Coarse-graining provides insights on the essential nature of heterogeneity in actin filaments

Supplemental Information

Jun Fan, Marissa G. Saunders, and Gregory A. Voth*

Department of Chemistry, Institute for Biophysical Dynamics, James Franck Institute, and Computation Institute, University of Chicago, 5735 S Ellis Ave., Chicago, IL 60637

*Corresponding author: Email: gavoth@uchicago.edu

SUPPLEMENTARY TEXT

Essential Dynamics Coarse-Graining

The ED-CG method is described mathematically as the following: The displacement of atom *i* at time *t* can be determined to be $\Delta r_i(t) = r_i(t) - \bar{r}_i$, where $r_i(t)$ is the location of atom *i* at time *t*, and $\bar{r}_i$ is the equilibrium position of atom *i*. Correlation between atoms is evaluated based on whether the two atoms move together. If atom *i* and atom *j* are correlated then the displacement difference, $\left|\Delta r_i(t) - \Delta r_j(t)\right|^2$, will be small. To correctly capture the slow motions of the system, instead of calculating the displacement difference using $\Delta r_i(t)$, we transform the atomistic coordinates into the essential subspace in which only the lowest frequency principle components are retained (1). In this subspace, the displacements are denoted $\Delta r_i^{ED}(t)$. For a system of *n* atoms, the mapping of these atoms to N_{CG} coarse-grained sites will optimally capture long timescale correlations between atoms when the residual

$$\chi^2 = \frac{1}{3N_{CG}} \sum_{I=1}^{N_{CG}} \frac{1}{n_t} \sum_{t=1}^{n_t} \left(\sum_{i \in I} \sum_{j>i \in I} \left| \Delta r_i^{ED}(t) - \Delta r_j^{ED}(t) \right| \right) \quad (1)$$

is minimized. In the CG representation, the position of a CG site is given by the center of mass (COM) of all the atoms belonging to this CG site. Using this representation, MD simulation data can be transformed into CG coordinates.

The Hetero-Elastic Network Model method

In this method, the interaction between each pair of CG atoms is represented by a harmonic spring. The equilibrium distance is the average distance between each CG pair $\bar{x}_{ij} = \overline{\left|x_i - x_j\right|}$, and the corresponding mean-square distance fluctuation is given by $\Delta x_{ij}^2 = \overline{(x_{ij} - \bar{x}_{ij})^2}$, The spring constant k_{ij} is obtained by matching the mean-square

distance fluctuation of CG sites, calculated from the normal-mode analysis ($\Delta x^2_{ij,NMA}$), with that from the MD simulations ($\Delta x^2_{ij,MD}$) through the following iterative algorithm:

$$\frac{1}{k^{m+1}} = \frac{1}{k^m} - 4\alpha(\Delta x^2_{ij,NMA} - \Delta x^2_{ij,MD}), \quad (2)$$

where *m* indicates the iteration. The normal mode analysis of the CG elastic network was performed with CHARMM version c32b2 (2). No distance cutoff was imposed to limit the interactions, thus for a system with N_{CG} CG sites, HENM would yield $N_{CG}(N_{CG}\text{-}1)/2$ spring constants. Initial values of the spring constants were set at 1 kcal/molÅ^2 and iterations of Eq. (2) continued until $\left|k_{ij}^{m+1} - k_{ij}^{m}\right| \le 10^{-3}$.

Twist angle per subunit

Each subunit was CG into a point by calculating its center of mass (COM). The COM of each strand was fitted into a smooth curve using the cubic spline function of MATLAB. The projection of monomer i onto the other strand i' is defined by the point of the other strand with equal curvilinear distance to monomer i-1 and monomer i+1. A third curve, connected by the center point of i and i', denoted by i", is the twist axis. The angle between vector i-i" and vector (i+1)-(i+1)" projected to the plane with the norm direction along the vector i"-(i+1)" is the twist angle of monomer i+1 reference to monomer i.

Calculation of the free energy of CG models

For each CG system, the free energy could be calculated analytically by

$$G = H - TS, \quad (3)$$

where the enthalpy

$$H = \sum_{i,j=1}^{N_{CG}} \frac{1}{2} k_{ij}(x_{ij} - \bar{x}_{ij})^2 + \sum_{i}^{N_{CG}} \frac{1}{2} m_i v_i^2 + \sum_{i=1}^{3N_{CG}-6} \left(\frac{1}{2} N_A \hbar \nu_i + \frac{N_A \hbar \nu_i e^{-\hbar\nu_i / k_B T}}{1 - e^{-\hbar\nu_i / k_B T}}\right) \quad (4)$$

contains the bonding energy (*V*) of all CG pairs, the kinetic energy, and the vibrational contribution. N_A is Avogadro's number, ν_i is the i^{th} frequency of the normal mode of the system, k_B is Boltzmann constant, *T* is temperature, and the entropy

$$S = \sum_{i=1}^{3N_{CG}-6} [-R\ln(1-e^{-\hbar\nu_i/k_BT}) + \frac{N_A\hbar\nu_i e^{-\hbar\nu_i/k_BT}}{T(1-e^{-\hbar\nu_i/k_BT})}] \quad (5)$$

based on Normal Mode Analysis (3). For harmonic systems, according to equipartition theorem, the bonding energy and the kinetic energy are both $3/2N_{CG}RT$. The vibration frequency ν_i is the square root of the eigenvalue of the mass-weighted Hessian Matrix $\sqrt{M}^{-1}\mathrm{H}\sqrt{M}^{-1}$, where M is mass and $\mathrm{H} \in R^{3n} \times R^{3n}$ is the Hessian Matrix

$$\mathrm{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & & & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix}. \quad (6)$$

The term $h_{ij} \in R^3 \times R^3$ is a super-element in H. The super-elements are given by

$$h_{ij} = \begin{bmatrix} \partial^2V/\partial x_{i_1}\partial x_{j_1} & \partial^2V/\partial x_{i_1}\partial x_{j_2} & \partial^2V/\partial x_{i_1}\partial x_{j_3} \\ \partial^2V/\partial x_{i_2}\partial x_{j_1} & \partial^2V/\partial x_{i_2}\partial x_{j_2} & \partial^2V/\partial x_{i_2}\partial x_{j_3} \\ \partial^2V/\partial x_{i_3}\partial x_{j_1} & \partial^2V/\partial x_{i_3}\partial x_{j_2} & \partial^2V/\partial x_{i_3}\partial x_{j_3} \end{bmatrix}, \quad (7)$$

when $i \neq j$

$$\partial^2V/\partial x_{i_1}\partial x_{j_2} = -k_{ij}\left.\frac{(x_{j_1}-x_{i_1})(x_{j_2}-x_{i_2})}{x_{ij}^2}\right|_{x_{ij}=\bar{x}_{ij}}, \quad (8)$$

and when $i = j$

$$\partial^2V/\partial x_{i_1}\partial x_{i_2} = \sum_{l\neq i} k_{il}\left.\frac{(x_{i_1}-x_{l_1})(x_{i_2}-x_{l_2})}{x_{il}^2}\right|_{x_{il}=\bar{x}_{il}}. \quad (9)$$

SUPPLEMENTAL TABLES AND FIGURES

Table S1. Intra-subunit CG interactions for 4 intuitive sites model (ATP-bound filament)

Intra-subunit pairs	$\bar{x}_{ij}$ (Å)	$\bar{k}_{ij}$ (kcal/molÅ^2)	$\sigma(k_{ij})$ (kcal/molÅ^2)	$\sigma(k_{ij})/\bar{k}_{ij}$
1-2	26.80 (0.56)	2.68	1.23	45.9%
1-3	25.93 (0.31)	8.31	1.89	22.7%
1-4	35.76 (0.66)	4.54	1.31	28.9%
2-3	39.64 (0.56)	3.04	1.53	50.3%
2-4	29.20 (1.26)	2.93	1.47	50.2%
3-4	25.18 (0.22)	9.75	3.05	31.3%

Figure S1 Changes in the morphology of the D-loop (circled) alter the contacts between subunits along a strand in the filament. Each frame shows the ending configuration of a different subunit. Note that the bottom subunit in (a) has a folded D-loop configurations while in all other subunits D-loop are unfolded. This figure shows that distinct configurations are coupled with different contacts.

Figure S2 Diagrammatic representation of the 13 six-site CG configurations of the (a) ATP-bound and (b) ADP-bound filament and the averaged six-site CG model superimposed on the all-atom backbone structure. The top row contains the 1st -5th subunits, the middle row contains subunits 6 -10, and bottom row contains subunits 11 - 13 and the averaged model.

Figure S3 Incorporating heterogeneity into CG models better reproduces (a) the persistence length and (b) torsional stiffness observed in MD simulations for ADP-bound filament.

Figure S4 Incorporating heterogeneity into CG models better represents (a) the twist per subunit and (b) its distribution observed in MD simulations for ADP-bound filament.

Figure S1 Changes in the morphology of the D-loop (circled) alter the contacts between subunits along a strand in the filament. Each frame shows the ending configuration of a different subunit. Note that the bottom subunit in (a) has a folded D-loop configurations while in all other subunits D-loop are unfolded. This figure shows that distinct configurations are coupled with different contacts.

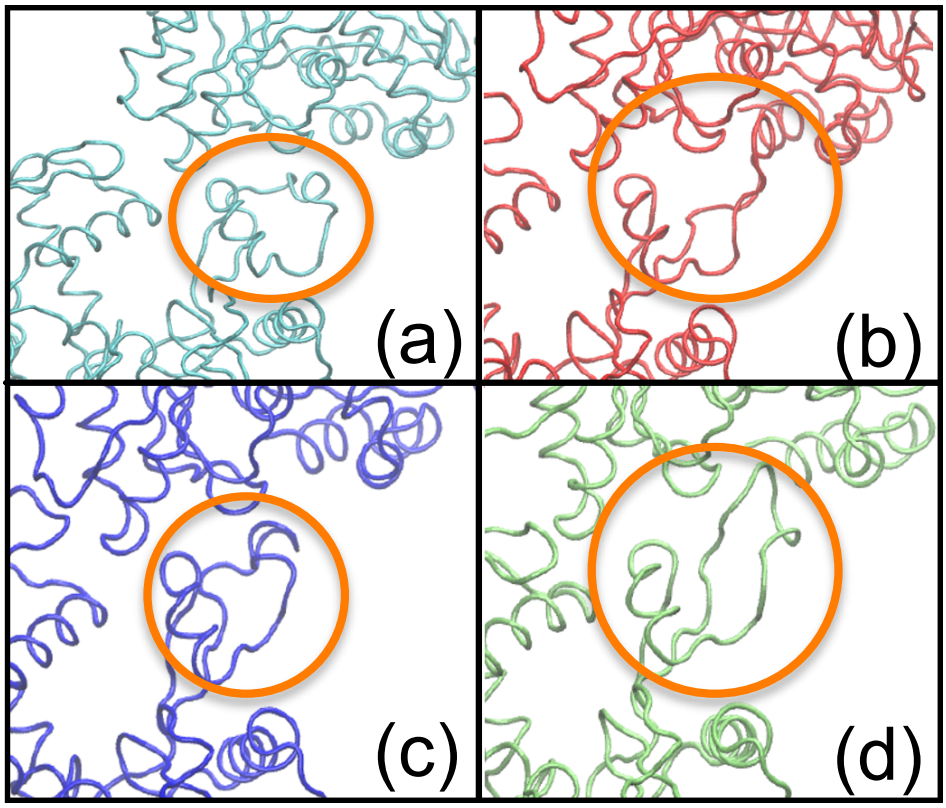
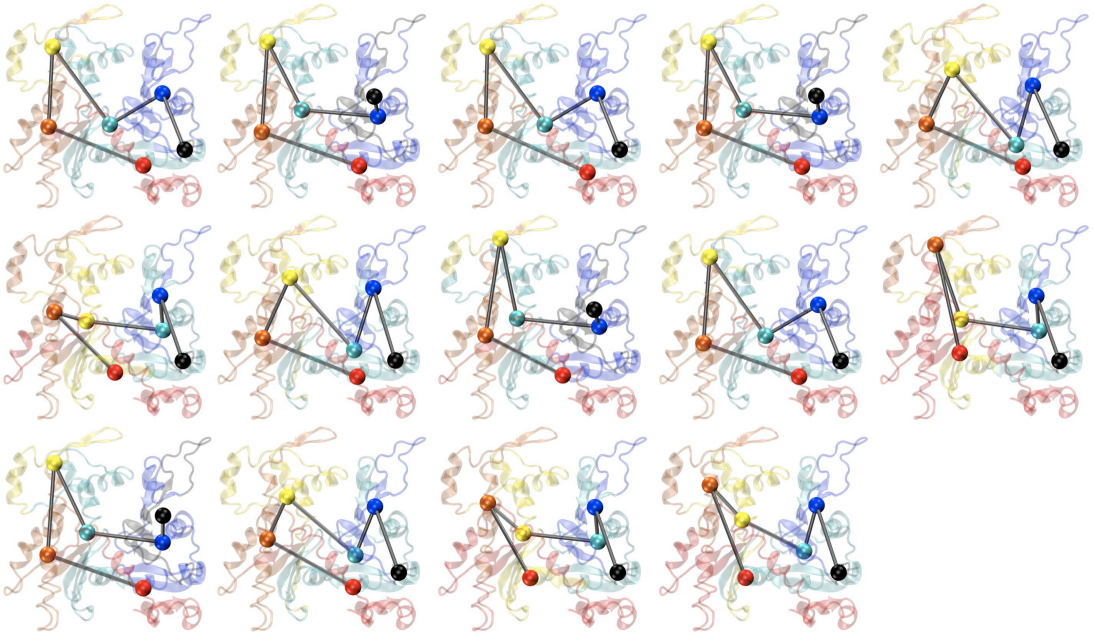


Figure S2 Diagrammatic representation of the 13 six-site CG configurations of the (a) ATP-bound (b) ADP-bound filament and the averaged six-site CG model superimposed on the all-atom backbone structure. The top row contains the 1st -5th subunits, the middle row contains subunits 6 -10, and bottom row contains subunits 11 -13 and the averaged model.

(a)



(b)

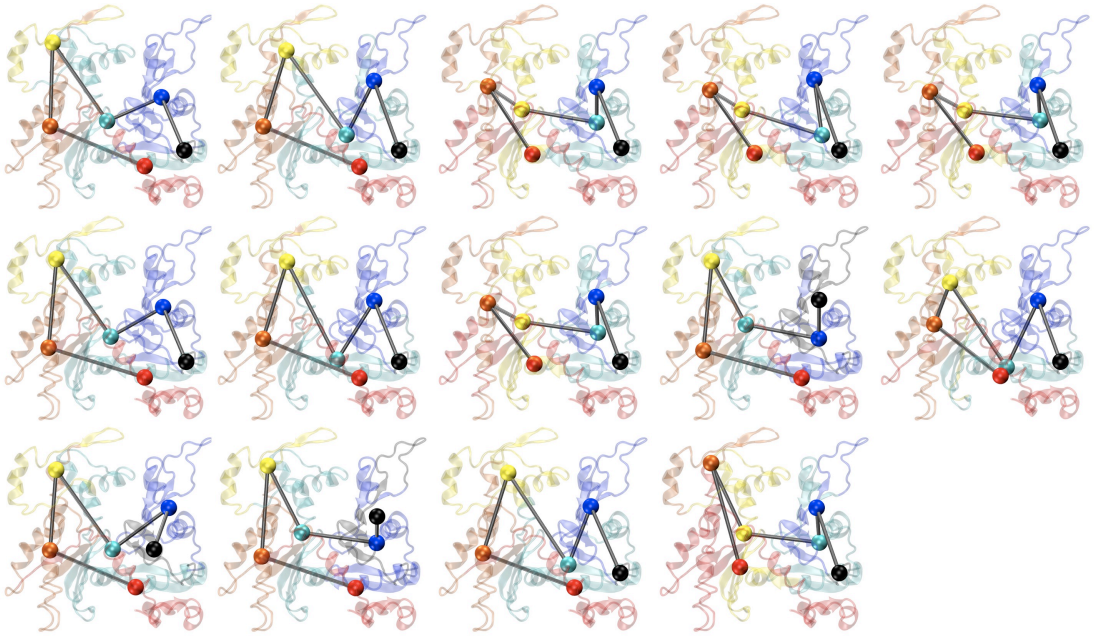


Figure S3 Incorporating heterogeneity into CG models better reproduces (a) the persistence length and (b) torsional stiffness observed in MD simulations for ADP-bound filament.

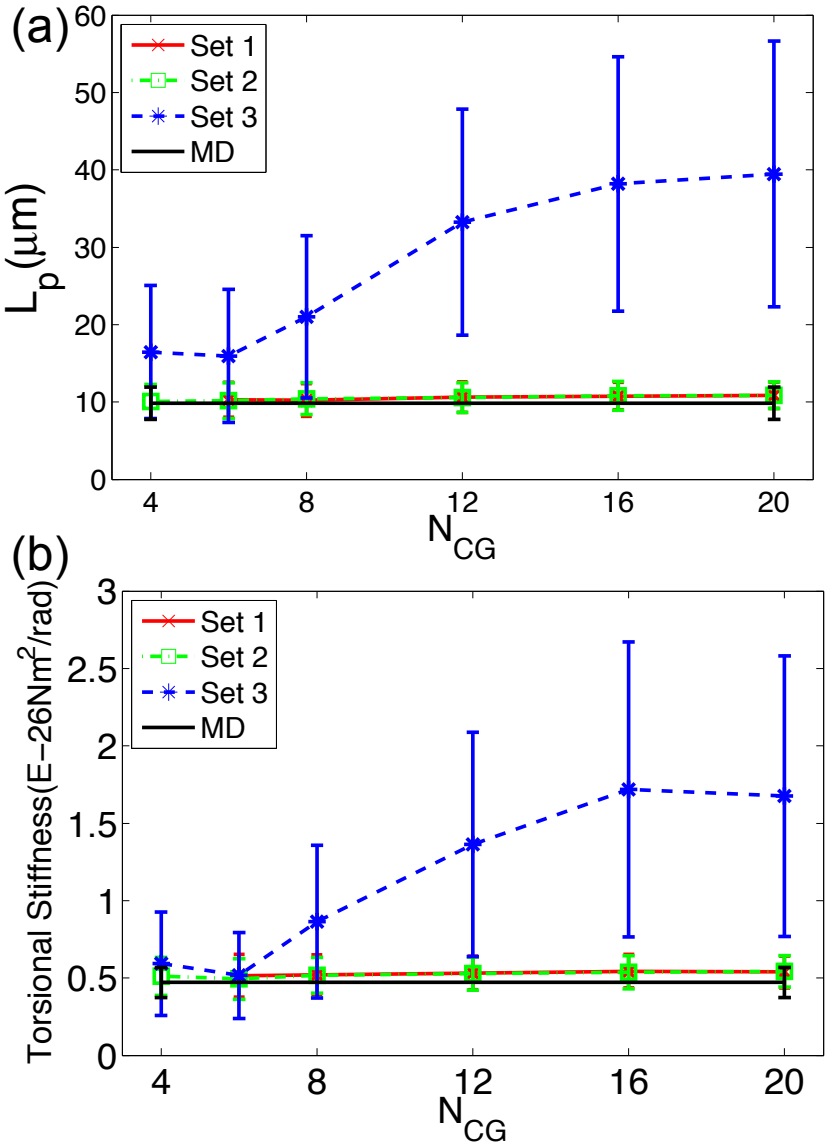
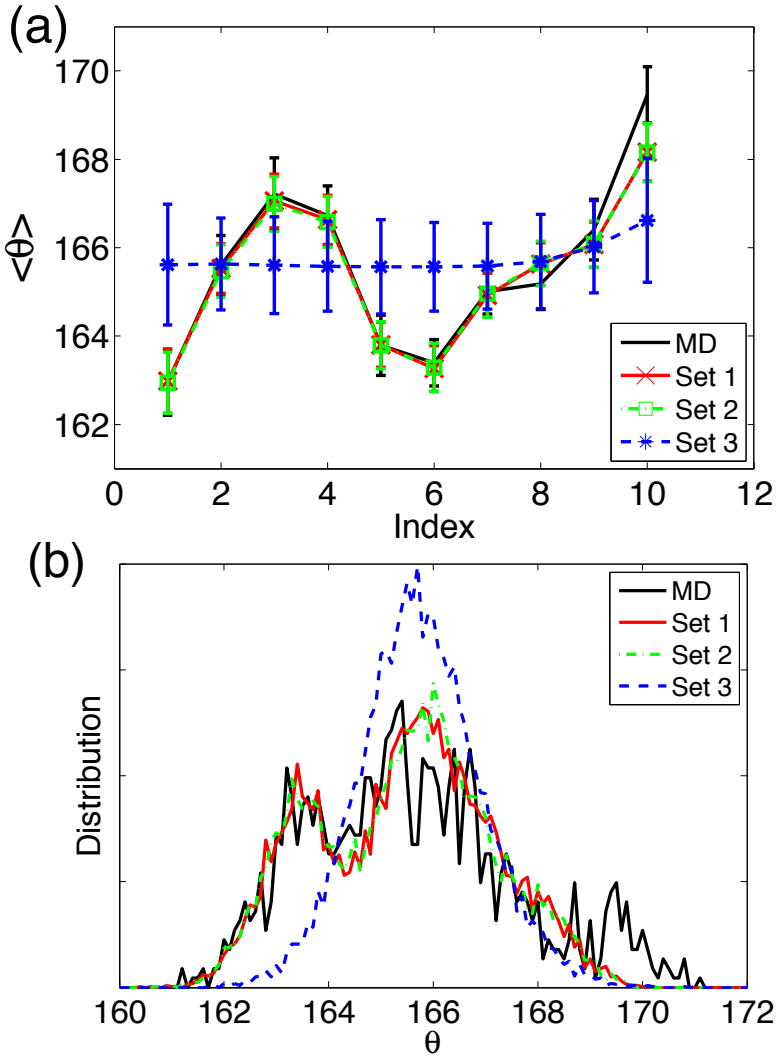


Figure S4 Incorporating heterogeneity into CG models better represents (a) the twist per subunit and (b) its distribution observed in MD simulations for ADP-bound filament.



References:

1. Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential Dynamics of Proteins. Proteins-Structure Function and Genetics 17:412-425.
2. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J Comput Chem 4:187-217.
3. McQuarrie, D. A. 1973. Statistical thermodynamics. Harper & Row, New York.