Flexible docking of the fragment of the troponin I to the troponin C

Jacek Wabik[1], Mateusz Kurciński[1] and Andrzej Kolinski[1*]

[1]Faculty of Chemistry, University of Warsaw, L. Pasteura 1, 02-093 Warsaw, Poland
{jwabik,mkurc,kolinski}@chem.uw.edu.pl

Abstract. Most of the current docking procedures are focused on fine conformational adjustment of assembled complexes and fail to reproduce large-scale motions. Our approach based on the coarse-grained modeling of flexible macromolecules overcomes these limitations. CABSDock procedure used in this work employs coarse-grained model CABS – efficient and versatile tool for modeling of proteins structure, dynamics and interactions [1-4].

In this work CABSDock was used to model assembly process of troponin C(TnC) with the N-terminal helix of the troponin I(TnI). TnC/TnI binding was investigated for both cardiac and skeletal troponin. TnI fragment was modeled allowing its unbiased movement. Entire structure of TnC domain was also treated as fully flexible, although its motion was restricted to near-native conformations. Binding of the TnI fragment changed the orientation between both domains. This picture provides valuable insight into mechanistic description of troponin function.

Keywords: flexible docking, troponin, CABS, coarse-grained modeling

1 Introduction

Protein-protein interactions are involved in great number of essential life processes. Experimental techniques encounter significant difficulties in obtaining information about protein interactions at atomic level. At the same time uprising theoretical models provide valuable aid in understanding this matter. Modeling of complex formation between proteins may explain functions of various cofactors, repressors and inhibitors, which alter behavior of their interacting partners [5, 6]. Investigation of protein complexes is also important for understanding the origins and mechanisms of many diseases e.g. Alzheimer's [7].

Typical protein complex consists of couple of thousands atoms, meaning tens of thousands degrees of freedom. Efficient algorithm for modeling of protein interactions must face a problem of sampling such a large conformational space. Traditional rigid docking methods sample only six-dimensional conformational space, referring to translation and rotation of one molecule (ligand) relative to the other (receptor). For obvious reasons such approach cannot account for any conformational changes, which both receptor and ligand molecules may undergo upon binding. More advanced approach was proposed for a number of "semi-flexible" methods. They challenge the

docking problem by allowing some level of flexibility for only part of the modeled system – either one of the molecules (smaller) or selected patches on the surface of the receptor (usually the bigger molecule). Such formulated approach requires some additional data about location or size of the binding size to define the flexible region [8].

In case of modeling protein-peptide complexes one needs to take backbone movement into account. It is especially important when the peptide induces structural transitions in the backbone of the receptor. Such movements are accounted for in FiberDock [9], ATTRACT [10] and Rosetta 'fold-and-dock' [11]. Backbone mobility may be achieved in various way [12]. For example, ATTRACT employs a coarse-grained model with Normal Mode Analysis, whereas Rosetta 'fold-and-dock' algorithm uses experimental data: nuclear magnetic resonance shifts and residual dipolar couplings.

CABSDock is a docking protocol based on the coarse-grained model CABS [13]. CABS has already been evaluated in de-novo structure prediction [1-3], homology modeling [4] and simulations of the dynamics of single proteins[1, 2]. It combines the computational efficiency of reduced models with realistic accuracy. CABSDock's unique feature is the ability to model proteins with variable level of flexibility – from stiff to completely flexible. At the same time very efficient sampling scheme of the CABS model enables modeling of large molecular systems undergoing significant conformational transitions. CABSDock has already been tested on a group of protein complexes [5, 14] and proved to be very effective in protein-peptide docking.

In our study it is used to simulate docking process of a flexible ligand to the receptor with fully flexible linkers between domains. This approach could be used not only for prediction of native conformations, but also to investigate a mechanism of folding-binding process through analysis of transition states. We investigated two systems: complex of the fragment of the skeletal troponin I and troponin C (pdb code: 1A2X [15]) and the cardiac complex of troponin C, I and T (pdb code:1J1D [16]).

Troponin is a complex of three proteins – troponin C (TnC), troponin I (TnI) and troponin T (TnT). It plays crucial role in the regulation of skeletal and cardiac muscle contraction [17]. Troponin binds at regular intervals along the actin polymer on the thin filament. TnT is responsible for binding the whole complex to the tropomyosin molecule. TnC forms the core of the complex – it binds to TnI with variable strength dependent on the Ca^{2+} concentration.

Main sites of TnC/TnI interaction are located in the hydrophobic N- and C-lobe of the TnC. TnC binds with high affinity two calcium ions. Another two calcium ions are bound at low affinity rate, triggering N-lobe exposure. During muscle contraction and relaxation cycle the Ca^{2+} concentration changes, which in turn triggers conformational transition in the regulatory N-domain of the TnC. When Ca^{2+} concentration is low, hydrophobic pocket on the N-lobe remains closed, blocking this interaction site for binding of the TnI. In this conformational state, the troponin prevents from the approach of the myosin heads towards the actin [15]. At high Ca^{2+} concentration, the N-lobe is accessible for TnI binding [18] which may be also co-induced by cardiac TnI molecule [19].

The second place of interaction between TnC and TnI is located in the hydrophobic pocket of the structural C-domain. TnI/TnC complex on this end is stable independently on the concentration of the calcium ions in sarcoplasm. Nevertheless, this interaction of TnC and TnI is important, because it is partly responsible for anchoring troponin to tropomyosin [20]. Mutations, which weaken these interactions can cause disorders in contraction and relaxation cycle of the muscles [20, 21].

N-terminal helix of the TnI interacts not only with C-domain of TnC, but also partly with the N-domain of TnC. TnC domains (Fig.1) are connected by a flexible linker [15, 22] and they change their relative positions in the process of complex formation. Movement of such scale could be impossible to capture by rigid or semi-flexible docking methods.

2 Methods

CABS [13] is a lattice model, which force field employs knowledge-based statistical potentials. It employs reduced representation of proteins: each amino acid is replaced by up to 3 pseudoatoms: Cα, Cβ and center of mass of the sidechain. Exceptions are glycine, with only the Cα atom and alanine, which is represented by the Cα and the Cβ. Additional pseudoatom is placed in the center of the virtual Cα-Cα bond for proper definition of the hydrogen bonds. Each Cα atom occupies a nod of the lattice with the grid size arbitrary set to 0.61Å. Sequential moves are sampled according to Monte Carlo scheme and controlled by the energy function calculated from knowledge-based potentials.

In simulations replica exchange Monte Carlo (REMC) technique was utilized. Number of replicas was optimized to twenty, as well as the temperature range (1.5-3.4 in CABS units; both temperature and energy in the CABS model are dimensionless. T=1 corresponds to the temperature of the crystal). Each simulation was repeated three times with different seeds for pseudo-random number generator.

We performed docking on two systems: fragment of the skeletal troponin (PDB: 1A2X) and complex of the cardiac troponin C, I and T (PDB: 1J1D), as shown in the Figure 1. Initially TnI was randomly placed around the troponin C.

In the first system we applied distance restraints on pairs of atoms within each domain of troponin C. No cross-domain constraints were used. The loops between domains and the TnI were fully flexible.

In the second system constraints were applied in a similar manner. No constraints across domains of the TnC were used. Furthermore, N-terminal fragment of the TnI (residues: 37-80), the loop between domains of the TnC and all loops of the TnI were modeled with full flexibility. TnT and the fragment of the C-terminal TnI in the N-lobe were constrained.

Because structure of the troponin I is incomplete (PDB:1J1D), missing loop (res. F138-F149) had to be modeled before simulation. It was done on the FALC-Loop web server [23].

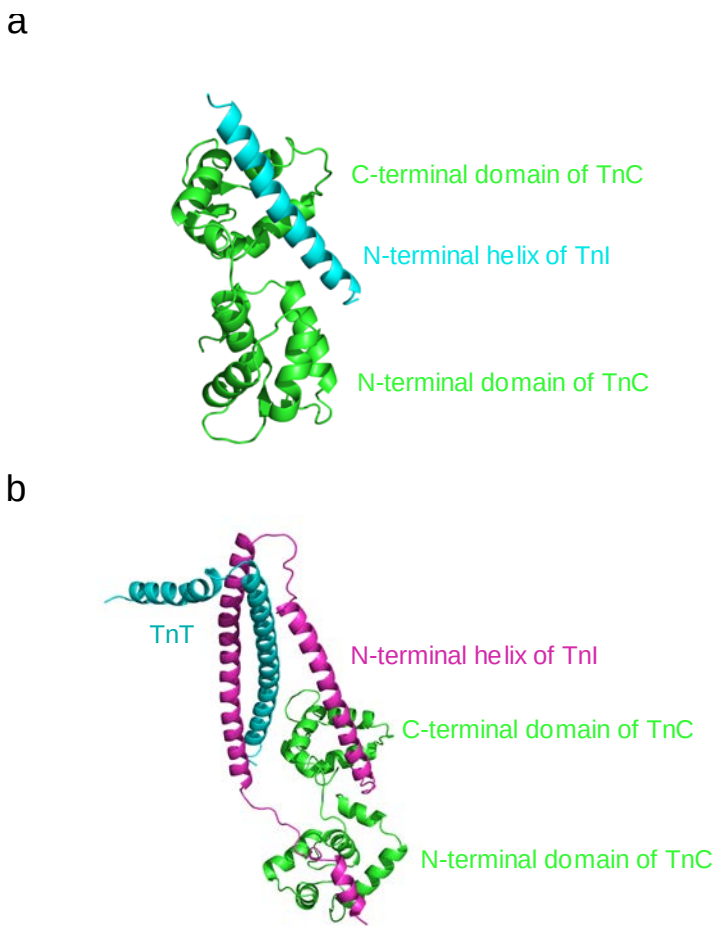

Fig. 1. PDB structures of a) 1A2X b)1J1D with rebuilt loop.

3 Results and discussion

Plots in the Figure 2 show fraction of near-native conformations present in the trajectories for each of the 20 replicas, averaged over three independent simulations. Sigmoidal shape of the transition curves indicates cooperative character of the modeled process.

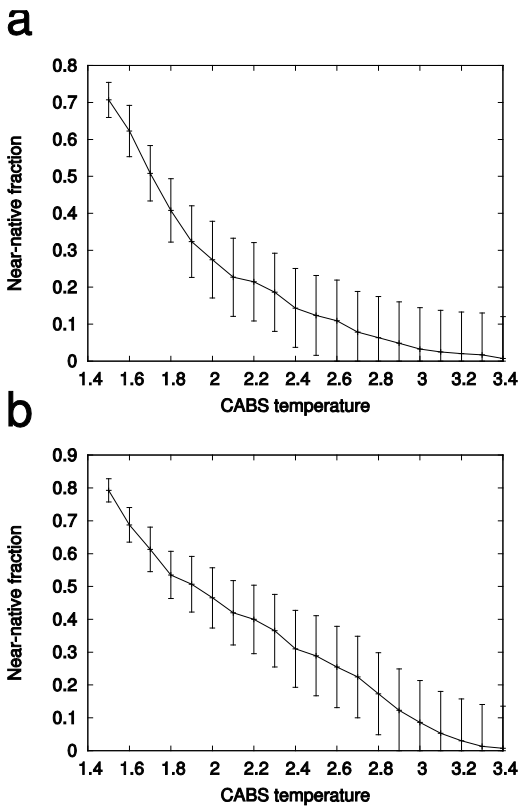

Fig. 2. Transition curve for a) fragment of the skeletal TnC with fragment of the TnI (pdb code: 1A2X) b) complex of troponin C,I and T (pdb code: 1J1D). Conformations for which RMSD was below 6Å were considered near-native. RMSD was calculated on Cα atoms of the troponin C complex with N-terminal helix of troponin I vs the crystallographic structures.

Trajectory of the lowest temperature replica contained ca. 70% of near-native conformations, for the fragment of the skeletal troponin (Fig. 2a). For this replica, map of conformation population was prepared (Fig. 3a). We introduced two measures: RMSDdomains (RMSD calculated on the Cα atoms of both domains of the troponin C vs crystallographic structure) and RMSDbound (RMSD calculated on the Cα atoms of only the C-terminal domain of the TnC and the N-terminal helix of the TnI: residues 3-33 for 1A2X and residues 37-80 for 1J1D). The plot shows how the orientation of the ligand in the C-terminal domain can impact relative position of both domains. It is especially evident in the skeletal troponin (Fig. 3a). At low calcium concentration linker between domains of skeletal troponin C is flexible, therefore allowing for broader range of domain shifts, when compared to the complete complex of cardiac troponin.

There are three clusters of points in the skeletal troponin map (Fig. 3a). RMSDbound fluctuates around 2.5A for most of the conformations close to the native state. This cluster contains structures of the ligand (TnI) properly docked to the receptor (open C-lobe of the TnC).

Second cluster at RMSDbound ~ 5Å and RMSDdomains ~ 2.5Å corresponds to the TnI shifting inside the hydrophobic lobe. It induces a change of the relative position of domains. Third cluster reflects the transition state (RMSDbound ~ 13Å) - the ligand is only partially bound and protrudes from the hydrophobic pocket. By interacting with the N-terminal domain it stabilizes the arrangement of TnC domains (RMSDdomains ~ 3.5Å).

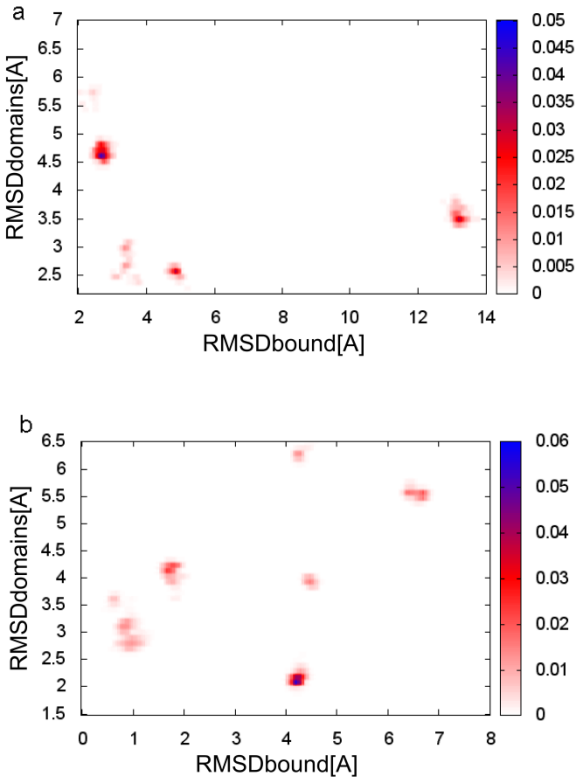

Fig. 3. Conformational density map plotted as a function of RMSD to the native state (at T=1.5) for a) fragment of the skeletal troponin C with the fragment of the troponin I (PDB:1A2X) b) complex of the troponin C, I and T (PDB: 1J1D).

Relative position of domains in the cardiac troponin C also depends on the interaction with troponin I. It is less mobile however, compared to the fragment of the skeletal troponin complex. Trajectory of the lowest temperature replica contained ca. 80%

of the near-native structures (Fig. 2b). Most populated cluster is characterized by proper relative position of both domains (RMSDdomains ~ 2Å) and slight shift from the native state in position of the ligand (RMSDbound ~ 4.1Å).

Second biggest cluster consists of conformations that are most similar to the native state and have RMSDbound around 1Å. That confirms that the fragment of the TnI was docked to the hydrophobic C-lobe the same way as it is in the native state. Correct position of the ligand implicates also correct relative position of TnC domains (RMSDdomains ~ 3Å).

Representative conformations of these two biggest clusters are presented in the Figure 4. One can notice that residues from the N-terminal part of the TnI interact with the N-terminal residues of the TnC. These interactions may alter the positions of both domains. Residues of the loop L2 of the TnI also interact with the TnC and have impact on the whole TnC/TnI complex, which agrees with the results of Varughese et al. [24].

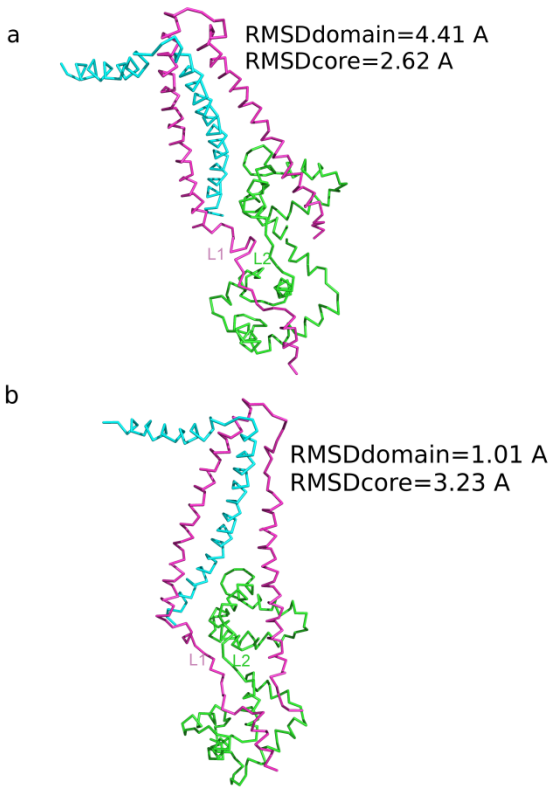

Fig. 4. Representative conformations of the two biggest clusters, at the lowest temperature, from the simulation of the cardiac troponin.

4 Conclusions

We present a method which can be used for efficient modeling of protein docking and folding. In the result of simulations for both systems (1A2X, 1J1D) cluster of near-native conformations is obtained. It is compared with available, crystallographic structures.[15, 16] Since for both systems linker between N-terminal and C-terminal domain is flexible [15, 22] analysis revealed significant oscillation around native states. The relative position of TnC domains is highly dependent on position of N-terminal TnI helix. Nevertheless, application of our method goes further than only description of near-native modeling. It can be utilized to analysis of transition states of complexes, which conformations of ligand and receptor differ from native state.

Therefore, our docking method takes into account also non-native contacts, which is a big advantage comparing to other flexible docking methodologies.

In the past CABS model has proven to be efficient tool for studies of protein folding mechanisms [2]. Here we show that it is also a good tool for predicting mechanism of protein-protein binding. When properly combined with all-atom refinement, CABS-based approach can be efficient tool for multiscale modeling of protein-protein interactions.

Acknowledgements. The authors acknowledge the support from the Foundation for Polish Science TEAM project (TEAM/2011-7/6) cofinanced by the European Regional Development Fund operated within the Innovative Economy Operational Program. Calculations were performed in the Interdisciplinary Center for Mathematical and Computational Modeling (ICM) of the University of Warsaw (Grant G43-9).

References

1. Kmiecik, S., Gront, D., Kouza, M., Kolinski, A.: From Coarse-Grained to Atomic-Level Characterization of Protein Dynamics: Transition State for the Folding of B Domain of Protein A. J Phys Chem B (2012)
2. Kmiecik, S., Kolinski, A.: Folding pathway of the B1 domain of protein G explored by multiscale modeling. Biophys. J. 94, 726-736 (2008)
3. Wabik, J., Kmiecik, S., Gront, D., Kouza, M., Koliński, A.: Combining Coarse-Grained Protein Models with Replica-Exchange All-Atom Molecular Dynamics. Int. J. Mol. Sci. 14, 9893-9905 (2013)
4. Kolinski, A., Bujnicki, J.M.: Generalized protein structure prediction based on combination of fold-recognition with de novo folding and evaluation of models. Proteins 61, 84-90 (2005)

5. Kurcinski, M., Kolinski, A.: Theoretical study of molecular mechanism of binding TRAP220 coactivator to Retinoid X Receptor alpha, activated by 9-cis retinoic acid. J. Steroid Biochem. Mol. Biol. 121, 124-129 (2010)
6. Frembgen-Kesner, T., Elcock, A.H.: Absolute Protein-Protein Association Rate Constants from Flexible, Coarse-Grained Brownian Dynamics Simulations: The Role of Intermolecular Hydrodynamic Interactions in Barnase-Barstar Association. Biophys. J. 99, L75-L77 (2010)
7. Strodel, B., Lee, J.W.L., Whittleston, C.S., Wales, D.J.: Transmembrane Structures for Alzheimer's Aβ1−42 Oligomers. J. Am. Chem. Soc. 132, 13300-13312 (2010)
8. Schneidman-Duhovny, D., Nussinov, R., Wolfson, H.J.: Predicting molecular interactions in silico: II. Protein-protein and protein-drug docking. Curr. Med. Chem. 11, 91-107 (2004)
9. Mashiach, E., Nussinov, R., Wolfson, H.J.: FiberDock: Flexible induced-fit backbone refinement in molecular docking. Proteins 78, 1503-1519 (2010)
10. May, A., Zacharias, M.: Energy minimization in low-frequency normal modes to efficiently allow for global flexibility during systematic protein-protein docking. Proteins 70, 794-809 (2008)
11. Das, R., Andre, I., Shen, Y., Wu, Y., Lemak, A., Bansal, S., Arrowsmith, C.H., Szyperski, T., Baker, D.: Simultaneous prediction of protein folding and docking at high resolution. Proc. Natl. Acad. Sci. USA 106, 18978-18983 (2009)
12. Tuffery, P., Derreumaux, P.: Flexibility and binding affinity in protein–ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches. J. R. Soc. Interface 9, 20-33 (2012)
13. Kolinski, A.: Protein modeling and structure prediction with a reduced representation. Acta Biochim. Pol. 51, 349-371 (2004)
14. Kurcinski, M., Kolinski, A.: Hierarchical modeling of protein interactions. J. Mol. Model. 13, 691-698 (2007)
15. Vassylyev, D.G., Takeda, S., Wakatsuki, S., Maeda, K., Maeda, Y.: Crystal structure of troponin C in complex with troponin I fragment at 2.3-A resolution. Proc. Natl. Acad. Sci. USA 95, 4847-4852 (1998)
16. Takeda, S., Yamashita, A., Maeda, K., Maeda, Y.: Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. Nature 424, 35-41 (2003)
17. Metzger, J.M., Westfall, M.V.: Covalent and noncovalent modification of thin filament action: the essential role of troponin in cardiac muscle regulation. Circ Res 94, 146-158 (2004)
18. Vinogradova, M.V., Stone, D.B., Malanina, G.G., Karatzaferi, C., Cooke, R., Mendelson, R.A., Fletterick, R.J.: Ca2+-regulated structural changes in troponin. Proc. Natl. Acad. Sci. USA 102, 5038-5043 (2005)
19. Sia, S.K., Li, M.X., Spyracopoulos, L., Gagné, S.M., Liu, W., Putkey, J.A., Sykes, B.D.: Structure of cardiac muscle troponin C unexpectedly reveals a closed regulatory domain. J. Biol. Chem. 272, 18216-18221 (1997)
20. Baryshnikova, O.K., Robertson, I.M., Mercier, P., Sykes, B.D.: The Dilated Cardiomyopathy G 159D Mutation in Cardiac Troponin C Weakens the Anchoring Interaction with Troponin I†. Biochemistry 47, 10950-10960 (2008)
21. Baryshnikova, O.K., Robertson, I.M., Mercier, P., Sykes, B.D.: The dilated cardiomyopathy G159D mutation in cardiac troponin C weakens the anchoring interaction with troponin I. Biochemistry 47, 10950-10960 (2008)
22. Abbott, M.B., Gaponenko, V., Abusamhadneh, E., Finley, N., Li, G., Dvoretsky, A., Rance, M., Solaro, R.J., Rosevear, P.R.: Regulatory domain conformational exchange and

linker region flexibility in cardiac troponin C bound to cardiac troponin I. J. Biol. Chem. 275, 20610-20617 (2000)
23. Ko, J., Lee, D., Park, H., Coutsias, E.A., Lee, J., Seok, C.: The FALC-Loop web server for protein loop modeling. Nucleic Acids Res 39, W210-214 (2011)
24. Varughese, J.F., Chalovich, J.M., Li, Y.: Molecular dynamics studies on troponin (TnI-TnT-TnC) complexes: insight into the regulation of muscle contraction. J. Biomol. Struct. Dyn. 28, 159-174 (2010)