

Motion Modes Analysis with Elastic Network Model from the p53 DNA-binding Domain

Hai Yan Li^{1,2*}

¹Shandong Provincial Key Laboratory of Functional Macromolecular Biophysics,
Dezhou University, 566 University Road West, Dezhou - 253 023, China.

²Department of Physics, Dezhou University, 566 University Road West, Dezhou - 253 023, China.

(Received: 27 September 2013; accepted: 04 November 2013)

p53 DNA-binding domain (p53DBD) plays crucial roles in domain stability, DNA binding and dimerization. In this article, we analyze the motion mode of p53DBD using the Gaussian network model and anisotropy elastic network model. For each structure, the calculated B-factors by two models obtain good agree with the experimentally determined B-factors in x-ray crystal structures. The results show that the four CRs (conserved regions) in the p53DBD have small fluctuation in the slowest mode of GNM, but the same regions lay the peaks of the fast modes. With the aid of the anisotropy elastic network mode, we analyze the motion directions of this domain. The first slowest mode of ANM for both structures mainly corresponds to the rotational motion. However, the CR V (N-terminal) shows an opposite direction with the other parts of structure. Through this motion, p53DBD will bind to its consensus sequence DNA in a highly cooperative manner. By analyzing cross-correlations between residue fluctuations of p53DBD, the results indicate that residues involved in similar functions are fluctuating in a cooperative manner.

Key words: p53 DNA-binding domain (p53DBD), Motion Mode, Elastic Network Model.

As a transcription regulator, p53 plays a fundamental role in cell cycle arrest, apoptosis and DNA repair¹. This function is necessary to help maintain healthy levels of cell proliferation. p53 with 393 residues contains an N-terminal transactivation domain, a DNA-binding core domain, a tetramerization domain and a C-terminal regulatory domain² Although the structure of full-length P53 has not been known yet, the crystal structures have been resolved for p53DBD in complex with DNA and in the absence of DNA. p53DBD has been widely studied by experimental and theoretical methods. Through many structural studies, residues within the p53DBD that are

crucial for domain stability, DNA binding and dimerization have been elucidated and correlated with known cancerous mutations³⁻⁵. Other studies also reveal that p53DBD exists in both a latent and active DNA-binding form that can be detected in vivo^{6,7}. Through molecular dynamics simulations, Lu et al. have investigated the influence of zinc ion on the p53 DNA recognition and the stability of the p53DBD⁸. Mey Khalili et al. have studied the dynamics and thermodynamics of two of the four evolutionarily conserved segments from the p53DBD⁹. MD simulation is an effective method to obtain the detailed microscopic dynamics of proteins, however, it is very time-consuming to study the conformational change for such a big protein, p53DBD.

Simplified coarse grained models have been established as valid and efficient means to probe large protein conformational motions and

* To whom all correspondence should be addressed.
Tel.: +86-534-8985879; Fax: +86-534-8985884;
E-mail: dzxylhy@163.com

flexibility^{10,11}. In this work, we focus on studying the motion modes for the bound and unbound forms of the p53DBD using the two coarse-grained elastic models: the Gaussian network model (GNM)¹² and the anisotropy elastic network model (ANM)¹³. GNM/ANM assumes an elastic network structure, formed by springs that connect the close neighboring C_α atoms in the 3-dimensional structure of proteins. The method is coarse-grained by using a single site per residue and assuming that all the residues in a cutoff distance are in contact. The two coarse-grained elastic models can effectively reveal conformational motions and flexibility of the bound and unbound forms of the p53DBD.

MATERIALS AND METHODS

The database of protein

Crystal structures have been solved for p53DBD in complex with DNA (bound form, PDB: ID 1TSR[3]) and in the absence of DNA (unbound form, PDB: ID 2OCJ¹⁴ (Fig.1). In this work, the isolated monomers A in bound form and unbound form are used in calculation. p53DBD is made up of a sandwich of antiparallel α -sheets with the H1 and H2 helices on one end of the molecule. The longer H2 helix, along with the proceeding loop, interacts with the DNA major groove of the pentamer sequence, while the shorter H1 helix, along with the preceding L2 loop and the L3 loop, ligate a zinc ion that together forms the p53DBD dimerization interface over the DNA minor groove. Another loop, the L1 region, shows a large amount of flexibility between species which has led to the suggestion that it is in no way involved with DNA binding.

Methods

The GNM models a protein as an elastic network with the nodes being the amino acids, usually represented by the, attached by Hookean springs where the atoms fluctuate about their mean positions¹². The Kirchhoff matrix of such a structure is constructed using Eq.1:

$$\Gamma = \begin{cases} -1 & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{i,i \neq j} \Gamma_{ij} & \text{if } i = j \end{cases} \quad \dots(1)$$

Where i and j , indices of C_α and r_c , is the cutoff

distance, 7 Å is adopted in this work.

The inverse of this Kirchhoff matrix is related to the magnitude of relative fluctuations of the i th and j th units in the network as shown in Eq. 2, and represents the mean-square fluctuation of each unit when . The intrinsic flexibility of the structure, which is reported in the crystallographic B-factors, is also directly related to the mean-square fluctuations by Eq. 3:

$$\langle \Delta R_i \cdot \Delta R_j \rangle = (3k_B T / \gamma) [\Gamma^{-1}]_{ij} \quad \dots(2)$$

$$B_i = 8\pi^2 \langle \Delta R_i \cdot \Delta R_i \rangle / 3 \quad \dots(3)$$

Where k_B is the Boltzmann constant, T is temperature, and γ is the harmonic force constant.

The mean-square fluctuation of the i th residue associating with the k th mode can be expressed as

$$\langle \Delta R_i \cdot \Delta R_j \rangle_k = (3k_B T / \gamma) \lambda_k^{-1} [u_k]_i [u_k]_j \quad \dots(4)$$

In the GNM, the cross correlation is normalized as

$$C_{ij} = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{[\langle \Delta R_i^2 \rangle \times \langle \Delta R_j^2 \rangle]} \quad \dots(5)$$

Although GNM provides the magnitudes of the displacements of atoms or chain units from their equilibrium positions for large scale motions, it does not provide any information on the direction of the motions. The ANM¹³ is an extension of GNM, which adds direction to the motions. A 3N-dimensional Hessian matrix H is adopted in the ANM.

$$H = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & h_{N3} & h_{NN} \end{bmatrix} \quad \dots(6)$$

Where the element h_{ij} is a submatrix with order . The detailed calculation of h_{ij} is:

$$h_{ij} = \begin{bmatrix} \frac{\partial^2 V}{\partial x_i \partial x_j} & \frac{\partial^2 V}{\partial x_i \partial y_j} & \frac{\partial^2 V}{\partial x_i \partial z_j} \\ \frac{\partial^2 V}{\partial y_i \partial x_j} & \frac{\partial^2 V}{\partial y_i \partial y_j} & \frac{\partial^2 V}{\partial y_i \partial z_j} \\ \frac{\partial^2 V}{\partial z_i \partial x_j} & \frac{\partial^2 V}{\partial z_i \partial y_j} & \frac{\partial^2 V}{\partial z_i \partial z_j} \end{bmatrix} \quad \dots(7)$$

When $i \neq j$, the analytic expression for the elements of h_{ii} is

$$\frac{\partial^2 V}{\partial x_i \partial y_j} = \frac{-\gamma (x_j - x_i)(y_j - y_i)}{R_y^2} \Big|_{R_y = R_y^0} \quad (8)$$

When $i = j$, the analytic expression for the elements of h_{ii} is

$$\frac{\partial^2 V}{\partial x_i \partial y_j} = \gamma \sum_{j \neq i} \frac{(x_j - x_i)(y_j - y_i)}{R_y^2} \Big|_{R_y = R_y^0} \quad \dots(9)$$

The meanings of γ and R are the same as in Eq.1. x , y and z represent the position coordinates of atoms. The cutoff distance in ANM has been adopted as 15 Å in this work.

RESULTS AND DISCUSSION

Comparison for theoretical and experimental B-factors

To test the validity of the Elastic Network Model, the experimental crystallographic B-factors were compared to the data calculated with the GNM and ANM methods. The relationship between an individual residue's fluctuations and its temperature factor according to Eq.3, the B-factors of the isolated monomers A in bound form and the isolated monomers A with DNA in the unbound form were calculated, respectively. The factor $k_B T / \gamma$ is essentially a force constant for the virtual springs connecting C_α and sets the overall scale factor. The resulting $k_B T / \gamma$ values used for the bound form in GNM and ANM are 0.789 Å² and 0.328 Å², and that for the unbound form are 0.683 Å² and 0.294 Å².

In Fig.2, the predicted temperature factors by the GNM (solid curves) and ANM (dashed curves) are compared to those experimental measured by X-ray crystallography (dotted curves). The correlation coefficient of the B-factor between the experiment and GNM is 0.695, 0.501 for the unbound and the bound forms, respectively. That between the experiment and ANM is 0.647 and 0.607 for the unbound and the bound forms, respectively. The results are similar

to those of recent studies for other proteins^{10, 15}. It is found that the unbound structure gives a higher correlation coefficient than the bound structure, which reflects the effect of the DNA on the residue fluctuations of the protein. The agreement between experiments and theory is excellent, lending support to the use of GNM and ANM for further investigation of the conformational changes of P53 core domain.

Mode shape analysis by GNM

The total residues fluctuations can be decomposed into high- and low-frequency fluctuations, namely fast and slow modes. The slow modes are reported to be associated with the collective dynamics of the overall tertiary structure of protein and relevant to biological function¹⁶. Fig.3 shows the first slowest mode of 2OCJ and 1TSR calculated by the GNM. It points out the functionally important sections in the p53 core domain structure, which correspond to the conserved regions (CRs) as described in literature¹⁷. There is four CRs (shown in figure 1 (a)) in the p53 core domain structure: L1 loop (residues 112–124) and S2 and S2' sheets (residues 124–141) together corresponding to CR II, part of L2 loop and H1 helix (residues 171–181) together corresponding to CR III, L3 loop (residues 236–251) corresponding to CR IV, end of S10 sheet (residues 271–274) and H2 helix (residues 278–286) together corresponding to CR V. It should be noted that CR I is not positioned in the p53 core domain (located in the N-terminus domain). Fig.3 shows that CR II, CR III, CR IV and CR V in the p53 core domain of 1TSR have small fluctuation during the domain movements. While CR I and CR V have large fluctuation in 2OCJ. This is because CR I and CR V in the p53 core domain involve the residues which make base specific contacts with the DNA major groove. The significant difference in two structures shows that the residues (LYS120 in CR II, ALA2800ARG281 and ARG 283 in CR V) make contacts with DNA reduce the fluctuation during the conformation changes, as well as the residues around them. Besides, residues 117-121 of the L1 Loop are disordered in each subunit of the human p53DBD dimer-DNA complex. It is known from Fig.3 the L1 Loop shows considerable flexibility in the p53DBD domain in the absence of DNA. These are six residues (Arg175, Gly245, Arg248, Arg249, Arg273 and Arg283) located

in the conserved regions have been found to be most frequently mutated in human cancer. These residues all have lower fluctuations in the two structures except for Arg283. Previously reported structural studies on the p53 core domain had noted a zinc ion that is ligated by three cysteine residues (Cys176, Cys238 and Cys242) and one histidine residue (His179) from the H1 helix and L2 and L3 loops. Fig.3 displays that the residues bonding with the zinc ion have small fluctuations both in each structure.

The fast modes of the motions correspond to the geometric irregularity in the local structure¹⁸. Previous studies have found that high frequency fluctuating residues are thought to be kinetically key residues and critically important for the stability of the tertiary fold. Figure 3 shows the fastest ten modes of the two structures. As shown

in Fig. 4, there are similar peak residues (i.e. Thr125, Phe134, Val157, Val216, Asn 235, Asn268 and Pro278) in the curves of the two forms. Most of the peak residues lie in the CRs except of the CR III. Residues Thr125, Phe134, Asn 235 and Pro278 are located at the interface between DNA and p53DBD. The two residues (i.e., Val157 and Val216) are the conserved hydrophobic ones. All of these peak residues are highly conserved and play a key role in the stability of the protein.

Mode shape analysis by ANM

GNM can only provide the magnitude of displacement of atoms from their equilibrium positions for large-scale motions. To obtain the directions of motion, ANM is applied to analyze the slow modes of the motions. As seen in Fig.5, the first slowest mode of 1TSR mainly corresponds to the rotation motion with an anticlockwise direction.

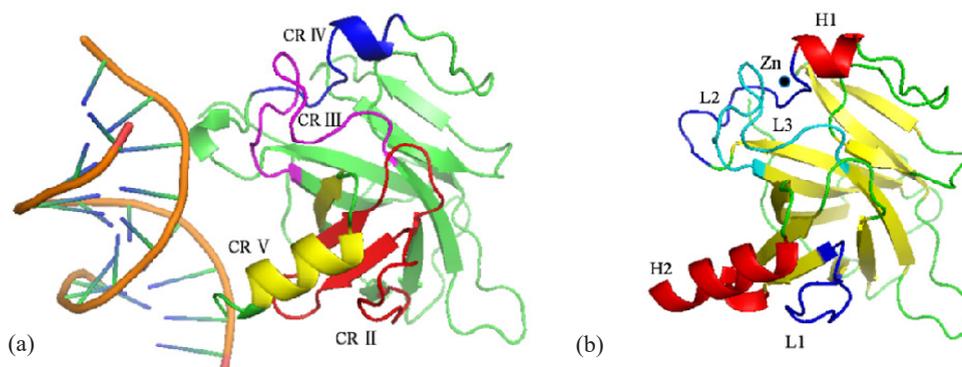


Fig. 1. P53 core domain of (a) the bound form (1TSR in PDB) and (b) the unbound form (2OCJ in PDB). The four conserved regions CR II, CR III, CR IV and CR V are colored in red, magentas, blue and yellow

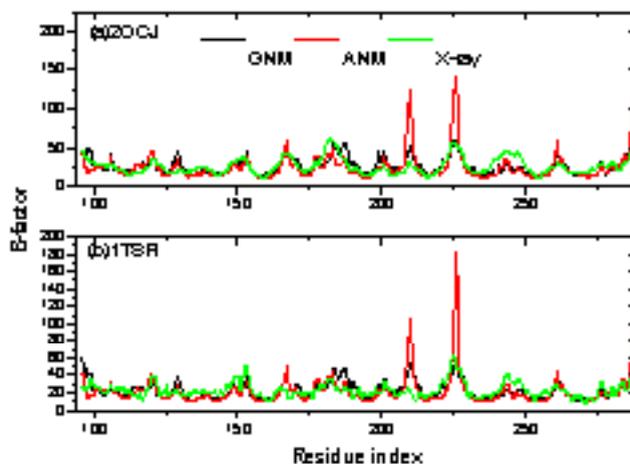


Fig. 2. Comparison of experimental and theoretical B-factors as a function of residue number along the chain
J PURE APPL MICROBIO, 7(SPL. EDN.), NOVEMBER 2013.

While, this mode of 2OCJ have the similar rotation motion with a clockwise direction. Through this motion, the predominantly monomeric p53DBD will bind to its consensus sequence DNA in a highly cooperative manner. However, the CR V (N-terminal) shows an opposite direction with the other parts of structure. This motion is helpful for CR V to have a direct contact with DNA major groove. In the two structures, the residues Gln167,

Asn210, Val225, Ser261 and Leu289 whose fluctuations increased significantly have been marked in the Fig.5. Four residues Gln167, Asn210, Val225 and Ser261 are located on the outside of top of p53DBD, which have large fluctuation to make the whole domain with an anticlockwise direction rotation. The motive magnitude of the N-terminal coil (residues 289) is much larger than those of the other parts.

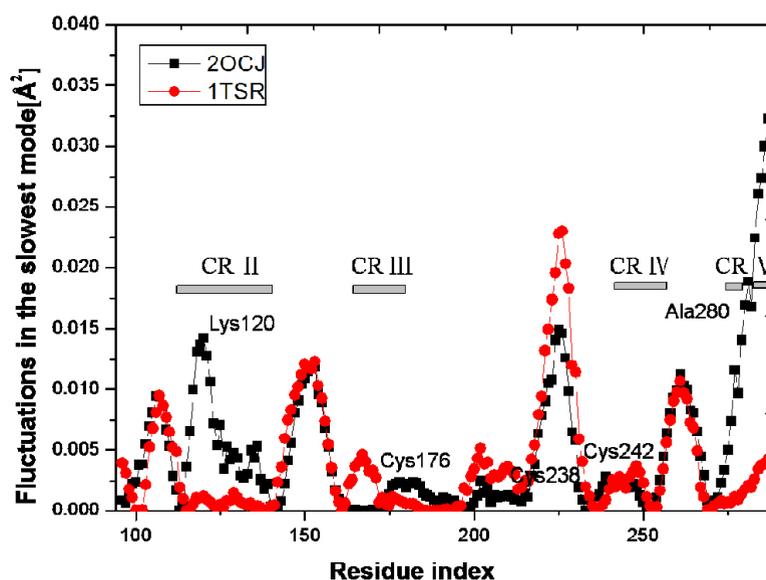


Fig. 3. Comparison between the slowest mode shapes of the two structures

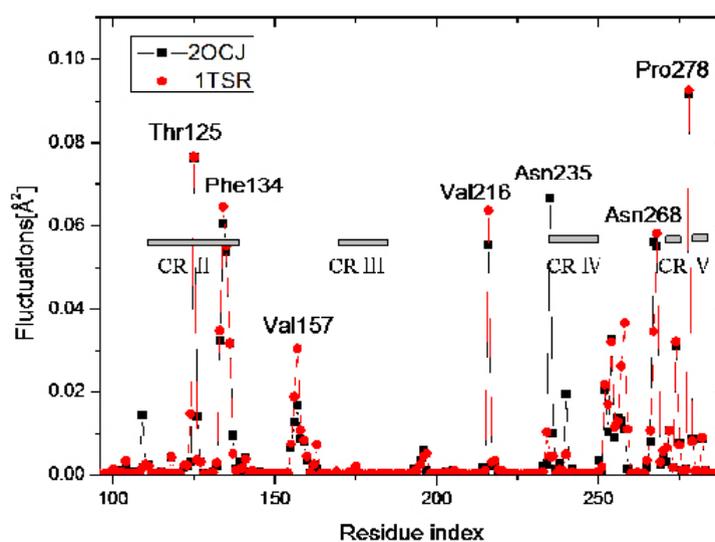


Fig. 4. Comparison between the fastest 10 mode shapes of the two structures

Cross-Correlations between atomic fluctuations

Similar to the analysis of Su *et al.*^{10, 15}, we explore the interresidue cross correlation of motions in all GNM modes of the two proteins. The cross-correlations between the fluctuations of residues are calculated with Equation 5. The cross-correlation value ranges from -1 to 1. The positive values represent that the motions of residues are in the same direction, and the negative values represent that they move in the opposite direction. The higher the absolute cross-correlation value is, the more the two residues are correlated (or anti-correlated). On the other hand, the cross-correlation

value $C_{ij} = 0$ means that the motions of residues are completely not correlated. Fig. 6 presents the cross-correlation maps of 1TSR. Similar cross-correlation maps are obtained for the isolated monomers A of 2OCJ. As shown in Figure 6, along the diagonal of the map, there are some light blocks with positive correlations, significantly, four regions of which correspond to the four conserved regions CR II, CR III, CR IV and CR V, respectively. The four CRs are indicated by a red circle. The Fig. 6 also shows that the high positive correlations between the CR II (residues 112-141) and CR V (residues 271-274 and residues 278-286). These regions have a direct contact with DNA major

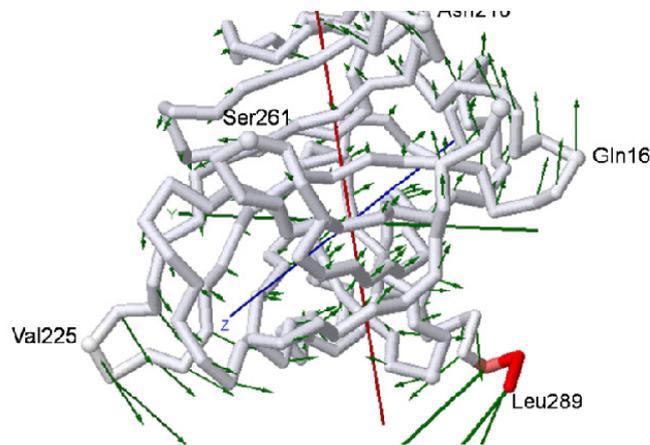


Fig. 5. The first slowest motive mode sketch maps of 1TSR

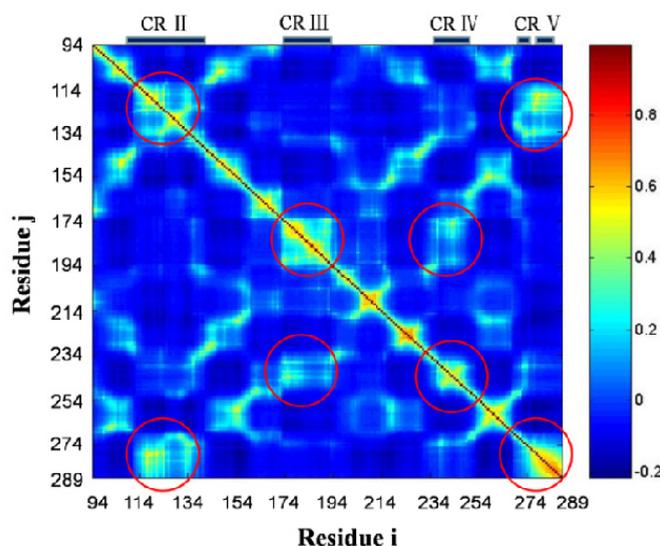


Fig. 6. Cross-correlation maps including the all slowest modes for 1TSR

groove. Similarly, the regions between CR III (residues 171-181) and CR IV (residue 236-251) have high positive correlation. As shown in section 2.1, the L2 and L3 loops in these regions ligate a zinc ion that together forms the p53DBD dimerization interface over the DNA minor groove. The result is consistent with previous study by Kantarci *et al.*¹⁷. The correlations in the map indicate that residues involved in similar functions are fluctuating in a cooperative manner.

CONCLUSION

In this work, simple coarse-grained methods are proposed to investigate the motion mode of p53DBD. The calculated B-factors in both the calculated models are in good agreement with the experimentally determined B-factors in X-ray crystal structures. The slowest mode analysis by GNM shows that the four CRs in the p53core domain have small fluctuation, but the same regions lay the peaks of the fast modes. The significant difference in two structures shows that the residues (Lys120, Ala280, Arg281 and Arg 283) make contacts with DNA reduce the fluctuation during the conformation changes, as well as the residues around them. With the aid of the anisotropy elastic network mode, we analyze the motion directions of this domain. The ANM calculation of the first slowest mode of the two structures mainly shows the rotational motion. However, the CR V (N-terminal) shows an opposite direction with the other parts of structure. Through this motion, p53DBD will bind to its consensus sequence DNA in a highly cooperative manner. By analyzing cross-correlations between residue fluctuations, the results indicate that residues involved in similar functions are fluctuating in a cooperative manner.

ACKNOWLEDGMENTS

We would like to thank JiGuo Su (Chinese Beijing University of Technology) for providing us with the GNM and ANM programs developed with Matlab. This work was financially supported by the Natural Science Foundation of China (31000324, 61271378) and the Shandong Provincial Natural Science Foundation (ZR2012CL09), which is gratefully acknowledged.

REFERENCES

1. Horn HF, Vousden KH. Coping with stress: multiple ways to activate p53. *Oncogene* 2007; **26**(9): 1306-16.
2. Malecka KA, Ho WC, Marmorstein R. Crystal structure of a p53 core tetramer bound to DNA. *Oncogene* 2009; **28**(3): 325-33.
3. Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* 1994; **265**(5170):346-55.
4. Kwon E, Kim DY, Suh SW, Kim KK. Crystal structure of the mouse p53 core domain in zinc-free state. *Proteins* 2008; **70**(1): 280-3.
5. Ho WC, Fitzgerald MX, Marmorstein R. Structure of the p53 core domain dimer bound to DNA. *J Biol Chem* 2006; **281**(29):20494-502.
6. Hupp TR, Lane DP. Allosteric activation of latent p53 tetramers. *Curr Biol* 1994; **4**(10):865-75.
7. Waterman JL, Shenk JL, Halazonetis TD. The dihedral symmetry of the p53 tetramerization domain mandates a conformational switch upon DNA binding. *EMBO J* 1995; **14**(3): 512-9.
8. Lu Q, Tan YH, Luo R. Molecular dynamics simulations of p53 DNA-binding domain. *J Phys Chem B* 2007; **111**(39):11538-45.
9. Khalili M, Wales DJ. Computer simulations of peptides from the p53 DNA binding domain. *Journal of Chemical Theory and Computation* 2009; **5**(5):1380-92.
10. Su JG, Jiao X, Sun TG, et al. Analysis of domain movements in glutamine-binding protein with simple models. *Biophys J* 2007; **92**(4):1326-35.
11. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J Mol Biol* 1999; **285**(3): 1023-37.
12. Hinsen K. Analysis of domain motions by approximate normal mode calculations. *Proteins* 1998; **33**(3):417-29.
13. Atilgan AR, Durell SR, Jernigan RL, et al. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 2001; **80**(1): 505-15.
14. Wang Y, Rosengarth A, Luecke H. Structure of the human p53 core domain in the absence of DNA. *Acta Crystallogr D Biol Crystallogr* 2007; **63**(Pt 3): 276-81.
15. Li HY, Cao ZX, Zhao LL, Wang JH. Analysis of Conformational Motions and Residue Fluctuations for *Escherichia coli* Ribose-Binding Protein Revealed with Elastic Network Models. *Int J Mol Sci* 2013; **14**(5):10552-69.

16. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* 1997; **2**(3): 173-81.
17. Kantarci N, Doruker P, Haliloglu T. Cooperative fluctuations point to the dimerization interface of p53 core domain. *Biophys J* 2006; **91**(2):421-32.
18. Bahar I, Jernigan RL. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms. *J Mol Biol* 1998; **281**(5): 871-84.