

## Motion Modes Analysis of Talin Rod Domain with Gaussian Network Model

Jiao Xiong<sup>1\*</sup> and Yang Lifeng<sup>2</sup>

<sup>1</sup>Institute of Applied Mechanics and Biomedical Engineering,  
Taiyuan University of Technology, Taiyuan - 030 024, China.

<sup>2</sup>College of Computer Science and Technology, College of Software,  
Taiyuan University of Technology, Taiyuan - 030 024, China.

(Received: 03 March 2013; accepted: 14 April 2013)

**Talin is an important linking between integrin and the actin cytoskeleton. In the process of signal transduction, talin play a critical role in the bidirectional signal transfer. In this paper, we analyze the motion modes of talin rod domain with the Gaussian network model. The results show that the two free terminals move in a negative correlated way, and this large-scale conformational change trend rest with the topological characteristics of the structure. For the peaks of the fast modes, it corresponds to some residues, and these residues are important to the stability of the structure. They may also act as key residues when they talin interacted with integrin or other signaling molecules. As an inherent feature of the structure, these motion modes are determined by the topology character of the structure. These results can provide a better understanding of the motion trend of talin in the process of signal transfer.**

**Key words:** Talin, Rod Domain, Motion Mode, Gaussian Network Model.

---

Talin is a ubiquitous cytosolic protein (Rees *et al.*, 1990; Gingras *et al.*, 2009; Critchley *et al.*, 2008). Its C-terminal rod domain contains bundles of alpha helices. The N-terminal FERM domain includes three sub domains: F1, F2, and F3. The F3 sub domain can interact with integrin, and the rod domain contains binding sites for vinculin. With these interactions, talin can link the integrin to the actin cytoskeleton either directly or indirectly by interacting with vinculin and alpha-actinin.

Talin plays an important role in mechanotransduction. From the extracellular matrix

to the transmembrane signal protein-integrin, and then to actin cytoskeleton, outer mechanical stimulation can be transmitted to biochemical signal. Some biochemical signaling pathways, activated <sup>1</sup> by mechanical stimuli, have been studied systematically.

However, the mechanisms of this process is still underway, and there are still many problems need to be studied. With some signal molecular structure data be solved, the molecular detail of the mechanotransduction process has now become a research hotspot.

The rod domain of talin provides an essential structure link between integrin and the actin cytoskeleton (Moes *et al.*, 2007). Several studies have examined the mechanical properties of other molecules with rod-like coiled-coil conformations. The results of these works suggest that the rod domain can contribute some

---

\* To whom all correspondence should be addressed.  
E-mail: jiaoxiong@tyut.edu.cn

conformational rigidity and flexibility to the molecular structure. For a linker between the integrin and actin cytoskeleton, less focus has been paid for the rod domain of talin.

In this work, the motion mode of the rod domain of talin is investigated from a novel point of view through a simple coarse-grained scheme-Gaussian network model (GNM) (Haliloglu *et al.*, 1997; Erman, 2006; Yang *et al.*, 2005). The GNM can investigate the large-scale conformation motions of the biomolecular systems, which does not require a high computational cost of MD. In numerous application studies, the GNM method has been proven to be a simple yet useful tool, and the results of GNM are in good agreement with that of MD simulation (Su *et al.*, 2007).

In this work, we choose the C-terminal dimerization rod domain of talin as our research object to investigate its large-scale conformation motion modes.



Fig. 1. The ribbon representation of the of talin rod domain

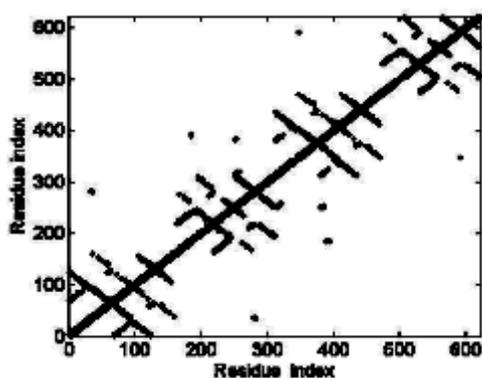


Fig. 2. The contact map of the rod domain

The rod domain of talin is an anti-parallel dimer. Figure 1 is the ribbon representation of the molecule. Each monomer contains 317 residues. There are ten helices in the chain A and nine helices in the chain B. Each of these helices can be divided into two groups. These two groups are linked by the longest helix of the monomer.

To show the structure of talin rod domain more clearly, we constructed the native contact map of its conformation, showed as Figure 2.

With the structure data got from protein data bank (PDB), we can calculate the distance between residues. Here, the carbon alpha is used to present residue. If the distance between two residues is less than a cut-off, then there will be a native contact between these two residues. In this work, we use the 7.3 angstrom is used as the cut-off.

Interestingly, as shown in Figure 2, the contact map presents an obvious regularity. Shown as in Figure 1, the second structure element of the rod domain is alpha-helix. So, the regular pattern of its structure is transferred to the regularity of its contact map. There are four helix bundles in the chain A and B, and each of them is composed with some alpha helices. The two bundles in the chain A is linked by a long helix. Due to the flexibility of this long helix, the bundle can retain its structural rigidity on a certain extent. The contact map is divided into four parts, corresponding to these four helix bundles.

## METHODS

In the Gaussian network model, the 3-D structure of protein is simplified to an elastic network, with the  $C_{\alpha}$  atoms as nodes. When the distance between two  $C_{\alpha}$  atoms are less than a certain cutoff (7.3 Å is adopted in this work), they will be connected by a harmonic springs link. The force constant of all springs is identical.

Considering all contacting residues, the internal Hamiltonian of the system can be written as

$$V = \frac{1}{2} \gamma [\Delta R^T (\Gamma \otimes E) \Delta R] \quad \dots(1)$$

where  $\gamma$  is the harmonic force constant; represents the 3N-dimensional column vector

of the X, Y, and Z components of the fluctuation vectors  $\Delta R_1, \Delta R_2, \dots, \Delta R_N$  of the  $C_\alpha$  atoms, N is the number of residues; the superscript T denotes the transpose; E is a third-order identity matrix;  $\otimes$  is the direct product; and  $\Gamma$  is the  $N \times N$  symmetric Kirchhoff matrix in which the elements are written as:

$$\Gamma = \begin{cases} -1 & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{k, k \neq i} \Gamma_k & \text{if } i = j \end{cases} \quad \dots(2)$$

where  $R_{ij}$  is the distance between the  $i$ th and  $j$ th  $C_\alpha$  atoms and  $r_c$  is the cutoff distance.

The mean-square fluctuations of each atom and the cross-correlation fluctuations between different atoms are in proportion to the diagonal and off-diagonal elements of the pseudo inverse of the Kirchhoff matrix. The inverse of the Kirchhoff matrix can be decomposed as

$$\Gamma^{-1} = U\Lambda^{-1}U^T \quad \dots(3)$$

where U is an orthogonal matrix whose columns  $u_i (1 \leq i \leq N)$  are the eigenvectors of  $\Gamma$ .

$\Lambda$  is a diagonal matrix of eigenvalues  $\lambda_i$  of  $\Gamma$ . The cross-correlation fluctuations between the  $i$ th and  $j$ th residues are given by

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ij} \quad \dots(4)$$

where  $k_B$  is Boltzmann constant,  $T$  is absolute temperature, and the meanings of  $\gamma$  and  $\Gamma$  are the same as that of Eq. (1). When  $i = j$ , the mean-square fluctuations of the  $i$ th residue can be obtained.

The mean-square fluctuation of the  $i$ th residue associating with the  $k$ th mode is given by

$$\langle \Delta R_i \cdot \Delta R_i \rangle_k = \frac{3k_B T}{\gamma} \lambda_k^{-1} [u_k]_i [u_k]_i \quad \dots(5)$$

The B-factor is correlated to the mean-square fluctuation. The calculation of the B-factor can be showed as the following expression:

$$B_i = 8\pi^2 \langle \Delta R_i \cdot \Delta R_i \rangle / 3 \quad \dots(6)$$

## RESULTS AND DISCUSSION

### Comparison of B-factors

In order to do an assessment of the feasibility of using the GNM method to calculate the rod domain of talin, we compared the B-factors got from GNM with the data got from X-ray in the experiment.

According to the principle of GNM, we calculate the B-factor of the rod domain of talin.

Figure 3 shows the comparisons between the calculated B-factor with the experimental data from X-ray crystallography.

In Figure 3, the calculated B-factor is presented with solid lines, marked with a label Bfactor-js. The data from the X-ray experiment is shown as a dotted line, and marked with a label Bfactor-sy.

The correlation coefficient between the experimental B-factor and the calculated one is about 0.5. This result is similar to those of the recent studies for other proteins.

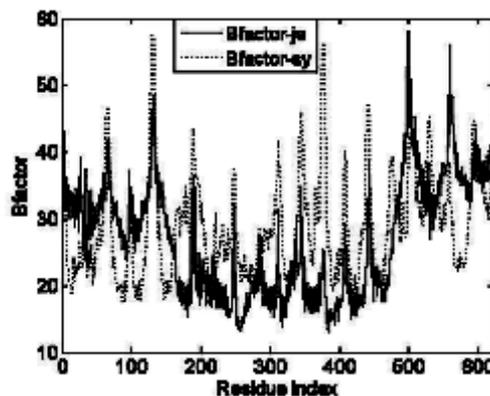


Fig. 3. Comparison between the calculated B-factor (Bfactor-js, showed as solid lines) with the experimental one (Bfactor-sy, showed as dotted lines)

### The slow modes of the motions

The slow and long wavelength collective modes represent functionally relevant motions of protein.

Figure 4 displays the first mode of the dimerization rod domain got from the GNM. The ordinates in Figure 4 show the normalized distribution of squared fluctuations driven by the first slowest modes.

From Figure 1, we can see the two monomers from a shape similar to V. The chain A, B

all have a free end, with the remained two terminals are adjacent. In Figure 4, the two monomers of the dimer can be divided into three groups. The first group is the N-terminal of the chain A, the third group is the C-terminal of the chain B. For the C-terminal of the chain A and the N-terminal of the chain B is spatially close, so they formed the second group with low fluctuation.

From Figure 4, it is obvious that the fluctuation of the C-terminal of the chain B is bigger than that of the N-terminal of the chain A.

The structural basis of this phenomenon is that the chain A is more curved than the chain B, so the N-terminal of the chain A will be closer to the other part of the molecule, and will have more interaction with other residues.

In contrast with the chain A, C-terminal of the chain B is freer, and there are fewer links between the C-terminal of the chain B with other residue.

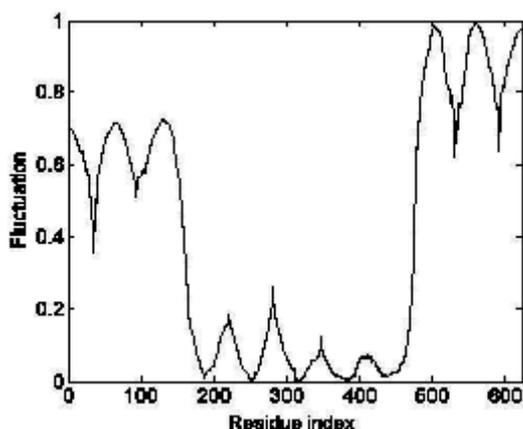


Fig. 4. The slowest mode shape of the structure for talin

The second group, including the C-terminal of the chain A and the N-terminal of the chain B, has a low fluctuation. These two parts of the second group are much closed spatially, so there are many interactions between them, and they nearly formed a rigid body that has a very low fluctuation.

#### The fast modes of the motions

The fast modes correspond to geometric irregularity in the local structure, and the fluctuations associated with these modes are accompanied by a larger decrease in entropy than that of slow modes.

So, residues acting in the fast modes are thought as kinetically hot residues, and they are critically important for the stability of the tertiary fold.

Figure 5 shows the fastest six modes of the structures. There are several peaks in the curve that correspond to the kinetically hot residues, which may play a key role in the stability of the protein.

For the chain A, there are ten  $\alpha$ -helices. As the longest helix in the chain A, the helix 6 contains 55 amino acids, with residue index from 134 to 188. So, residues acting in the fast modes are thought as kinetically hot residues, and they are critically important for the stability of the tertiary fold. So, residues acting in the fast modes are thought as kinetically hot residues, and they are critically important for the stability of the tertiary fold.

Figure 5 shows the fastest six modes of the structures. There are several peaks in the curve that correspond to the kinetically hot residues, which may play a key role in the stability of the protein.

For the chain A, there are ten  $\alpha$ -helices. As the longest helix in the chain A, the helix 6 contains 55 amino acids, with residue index from 134 to 188.

For the chain B, there are nine helices. The longest helix (helix 15) contains 54 residues, with residue index from 451 to 504. The helix 6 is a link between the N-terminal and the c-terminal of the chain A, so is the helix 15.

It is interesting to find that few peak appeared in the two sections corresponded to the helix 6 and helix 15.

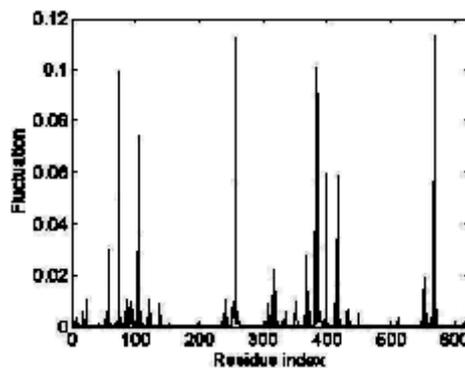


Fig. 5. The fastest six mode shapes of the structure

This showed that these two helixes are used mainly as a link to build the structure of the protein, less functional key residues is located in these parts.

It is also found that these hot residues, almost all of them are belonged to some helixes other than some random coil, are tightly packed. A large part of them is located in an exposed region.

It has been proposed that there are some integrin binding sites on these two chains. So these hot residues, with big fluctuation in their fast motion modes, may be involved in the interactions with integrin.

#### Cross-correlations between atomic fluctuations

The cross-correlations between the fluctuations of  $C_{\alpha}$  atoms are calculated with Eq. (4).

Since the modes with low frequency correspond to functional motions, and the modes with high frequency correspond to localized motions, here only low frequency modes are used to improve the signal-to-noise ratio. We used the first 40 modes in our calculations. The results are shown in Figure 6.

The cross-correlation value ranges from -1 to 1, in which the positive values represent that the residues move in the same direction, and the negative values represent that they move in the opposite direction. As shown in the color bar, we use red color to present a positive cross-correlation value, and we use blue to show a negative value. The higher the absolute cross-correlation value, the well the two residues correlated (or anti-correlated).

As shown in Figure 6, there is negative correlation in the dark regions and positive correlation in the light regions. It is obvious that the cross-correlation map of this dimmer can be divided into nine areas. There are four light portions anti-diagonal. These four parts correspond to the four groups of the dimmer, which are the N-terminal and the C-terminal of two chains.

The corners along the diagonal are two dark portions. The negative correlation of this portion indicates that the relative motion between the N-terminal of the chain A and the C-terminal of the chain B is negatively correlated, or they are moving in an opposite direction.

Between the light and dark portions are

some gray portions, corresponding to the cross-correlation of the relative movements between the four groups.

In the four light portions, corresponding to the movement in the inner of each group, some light band appeared. Each light band is corresponding to a helix in the dimmer.

It is notable that the cross-correlation of the longest helix (helix 6) in the chain A, residue index from 134 to 188, is divided into two sections. From the structure data, we can find that this helix can be parted as two sections. One is belonged to the N-terminal with residues from 134 to 164, and the other is belonged to the C-terminal with residues 168 to 188.

The cross-correlation in these two sections is positive, but the cross-correlation between these two sections is changed to be negative, as marked out by a rectangle in the figure.

This phenomenon shows that the residues within the two sections move in a highly coupled way, and they will keep their structures stable during movements.

The two sections will move in an opposite direction, so the helix will be bent as shown in the structure representation (Figure 1).

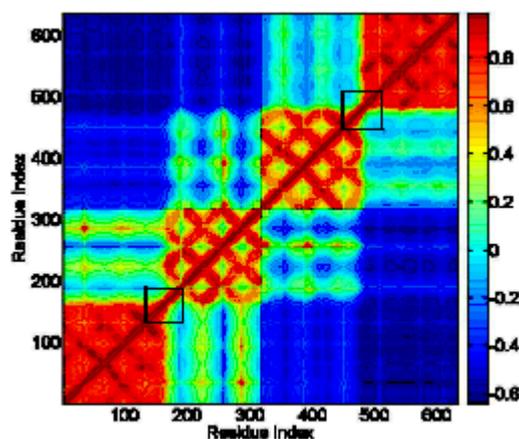


Fig. 6. The cross-correlation maps calculated using the first 40 modes of the structure

The situation of the helix 16, the longest helix in the chain B, is similar to that of the helix 6, also be marked out by a rectangle in the Figure 6.

## CONCLUSIONS

The rod domain is an essential structural

link between integrin and the actin cytoskeleton (Wegener *et al.*, 2007; Calderwood *et al.*, 2002; Calderwood *et al.*, 1999). Its motion modes are investigated in this work. The results showed that the protein can be divided into three groups with the N-terminal of the chain A as a group and the C-terminal of the chain B as a group, the C-terminal of the chain A and the N-terminal of the chain B from the other group. The two free terminals move in an opposite direction. This result is consistent with the view that the general shape instead of the local interactions plays an important role in low-frequency motions (Lu *et al.*, 2005). The peaks in the fastest six modes are the most packed residues which may be involved in the interactions with integrin.

#### ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 31070828), China Postdoctoral Science Foundation funded project (Grant No. 2012T50247, 20100471587), Program for the Innovative Talents of Higher Learning Institutions of Shanxi, and Natural Science Foundation of Shanxi (Grant No. 2009021018-2). An earlier version of this paper was presented at the International Conference on BMEI 2010.

#### REFERENCES

1. Calderwood D. A., Yan B., De Pereda J. M., Alvarez B. G., Fujioka Y., Liddington R. C., and Ginsberg M. H., The Phosphotyrosine Binding-like Domain of Talin Activates Integrins. *Journal of Biological Chemistry*. 2002; **277**: 21749-21758.
2. Calderwood D. A., Zent R., Grant R., Rees D. J. G., Hynes R. O., and Ginsberg M. H., The Talin Head Domain Binds to Integrin  $\beta$  Subunit Cytoplasmic Tails and Regulates Integrin Activation. *Journal of Biological Chemistry*. 1999; **274**: 28071-28074.
3. Critchley D. R. and Gingras A. R., Talin at a glance. *Journal of cell science*. 2008; **121**: 1345-1347.
4. Erman B., The Gaussian Network Model: Precise Predictions of Residue Fluctuations and Application to Binding Problems. *Biophysical journal*. 2006; **91**: 3589-3599.
5. Gingras A. R., Ziegler W. H., Bobkov A. A., Joyce M. G., Fasci D., Himmel M., Rothmund S., Ritter A., Grossmann J. G., and Patel B., Structural Determinants of Integrin Binding to the Talin Rod. *Journal of Biological Chemistry*. 2009; **284**: 8866-8876.
6. Haliloglu T., Bahar I. and Erman B., Gaussian Dynamics of Folded Proteins. *Physical review letters*. 1997; **79**: 3090-3093.
7. Lu M. and Ma J., The Role of Shape in Determining Molecular Motions. *Biophysical journal*. 2005; **89**: 2395-2401.
8. Moes M., Rodius S., Coleman S. J., Monkley S. J., Goormaghtigh E., Tremuth L., Kox C., van der Holst P. P. G., Critchley D. R., and Kieffer N., The Integrin Binding Site 2 (IBS2) in the Talin Rod Domain Is Essential for Linking Integrin  $\beta$  Subunits to the Cytoskeleton. *Journal of Biological Chemistry*. 2007; **282**: 17280-17288.
9. Rees D. J., Ades S. E., Singer S. J. and Hynes R. O., Sequence and domain structure of talin. *Nature*. 1990; **347**: 685-689.
10. Su J. G., Jiao X., Sun T. G., Li C. H., Chen W. Z., and Wang C. X., Analysis of domain movements in glutamine-binding protein with simple models. *Biophysical journal*. 2007; **92**: 1326-1335.
11. Wegener K. L., Partridge A. W., Han J., Pickford A. R., Liddington R. C., Ginsberg M. H., and Campbell I. D., Structural basis of integrin activation by talin. *Cell*. 2007; **128**: 171-182.
12. Yang L. W., Liu X., Jursa C. J., Holliman M., Rader A. J., Karimi H. A., and Bahar I., iGNM: a database of protein functional motions based on Gaussian Network Model. *Bioinformatics*. 2005; **21**: 2978-2987.