http://dx.doi.org/10.5012/bkcs.2012.33.11.3597

Potential of Mean Force Simulation by Pulling a DNA Aptamer in Complex with Thrombin

Changwon Yang, Eunae Kim,[†,‡] and Youngshang Pak[*]

Department of Chemistry and Institute of Functional Materials, Pusan National University, Busan 609-735, Korea
[*]E-mail: ypak@pusan.ac.kr
[†]BK-21 Project Team, College of Pharmacy, Chosun University, Gwangju 501-759, Korea
[‡]Supercomputing Center, Korea Institute of Science and Technology Information, Daejeon 305-806, Korea
Received July 15, 2012, Accepted August 7, 2012

Thrombin binding aptamter (TBA-15) is a 15-mer guanine-rich oligonucleotide. This DNA apamer specifically binds to the thrombin protein involved in blood coagulation. Using extensive umbrella sampling molecular dynamics simulation method at all atom level, we investigated the potential of mean force (PMF) upon pulling the DNA aptamer from the binding mode of aptamer/thrombin complex. From this calculation, the free energy cost for a full dissociation of this aptamer/protein complex is 17 kcal/mol, indicating a substantial binding affinity of TBA-15. Interestingly, this PMF reveals noticeable plateau regions along the pulling coordinate. Possible structural changes of this complex in the plateau were investigated in details.

Key Words : TBA-15, Thrombin binding aptamer, Umbralla sampling, Potential of mean force

Introduction

A single stranded guanine (G)-rich oligonucleotide is of fundamental importance, since this G-rich strand constitutes an overhang of telomeres by adopting G-quadruplexes *via* non-canonical base pair interactions.[1] Each quadruplex displays a planar structure by intramolecular stacking of the four-guanine bases. This G-quadruplex structure was known to inhibit enzymatic activity of telomerase. DNA aptamers, which are also made of a single stranded DNA segment, can fold into their stable three-dimensional structure of the G-stem. Due to their highly specific biological activities, DNA aptamers have been used for drug designs and for bio-medical applications.

Among many DNA aptamers, the thrombin binding aptamer (TBA-15) is one of the well-known oligonucleotides with a sequence of GGTTGGTGTGGTTGG. In particular, TBA-15 exhibits a highly specific binding affinity for the serine protease thrombin. As a G-quadruplex-based aptamer, TBA-15 is composed of two planar G-quartets connected by two intervening TT and one TGT loops. In particular, the G-quadrants in TBA-15 are stabilized by the cation binding in its central binding site located between the two-quadruplex planes. In general, two binding sites of exosite I and exosite II were identified in thrombin.[2] The exosite I is related to the fibrinogen-recognition site and the exosite II is known to be the heparin binding site. In TBA-15, the two lateral TT loops are associated with the binding site (exosite I) of thrombin and the TGT loop at the center of the sequence is involved with another binding site (exosite II) of the thrombin. Thus, TBA-15 molecules can bind to both binding sites of a single thrombin molecule. Because of the importance of TBA-15 for potential application of bio-medical applications, several experimental studies were carried out to probe many out-standing issues of the TBA-15 and other aptamer sequences.[3-6]

Furthermore, extensive molecular dynamics (MD) simulations have been performed to investigate structural and dynamic aspect of the 15-TBA in the presence of monovalent metal cations.[7-10] Although previous MD simulation studies helped to understand the detailed structural changes of the G-stem and the role of the intervening loops upon cation binding, its binding characteristics with the thrombin has not been studied. In this study, we endeavored to compute the potential of mean force (PMF) upon pulling the aptamer molecule from the binding mode (exosite I) of the TBA:Thrombin complex. This PMF can allow us to gain some physical insight into detailed energetics and structural changes as a result of the TBA binding to the thrombin protein.

Method

The binding mode of TBA-15:Thrombin was obtained from the protein data bank (PDB entry 1HAO)[2] (See Figure 1). For the calculation of PMF profile upon pulling the aptamer segment from the binding mode, the pulling coordinate ξ was defined as a bond distance between two ends, such that one end is the center of mass (COM) of TBA-15 and the other end is the COM of the protein residues (Glu77-Lys81) enclosed by the TT-loops of the aptamer. At the central binding site of the G-quadruplexes, a single K^+ ion was inserted. The amber03* force field[11] was used for the thrombin as the protein part and the param99bsc0 force field[12] was used for the aptamer as the DNA part. The present simulation system consists of the aptamer/thrombin complex, which is solvated by 23858 TIP3P water molecules.[13] The total charge of this system is −6. Thus, 51 Na^+ ions and 46 Cl^- ions were inserted to the system, which then maintains a 100 mM salt concentration. For this system, a

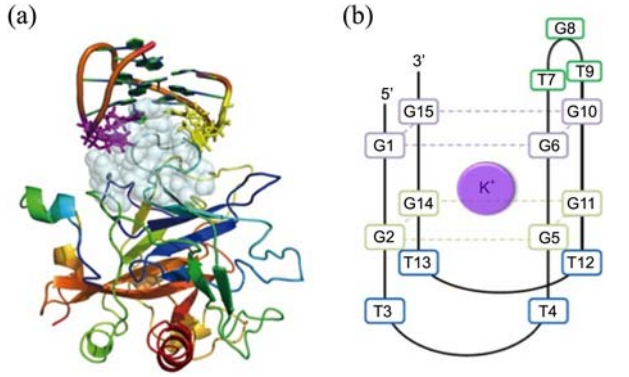

Figure 1. (a) The binding mode of TBA-15 with thrombin. This structure was taken from the protein data bank (PDB entry 1HA0). The upper part is the TBA-15 molecule as the cartoon representation. In the TT loops of the aptamer, T3 and T4 are colored by yellow and T12 and T13 are colored by magenta. The thrombin molecule is represented by a ribbon model. The van der Waal's model represents the hydrophobic pocket of thrombin (ILE24, HIS71, ILE79, TYR117), the polar residues of thrombin (Arg77A, Arg75, Asn78, and Glu77). (b) The schematic figure in the right is the TBA-15 structure.

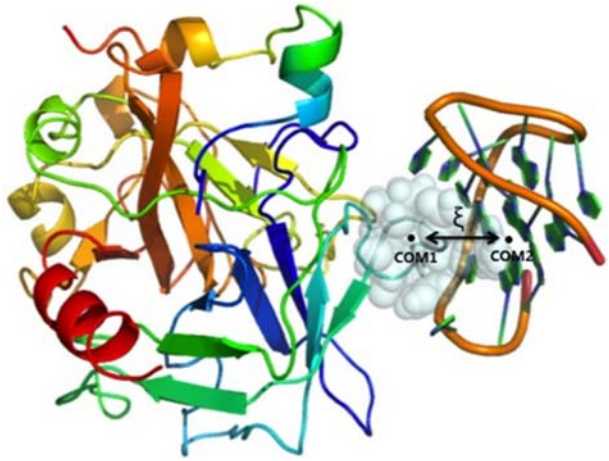

Figure 2. The reaction coordinate used for the PMF simulation. The reaction coordinate is defined to be a distance between the DNA and the protein segment. One end is the center of mass (COM) of TBA-15 and the other end is the COM of the protein residues region enclosed by the TT loops of the aptamer.

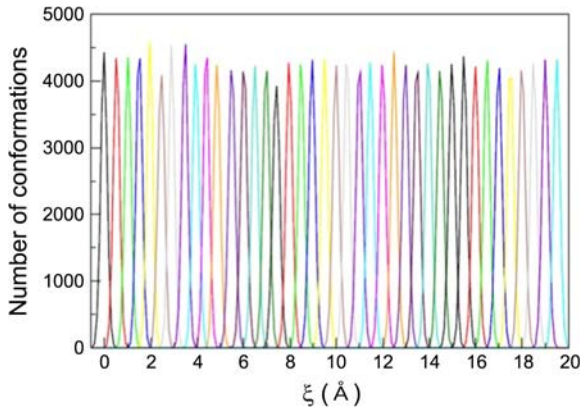

Figure 3. The histogram of 40 umbrella windows. The first 5 ns data were discarded for producing the umbrella distribution.

simulation box of 80 × 80 × 120 Å is needed. The protonation states of Asp, Glu, and His in the protein part were employed by referring to the experimental result.[14] For the PMF calculations, a total of 40 umbrella windows were used with the restraining force constant of 30 kcal/mol/Å^2. For each window, the MD simulations were performed for 20 ns at 300 K using the modified Berendsen method of Bussi *et al.*[15] We discarded the first 5 ns trajectory for the PMF calculation. The particle mesh Ewald (PME) method was employed with a cutoff distance of 10 Å. From the trajectories of 40 windows, the PMF was calculated using the weighed histogram analysis method (WHAM),[16] and its error estimation was performed by the Bayesian bootstrap analysis of the raw simulation data. For the current simulation, the GROMACS 4.5.1 program was used.[17]

Result and Discussion

Figure 2 illustrates the reaction coordinate for the umbrella windows. Figure 3 shows the umbrella histogram collected for these windows. A total of 40 windows cover from the native complex state (ξ = 0.0 Å) to a fully dissociation state (ξ = 20.0 Å). Here, the x-axis is the reaction coordinate and the y-axis is the total number of the conformations sampled at each window. The conformational distributions of adjacent windows are well overlapped and smooth.

The resulting PMF as a function of the pulling distance is given in Figure 4(a). The histogram data were used for the construction of the PMF using WHAM. In comparison, the corresponding mean forces are given in Figure 4(b) after taking the derivative of the PMF with ξ. As shown in Figure 4(a), the full dissociation free energy of this complex is 17 kcal/mol, indicating a significant binding affinity of TBA-15 with the thrombin. This value is somewhat larger than the experimental binding affinity of 9 kcal/mol. However, this deviation can be well expected, since our predicted value is the free energy cost for a full separation of the DNA from the protein substrate, while the experimental value[18] was measured based on the relative populations of the complex and the non-complex forms without considering energetic and structural differences between two binding sites (exosites I and II). Thus, in terms of dissociation energetics, a direct comparison between theory and experiment is not yet feasible. As shown in Figure 4(a), the PMF profile increases as the DNA segment is gradually displaced from the binding pocket. One interesting point is the existence of some flat PMF regions ($0 \le \xi \le 2$ Å and $5 \le \xi \le 7$ Å) along the reaction coordinate. As can be seen in Figure 4(b), in those plateau regions, the net mean forces are fluctuating around zero. Before the DNA is entirely removed from the protein substrate, the corresponding mean force profile displays quite a complicated peak pattern, indicating notable structural changes in atomistic level.

In an attempt to obtain more insight into detailed changes of the protein/DNA interactions, we traced the simulation

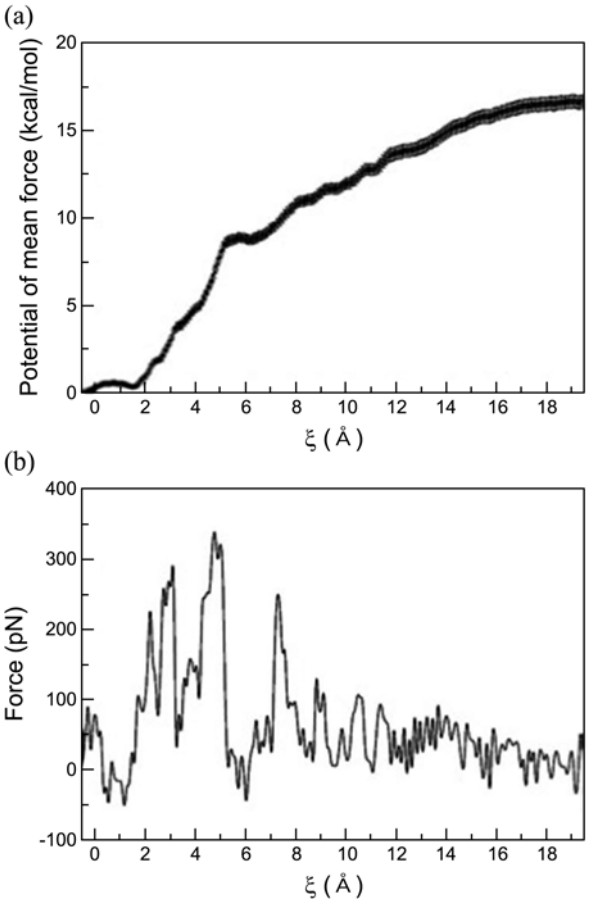

Figure 4. (a) The potential of mean force along the reaction coordinate. (b) The corresponding mean force, which is obtained by taking the first order derivative with respect to the reaction coordinate.

trajectories of all the windows. In the native complex structure, Arg75 interacts with T4, T12, and T13 in TBA-15 *via* hydrogen bonds and Arg77A interacts with T13. In addition, T3 is in a direct contact with Tyr76 *via* a face-to-face stacking interaction and T12 is near the hydrophobic pocket of thrombin (ILE24, HIS71, ILE79, TYR117). As shown in Figure 4(a), in the first plateau region, one of the notable changes is the disruption of the hydrogen bond (HB) between T12 and Arg75. As the aptamer center is further displaced up to $\xi = 5$ Å, one can see another plateau, where the T12 base gets loose by escaping from the aforementioned hydrophobic pocket and only the HBs between T13 and Arg75/Arg77A are observed. In a substantially pulled state ($\xi = 8$ Å), TBA-15 is about to leave the protein substrate. In this extension, both Arg75 and Arg77A lost their HBs with T4 and T12 bases. Instead, they find a way to interact with the DNA backbone by way of overall rotation of the aptamer part. For an illustration of observed structural changes during the PMF simulation, representative structures along the reaction coordinate are given in Figure 5.

Conclusion

TBA-15 is one of the simplest G-quaduplex. Due to its biological activities, this has been a subject of extensive researches for its potential applications in bio-medical sciences of many disciplines. In particular, TBA-15 can specifically bind to the thrombin protein involved in the blood coagulation.

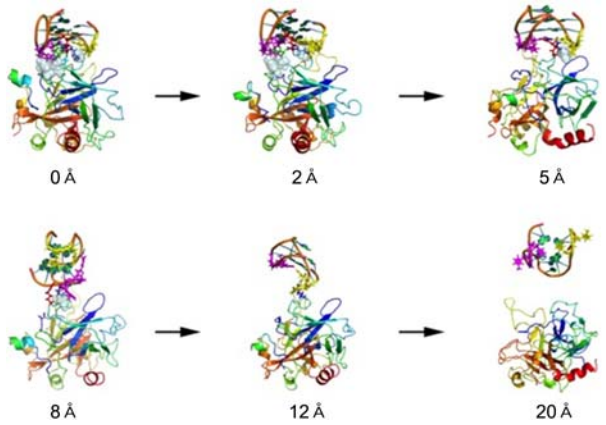

Figure 5. The representative snapshots along the reaction coordinate. The cartoon representation is the same as that of Figure 1(a). The transparent sphere model represents the binding residues of thrombin.

In order to provide molecular insights into the aptamer pulling process from such binding mode of the complex, the PMF profile was computed using an all atom level umbrella sampling molecular dynamics simulation method. The simulated PMF profile enables us to determine the free energy cost for a complete dissociation of the DNA/protein complex. According to the PMF simulation, the free energy landscape for detaching the aptamer from the binding mode displays at least two major plateau regions along the reaction coordinate describing an extent of pulling.

Although this would be an interesting phenomenon, the molecular origins of these plateau regions have not yet been revealed by this work due to its inherent complexity of analyzing the results. In the fibrinogen recognition site, the TT loops tightly grip the protein residues *via* hydrophobic interactions. The other notable interactions are the hydrogen bonds between the thymine residues of the TT loops and Arg75/Arg77A. In general, the Arg residues are relatively long and flexible in the binding site of the protein, so they can display much diverse interactions with the departing DNA segment from the protein substrate. In the native bound state, in addition to the aforementioned hydrophobic interactions, these arginine residues stabilize the complex *via* several hydrogen bonds with the TT loops of the DNA. As the DNA segment is gradually removed from the protein binding sites, Arg75 and Arg77A lose their connection with the TT loops due to disruption of hydrogen bonds, but they find a way to interact with the DNA backbone part. This adjustment event sometimes accompanies an overall rotation of the aptamer molecule.

In this work, we investigated detailed interaction pictures of the complex of TBA-15 and thrombin by pulling the DNA part from complex using the MD simulation. How-

ever, there are several issues to be further clarified. One imminent issue is the choice of the pulling end. We chose the COM of the total DNA as a pulling end, but other choices may be possible. Thus, how much the PMF is affected by a different choice of pulling end will be an interesting subject in near future. The other concern is the validity of the force field. In principle, the mean force from the PMF calculation can be directly compared to the force profile from atomic force microscopy (AFM). Unfortunately, the AFM force-extension profile has not been available for this complex. Once the AFM force profile is available, the accuracy of force field can be tested and further improved by direct comparisons.

Acknowledgments. This work was supported by a 2-Year Research Grant of Pusan National University.

References

1. Mergny, J. L.; Helene, C. *Nat. Med.* **1998**, *4*, 1366.
2. Padmanabhan, K.; Tulinsky, A. *Acta Crystallogr. D* **1996**, *52*, 272.
3. Kankia, B. I.; Marky, L. A. *J. Am. Chem. Soc.* **2001**, *123*, 10799.
4. Mao, X..; Marky, L. A.; Gmeiner, W. H. *J. Biomol. Struct. Dyn.* **2004**, *22*, 25.
5. Olsen, C. M.; Gmeiner, W. H.; Marky, L. A. *J. Phys. Chem. B* **2006**, *110*, 6962.
6. Marathias, V. M. B.; Marathias, V. M.; Bolton, P. H. *Nucleic. Acids Res.* **2000**, *28*, 1969.
7. Reshetnikov, R.; Golovin, A.; Spiridonova, V.; Kopylov, A. *J. Chem. Theory Comput.* **2010**, *6*, 3003.
8. Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E.; Kulinski, T.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 2514.
9. Reshetnikov, R. V.; Sponer, J.; Rassokhina, O. I.; Kopylov, A. M.; Tsvetkov, P. O.; Makarov, A. A.; Golovin, A. V. *Nucleic. Acids Res.* **2011**, *39*, 9789.
10. Yang, C.; Jang, S.; Pak, Y. *J. Chem. Phys.* **2011**, *135*.
11. Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004.
12. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817.
13. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
14. Ahmed, H. U.; Blakeley, M. P.; Cianci, M.; Cruickshank, D. W. J.; Hubbard, J. A.; Helliwell, J. R. *Acta Crystallogr. D* **2007**, *63*, 906.
15. Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*.
16. Hub, J. S.; de Groot, B. L.; van der Spoel, D. *J. Chem. Theory Comput.* **2010**, *6*, 3713.
17. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Comput. Chem.* **2008**, *4*, 435.
18. Pagano, B.; Martino, L.; Randazzo, A.; Giancola, C. *Biophys J.* **2008**, *94*, 562.