Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

SUPPLEMENTARY INFORMATION:
Generalization of the Elastic Network model for the study of large conformational changes in biomolecules

Adolfo B. Poma, Mai Suan Li and Panagiotis E. Theodorakis

[1]*Institute of Physics, Polish Academy of Sciences,*

Al. Lotników 32/46, 02-668 Warsaw, Poland

I. HARMONIC APPROXIMATION FOR THE LENNARD–JONES POTENTIAL

The Lennard–Jones (LJ) potential used to model native contacts is given by:

$$V_{\rm LJ}(r_{\rm ij}) = 4\epsilon_{\rm ij}\left[\left(\frac{\sigma_{\rm ij}}{r_{\rm ij}}\right)^{12} - \left(\frac{\sigma_{\rm ij}}{r_{\rm ij}}\right)^{6}\right] \tag{1}$$

where $\epsilon_{\rm ij}$ and $\sigma_{\rm ij}$ are the parameters of the LJ potential.

Now we will use an explicit relationship between the force constant C of a harmonic potential and $\epsilon_{\rm harm}$ known as the harmonic approximation as follows:

Let us define: $x = r_{\rm ij} - r^0_{\rm ij}$, where $r^0_{\rm ij}$ corresponds to the minimum of the potential. In the case of LJ 12–6, this is equal to

$$r^0_{\rm ij} = 2^{1/6}\sigma_{\rm ij}. \tag{2}$$

By using the Taylor expansion of $V_{\rm LJ}$ up to second order in Eq.(1), we obtain,

$$V(x) = V(0) + x\frac{dV(0)}{dx} + \frac{x^2}{2!}\frac{d^2V(0)}{dx^2}, \tag{3}$$

$$= -\epsilon + 0 + \frac{1}{2}Cx^2 \tag{4}$$

Note that $V(x) = V_{\rm LJ}(r_{\rm ij})$ and where $C = \frac{d^2V(0)}{dx^2}$. Here the first derivative vanishes at $x = 0$ (or $r_{\rm ij} = r^0_{\rm ij}$), which corresponds to the minimum of $V(x)$. Now we can work out the second derivative as follows:

$$\frac{d^2V(0)}{dx^2} = \frac{d^2U(r^0_{\rm ij})}{dr^2_{\rm ij}} = 4\epsilon_{\rm ij}\left(12\cdot 13\left(\frac{\sigma^{12}_{\rm ij}}{(r^0_{\rm ij})^{14}}\right) - 6\cdot 7\left(\frac{\sigma^{6}_{\rm ij}}{(r^0_{\rm ij})^{8}}\right)\right) \tag{5}$$

and using Eq.(2) we have that,

$$\frac{d^2V(0)}{dx^2} = \frac{4\epsilon_{\rm ij}}{(2^{1/6}\sigma_{\rm ij})^2}\left(12\cdot 13\left(\frac{1}{2}\right)^2 - 6\cdot 7\left(\frac{1}{2}\right)\right) \tag{6}$$

$$\frac{d^2V(0)}{dx^2} = \frac{36\epsilon_{\rm ij}}{2^{2/3}\sigma^2_{\rm ij}} \tag{7}$$

From this expression we have that, $C = \frac{36\epsilon_{\rm ij}}{2^{2/3}\sigma^2_{\rm ij}}$. Here we define $\epsilon_{\rm ij} = \epsilon_{\rm harm}$. The relatioship between $\epsilon_{\rm harm}$ and C is given by,

$$\epsilon_{\rm harm} = C\sigma^2_{\rm ij}36^{-1}(2^{2/3}) \tag{8}$$

Finally, we employ this last equation to describe the EN contacts with $|i-j| > 3$ in the GEN model.

II. PULLING RESULT FOR 1TIT

In Fig. S1 we show the F-d profile obtained for the GEN model and other 'breakable' EN models for titin.

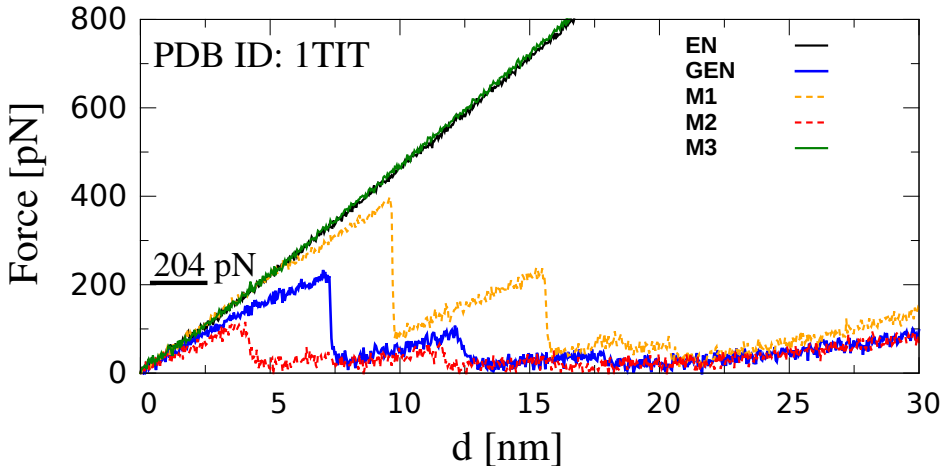

FIG. S1: We plot the force vs. cantilever displacement, d, for titin with PDB ID: 1TIT (I27 domain) for the GEN model and other breakable EN models, namely M1, M2, and M3 models. The experimental value for the maximum force is indicated by the horizontal black line along its corresponding value, which is 204±30 pN.

III. SEQUENCE OF TRANSITION STATES DURING PULLING SIMULATION

The native state of the protein studied here is mainly consisting of β-strands. For instance, the I27 domain of titin (PDB ID: 1TIT) is made by 8 β-strands: β_{1a} (4-7), β_{1b} (11-15), β_2 (18-25), β_3 (32-36), β_4 (47-52), β_5 (55-61), β_6 (69-75), and β_7 (78-88). For the type I cohesin domain (PDB ID: 1AOH) we have 9 β-strands: β_{1a} (6-11), β_{1b} (12-15), β_2 (19-28), β_3 (36-44), β_4 (48-57), β_5 (69-74), β_6 (78-86), β_7 (99-109), β_8 (115-128), β_{9a} (136-140) and β_{9b} (142-147) (see Fig. S2).

Fig. S3 we show the fraction of native contacts (NC) for each pair of secondary structures involved in all breakable models as a function of cantilever distance (d) in our pulling simulation. Our results for titin agree with previous computational studies done by Kouza et al.[1] and experimental studies [2, 3] showing that the unfolding starts by detachment of the β_{1a} from β_7, and then β_{1b} from β_7. This process plays an important role during the AFM pulling process and it requires a force of about 200 pN. We can see how the total number of NC in the GEN model decreases at the position of maximum force. It also indicates the sequence of rupture as follows: $(\beta_{1a} - \beta_7) \rightarrow (\beta_{1b} - \beta_7) \rightarrow (\beta_{1a} - \beta_2)$ at $d = 7.5$ nm for GEN

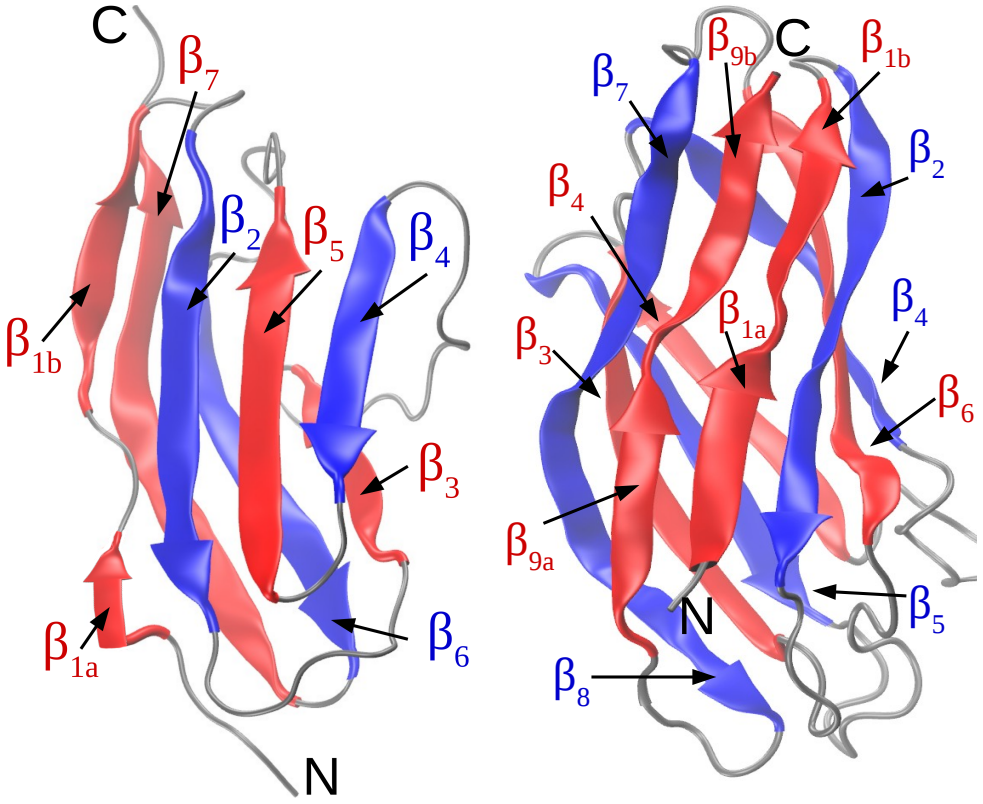

FIG. S2: PDB structure in cartoon representation for 1TIT (left panel) and 1AOH (right panel). Red and blue β-strands belong to different β-sheets. N- and C-terminal residues are marked by N and C, respectively.

model. And at distance $d = 10$ nm and $d = 5$ nm for $M1$ and $M2$ models, respectively. Overall, unfolding pathways of 1TIT are the same in GEN and m1 models (Fig. S3). The most notable difference between these two models and m2 model is in sequencing of $\beta_3 - \beta_6$ and $\beta_6 - \beta_7$ contacts. In M2 the $\beta_3 - \beta_6$ contact break after the $\beta_6 - \beta_7$ contacts, while the opposite happens in GEN and M1.

For 1AOH protein it is clear that (Fig. S3) for all breakable models the detachment of β_1 from β_9 occurs first at the same position of the maximum force signal in the F-d plot (see Fig. 3 in the main text). This result is in agreement with computational [4] and experimental studies [5]. The largest force peak corresponds to detachment of $(\beta_{1a} - \beta_{1b})$ from $(\beta_{9a} - \beta_{9b})$ as we show in Fig. S3 and in agreement with ref. [4]. In all breakable models the mechanical unfolding pathway of 1AOH is $(\beta_1 - \beta_9) \rightarrow (\beta_2 - \beta_7) \rightarrow (\beta_3 - \beta_7) \rightarrow (\beta_6 - \beta_7)$ (Fig. S3).

[1] Kouza, M., C.-K. Hu, M. S. Li, and A. Kolinski, 2013. A structure-based model fails to probe the mechanical unfolding pathways of the titin I27 domain. *J. Chem. Phys.* 139:08B615_1.

[2] Fowler, S. B., R. B. Best, J. L. T. Herrera, T. J. Rutherford, A. Steward, E. Paci, M. Karplus, and J. Clarke, 2002. Mechanical unfolding of a titin Ig domain: structure of unfolding intermediate revealed by combining AFM, molecular dynamics simulations, NMR and protein engineering. *J. Mol. Biol.* 322:841–849.

[3] Best, R. B., S. B. Fowler, J. L. T. Herrera, A. Steward, E. Paci, and J. Clarke, 2003. Mechanical unfolding of a titin Ig domain: structure of transition state revealed by combining atomic force microscopy, protein engineering and molecular dynamics simulations. *J. Mol. Biol.* 330:867–877.

[4] Wojciechowski, M., P. Szymczak, M. Carrión-Vázquez, and M. Cieplak, 2014. Protein unfolding by biological unfoldases: Insights from modeling. *Biophys. J.* 107:1661–1668.

[5] Valbuena, A., J. Oroz, R. Hervás, A. M. Vera, D. Rodríguez, M. Menéndez, J. I. Sulkowska, M. Cieplak, and M. Carrión-Vázquez, 2009. On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl. Acad. Sci. U.S.A.* 106:13791–13796.

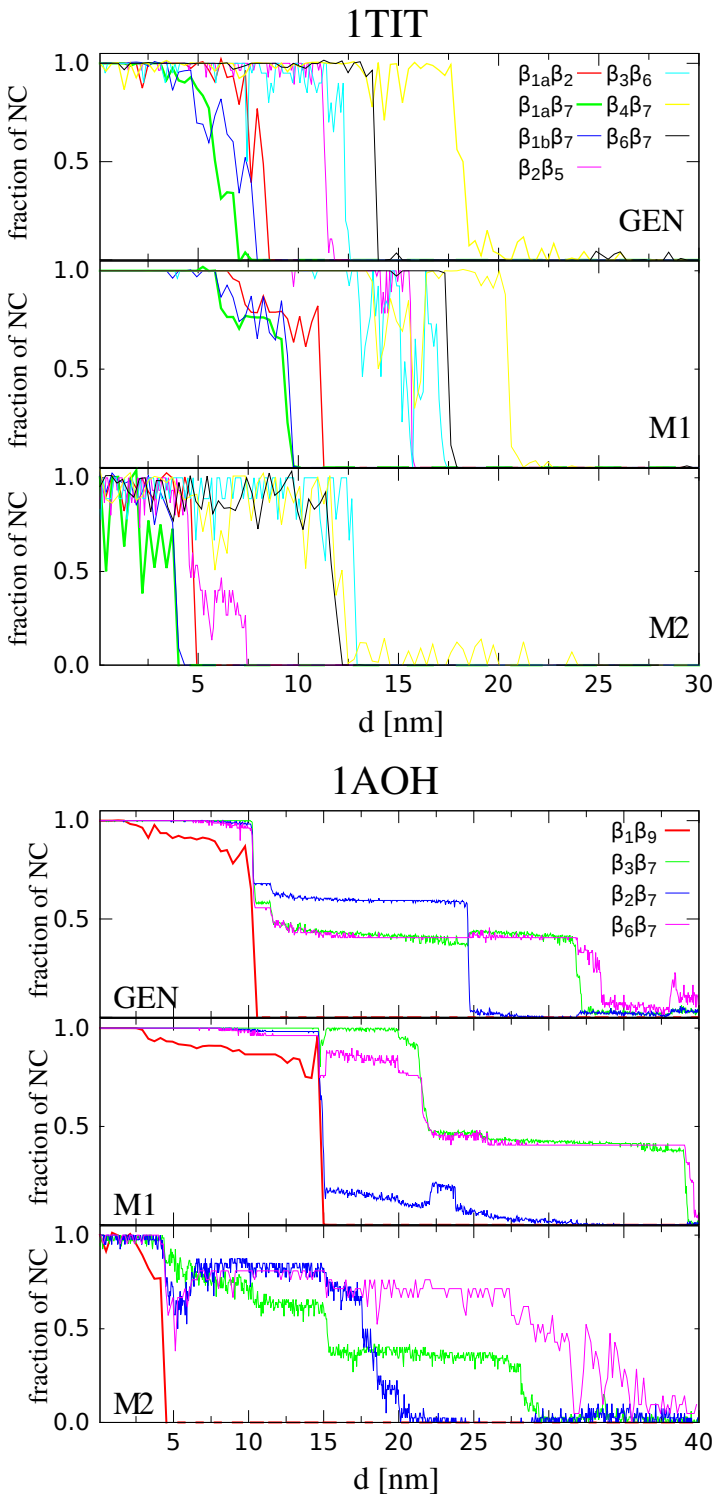

FIG. S3: Cantilever distance dependence of averaged fractions of native contacts (NC). Native contacts are formed by 8 β-strands for titin (PDB ID: 1TIT) and 9 β-strands for cohesin domains (PDB ID: 1AOH). Clearly, in the case of titin the unfolding process starts from the N-terminus by detaching β_{1a} and β_{1b} from β_7 for all breakable models. For 1aoh protein, we observe that the detachment of β_1 from β_9 is the first event and thus it is responsible for the main force peak.