PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Coarse-grained dynamics of the receiver domain of NtrC: Fluctuations, correlations and implications for allosteric cooperativity

Ming S. Liu,[1]* B. D. Todd,[1] Shenggen Yao,[2] Zhi-Ping Feng,[2] Raymond S. Norton,[2] and Richard J. Sadus[1]

[1] Centre for Molecular Simulation, Swinburne University of Technology, Hawthorn 3122, Australia
[2] The Walter and Eliza Hall Institute of Medical Research, Parkville 3050, Australia

ABSTRACT

Receiver domains are key molecular switches in bacterial signaling. Structural studies have shown that the receiver domain of the nitrogen regulatory protein C (NtrC) exists in a conformational equilibrium encompassing both inactive and active states, with phosphorylation of Asp54 allosterically shifting the equilibrium towards the active state. To analyze dynamical fluctuations and correlations in NtrC as it undergoes activation, we have applied a coarse-grained dynamics algorithm using elastic network models. Normal mode analysis reveals possible dynamical pathways for the transition of NtrC from the inactive state to the active state. The diagonalized correlation between the inactive and the active (phosphorylated) state shows that most correlated motions occur around the active site of Asp54 and in the region Thr82 to Tyr101. This indicates a coupled correlation of dynamics in the "Thr82-Tyr101" motion. With phosphorylation inducing significant flexibility changes around the active site and α3 and α4 helices, we find that this activation makes the active-site region and the loops of α3/β4 and α4/β5 more stable. This means that phosphorylation entropically favors the receiver domain in its active state, and the induced conformational changes occur in an allosteric manner. Analyses of the local flexibility and long-range correlated motion also suggest a dynamics criterion for determining the allosteric cooperativity of NtrC, and may be applicable to other proteins.

Proteins 2008; 73:218–227.

Key words: nitrogen regulatory protein C (NtrC); coarse-grained dynamics; normal mode analysis; intrinsic fluctuation; flexibility; dynamical correlation; allosteric cooperativity; two-component system.

INTRODUCTION

Nitrogen regulatory protein C (NtrC) is a signal transcription switching factor that helps to regulate the nitrogen metabolism and activate transcription by the σ^{54}-holoenzyme form of RNA polymerase.[1–4] The N-terminal receiver domain of NtrC, also called the switch domain, exerts a dynamical switch function for NtrC as a "two-component" regulatory protein. The first component of a kinase is activated by phosphorylation in response to an environmental stimulus; the phosphorylated N-terminal receiver domain then acts on the central output domain to promote oligomerization of the otherwise dimeric protein. Formation of NtrC oligomers is required for ATP hydrolysis, which in turn leads to transcriptional activation through NtrC interaction with the σ^{54}-holoenzyme form of RNA polymerase.[1,4] Nuclear magnetic resonance (NMR) studies have revealed that the receiver domain of NtrC undergoes large structural re-arrangements upon phosphorylation.[2] It was also found that both the active (phosphorylated) and inactive (unphosphorylated) conformations of NtrC exist in an equilibrium population, with phosphorylation promoting a conformational population shift toward the fully active state.[3] This provided the first evidence that dynamics plays a pivotal role in allosteric functions of proteins.[5] However, mechanisms and dynamical pathways underlying the long-range correlated motions and allosteric cooperativity of NtrC remain unclear at the molecular level.

Two-component systems regulate a wide array of biological processes in all forms of life.[6] Two-component signaling transduction normally integrates signal sensing with other regulation mechanisms to yield global regulation, such as, by metabolites, uridylylation, phosphorylation and protein–protein contacts. Two-component functions involve complex cooperativity schemes and coupled dynamical correla-

Additional Supporting Information may be found in the online version of this article.
Abbreviations: GNM, gaussian network model; MD, molecular dynamics; NMA, normal mode analysis; NMR, nuclear magnetic resonance; NtrC, nitrogen regulatory protein C; PDB, protein data bank; RMSD, root-mean-square deviation.

Grant sponsor: Australian Research Council (ARC Discovery Project); Grant number: DP0557991.
*Correspondence to: Ming S. Liu, Centre for Molecular Simulation, Swinburne University of Technology, P. O. Box 218, Hawthorn 3122, Australia. E-mail: mliu@swin.edu.au
Received 26 November 2007; Revised 7 February 2008; Accepted 19 February 2008
Published online 15 April 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22056

tions between different domains or subunits of the system. Normally these dynamic phenomena take place over large distance and timescales in an allosteric manner.[7] For example, transduction modules in NtrC proteins have separate but highly conserved protein domains for sensor and receiver functions. Once an environmental stimulus phosphorylates the sensor domain, transfer of the phosphate to the aspartate site of the receiver domain dynamically triggers a conformational change that activates the signal transduction process.[4] As internal mobility and correlated dynamics of these proteins intrinsically regulate their biological activities, an understanding of dynamical collectivity, transitions and correlations between different domains or different conformational states is vital for the understanding of two-component systems, and more generally to predict signaling networks and enzyme allosteric sites in allostery-based drug design.[8]

Here, we use NtrC as a model system for a coarse-grained dynamics investigation of protein dynamics, particularly to identify long-range dynamic correlations and determine the dynamics underlying allosteric cooperativity and the "two-component" transduction. For determining protein dynamics, experimental methods such as NMR[9,10] have both size and timescale limitations. Moreover, protein conformational changes usually occur on timescales currently inaccessible to explicit molecular dynamics simulation, e.g. μs–ms. Therefore new computational approaches are required to overcome the size and timescale difficulties associated with probing protein dynamics. One effective approach is to coarse grain the atomistic details of protein with a Go-network model to speed up the dynamics simulation.[11–14] Meanwhile, normal mode analysis (NMA) using elastic network models[15–18] has been adopted to describe the flexibility and collective large-amplitude motions of proteins interconverting among different conformations. We have applied the NMA-elastic network model to the SPRY domain-containing SOCS proteins[19] and multi-subunit motor proteins,[20] where collective dynamics of the functional domains/motifs were probed at a timescale and spatial resolution at which either molecular dynamics experiments or explicit simulations encounters limitations.

In this work, we have applied coarse-grained dynamics to analyze the intrinsic fluctuations, elasticity, flexibility, and correlations of the NtrC protein and compared the results with those of NMR experiments[2,3,21] and explicit molecular dynamics simulations.[22] We have obtained a quantitative description of dynamical stability, residue fluctuations, and correlated motions in the receiver domain of NtrC upon phosphorylation. Possible dynamical pathways associated with the structural transitions in NtrC are identified using overlapped NMA. These dynamical analyses lead to new insights into how different flexibility and long-range correlated motion induce the allosteric cooperativity of NtrC. The general application of coarse-grained dynamics to elucidating allostery in protein conformational changes is also discussed.

METHODS

Structural models of the receiver domain of NtrC

The N-terminal receiver domain of NtrC is homologous to a large family of signal transduction receiver domains.[23] This superfamily has highly conserved active site residues. Structures of both the inactive state (unphosphorylated, e.g. PDB:1dc7)[2] and the activated state (phosphorylated, e.g. PDB:1dc8)[2] of the receiver domain of NtrC were determined in solution, and they were taken as the modeling structures for different conformational states. A BeF_3^--activated NtrC receiver domain (PDB:1krw, whose structure is in fact the same as the phosphorylated state of PDB:1dc8 with a backbone RMSD of 0.57 Å[21]) was also used as a model of the active state for comparison. Figure 1 shows a superimposition of the inactive and active structures of the receiver domain of NtrC. The receiver domain contains one α, β twisted open-sheet structure, comprising two β-α-β motifs that are connected by α helices. Two adjacent β strands form a binding crevice lined by two adjacent loop regions. Positioned in the crevice is the active site aspartate (Asp54) where phosphorylation mostly occurs. Key structural changes from the inactive to active state are the re-arrangements of the α3 and α4 helices and the β3/α3 and α4/β5 loops leading out from the active site of Asp54 (in the supplemental Figure S1, these structural re-arrangements are represented by a distance-difference matrix between the two conformations).

Dynamical correlations in proteins

Dynamical correlations in a protein can be interpreted by the cross-correlations among residue fluctuations when the protein undergoes conformational changes. As the covariance matrix calculates the correlated fluctuations of atoms throughout the protein, to determine coupling between intra-and inter-domain changes, we define a difference correlation matrix of motions between two discrete conformations of protein, e.g. upon ligand binding, as,

$$C_{AB} = \left(C_{ij}^{A} - C_{ij}^{B}\right)\Big/2 = \left(\left\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \right\rangle^{A} - \left\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \right\rangle^{B}\right)\Big/2 \quad (1)$$

where A and B refer to two dynamically distinguished conformations. In the case of NtrC, they refer to the active and inactive states. The diagonalized and normalized C_{AB} should have values varying from −1, 0 to +1 showing the strongly-correlated, noncorrelated, and

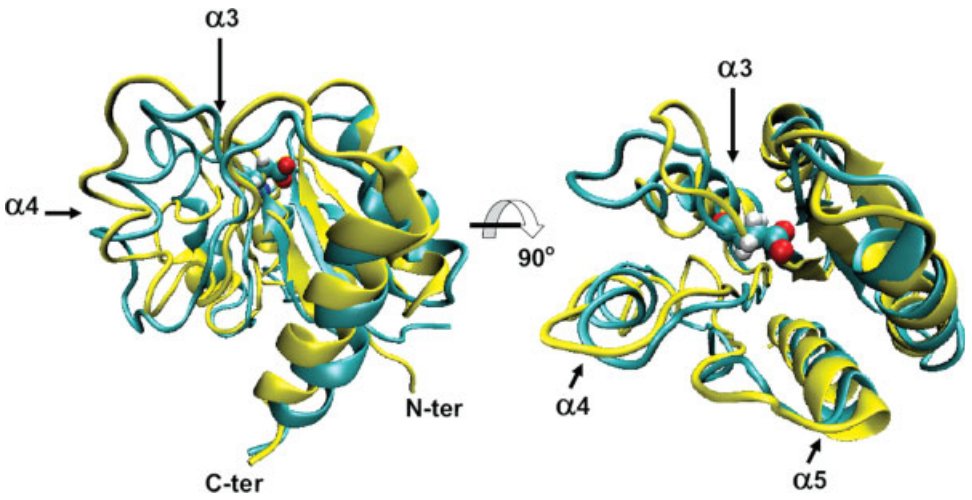

Figure 1

Superimposed structures of the inactive state (cyan, PDB:1dc7) and the active state (yellow, PDB:1dc8) of the receiver domain of NtrC. Two structures were aligned over residues 4-9, 14-53, and 108-121. The phosphorylation active site of Asp54 is highlighted by CPK atoms, and termini and α3-5 are labeled.

strongly-anticorrelated motions between conformations of A and B. $\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle$ refers to the variance-covariance matrix of the atomic fluctuations between atom i and atom j. This matrix can be derived from dynamics data from either experimental measurements or molecular simulations. However, both experimental methods and molecular dynamics simulations currently have great difficulty in handling large biomolecules over long timescales. Here, an alternative approach is used for an accurate, and computationally fast, determination of $\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle$. Using the dynamics algorithm of an improved Gaussian network model (GNM),[15] the atomic details of a protein can be coarse-grained by a Gaussian elastic network of its backbone C^{α} atoms. Mean-square fluctuations of residues, or the cross-correlations, can then be determined via,[18]

$$\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle = \frac{3k_B T}{\gamma} \left[\Gamma^{-1}\right]_{ij} \qquad (2)$$

where γ is a theoretical parameter mimicking the strength of the harmonic potential of the elastic network of C^{α} atoms. $\left[\Gamma^{-1}\right]_{ij}$ is the ijth element of the inverse Kirchhoff matrix Γ, a symmetric matrix also known as the valency adjacency matrix in graph topology theory. In the GNM approximation, the elements of Γ are given by,[15]

$$\Gamma_{ij} = \begin{cases} -1, & \text{if } i \neq j \text{ and } \mathbf{r}_{ij} \leq r_C \\ 0, & \text{if } i \neq j \text{ and } \mathbf{r}_{ij} > r_C \\ -\sum_{i \neq j} \Gamma_{ij} & \text{if } i = j \end{cases} \qquad (3)$$

The summation for evaluating Γ_{ij} is performed over all off-diagonal elements on the ith residue at a cutoff distance of r_C, which defines the range of interactions between residues. The mean-square fluctuations of C^{α} atoms from Eq. (2) also lead to theoretical temperature factor B-values, i.e. $B_i = 8\pi^2 \cdot \langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_i \rangle / 3$, which compare well with X-ray crystallography measurements and other methods.[15]

Normal mode analysis using elastic network model

NMA provides an approximation of protein motions by expanding the motion into a superposition of different normal modes of vibrations, where each normal mode is characterized by a frequency of vibration of overdamped dynamics. Low frequency NMA modes using elastic network models have been applied widely to globular proteins[16,17] to describe their collective dynamics, dynamic domains, hinge-bending motion, and other functionally important motions. As proteins mostly carry out their biological functions dynamically starting near the native states, NMA using elastic network model of proteins is based on an harmonic perturbation approximation of the potential energy function around a global minimum conformational state. Therefore, NMA captures the intrinsic modes of motions encoded in the protein structures. Conformational transitions in proteins, such as open/closed conformers upon ligand binding or oligomerization due to protein–protein interaction, can be well captured via NMA.

NMA can discriminate conformation changes between large-scale, long-range amplitude motions (via low frequency modes) and confined, local motions (via high frequency modes). The relative importance of normal modes in a particular conformational change is evaluated

in terms of collectivity and overlap coefficients.[16] Collectivity measures the collective degree of protein motion in a certain mode, i.e. the number of residues significantly involved in that mode. The "overlap coefficient," defined as O_k in Eq. (4),[16,24] quantifies the degree of agreement between the movement predicted from a particular normal mode sampling (e.g. $\mathbf{u}_k$) and the "observed" movement (in this case, the interconversion between the inactive state and the active state of NtrC), through the dot product of the difference vector $\Delta\mathbf{r}$ between two known structures of the same macromolecule and each normal mode $\mathbf{u}_k$,

$$O_k = \sum_n \Delta\mathbf{r}_n \cdot \mathbf{u}_k^n \Big/ \left[\sum_n (\Delta\mathbf{r}_n)^2 \cdot \sum_n (\mathbf{u}_k^n)^2\right]^{1/2} \quad (4)$$

where $\Delta\mathbf{r}_n = \Delta\mathbf{r}_n^{\mathrm{A}} - \Delta\mathbf{r}_n^{\mathrm{IA}}$, $\Delta\mathbf{r}_n^{\mathrm{A}}$ and $\Delta\mathbf{r}_n^{\mathrm{IA}}$ are the nth atomic coordinate of the protein in the active and inactive structures, respectively. A value of one for the overlap coefficient means that the direction given by normal mode k is identical to $\Delta\mathbf{r}$, whereas a value of zero means the direction is uncorrelated with $\Delta\mathbf{r}$. Thus high overlap coefficients help lead to identification and prediction of the modes that can be used to describe the possible pathways of a conformational transition between different states of a protein.[16,25–27]

In this work, NMA was performed mainly with the GNM algorithm[15] where the protein structure details were coarse-grained by a Gaussian elastic network of the backbone C^{α} atoms. Fluctuations (interpreted as the residue temperature factor) were derived by Eq. (2). Dynamical correlations in conformational transitions were determined by Eq. (1). The residue stability/flexibility constant of NtrC can also be determined from the dynamical ensembles of conformational states. These conformations can be derived from different NMA modes. To do this, a COREX/BEST algorithm[28] was employed where the fluctuations for different conformational states, e.g. the inactive and active states, were sampled by statistical thermodynamics rather than structural states.[29] COREX/BEST produces a stability constant at each residue as thermodynamic metric wherein each of the static structural properties is weighted according to its energetic impact (e.g. surface area, polarity, and packing are implicitly considered). The dynamical nature of stability[28,29] means that residues with high stability constants will be less flexible and will fold in the majority of highly probable states. On the other hand, residues with low stability constants will be more flexible and less likely to be folded in many of the highly probable states.

Modeling, simulation, and visualizations

The atomic coordinates of PDB:1dc7,[2] PDB:1dc8[2] and PDB:1krw[21] were used as the initial model structures for the inactive state and active states of NtrC, respectively. They were first relaxed by molecular mechanics (MM) minimization before the NMA/GNM and COREX/BEST computations. The MM minimization was performed using approximately 1000 steps by NAMD[30] with threshold energy less than 10^{-3} kcal mol^{-1} Å^{-1}. The dynamics minimization and the normal modes analysis were carried out on a Linux cluster of Pentium-4 workstations. Typically it takes only a few CPU minutes (<5 min) to accomplish a GNM calculation for the receiver domain of NtrC (of 124 residues). The residues stability constants were calculated within 24 CPUh[28,29] for about 150,000 Monte Carlo ensembles of conformations of the input structures with a window-size of seven residues. Structural alignment was done by SUPERPOSE[31] and visualization graphics were processed by the virtual molecular dynamics (VMD) packages.[32]

RESULTS

Dynamical fluctuations and correlations in NtrC

Theoretical temperature factors of the inactive (PDB:1dc7) and the active (PDB:1dc8) states of the receiver domain of NtrC are presented in Figure 2. The residue temperature factor is a measure of fluctuations and flexibility of backbone C^{α}. Except for the termini, major fluctuation changes upon phosphorylation occur throughout the β3/α3 loop and α3-β4-α4-β5 region (also referred as the "3445" face). Significant changes of fluctuations appear in the β3/α3 loop (where the activa-

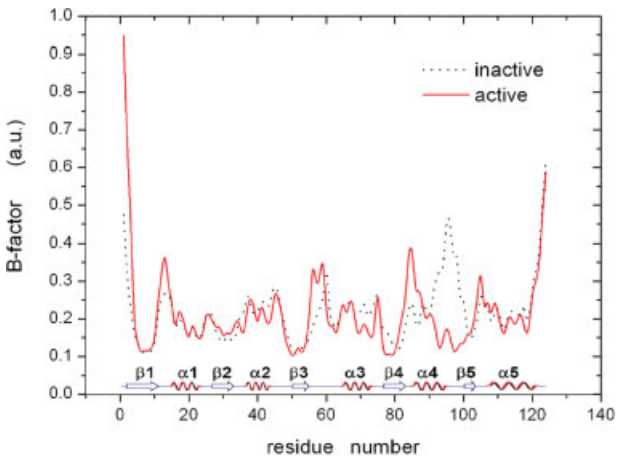

Figure 2

Residue temperature factors (B-factors) of the active and inactive states predicted by the elastic network model of backbone C^{α} using the GNM algorithm. A cutoff value for the harmonic potential energy was set at r_c = 10 Å. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

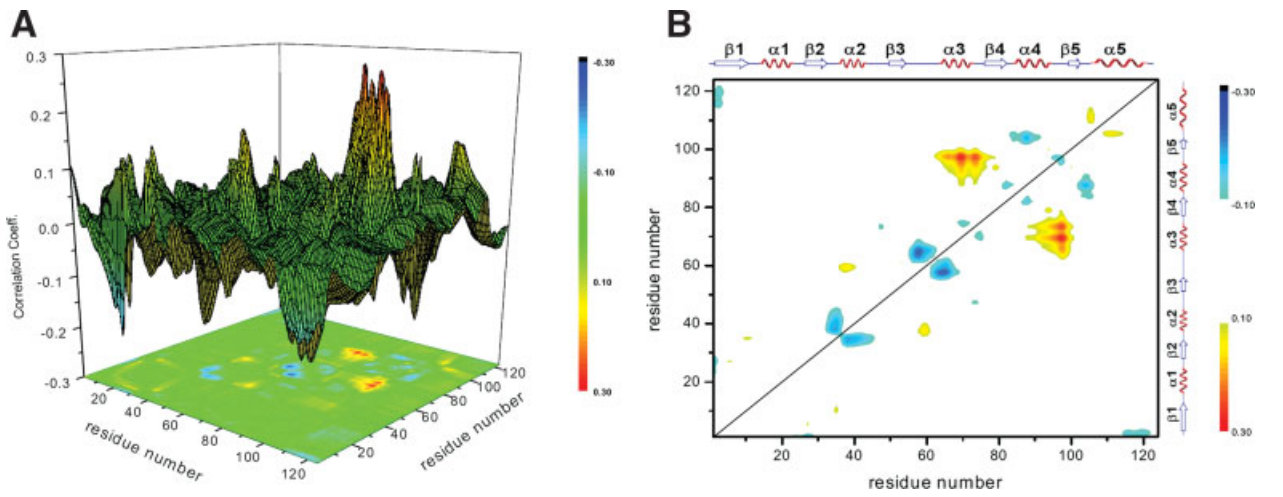

Figure 3

*Difference correlation matrix of the intra-domain motions of NtrC, with color coding in (**A**) 3D surface map mode and (**B**) 2D Contour map mode, before/after the phosphorylation at the active site Asp54. An anticorrelated motion (red) of α3-helix and α4-helix occurs along with a correlated motion (blue) of the segments upstream and downstream of the active site region. Red (blue) regions correspond to opposite (same) direction distortions.*

tion site Asp54 is located), the β4/α4 loop, α4-helix and the α4/β5 loop. These are the dynamical consequence of structural transitions between the inactive and the active conformers. Our data generally agree very well with the backbone dynamics derived from NMR relaxation measurements,[2,21] as well as structural fluctuations from targeted molecular dynamics simulations.[22] In both cases, large fluctuation changes were observed in the β3/α3 loop, α4-helix and the α4/β5 loop. The coarse-grained fluctuations also predict flexibility of certain residues that was undetectable from NMR measurements,[3,21] e.g. the residues around the active site and Ser85-Asp86-Leu87 (part of β4/α4 loop). From the coarse-grained dynamics, no difference in fluctuations or correlations was detected between the BeF_3^--actived NtrC (e.g. PDB:1krw) and the phosphorylated NtrC (e.g., PDB:1dc8). This is consistent with the NMR data[21] showing that they are basically the same structural state. Consequently, we take the phosphorylated form PDB:1dc8 as the active state in this work.

Figure 3 shows the color-coded difference correlation matrix of motions in the receiver domain of NtrC upon phosphorylation. Phosphorylation induces major fluctuation changes and correlations before and after the active site Asp54 and throughout Thr82-Tyr101 (α4 and β5 region) of NtrC [Fig. 3(A)]. The changes of the correlation coefficient C_{ij} matrix are remarkable as shown in [Fig. 3(B)], and the largest anticorrelated motion (red spots in [Fig. 3(B)], where residues move with comparable amplitude but opposite direction) of α3-helix and α4-helix occurs with a correlated motion of the segments upstream and downstream of the active site region (blue spots in [Fig. 3(B)], where residues move with comparable amplitude and same direction). This implies that the concerted changes between Thr82 and Tyr101 (the so-called "Y-T" coupling,[2,21]) are actually a direct dynamic effect of correlated motions. Apart from the fact that helices α3 and α4 move the most in the structural rearrangement of NtrC (see Fig. 1), a positive correlation indicates that the concerted motions of helices α3 and α4 are mutually coupled to each other. On the other hand, phosphorylation activation does not produce significant dynamical changes of the active site Asp54, and the correlated motions of loops (i.e. β2 vs. α2 and β3 vs. α3) may provide a dynamic buffering zone for the binding groove around Asp54. This has not been previously reported in either the NMR experiments[2,3,21] or MD simulations.[22] For the remaining segments of NtrC, only minor correlated movements of other secondary structures are detected from this correlation map.

The presence of both positive and negative correlation modes that best couple to the conformational changes of NtrC implies that, as the phosphorylation proceeds, the whole α3- and α4-helices move collectively in an antiparallel fashion, as shown in the correlations contour maps. The correlated binding groove (i.e. β2 vs. α2 and two ends of loop β3/α3) around active site Asp54 moves in parallel to helices α3 and α4. This clearly depicts how different regions of NtrC facilitate the conformational changes upon phosphorylation via correlated dynamical pathways.

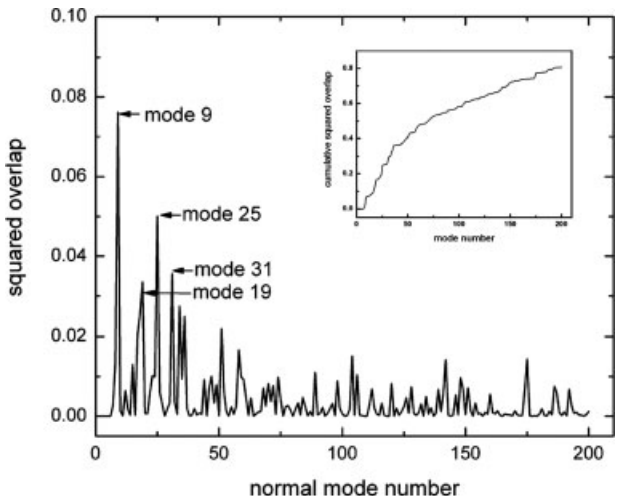

Figure 4

Overlap for normal modes of the receiver domain of NtrC as calculated from the inactive state. Four modes (modes 9, 19, 25, and 31) that most likely account for dynamical transitions between the active and inactive state are marked. Inset is the cumulative squared overlap from mode 7 up to mode 200.

Normal mode analysis of conformational transitions of NtrC

Superimposition of the inactive and active states of NtrC (Figure 1 and the supplemental Figure S1) gave a pairwise RMSD of 2.7 Å between the two conformations. The normal modes of up to 200 modes were determined for both the inactive and the active states and the overlap coefficients of each normal mode were also derived. Among these normal modes, Figure 4 shows that normal modes 9, 19, 25, and 31 (as computed from the inactive state) are the most coupled modes and have the highest overlaps with the conformational transition from the inactive to the active state. The elastic network energy appears larger than expected (relevant to the dominant normal mode seven). Of course, no single normal mode can represent the observed conformational changes, an accurate representation of which requires the combination of a few key normal modes, e.g. mode 9, 19, 25, and 31.[24,33]

Considering possible pathways of conformational transition in NtrC, normal mode 9 and mode 25 (see Fig. 4) have the largest mode-overlap coefficients and are the two most probable schemes for the inactive ↔ active transition: their conformational changes cover about 30% of the overall structural difference between the inactive and active structures. For a NMA interpretation of these dynamical changes, Figure 5 illustrates the net force vectors on residues in mode 9 and mode 25 and corresponding RMSD to make the inactive ↔ active transition. The pathways indicate that the active site region does not move very much, rather the "3445-face" (the α3-β4-α4-β5 region) contributes mainly to the conformational transition.

Intrinsic flexibility and stability of NtrC

Although activation of NtrC results in structural re-arrangement of the α3- and α4-helices and the loops surrounding the active site (see Fig. 1), the dynamics nature of this re-arrangement is unclear. From normal mode overlaps, we have identified some possible pathways for this re-arrangement (see above), but these pathways may be not exclusive. To investigate dynamically important residues or segments of NtrC possibly responsible for structural transitions, ensembles of conformational changes were sampled and dynamic properties for every residue were determined and interpreted by residue flexibility/stability. Residue flexibility/stability of NtrC, evaluated via the COREX/BEST algorithm,[28] is presented in Figure 6 as the color-coded diagrams. The red color indicates the largest flexibility and blue shows the least. Clearly the active and inactive states inherit great difference in flexibility/stability properties from their intrinsic network. Flexibility changes, upon activation, in helices α3 and α4 and in loops leading to/from the active site are clearly visible.

Figure 7 interprets Figure 6 into numerical values. It compares the residue stability constants, K_f,[28] of the receiver domain of NtrC when NtrC is inactive (the dashed lines) or activated by phosphorylation (the solid lines). Phosphorylation introduces significant stability changes in the α3-β4-α4-β5 region (i.e. the "3445" face) and regions leading into the active site (i.e., α2, the α2/β3 loop). The most pronounced stability changes are the switching of stability constants in the α3/β4 and α4/β5 loops, which become much more stable upon the phosphorylation at Asp54. The two conserved residues of Thr82 and Tyr101 appear to play critical roles in these stability changes. From the changes of stability constants, conformational change of α4-helix (as well as the twist of the "3445" face) favors the active state of NtrC over the inactive state. This is consistent with NMR observations[2,3] that the β4/α4 loop and part of β5 have persistent motions (i.e. greater instability) in both inactive and active states. While the stability changes indicate that the α3/β4 and α4/β5 loops slow down after phosphorylation, they may be responsible for the relaxation dynamics at micro- and millisecond timescales which disappeared after activation.[3]

DISCUSSION

Dynamical correlations induce the allosteric cooperativity

In the case of NtrC, good agreement of coarse-grained dynamics with NMR relaxation dynamics[2,3] and MD

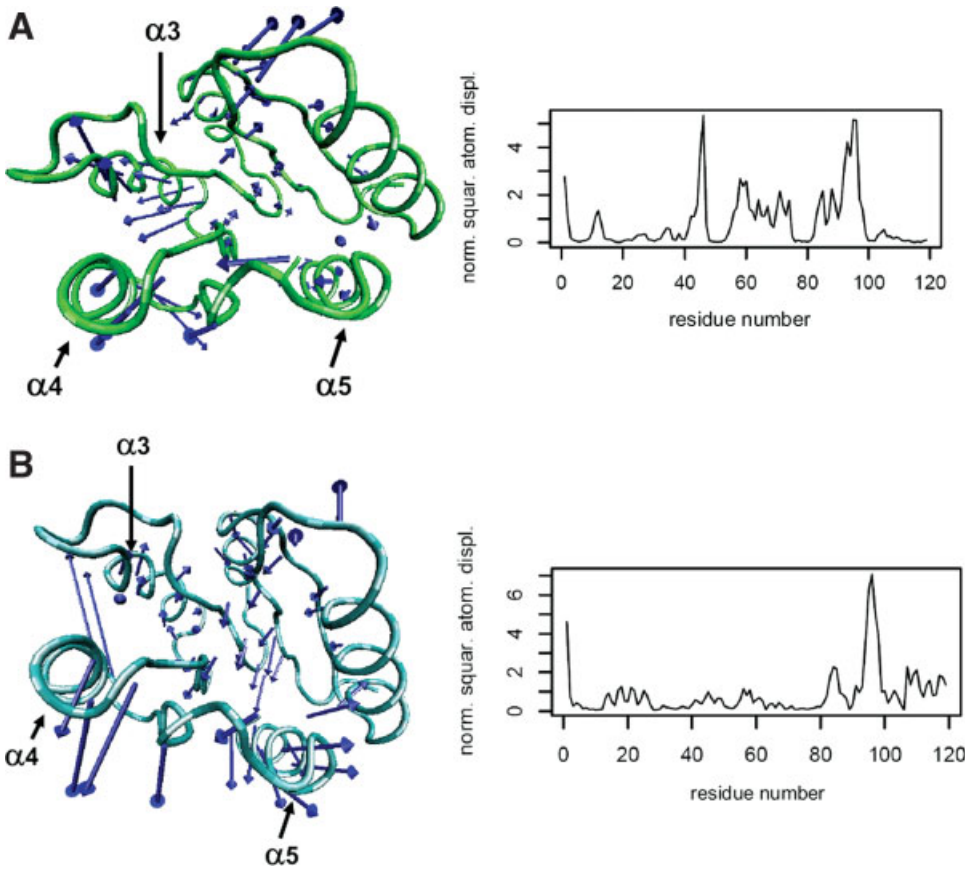

Figure 5

*Two most likely dynamical pathways from the inactive state (PDB:1dc7) to the active state (PDB:1dc8): (**A**) mode 9, colored in green, and (**B**) mode 25, colored cyan. Force vectors (blue arrows) as derived from NMA indicate the direction and amplitude of forces acting on residues. Inserts right of the ribbon structures are the RMSD of individual residues in the corresponding normal mode of motion. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]*

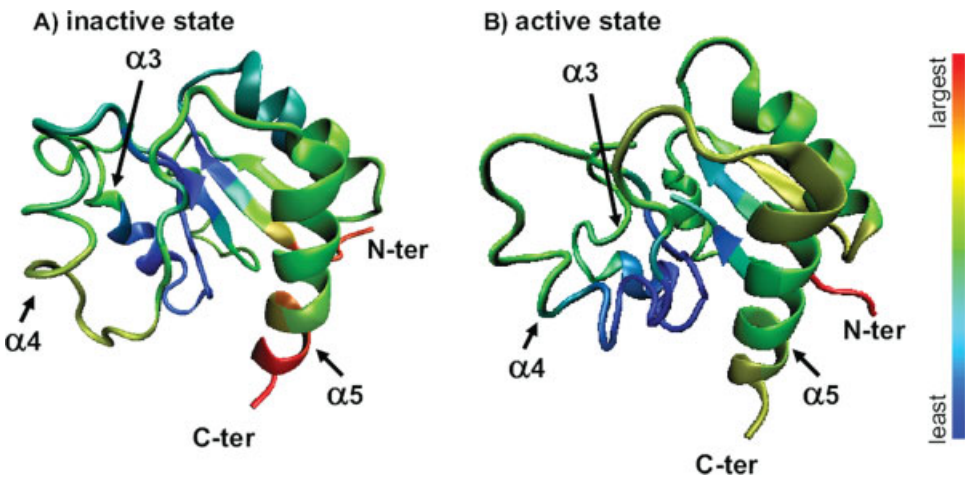

Figure 6

*Flexibility/stability changes (**A**) before and (**B**) after phosphorylation. The red color indicates the largest flexibility and blue the least, with the same absolute scales of the color bar for both cases.*

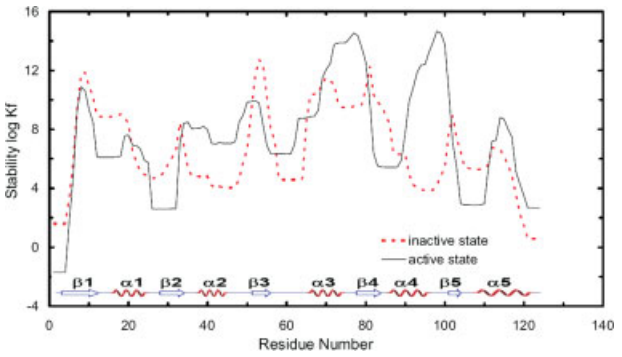

Figure 7

Dynamic stability changes as indicated by the stability constant of K_f[28] Activation induces the stability changes in the segment before the active site (α2-β3) and loop regions of α3/β4 and α4/β5. The phosphorylation makes the loop regions of α3/β4 and α4/β5 more stable. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com]

simulations[22] suggests a prominent role of the topology of NtrC structures in regulating the intrinsic dynamics and associated functions. A key question remaining for NtrC is how the conformational changes at the phosphorylation site are propagated, in a dynamical and allosteric way, to the remote site for promotion of transcription. This question has fundamental significance when we attempt to define dynamic criteria to elucidate the allosteric cooperativity of proteins.

For the receiver domain of NtrC, NMR relaxation dynamics[3] showed that the inactive state and active state co-exist in the unphosphorylated NtrC, but the active state prevails after phosphorylation. As seen in Figure 1, activation and subsequent re-arrangements of the receiver domain allows the reorganization of the "3445" face of NtrC which tilts to the left away from the active site. This suggests that the active site and the "3445" face may interact with each other. Further we have a dynamic picture that a mechanical "clutch"-like mechanism is generated at α4-helix when the α3 and α4 helices are pushed away from the active site (due to activation). This structural re-arrangement is achieved with strongly correlated and anticorrelated dynamics over large distances (about 10.4–21.7 Å) and between distant tertiary motifs (as indicated in the distance-difference matrix in the supplemental Figure S1). Phosphorylation induces both correlated motions upstream and downstream of the loops of the active site and anticorrelated motions between the α3 and α4 helices (see Fig. 3). Therefore, different dynamical correlations across NtrC induce an allosteric cooperativity between the active site region and the α3- and α4-helices of the binding units, and the allostery may be propagated by α3-helix. Of course, these dynamical correlations are driven over large distances and controlled by entropy equilibrium where activation stabilizes the conformation in favor of the active state (see Fig. 7). Dynamically, this allosteric coupling plays a key step for the receiver domain to switch functions from phosphorylation to the downstream transcription factor.[4]

A dynamic criterion to determine the allostery

Apart from NtrC, dynamic coupling, correlated motion and allosteric behavior appears to be conserved in the long range communication and conformational transitions of globular proteins.[6,34,35] For example, a simple mutation can produce marked effects at distal sites via undefined pathways for a conventionally nonallosteric protein.[34] CheY, a signaling protein like NtrC, shows an allosteric mechanism[35] reconciling the "induced-fit" scheme[7,36] with the "population-shift" theory[5,37] for allostery.

Given that pre-existing pathways and allosteric behavior seems to be intrinsic for globular proteins,[34,35,37] collective motion is a possible general cause of allostery. Dynamics certainly plays an essential role in allosteric regulations.[5,35] Our coarse-grained dynamics analysis of NtrC suggests a possible dynamic criterion to determine the allostery in general. Given two distinctive conformational states, dynamical fluctuations and correlations, either amongst the distant functional motifs (e.g. in NtrC) or different subunits,[38] can be determined to account for the conformational transitions between them. If these correlations result in both correlated and anticorrelated modes of motions, allosteric cooperativity will occur simultaneously. Of this mechanism, allostery prevails upon dynamics.[35,39]

Limitation of using coarse-grained dynamics on investigation of fluctuation and allostery

Structural transitions in proteins typically occur on micro- to millisecond timescales. These large-scale events of proteins would generally be inaccessible to conventional dynamics experiments[9,10] and explicit simulations. According to our analysis of NtrC, the residue fluctuations, domain collectivity, internal correlations and even dynamical pathways associated with structural transitions are well captured by coarse-grained dynamic analysis using the elastic network model. The coarse-grained NMA of NtrC predicted dynamics and allostery that generally compares well with NMR relaxation dynamics with no computing limitation in timescale and/or system size.

NMA handles large-scale motions of proteins with a linear elastic approximation, but for some biological functions (such as ligand binding) nonlinearity is also essential. These nonlinear events can only be determined precisely by coarse-grained dynamical correlations with two distinct states available,[18,40] e.g. the active and the inactive state of NtrC or more generally the open and closed states of a protein. In practice, allostery that nor-

mally involves large-scale motion and possible partial unfolding can be interpreted by structural fluctuations [by Eq. (2)] and dynamical correlation differences [by Eq. (1)] via normal modes of the starting and reference structures. Allosteric cooperativity can be verified by the existence of both correlated and anticorrelated dynamics as the system switches from one conformation state to the other.

It should be noted that, whenever protein function is related to major allosteric transitions of a given structure, NMA based on only one state of the structure is not appropriate to describe the full spectrum of dynamics. In such circumstances, all-atom NMR measurements and MD simulations remain a better choice, and normal modes may be the precursors for mimicking larger anharmonic motions.[18,39] In such cases, the allosteric conformational changes may be described by algorithms that can unite a free energy landscape description with the elastic-model-based normal modes, i.e. to find the free energy barriers between different energetic minima of allosteric structural changes.[40] Beyond these theoretical analyses, defining functionally important dynamics in proteins requires combined investigations of static structures, molecular dynamics (experiments and multi-scale simulations) and biological functions.

CONCLUSIONS

In summary, we have investigated dynamic fluctuations and correlations in the receiver domain of NtrC using a coarse-grained dynamics algorithm. The diagonalized correlations between the inactive and active states indicate that, upon activation, most fluctuations in the NtrC domain occur around the active site of Asp54 and in the region Thr82-Tyr101. Phosphorylation induces flexibility changes in the active site with the "3445" face being pushed away from the active site. The changes make the $\alpha 3/\beta 4$ and $\alpha 4/\beta 5$ loops more stable, and may introduce the axial rotation of $\alpha 4$-helix. This leads to the conformation of the receiver domain toward the entropically-favored active state which can promote binding capability for transcription at the allosteric site. These structural and dynamical analyses suggest that allostery and cooperativity of proteins are intrinsically related to dynamic fluctuations and correlations. This work also demonstrates that coarse-grained dynamics represents a robust and powerful tool for probing protein dynamics and its roles in biological functions.

ACKNOWLEDGMENTS

Generous allocations of computing resources from the Australian and Victorian Partnerships for Advanced Computing are acknowledged.

REFERENCES

1. Wyman C, Rombel I, North AK, Bustamante C, Kustu S. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. Science 1997;275:1658–1661.
2. Kern D, Volkman B, Luginbuhl P, Nohaile M, Kustu S, Wemmer DE. Structure of a transiently phosphorylated switch in bacterial signal transduction. Nature 1999;402:894–898.
3. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. Science 2001;291: 2429–2433.
4. de Carlo S, Chen B, Hoover TR, Kondrashkina E, Nogales E, Nixon BT. The structural basis for regulated assembly and function of the transcriptional activator NtrC. Genes Dev 2006;20:1485–1495.
5. Kern D, Zuiderweg ER. The role of dynamics in allosteric regulation. Curr Opin Struct Biol 2003;13:748–757.
6. Stock AM, Robinson VL, Goudreau PN. Two component signal transduction. Annu Rev Biochem 2000;69:183–215.
7. Changeux J-P, Edelstein SJ. Allosteric mechanisms of signal transduction. Science 2005;308:1424–1428.
8. Hardy JA, Wells JA. Searching for new allosteric sites in enzymes. Curr Opin Struct Biol 2004;14:706–715.
9. Wand AJ. Dynamic activation of protein function: a view emerging from NMR spectroscopy. Nature Struct Biol 2001;8:926–931.
10. Mittermaier A, Kay LE. New tools provide new insights in NMR studies of protein dynamics. Science 2006;312:224–228.
11. Clementi C, Nymeyer H, Onuchic JN. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. J Mol Biol 2000;298:937–953.
12. Ding F, Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Direct molecular dynamics observation of protein folding transition state ensemble. Biophys J 2002;83:3525–3532.
13. Okazaki K-I, Koga N, Takada S, Onuchic JN, Wolynes PG. Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: structure-based molecular dynamics simulations. Proc Natl Acad Sci USA 2006;103:11844–11849.
14. Sharma S, Ding F, Dokholyan NV. Multiscale modeling of nucleosome dynamics. Biophys J 2007;92:1457–1470.
15. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Folding Des 1997;2:173–181.
16. Delarue M, Dumas P. On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. Proc Natl Acad Sci USA 2004;101:6957–6962.
17. Ma J. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. Structure 2005;13:373–380.
18. Tobi D, Bahar I. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. Proc Natl Acad Sci USA 2005;102:18908–18913.
19. Yao S, Liu MS, Masters SL, Zhang J-G, Babon JJ, Nicola NA, Nicholson SE, Norton RS. Dynamics of the SPRY domain-containing SOCS protein 2: flexibility of key functional loops. Protein Science 2006;15:2761–2772.
20. Liu MS, Todd BD, Sadus RJ. Dynamical and coordinating domain motions in the active subunits of the F1-ATPase molecular motor. Biochimica et Biophysica Acta 2006;1764:1553–1560.
21. Hastings CA, Lee SY, Cho HS, Yan D, Kustu S, Wemmer DE. High-resolution solution structure of the beryllofluoride-activated NtrC receiver domain. Biochemistry 2003;42:9081–9090.
22. Hu X, Wang Y. Molecular dynamics simulations of the N-terminal receiver domain of NtrC reveal intrinsic conformational flexibility in the inactive state. J Biomol Struct Dynam 2006;23:509–517.
23. Volkman BF, Nohaile MJ, Amy NK, Kustu S, Wemmer DE. Three-dimensional solution structure of the N-terminal receiver domain of NtrC. Biochemistry 1995;34:1413–1424.

24. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. Protein Eng 2001;14:1–6.
25. Maragakis P, Karplus M. Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase. J Mol Biol 2005;352:807–822.
26. Zheng W, Brooks B. Identification of dynamical correlations within the myosin motor domain by the normal mode analysis of an elastic network model. J Mol Biol 2005;346:745–759.
27. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D. Intrinsic motions along an enzymatic reaction trajectory. Nature 2007;450:838–844.
28. Hilser VJ, Dowdy D, Oas TG, Freire E. The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. Proc Natl Acad Sci USA 1998;95:9903–9908.
29. Pan H, Lee JC, Hilser VJ. Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble. Proc Natl Acad Sci USA 2000;97:12020–12025.
30. Kalé L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K. NAMD2: greater scalability for parallel molecular dynamics. J Comput Phys 1999; 151:283–312.
31. Maiti R, Van Domselaar GH, Zhang H, Wishart DS. SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res 2004;32:W590–W594.
32. Humphrey W, Dalke A, Schulten K. VMD—visual molecular dynamics. J Mol Graph 1996;14:33–38.
33. van Wynsberghe AW, Cui Q. Interpreting correlated motions using normal mode analysis. Structure 2006;14:1647–1653.
34. Clarkson MW, Gilmore SA, Edgell MH, Lee AL. Dynamic coupling and allosteric behavior in a nonallosteric protein. Biochemistry 2006;45:7693–7699.
35. Formaneck MS, Ma L, Cui Q. Reconciling the "old" and "new" views of protein allostery: a molecular simulation study of chemotaxis Y protein (CheY). Proteins 2006;63:846–867.
36. Koshland DEJ, Hamadani K. Proteomics and models for enzyme cooperativity. J Biol Chem 2002;277:46841–46844.
37. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Protein Struct Funct Genet 2004;57:433–443.
38. Liu MS, Todd BD, Sadus RJ. Cooperativity in the motor activities of the ATP-fueled molecular motors. Biochimica et Biophysica Acta 2005;1752:111–123.
39. Zheng W, Brooks BR, Thirumalai D. Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. Proc Natl Acad Sci USA 2006;103: 7664–7669.
40. Miyashita O, Onuchic JN, Wolynes PG. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. Proc Natl Acad Sci USA 2003;100:12570–12575.