Protein conformational changes: What can geometric nonlinear analysis tell us?

AIMETA 2019
XXIV Conference
The Italian Association of Theoretical and Applied Mechanics
Rome, Italy, 15-19 September 2019

PROTEIN CONFORMATIONAL CHANGES: WHAT CAN GEOMETRIC NONLINEAR ANALYSIS TELL US?

Domenico Scaramozzino, Giuseppe Lacidogna and Alberto Carpinteri

Politecnico di Torino, Department of Structural, Geotechnical and Building Engineering, Corso Duca degli Abruzzi 24 – 10129 Torino, Italy

E-mail: domenico.scaramozzino@polito.it, giuseppe.lacidogna@polito.it, alberto.carpinteri@polito.it

Keywords: Protein Conformational Change, Geometric Nonlinear Analysis, Mechanical Instability, HIV-1 Protease.

Abstract. *Protein functioning is usually associated to conformational changes. In the last decades, several researchers have made use of elastic network models to investigate such shape changes, showing that these depend on the protein intrinsic flexibility. Moreover, it has been indicated that low-frequency modes, arising from the application of modal analysis, are strictly related to the conformational transition. Several efforts have also been made in order to generate feasible pathways as well as to investigate the active forces applied at specific locations driving the conformational change. The problem has usually been addressed by means of linear theories, under the assumption of small displacements. In this contribution, we question whether such assumption is reliable from a mechanical viewpoint. In particular, we investigate the influence of geometric nonlinearities, by applying both linear and (geometric) nonlinear analysis to the protein elastic network model and comparing the outcomes in terms of force profiles. Eventually, from the results regarding the conformational change of HIV-1 protease subunit, we show that the displacements should not be considered small a priori.*

1 INTRODUCTION

Protein activity has fascinated researchers for decades since many fundamental biological processes rely on it, e.g. oxygen transportation along the body due to hemoglobin performance, nutrients carriage inside the cell via molecular motors, ion exchange between the inside and the outside of the cell due to the activity of transmembrane proteins such as Na^+/K^+-ATPase and Ca^{2+}-ATPase, etc. Protein functioning is known to be related to the three-dimensional structure, which in turns depends on the specific amino acid sequence and environmental conditions (pH, temperature, applied forces, etc.) [1]. Moreover, protein activity is usually associated to conformational changes, which imply a modification of the three-dimensional shape, e.g. hemoglobin exhibits a structural rearrangement when switching from the deoxy- to the oxy-state, a large conformational transition (triggered by ATP-binding) occurs in kinesin heads when it walks along microtubules, etc.

Although proteins are complex systems interacting among each other and with the external environment, simple elastic network models based on a purely mechanical description of the structure (point masses connected by linear springs) have proven their efficacy in describing protein dynamics and flexibility [2]. It has also been shown that even coarse-grained representations, based only on C_α atoms, are able to capture the essential characteristics of protein activity [3-5].

Elastic network models were also useful for investigating protein behavior under several perspectives. Tama and Sanejouand [6] investigated the contribution of the eigenmodes arising from normal mode analysis to the conformational change of several proteins, and observed that the low-frequency modes, generally implying large collective motions, are strictly related to the conformational transition. Moon *et al.* [7,8] made use of elastic network models in order to generate realistic pathways for protein conformational changes. Eyal and Bahar [9] investigated the mechanical behavior of proteins subject to forces applied at different locations and provided numerical insights on the experimental measurements performed with single-molecule manipulation techniques, like atomic force microscopy (AFM) and optical tweezers (OT). C. Atilgan and A. R. Atilgan [10] proposed the perturbation-response scanning (PRS) method to investigate the conformational change of ferric binding protein, by applying directional forces at single residues. The PRS was also subsequently applied by Atilgan *et al.* [11] to a set of 25 proteins in order to evaluate the residues which had to be perturbed to generate an accurate description of the observed conformational change. More recently, Liu *et al.* [12] made use of an elastic network model in order to evaluate the active forces, originating from ATP hydrolysis, which drive the GroEL conformational change.

The calculations performed by the abovementioned authors were based on the linear response theory, which implies considering small displacements from a Structural Mechanics viewpoint. However, dealing with elastic networks (which can be considered as space truss structures made up of bars connected together with spherical hinges), it is likely that geometric nonlinearities can affect the structural response if the displacements are not so small. As a matter of fact, the presence of nonlinear effects in relaxation dynamics of motor proteins has been suggested in [13,14]. Therefore, in this contribution, we investigate the influence of geometric nonlinearities on protein conformational transitions, questioning whether the assumption of small displacements is reliable. HIV-1 protease is selected as a case study and a coarse-grained elastic network model is considered to analyze the open-to-closed transition.

In particular, starting from the displacement field of the observed conformational change, we evaluate the forces on the residues required to fulfill the equilibrium conditions, both referring to the undeformed (linear analysis) and deformed configuration (geometric nonlinear analysis). The calculations are performed also by changing some fundamental model

parameters, such as the cutoff value and the stiffness variation law of the connections. Finally, by comparing the results in terms of force values and orientations, we conclude that the displacements involved within the conformational change should not be considered small *a priori* from a mechanical viewpoint.

2 METHODOLOGY

As mentioned in the Introduction, the open-to-closed conformational transition of HIV-1 protease subunit was analyzed here by means of a coarse-grained elastic network model.

2.1 Protein elastic network model

The elastic network was built considering only the coordinates of C_α atoms. These were taken from the Protein Data Bank [15] for the "open" (pdb code: 1hhp) and "closed" (1ajx) configuration. Three different cutoff values r_c (8, 10 and 12 Å) were considered to create the connections among the nodes (Fig. 1).

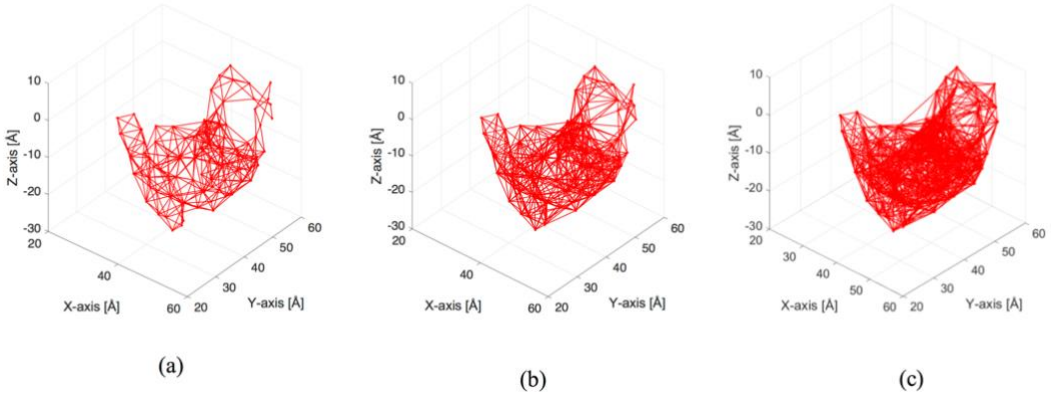

Figure 1: HIV-1 protease subunit open configuration (1hhp): (a) r_c = 8 Å; (b) r_c = 10 Å; (c) r_c = 12 Å.

Moreover, three different stiffness variation laws were considered for the connections, which are usually employed within elastic network models, i.e. constant (Eq. 1a), exponential decay (Eq. 1b) and inverse decay (Eq. 1c), as follows:

$$k_m = k_0, \tag{1a}$$

$$k_m = k_0 \exp\left(-\frac{{L_m}^2}{{L_0}^2}\right), \tag{1b}$$

$$k_m = \frac{{k_0}^*}{L_m}, \tag{1c}$$

where k_m is the stiffness of the mth connection, L_m its length, and k_0, ${k_0}^*$ and L_0 represent the parameters of the stiffness laws. In the present analysis, k_0 and ${k_0}^*$ were set equal to 1, since their values do not affect the distribution of the forces among the protein chain, but cause only a uniform scaling of such distribution. For this reason, the forces arising from the analyses will be expressed in "arbitrary units" (Section 3). The parameter L_0 was set equal to 7 Å in agreement with Hinsen's work [4]. Summarizing, the nine models reported in Tab. 1 were considered by varying the cutoff value and the stiffness variation law.

Model	Cutoff [Å]	Stiffness variation law
M8C	8	Constant
M8E	8	Exponential decay
M8I	8	Inverse decay
M10C	10	Constant
M10E	10	Exponential decay
M10I	10	Inverse decay
M12C	12	Constant
M12E	12	Exponential decay
M12I	12	Inverse decay

Table 1: Nine considered models for the protein elastic network.

2.2 Evaluation of the open-to-closed conformational change

As for the displacement field of the conformational change, this was computed by vector difference between the two end configurations, after they have been superimposed. Note that if the number of residues is N, the size of the conformational change vector $\boldsymbol{\delta}$ is $3N \times 1$ ($N = 99$ for HIV-1 protease subunit). The obtained absolute displacements are shown in Fig. 2. As can be seen, the conformational change is very localized in the central portion of the protein chain (the "flap" region). The maximum displacement is ~ 4 Å.

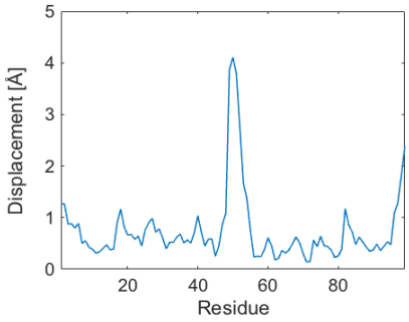

Figure 2: Absolute displacements in the open-to-close transition of HIV-1 protease subunit.

2.3 Equilibrium equations in the undeformed structure (linear analysis)

Under the assumption of small displacements and linear elasticity, i.e. considering the equilibrium conditions in the initial (undeformed) configuration, the structural problem could be formulated as a linear relationship as follows:

$$\boldsymbol{F_L} = \boldsymbol{K\delta}, \tag{2}$$

where $\boldsymbol{F_L}$ represents the $3N \times 1$ vector containing the forces applied to the residues (obtained via linear analysis), $\boldsymbol{\delta}$ is the $3N \times 1$ conformational change vector (Section 2.2) and $\boldsymbol{K}$ is the $3N \times 3N$ stiffness matrix calculated in the open configuration.

The stiffness matrix was computed by means of the displacements' method, i.e. each stiffness coefficient $K_{ij}^{\alpha\beta}$ (which refers to the nodes i and j, and to the directions α and β) could be calculated by imposing a unitary displacement on node j along the direction β and evaluating the reaction force along the direction α on node i. This procedure was automatized into MatLab environment so that, given the information on the structure (coordinates of the points and stiffness values of the connections), the global stiffness matrix could be automatically

computed. Obtained the stiffness matrix $\boldsymbol{K}$ and known the vector of the observed conformational change $\boldsymbol{\delta}$ (Section 2.2), the force vector $\boldsymbol{F_L}$ could be obtained via Eq. 2.

2.4 Equilibrium equations in the deformed structure (geometric nonlinear analysis)

Removing the hypothesis of small displacements, in order to take into account geometric nonlinearities, the equilibrium equations were written in the deformed (final) configuration. The relationship between the force and displacement vector took now the following general form:

$$\boldsymbol{F_{NL}} = \boldsymbol{f}(\boldsymbol{\delta}), \tag{3}$$

where $\boldsymbol{F_{NL}}$ represents the $3N \times 1$ vector containing the forces applied to the residues (obtained via geometric nonlinear analysis) and $\boldsymbol{f}$ represents the system of nonlinear equations linking the force and displacement vectors.

The nonlinear system reported in Eq. 3 could be written analytically by considering the equilibrium conditions in the final (deformed) configuration, through the following steps: (1) the elongation of the m^{th} connection was computed based on the effective positions of the nodes in the initial and final configuration; (2) the internal force within the m^{th} connection was computed based on the stiffness value of the connection (Eq. 1) and calculated elongation; (3) the total force on the i^{th} node was finally computed by equilibrating all the internal forces acting on the i^{th} node, referring to the final (deformed) configuration.

2.5 Comparison between linear and nonlinear analysis

It is clear that, if the hypothesis of displacements were valid, linear and nonlinear analysis should provide the same results, i.e. the force vectors evaluated via Eqs. 2 and 3 should be identical ($\boldsymbol{F_{L,i}} = \boldsymbol{F_{NL,i}}$ for each residue *i*). Therefore, by comparing the obtained vectors $\boldsymbol{F_{L,i}}$ and $\boldsymbol{F_{NL,i}}$, the influence of geometric nonlinearities could be properly investigated.

In particular, the modules of the force vectors $\boldsymbol{F_{L,i}}$ and $\boldsymbol{F_{NL,i}}$ at the i^{th} residue were compared, as well as the difference in their orientation. The latter was analyzed by means of the following cosine operator:

$$\cos\theta_i = \frac{\boldsymbol{F_{L,i}}^T \boldsymbol{F_{NL,i}}}{\sqrt{(\boldsymbol{F_{L,i}}^T \boldsymbol{F_{L,i}})}\sqrt{(\boldsymbol{F_{NL,i}}^T \boldsymbol{F_{NL,i}})}}, \tag{4}$$

which provides a numerical estimate of the different orientation between the vectors $\boldsymbol{F_{L,i}}$ and $\boldsymbol{F_{NL,i}}$. If linear and nonlinear analysis provided the same results, i.e. the influence of geometric nonlinearities were negligible, we should obtain for each residue $|\boldsymbol{F_{L,i}}| = |\boldsymbol{F_{NL,i}}|$ and $\cos\theta_i = 1$. Therefore, the larger the deviation from these conditions, the greater the influence of geometric nonlinearities.

3 RESULTS AND DISCUSSION

In Figs. 3-8 the results arising from the calculations described in Section 2 are reported. In particular, Figs. 3, 5 and 7 show the comparison of the force modules' distributions (in a.u.), whereas the orientation difference between the force vectors is shown in Figs. 4, 6 and 8.

As can be observed from the force comparisons, the profiles of the vector force modules $|\boldsymbol{F_{L,i}}|$ and $|\boldsymbol{F_{NL,i}}|$ show several common features for large portions of the protein chain; however, significant differences are evident in the central region, specifically between residues 45 and 55 (flap region), which in turn is the portion which exhibits the largest displacements (Fig. 2). For this reason, we can conclude that the displacements involved within the open-to-

closed transition of HIV-1 protease subunit cannot be considered small from a mechanical point of view, and that the influence of geometric nonlinearities is not negligible. Note that this conclusion can be drawn from all the investigated models (Figs. 3, 5 and 7), making the model parameters not so critical for drawing such conclusion.

The difference between the force vectors $\boldsymbol{F_{L,i}}$ and $\boldsymbol{F_{NL,i}}$ in the central region of the protein chain is not only observable in terms of modules, but also in terms of spatial orientations. In fact, in this region the cosine operator (Eq. 4) exhibits values less than 1 (Figs. 4, 6 and 8). This indicates that the forces applied to the nodes of the flap region show different orientations, if they are computed via linear or nonlinear analysis. It is worth noting that also other protein portions are characterized by cosine values other than unity.

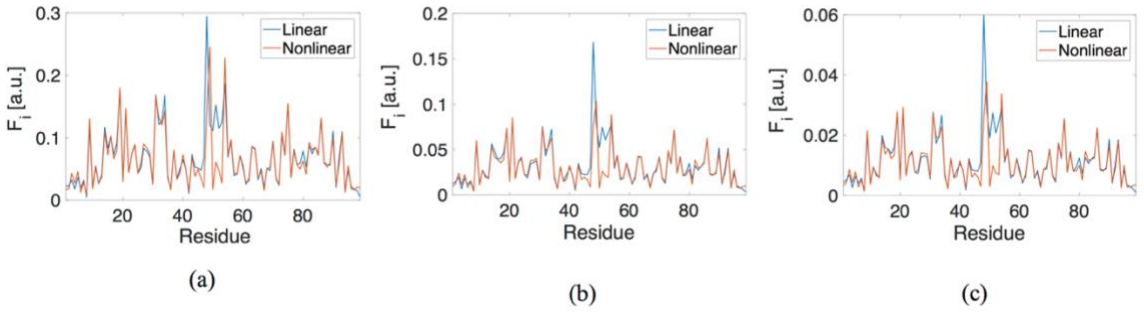

Figure 3: $|\boldsymbol{F_{L,i}}|$ vs $|\boldsymbol{F_{NL,i}}|$: (a) M8C; (b) M8E; (c) M8I.

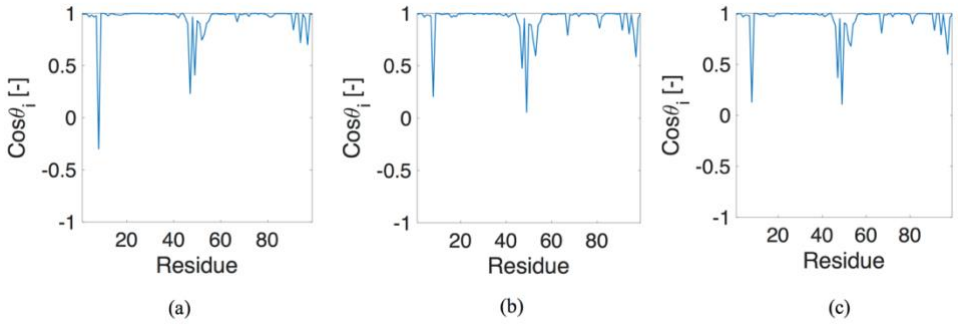

Figure 4: $\mathrm{Cos}\,\theta_i$: (a) M8C; (b) M8E; (c) M8I.

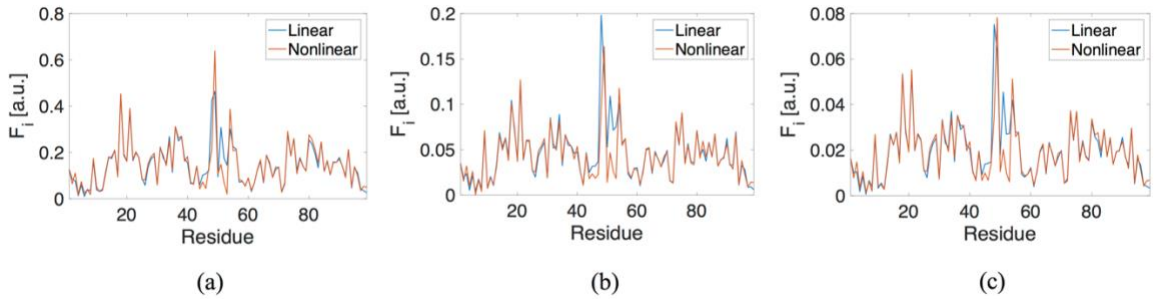

Figure 5: $|\boldsymbol{F_{L,i}}|$ vs $|\boldsymbol{F_{NL,i}}|$: (a) M10C; (b) M10E; (c) M10I.

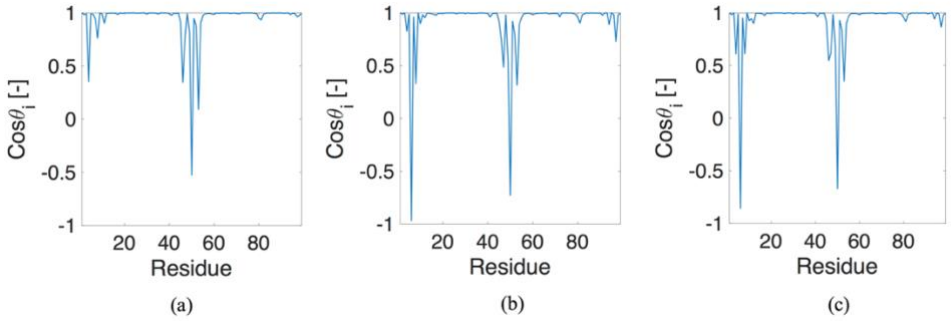

Figure 6: Cos θ_i: (a) M10C; (b) M10E; (c) M10I.

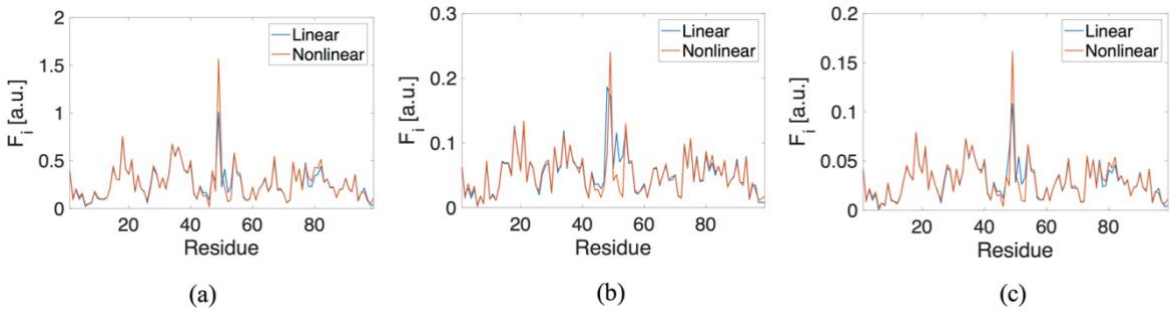

Figure 7: $|\boldsymbol{F_{L,i}}|$ vs $|\boldsymbol{F_{NL,i}}|$: (a) M12C; (b) M12E; (c) M12I.

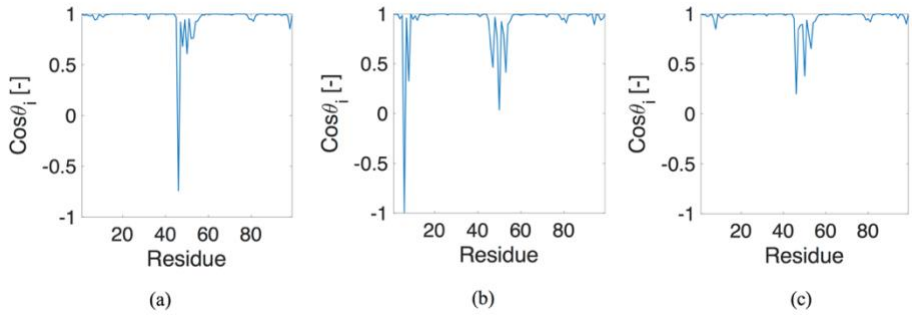

Figure 8: Cos θ_i: (a) M12C; (b) M12E; (c) M12I.

Since geometric nonlinearities have been shown to have a not negligible influence when investigating conformational changes through elastic network models, it is believed that some (snap-through-like) mechanical instability may also occur along the transition. This will be investigated in future research efforts by adopting a step-by-step approach, i.e. by writing the equilibrium equations for each of the intermediate configurations of a given (feasible) pathway.

4 CONCLUSIONS

In this contribution, we performed linear and geometric nonlinear analysis in order to investigate the influence of geometric nonlinearities on protein conformational transitions. In particular, a coarse-grained elastic network model was used for the analysis of HIV-1 protease subunit conformational change. Equilibrium equations were written both referring to the undeformed and deformed configuration, i.e. by neglecting or taking into account geometric nonlinearities, respectively. Different model parameters, such as cutoff values and stiffness variation laws, were also considered. From the comparison of the outcomes in terms of force values and orientation, we concluded that the displacements involved within the conformational

change should not be considered small from a mechanical perspective, and that the influence of geometric nonlinearities is not negligible. Moreover, it is also possible that some mechanical instabilities may subtend the conformational transition.

REFERENCES

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell, 5th Edition*. Garland Science, New York, 2002.

[2] M.M. Tirion, Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Physical Review Letters*, **77**, 1905-1908, 1996.

[3] I. Bahar, A.R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*, **2**, 173-181, 1997.

[4] K. Hinsen, Analysis of Domain Motions by Approximate Normal Mode Calculations. *Proteins: Structure, Function, and Genetics*, **33**, 417-429, 1998.

[5] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, I. Bahar, Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophysical Journal*, **80**, 505-515, 2001.

[6] F. Tama, Y.H. Sanejouand, Conformational change of proteins arising from normal mode calculations. *Protein Engineering*, **14**, 1-6, 2001.

[7] K.K. Moon, R.L. Jernigan, G.S. Chirikjian, Efficient Generation of Feasible Pathways for Protein Conformational Transitions. *Biophysical Journal*, **83**, 1620-1630, 2002.

[8] K.K. Moon, G.S. Chirikjian, R.L. Jernigan, Elastic models of conformational transitions in macromolecules. *Journal of Molecular Graphics and Modelling*, **21**, 151-160, 2002.

[9] E. Eyal, I. Bahar, Toward a Molecular Understanding of the Anisotropic Response of Proteins to External Forces: Insights from Elastic Network Models. *Biophysical Journal*, **94**, 3424-3435.

[10] C. Atilgan, A.R. Atilgan, Perturbation-Response Scanning Reveals Ligand Entry-Exit Mechanisms of Ferric Binding Protein. *PLoS Computational Biology*, **5**, e1000544, 2009.

[11] C. Atilgan, Z.N. Gerek, S.B. Ozkan, A.R. Atilgan, Manipulation of Conformational Change in Proteins by Single-Residue Perturbations. *Biophysical Journal*, **99**, 933-943, 2010.

[12] J. Liu, K. Sankar, Y. Wang, K. Jia, R.L. Jernigan, Directional Force Originating from ATP Hydrolysis Drives the GroEL Conformational Change. *Biophysical Journal*, **112**, 1561-1570, 2017.

[13] Y. Togashi, A.S. Mikhailov, Nonlinear relaxation dynamics in elastic networks and design principles of molecular machines. *Proceedings of the National Academy of Sciences USA*, **104**, 8697-8702, 2007.

[14] Y. Togashi, T. Yanagida, A.S. Mikhailov, Nonlinearity of Mechanochemical Motions in Motor Proteins. *PLoS Computational Biology*, **6**, e1000814, 2010.

[15] https://www.rcsb.org