is interesting that there is a high sequence homology (38% identity and 53% similarity) between NtrC, a positively regulated σ^{54} activator, and DctD (Di-Carboxylate Transport protein D), a negatively regulated σ^{54} activator. In *A. aeolicus* NtrC1, control is negative and appears to be very similar to that in DctD. Interestingly, calculation of buried surface in the heptamer shows that most of the residues that are involved in the dimerization are also involved in heptamerization (Fig. 5.4). These regions are mostly localized in H8, the GAFTGA insertion, β-hairpin insertion, β2, and sensor II helix and are well conserved among σ^{54}-activators. The fact that the dimeric interface indeed coincides with the heptameric interface tells us how receiver domains regulate the central ATPase domain. Contacts between the receiver domain [including C-terminal extension of H5 (125-135) of $NtrC1^{RC}$] and regions of the central ATPase domain of the opposite monomer in the dimer bury one of the two surfaces required for the central domain to form heptamers (Fig. 5.4). Thus in the inactive state, the C-terminal extension of H5 is key to prevent oligomerization by way of burying the surface required for heptamerization (Fig. 5.3 and Fig. 5.4). When phosphorylation changes the way receiver domains interact to form the alternative dimer in which crossed helices H5 no longer exist, the dimerization interfaces between the receiver and ATPase domains are disrupted [26,27] (Fig. 5.7). Although we cannot yet establish whether the active form of NtrC1 is a heptamer rather hexamer or octamer, both gel filtration study and light dynamic scattering studies show that $NtrC1^{C}$

exist as an oligomer bigger than the dimer (M.W. ~250kDa by the gel filtration study). A homoheptamer (ClpB), a heteropentamer (DNA clamp loader), and a mixture of homoheptamer and homohexamer (HslU) have been observed among AAA^+ ATPases [41-44]. Regardless of the exact number of subunits in the oligomeric ring the mechanism of regulation applies because the buried surface in the oligomeric ring is less likely to change significantly with the configurations of a hexamer or an octamer.

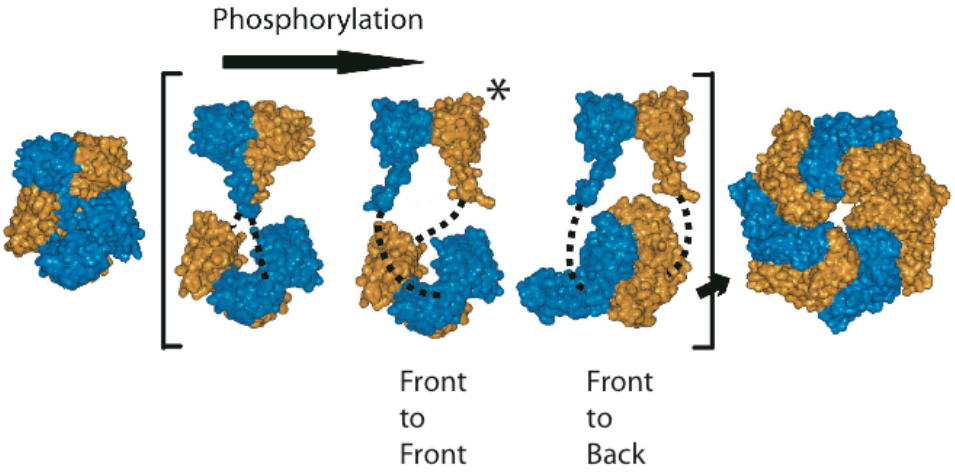


Figure 5.7 Oligomerization of NtrC1. Surface representations of NtrC1 changing from a dimer to a heptamer. In brackets, the surface representations for the receiver and central domains are separated to illustrate how they adopt different conformations upon phosphorylation. The dashed lines indicate the connections between the receiver and central domains for which there is currently no structural information. * indicates the transient state in which phosphorylation altered the dimeric configuration in the receiver domains but not the central ATPase domains, showing how the dimer interface between the receiver and the central ATPase domains can be disrupted. As a model for the phosphorylated receiver domain, we used the structure of BeF_3^--$DctD^R$ (PDB code:1L5Y), the activated form (see Fig 5.3b).

5.4. Novel structural motif and energy coupling

The GAFTGA sequence that comprises the inserted loop in H8 is highly conserved throughout σ^{54} activators and has long been known to be involved in energy coupling between ATP hydrolysis in σ^{54} activators and RNAP isomerization [32,34,37,38,40]. Mutations in the GAFTGA region abolish transcriptional activity but not ATP binding and ATP hydrolysis in general. It forms an unusual structural motif consisting of two extended strands (containing two H-bonds typical of anti-parallel β sheets) and an ordered loop projecting from the middle of H8 in the α/β subdomain of the ATPase. The ordered loop comprises two consecutive type II β-turns. In general, β-turns require four residues to complete the turn. Unlike type I β-turn, the type II β-turn requires a glycine at the third residue. There are two conserved glycines in the GAFTGA insertion to form consecutive type II β-turns (EKGAFTGA) (Fig. 5.2b). One interesting idea is that σ^{54} specifically recognizes and binds to the two consecutive type II β-turns in the activator, not a flexible and disordered loop. Unfortunately these regions in the heptamer have poor electron density (high B factors) so the model is in general not well defined. Therefore, it is not clear whether there are any changes in these type II β-turns in going from the dimer to the heptamer. The GAFTGA loop is found in the central pore of the NtrC1^{C} heptamer. Loops from different subunits form a surface that could contact σ^{54} and mediate the energy coupling between NtrC1 and σ^{54}-holoenzyme that allows the polymerase to initiate transcription.

How ATP hydrolysis can be translated into conformational change in σ^{54} by way of GAFTGA loop in the heptamer is not clear. However, there are three interesting examples that have some analogy to our structure, from which we can infer how energy

coupling might occur in NtrC1. The first example stems from HslU, part of HslUV protease complex (homologous to 26S proteosome), which contains a long insertion (I-domain) after H8 in the α/β subdomain (Fig. 5.8) [15]. HslU binds to a partially unfolded protein, completes the unfolding process, and then translocates the peptide chains into the peptidase, HslV. Domain motion between the α/β and helical domains is generated by ATP hydrolysis and propagated to the I-domain, and causes the substrate protein to unfold [29,30]. A second example is the DNA clamp loader γ-complex, for which the AAA^+ ATPase structure has been determined both with and without its target [43,45]. Both H4 of the catalytic γ-subunit (which corresponds to H8 of NtrC1) and the δ-subunit (the target of the clamp loader) undergo drastic conformational changes relative to the free proteins (Fig. 5.7) [43,45]. The third example comes from the structures of NSF D2 and p97, comparison of which lead to the proposal that H8 in p97 relocates upon ATP hydrolysis, again through a domain motion involving the α/β and helical subdomains [31]. Recently a cryo-electron microscopy study of p97 showed dramatic changes in the central of pore of the oligomer through the ATPase cycle [46]. There are key features observed among these AAA^+ ATPases: 1) There are hinge motions between the α/β and the helical domains during ATP hydrolysis (HslU and p97). 2) This motion is propagated to the H8, which leads to the change in the central pore (HslU and p97). 3) Insertions around H8 or H8 itself (H4 for clamp loader) are the substrate binding sites (σ^{54} activators, HslU, and DNA clamp loader). 4) Both the substrate binding region in the AAA^+ ATPase and the substrate undergo a conformational change (DNA clamp loader) upon complexation.

These features together with our activated $NtrC1^C$ structure suggest that the GAFTGA insertion in H8 of NtrC1 binds to σ^{54} and causes the conformational change in

σ^{54} upon ATP hydrolysis by way of a subdomain motion in the ATPase domain. The details of the mechanism by which this coupling takes place and what specific motions can be attributed to this action can be revealed by further σ^{54} activator structures in different nucleotide states and a complex of an activator with σ^{54}.

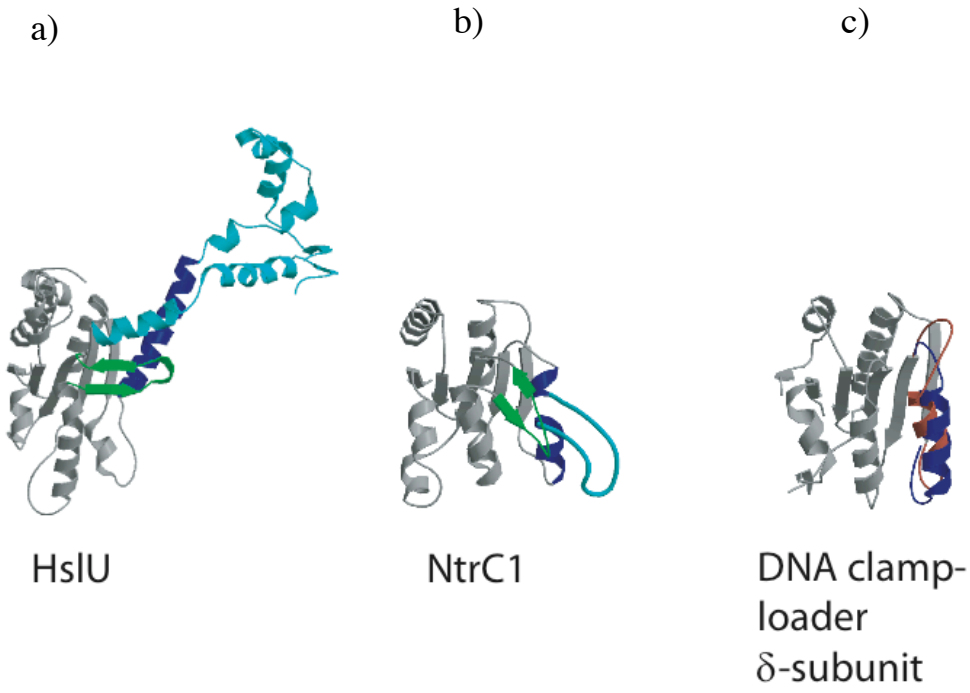


Figure 5.8 Comparison with other AAA$^+$ ATPase domains. a) □/□ domain of HslU. H8, I-domain (inserted), and □-hairpin were colored blue, cyan, and green respectively. b) □/□ subdomain of the central ATPase from NtrC1. H8, inserted GAFTGA structural motif, and □-hairpin were colored blue, cyan, and green respectively. c) □/□ domain of □ subunit of the DNA clamp-loader. Two conformations of H4 (corresponding to H8 in both NtrC1 and HslU) were shown for both with (red) and without its target (blue).

5.5 References

1. Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P. & Hendrickson, W. A. Crystallization of a fragment of human fibronectin - introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. *Protein-Struct. Funct. Genet.* **19**, 48-54 (1994).
2. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307-326 (1997).
3. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D-Biol. Crystallogr.* **55**, 849-61 (1999).
4. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472-494 (1997).
5. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D-Biol. Crystallogr.* **56**, 965-72 (2000).
6. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47 (Pt 2)**, 110-9 (1991).
7. Wang, B. C. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90-112 (1985).
8. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D-Biol. Crystallogr.* **54**, 905-21 (1998).

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

10. Vagin, A. A. & Isupov, M. N. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D-Biol. Crystallogr.* **57**, 1451-6 (2001).

11. Nixon, B. T., Ronson, C. W. & Ausubel, F. M. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. *Proc. Natl. Acad. Sci. U S A* **83**, 7850-4 (1986).

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

13. Meyer, M. G. et al. A dimeric two-component receiver domain inhibits the sigma54-dependent ATPase in DctD. *Faseb J* **15**, 1326-8 (2001).

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

15. Bochtler, M. et al. The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature* **403**, 800-5 (2000).

16. Han, Y. W. et al. A unique beta-hairpin protruding from AAA^+ ATPase domain of RuvB motor protein is involved in the interaction with RuvA DNA recognition protein for branch migration of Holliday junctions. *J. Biol. Chem.* **276**, 35024-8 (2001).

17. Yamada, K. et al. Crystal structure of the Holliday junction migration motor protein RuvB from *Thermus thermophilus* HB8. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1442-7 (2001).

18. Bubb, M. R. et al. Polylysine induces an antiparallel actin dimer that nucleates filament assembly: crystal structure at 3.5-A resolution. *J. Biol. Chem.* **277**, 20999-1006 (2002).

19. Vale, R. D. AAA proteins. Lords of the ring. *J. Cell. Biol.* **150**, F13-9 (2000).

20. Wyman, C., Rombel, I., North, A. K., Bustamante, C. & Kustu, S. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* **275**, 1658-61 (1997).

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

22. Weiss, D. S. et al. *Prokaryotic transcriptional enhancers.* (eds. McKnight, S. L. & Yamamoto, K. R.) (Cold Spring Harbor Laboratory, NY, 1992).

23. Simms, S. A., Keane, M. G. & Stock, J. Multiple forms of the CheB methylesterase in bacterial chemosensing. *J. Biol. Chem.* **260**, 10161-8 (1985).

24. Huala, E., Stigter, J. & Ausubel, F. M. The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. *J. Bacteriol.* **174**, 1428-31 (1992).

25. Gu, B., Lee, J. H., Hoover, T. R., Scholl, D. & Nixon, B. T. *Rhizobium meliloti* DctD, a sigma 54-dependent transcriptional activator, may be negatively

controlled by a subdomain in the C-terminal end of its two-component receiver module. *Mol. Microbiol.* **13**, 51-66 (1994).

26. Park, S. et al. Two-component signaling in the AAA^+ ATPase DctD: binding Mg^{2+} and BeF_3^- selects between alternative dimeric states of the receiver domain. *Faseb J* **16**, 1964-6 (2002).

27. Park, S., Zhang, H., Jones, A. D. & Nixon, B. T. Biochemical evidence for multiple dimeric states of the *Sinorhizobium meliloti* DctD receiver domain. *Biochemistry* **41**, 10934-41 (2002).

28. Sousa, M. C. et al. Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* **103**, 633-43 (2000).

29. Wang, J. et al. Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure (Camb)* **9**, 177-84 (2001).

30. Wang, J. et al. Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU. *Structure (Camb)* **9**, 1107-16 (2001).

31. Zhang, X. et al. Structure of the AAA ATPase p97. *Mol Cell* **6**, 1473-84 (2000).

32. Rombel, I., North, A., Hwang, I., Wyman, C. & Kustu, S. The bacterial enhancer-binding protein NtrC as a molecular machine. *Cold Spring Harb Symp Quant Biol* **63**, 157-66 (1998).

33. Perez-Martin, J. & de Lorenzo, V. ATP binding to the sigma 54-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. *Cell* **86**, 331-9 (1996).

34. Rombel, I. et al. MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. *J. Bacteriol.* **181**, 4628-38 (1999).

35. Yu, R. C., Hanson, P. I., Jahn, R. & Brunger, A. T. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat. Struct. Biol.* **5**, 803-11 (1998).

36. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**, 525-36 (1998).

37. Wang, Y. K., Lee, J. H., Brewer, J. M. & Hoover, T. R. A conserved region in the sigma54-dependent activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. *Mol. Microbiol.* **26**, 373-86 (1997).

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

39. Zhang, X. et al. Mechanochemical ATPases and transcriptional activation. *Mol. Microbiol.* **45**, 895-903 (2002).

40. Chaney, M. et al. Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. *Genes Dev.* **15**, 2282-94 (2001).

41. Miyata, T. et al. Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J. Struct. Biol.* **131**, 83-9 (2000).

42. Kim, K. I. et al. Heptameric ring structure of the heat-shock protein ClpB, a protein-activated ATPase in *Escherichia coli*. *J. Mol. Biol.* **303**, 655-66 (2000).

43. Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell* **106**, 429-41 (2001).

44. Rohrwild, M. et al. The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nat. Struct. Biol.* **4**, 133-9 (1997).

45. Jeruzalmi, D. et al. Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* **106**, 417-28 (2001).

46. Rouiller, I. et al. Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. *Nat. Struct. Biol.* **9**, 950-7 (2002).