Supplementary Information for:Dynamics of allosteric transitions in GroEL

Changbong Hyeon[1], George H. Lorimer[1,2] and D. Thirumalai[1,2]

[1]*Biophysics Program*

Institute for Physical Science and Technology

[2]*Department of Chemistry and Biochemistry*

University of Maryland,

College Park, MD 20742

(Dated: October 3, 2006)

SUPPLEMENTARY INFORMATION

Rationale for using the Self-organized Polymer (SOP) model for allosteric transitions: Several studies have shown that the gross features of protein folding [1, 2], mechanical unfolding of proteins and RNA [3, 4] and the complicated global motions of large systems can be captured using native structure-based models [5]. More recently, such native structure-based methods [6–10] have also been used to probe transitions between two given end-point structures and to ascertain the nature of robust modes that mediate the allosteric transitions [11]. In order to realistically simulate transitions between distinct allosteric states, it is necessary to use simple coarse-grained models that can be used to capture, at least qualitatively, the inherent dynamics connecting two or more states. Although simple elastic network models have given insights into the nature of low frequency dynamics of a number of systems [5], the inherent linearity of the model limits their scope when dealing with potential non-linear motions that must be involved in the allosteric transitions [6]. We have recently introduced a novel class of versatile structure-based models that incorporates the fundamental polymeric aspects of proteins and RNA. The SOP model, which is easily adopted for use in very large systems, has already proven to be successful in obtaining a number of totally unanticipated results for forced-unfolding and force-quench refolding of RNA and proteins [12, 13].

Building on the successful application to single molecule force spectroscopy of biomolecules we used a variant of the SOP model to probe the complex allosteric transitions in a prototypical biological nanomachine, namely, the *E. Coli.* chaperonin GroEL. Structures of GroEL (T, R, and R''), that are populated in the reaction cycle, have been determined (see Fig. S1 for a side view). Given the two allosteric states (say T and R of GroEL) we induce transitions between the two states by switching the energy function representing one structure to another. A few methods for achieving the switch smoothly have been recently proposed [7–10]. We have advocated a Langevin dynamics based method in which the switch is accomplished using Eq. 2 of the main text (see Fig. S2 for a pictorial view). In writing the equations of motions given by Eq. 2 in the main text, we assume that initially the ensemble of conformations are in the T state so that the conformations obey the Boltzmann distribution i.e, $P(\{\vec{r}_i\}) \simeq e^{-\beta H(\{\vec{r}_i\}|T)}$ where $H(\{\vec{r}_i\}|T)$ is the SOP energy for the T-state GroEL (see Eq. 2 in the main text). Upon switching the

Hamiltonian (see below) over a time interval $N_t h^*$ the dynamics follows the equations of motion given in step (iii) of Eq. 2 in the main text. As long as fluctuation dissipation theorem is satisfied, so that the standard potential conditions are met [14], it can be rigorously shown that at long times the biomolecule will reach the R state with the conformations given by the equilibrium Boltzmann distribution governed by the Hamiltonian $H(\{\vec{r}_i\}|R)$. From this perspective the procedure we have used is rigorous.

The assumption that switch occurs over a predetermined duration requires a few comments. The allosteric transitions occur as a result of ligand-binding or interactions with other biomolecules. As a result of binding, local strain is induced at the interaction site or sites. As long as the rate of strain propagation is larger than the rate of conformational change of the molecule then the switch over a reasonable time is justified. Extremely rapid switching, which is tantamount to very large local loading rates, is unphysical as is adiabatic change in the energy function. Given these extreme situations we choose a value of N_t that falls in between the extreme conditions. In our procedure N_t can be adjusted to mimic the potential rate of strain propagation, which induces the allosteric transitions. The efficacy of the procedure has been demonstrated by successful applications to describe the multiple allosteric transitions in GroEL.

Implementation of the Hamiltonian switch: Here we give details of the algorithm for executing the second step in the equations of motion (See subsection *Additional details* in the Methods section in the main text). During the transition interval we define $r_{ij}^o(T \rightarrow R)$ using linear combination of $r_{ij}^o(T)$ and $r_{ij}^o(R)$ where $r_{ij}^o(X)$ is the distance between the residues i and j in the structure X with $X = T$ or R. The switch between $r_{ij}^o(T)$ and $r_{ij}^o(R)$ is carried out slowly every 100 time steps. In the initial stage of the $T \rightarrow R$ transition (0 to 100 time steps) we let $r_{ij}^o(T \rightarrow R) = r_{ij}^o(T)$. Subsequently, we set $r_{ij}^o(T \rightarrow R) = (1 - 0.01k)r_{ij}^o(T) + 0.01k r_{ij}^o(R)$ where $k = 1, 2, 3....100$ is changed every 100 time steps. Thus, the switch in $r_{ij}^o(T \rightarrow R)$ occurs over 10,000 time steps. The loading condition can be varied by changing the number of time steps used to achieve the switch in the distances between the native contacts. For convenience we used a linear combination of $r_{ij}^o(T)$ and $r_{ij}^o(R)$ during the switch process. More generally, one can use non linear combinations, i.e., $r_{ij}^o(T \rightarrow R) = g(t)r_{ij}^o(T) + (1 - g(t))r_{ij}^o(R)$ where $g(t)$

is an arbitrary function (exponential for example).

At the end of each interval, the new value of $r_{ij}^o(T \rightarrow R)$ is substituted into the SOP Hamiltonian to compute the forces needed to solve the equations of motion in step (ii) (see Eq. 2 of the main text). In the present application, the procedure for using $r_{ij}^o(T \rightarrow R)$ lasts only $< 50\ ns$. As a result, the dynamics of distant pair is not affected. Only the equilibrium distances of native pairs, that are already in contact in the T state but lead to instability in the intergration of equations of motion due to rapid switching from T to R, are corrected to the equilibrium distances at R state (see Fig. S2).

Time scales and their relevance: The characteristic time scale of the Brownian dynamics in the overdamped limit is $\tau_H = \frac{\zeta \epsilon_h h}{k_B T} \tau_L$ where we used the friction coefficient $\zeta = 50\tau_L^{-1}$, $\epsilon_h = 2.0 kcal/mol$ and $\tau_L = (\frac{ma^2}{\epsilon_h})^{1/2} \sim 3\ ps$ for proteins. The simulations are performed at $T = 300\ K$ ($k_B T = 0.6 kcal/mol$). We chose the integration time step $h = 0.1\tau_L$ for (i) and (iii) while $h^* = (0.001 - 0.01)\tau_L$ for (ii). Thus, 10^6 integration time steps with $h = 0.1\tau_L$ in our Brownian dynamics simulations correspond to 50 μs. Because we have used a minimal model for GroEL the time scales quoted in the main text should be viewed as lower limit for the various processes. The actual time scales are expected to be much longer. The relatively time scales for different aspects of the allosteric transitions are likely to be correct. For example, we find that the tilt of K and L helices occur four times more slowly than the F and M helices (see Fig. 3 in the main text). Our prediction of factor of four is, in all likelihood, an accurate estimate. During simulations we collected the structures every 0.5 μs to analyze the allosteric transition dynamics.

Analysis of dynamics: To perform a quantitative analysis on the salt bridge or contact pair dynamics we averaged over the time traces of all the trajectories. For the contact dynamics of two-subunit GroEL we generated $N = 50$ trajectories in total and computed dynamic changes in specific residue pairs using $\langle d(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} d_i(t)$. In general $\langle d(t) \rangle$ is fit using $\langle d(t) \rangle = \langle d(t^*) \rangle + \Delta(f e^{-(t-t^*)/\tau_1} + (1-f) e^{-(t-t^*)/\tau_2})$, where $t^* = 50 \mu s$, Δ is the average decrease in the contact distance, and τ_1, τ_2 are the relaxation times for the pathways partitioned into f and $1 - f$.

[1] Onuchic, J. N. & Wolynes, P. G. (2004) *Curr. Opin. Struct. Biol.* **14**, 70–75.

[2] Thirumalai, D. & Hyeon, C. (2005) *Biochemistry* **44**(13), 4957–4970.

[3] Klimov, D. K. & Thirumalai, D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 7254–7259.

[4] Hyeon, C. & Thirumalai, D. (2005) *Proc. Natl. Acad. Sci.* **102**, 6789–6794.

[5] Bahar, I. & Rader, A. J. (2005) *Curr. Opin. Struct. Biol.* **15**, 586–592.

[6] Miyashit, O., Onuchic, J. N., & Wolynes, P. G. (2003) *Proc. Natl. Acad. Sci.* **100**, 12570–12575.

[7] Maragakis, P. & Karplus, M. (2005) *J. Mol. Biol.* **352**, 807–822.

[8] R. B. Best, Y. G. Chen & Hummer, G. (2005) *Structure* **13**, 1755–1763.

[9] Koga, N. & Takada, S. (2006) *Proc. Natl. Acad. Sci.* **103**, 5367–5372.

[10] Okazaki, K., Koga, N., Takada, S., Onuchic, J. N., & Wolynes, P. G. (2006) *Proc. Natl. Acad. Sci.* **103**, 11844–11849.

[11] Zheng, W., Brooks, B. R., Doniach, S., & Thirumalai, D. (2005) *Structure* **13**, 565–577.

[12] Hyeon, C., Dima, R. I., & Thirumalai, D. (2006) *Structure (in press).*

[13] Hyeon, C. & Thirumalai, D. (2006) *Biophys. J. (in press).*

[14] Gardiner, C. W. (1985) *Handbook of Stochastic Methods* (Springer-Verlag, Berlin Heibelberg), Second edition.

Figure Captions

Figure S1: The columns from left to right show the side view of GroEL structure in the T, R, and R'' states.

Figure S2: Illustration of the procedure to switch SOP Hamiltonian from T to R state. To avoid the computational instability caused by instantaneous switch of equilibrium distance, the equilibrium distance from T to R state ($r_{ij}^o(T) \rightarrow r_{ij}^o(R)$) is gradually switched using a series of transient potentials defined with $r_{ij}^o(T \rightarrow R)$ (see *Additional details* in Methods section).

Figure S3: TSEs represented in terms of distribution $P(\Delta^{\ddagger})$ where $\Delta^{\ddagger} = 1/2 \times |(RMSD/T)(t_{TS}) + (RMSD/R)(t_{TS})|$ for $T \rightarrow R$ transition. $\Delta^{\ddagger}$ for $R \rightarrow R''$ transition is similarly defined.

Figure S4: The dynamical changes in the distances between a number of residues in a single trajectory during $T \rightarrow R$ and $R \rightarrow R''$ are plotted on **A** and **B**, respectively.

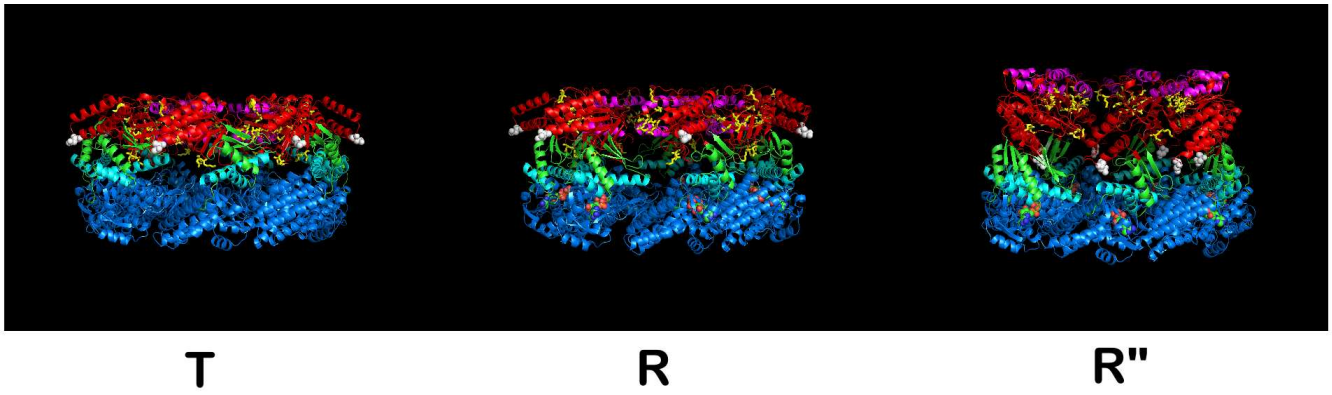

FIG. S1:

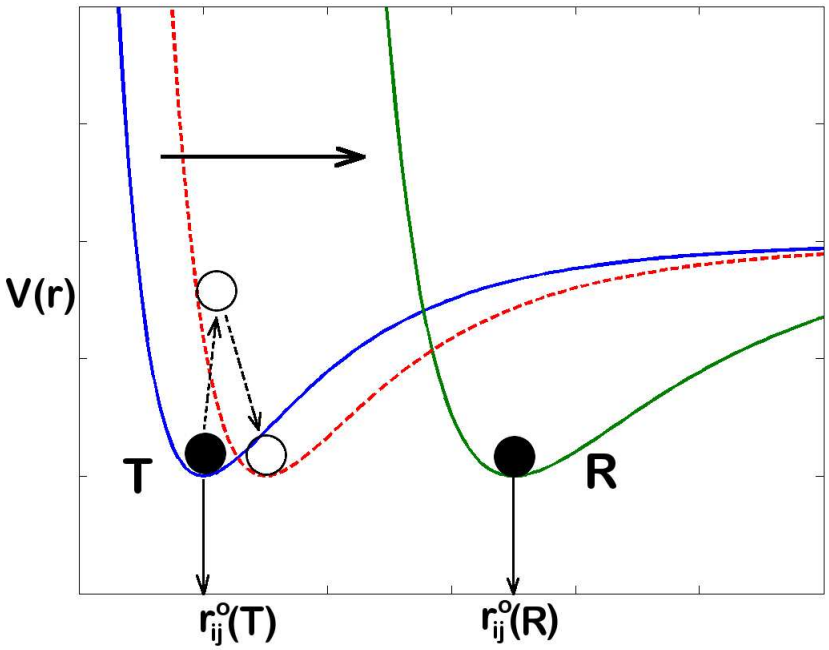

FIG. S2:

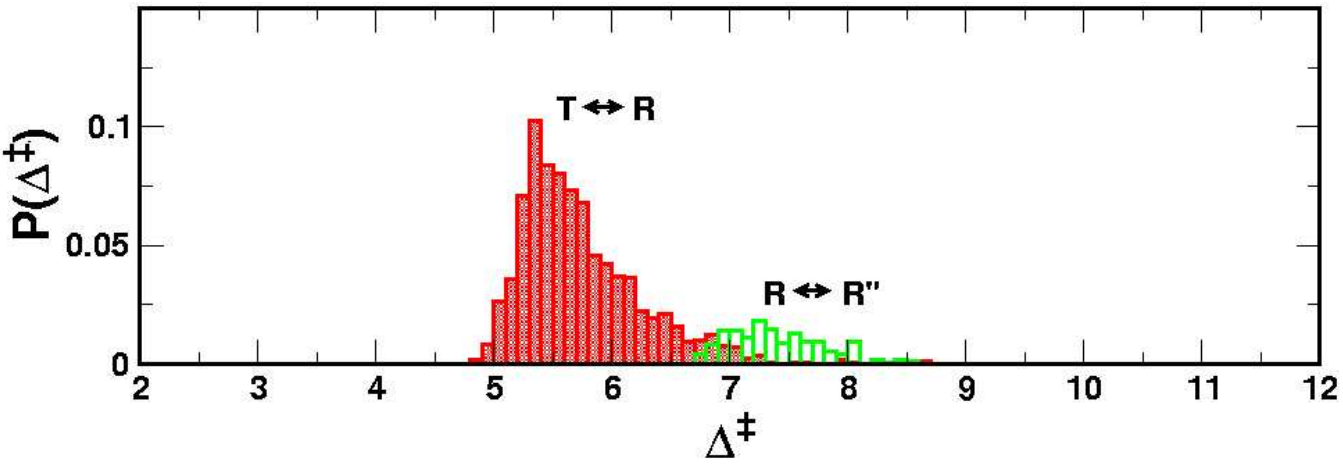

FIG. S3:

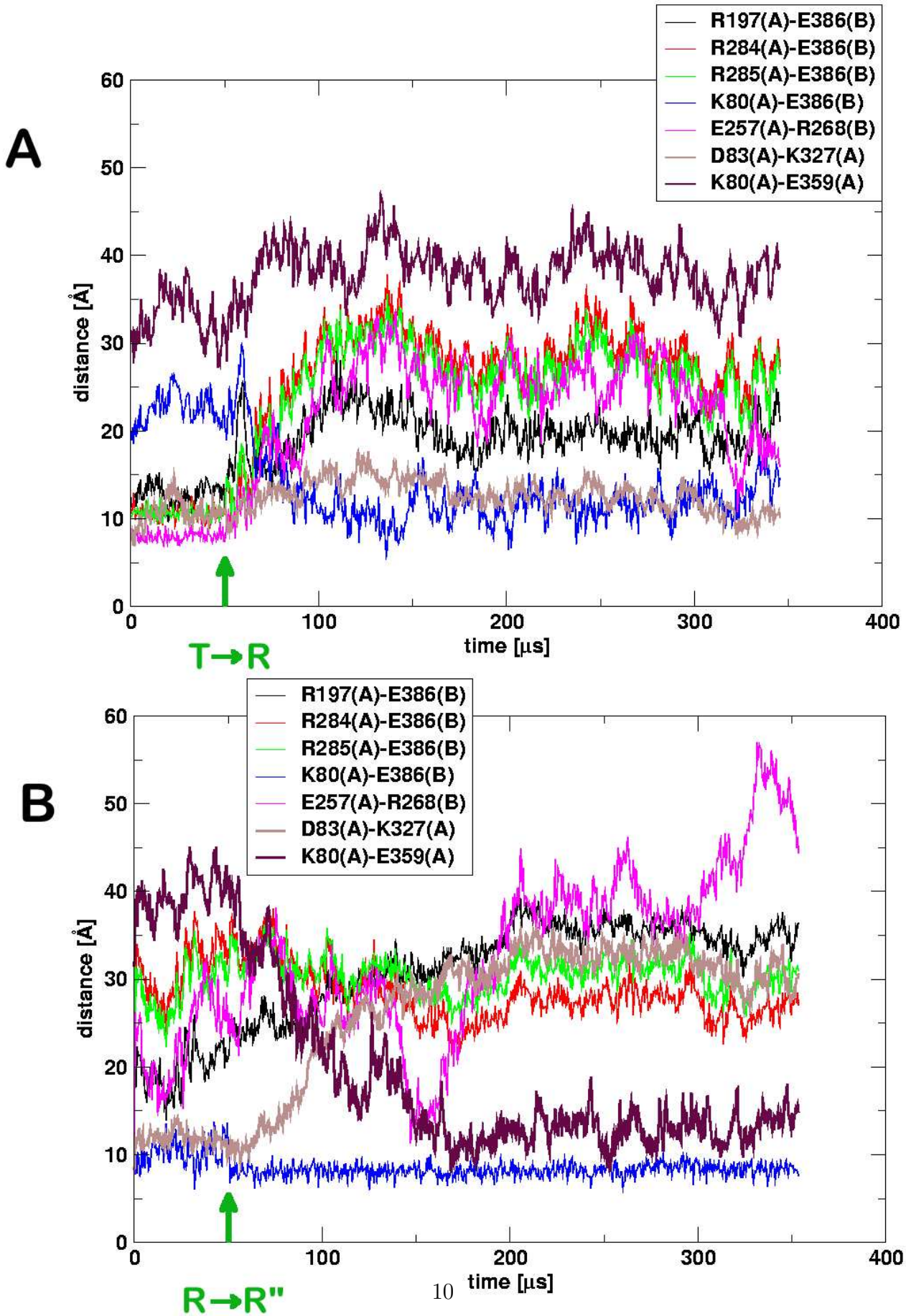
A
R197(A)-E386(B)
R284(A)-E386(B)
R285(A)-E386(B)
K80(A)-E386(B)
E257(A)-R268(B)
D83(A)-K327(A)
K80(A)-E359(A)
distance [Å]
time [μs]
T→R
B
R197(A)-E386(B)
R284(A)-E386(B)
R285(A)-E386(B)
K80(A)-E386(B)
E257(A)-R268(B)
D83(A)-K327(A)
K80(A)-E359(A)
distance [Å]
time [μs]
R→R"

FIG. S4: