Supporting Information

Study of the Binding Mechanism between Aptamer GO18-T-d and Gonyautoxin 1/4 by Molecular Simulation

Shunxiang Gao[a, §], Bo Hu[a, b, §], Xin Zheng[c, §], Dejing Liu[a], Mingjuan Sun[a], Jiaxiang Qin[d], Hao Zhou[e], Binghua Jiao[a, b,*], Lianghua Wang[a,*]

[a]*Department of Biochemistry and Molecular Biology, College of Basic Medical Sciences, Second Military Medical University, Shanghai 200433, China.*

[b]*Marine Biological Institute, College of Marine Military Medicine, Second Military Medical University, Shanghai 200433, China.*

[c]*Department of Laboratory Diagnosis, Changhai hospital, Second Military Medical University, Shanghai 200433, China.*

[d]*Chengdu FenDi Technology Co., Ltd.*

[e]*Key Laboratory of Industrial Ecology and Environmental Engineering (Ministry of Education), School of Food and Environment, Dalian University of Technology.*

[§]These authors contributed equally to this work.

[*]Correspondence authors: Binghua Jiao, Lianghua Wang.

Department of Biochemistry and Molecular Biology, College of Basic Medical Sciences, Second Military Medical University, Shanghai 200433, China

E-mail: bhjiao@smmu.edu.cn (B.H. Jiao), lhwang@smmu.edu.cn (L.H. Wang)

Materials and methods

1. Materials and Reagents

All nucleic acid sequences were custom-synthesized by Sangon Biotech (Shanghai, China), and GTX1/4 aptamer GO18-T-d: 5′-Biotin-AACCTTTGGTCGGG CAAGGTAGGTT-3′. GTX1/4, GTX2/3, STX and neoSTX were obtained from Taiwan Algal Science, Inc. (Taiwan). Binding buffer (pH 7.5, 20 mM Tris-HCl, 100 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl) were procured from Tiandz (Beijing, China) and was used for biolayer interferometry (BLI), and circular dichroism (CD) experiments.

2. Biolayer interferometry assay

The affinity and specificity of aptamer GO18-T-d were determined by BLI using an OctetRED 96 system (ForteBio, Shanghai). The principle and analysis procedures used herein were as detailed in Concepcion et al.[1] As shown in Fig. S2, the assay process includes five steps: (1) baseline (2 min); (2) loading (3 min); (3) washing (2 min); (4) association (4 min); (5) baseline (4 min). The response data obtained from the reaction surface were normalized by subtracting the signal simultaneously acquired from the reference surface to eliminate nonspecific binding and buffer-induced interferometry spectrum shift using the Octet Data Analysis Software CFR Part 11 Version 6.x; the affinity parameter K_d was then obtained. A 1:1 binding mode with mass transfer fitting was used to obtain the kinetic data.

3. Circular dichroism assay

The structural conformation of DNA aptamer before and after toxin binding was investigated using J-715 circular dichroism spectropolarimeter. The concentration of 2 μM aptamer GO18-T-d, and the spectrum was measured with different concentration of GTX1/4 in a 1 cm path length, quartz cuvette in an optical chamber. Background signals of binding buffer and 2 μM GTX1/4 in binding buffer were measured and subtracted from the CD spectra. The chamber was deoxygenated with dry purified nitrogen (99.99%) before use and kept in the nitrogen atmosphere during experiments. All CD spectrums was collected from 230 to 320 nm at 0.1 nm intervals, the accumulation of two scans at 20 nm/min, with a 1 nm bandwidth and a time constant of 1 s.

4. Generation of G-quadruplex structure for the Aptamer GO18-T-d

The structure prediction of GO18-T-d was performed using QGRS (quadruplex-forming G-rich sequences, http://bioinformatics.ramapo.edu/QGRS/ index.php). The minimum G-Group Size was set as 2. The results of QGRS prediction are shown in Table S1. Notably, the G-score of GO18-T-d was 20, indicating that GO18-T-d had a high probability of adopting a G-quadruplex structure.

The DNA sequences of G-quadruplex structure extracted for each chain in the corresponding Nucleic Acid Database (NDB)[2] and Protein Data Bank (PDB)[3] files were aligned to GO18-T-d. The structure 2HY9[4] was found to correspond to intrastrand G-quadruplexes with similar size. In addition, the arrangement of guanine

in 2HY9 was somewhat similar to GO18-T-d. Based on the atomic models of 2HY9, a 3D model of GO18-T-d with G-quadruplex structure was generated by Discovery Studio2.5 Client[5] through nucleic acid substitution, insertion and deletion.

Care was taken to make the coordination geometry as favorable as possible. The conformations of a few nucleic acids were therefore adjusted manually within well allowed ranges. Then, the model was optimized at the high-performance computing facility with the YASARA package,[6, 7] version 13.12 using the Amber99 force field[8] and the TIP3P water model. The temperature coupling of the model system was ascertained by the Berendsen thermostat method, while the manometer method was used for pressure coupling. Moreover, the starting structure was immersed in a periodic rectangular simulation cubic cell of water. For optimization in the simulated water condition, the backbone was first fixed and the side chain optimized for 5,000 steps, and then, the whole structure was optimized for 5,000 steps. Finally, the 3D model of GO18-T-d with G-quadruplex structure was generated (Fig. S3). After optimization, quantitative analysis was performed on the G-quadruplex structure of GO18-T-d. The results (Fig. S4) showed that due to the strong hydrogen bond interactions among G8, G12, G19 and G22 in GO18-T-d, which were the same as for G9, G13, G18 and G23; GO18-T-d formed a stable G-quadruplex structure.

5. Molecular docking of GO18-T-d aptamer with gonyautoxin group

For aptamer-ligand docking, the 3D structure of GO18-T-d was obtained from the model and initialized as receptor molecules with AutodockTools. Subsequently,

GO18-T-d was endowed with AD atomic type, and hydrogens and charges were added, followed by the mergence of nonpolar hydrogen. The binding sites of GO18-T-d recognizing the ligands were obtained based on crystal structures of DNA-ligand complexes with G-quadruplex structures,[9-12] which showed that the binding sites of DNA with a G-quadruplex structure were in the grooves at the top and the bottom. To our knowledge, GO18-T-d with G-quadruplex structure shows no groove at the bottom. In addition, the experiment suggested that the ligand could induce the formation of a G-quadruplex structure. Therefore, the binding site of GO18-T-d with the G-quadruplex structure was at the top (Fig. S5). The grid parameter file was built using AutoGrid 4.0. Simultaneously, the molecular conformations of GTX1/4, GTX2/3, STX and neoSTX were drawn by PubChem Compound database with CID 46173840, 11593018, 37165 and 104753, respectively. The small molecule was optimized using the MM2 method. Then, the number of rotatable bonds of ligands was set in AutodockTools.

The molecular docking analysis was then conducted using AutoDock 4.0. The number of Lamarckian genetic algorithm[13] runs and population size were set to 10 and 150, respectively. The maximum number of evaluations was set at 2,500,000, while the remaining default parameter settings were retained. For each of the docking cases, the lowest energy conformation, according to the AutoDock scoring function, was selected as the binding mode. The output from AutoDock was rendered with Chimera and PyMol.[14, 15]

6. Molecular dynamics simulations

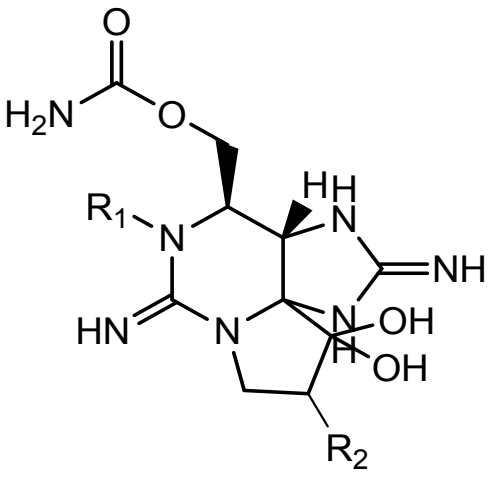
The simulations conducted were based on the complex obtained from docking in the AutoDock program. All simulations were performed using the molecular dynamics program YASARA V13.12 and the amber99 force field. The complex used in the simulation came from molecular docking with hydrogen generated by the YASARA program. All simulations were conducted with an integration step of 1 fs, and the coordinates of the simulation model were recorded every 1 ps. The starting structures were immersed in a periodic rectangular simulation cubic cell of water. The box dimensions were chosen to provide at least a 10 Å buffer of solvent molecules around the solute. To neutralize the charges of the systems, appropriate Na^+ ions were added to each system.

The fully solvated systems were then subjected to 5000 step steepest descent minimization runs to remove clashes between atoms. An 80 ps position-restrained MD simulation was performed for each system at constant pressure (1 atm) and temperature (300 K). The temperature and pressure were kept constant during the simulations. Temperature coupling was performed using the Berendsen thermostat with a temperature coupling constant of 0.1 ps, while the manometer method was used for pressure coupling with a reference pressure of 1 atm. A particle mesh Ewald scheme was used to calculate the long range electrostatic interactions, with a 10 Å cutoff for the real space.[16, 17] A cutoff of 14 Å was used for the van der Waals interactions (Lennard-Jones terms).Translation and rotation corrections were enabled during MD simulations to ensure that the structures in the trajectory were well

superimposed, which is convenient for structural analysis. The chemical bond lengths involving hydrogen atoms were fixed using the SHAKE algorithm[9]. By 1 ns, the simulated system had reached an equilibrium state; thus, the system was subjected to conventional MD (CMD) simulation for 20 ns.

Supporting Results

1. Chemical structures of gonyautoxin group



Substance	R_1	R_2
GTX1/4	OH	OSO_3^-
GTX2/3	H	OSO_3^-
STX	H	H
neoSTX	OH	H

Fig. S1 Chemical structures of GTX1/4, GTX2/3, STX, and neoSTX.

2. BLI technology assay process

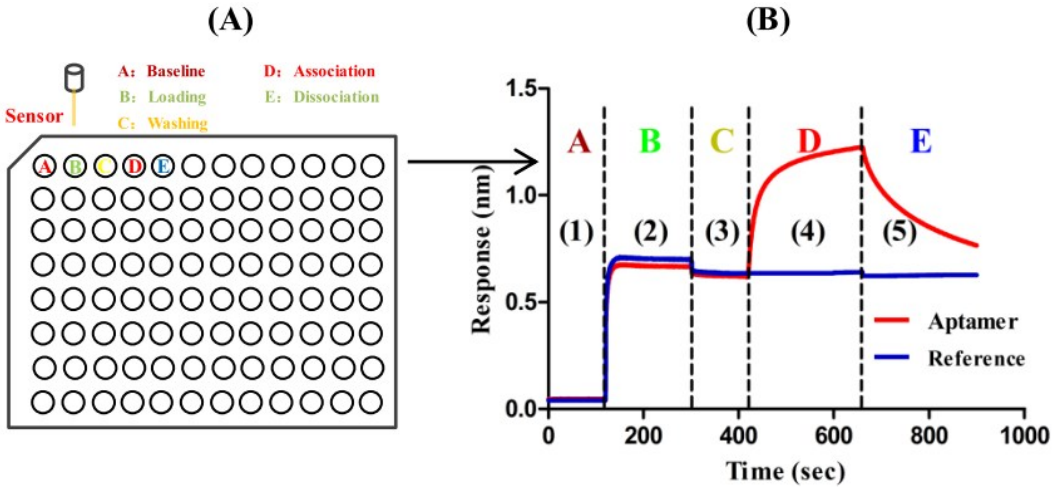


Fig. S2 BLI assay process includes five steps: (1) baseline (2 min); (2) loading (3 min); (3) washing (2 min); (4) association (4 min); (5) baseline (4 min). A reference sensor is always required as a control in every assay.

3. G-quadruplex structure prediction of the aptamer GO18-T-d by QGRS

Table S1. Results of QGRS prediction

Position	Length	QGRS	G-Score
8	16	GGTCGGGCAAGGTAGG	19
8	16	GGTCGGGCAAGGTAGG	20

4. G-quadruplex structure of the aptamer GO18-T-d

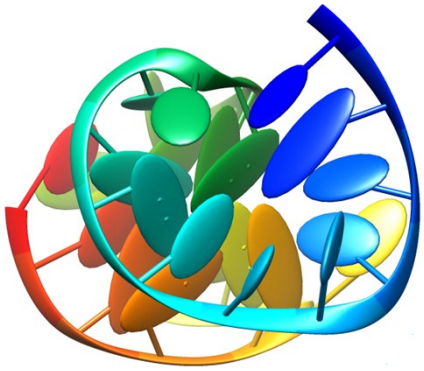


Fig. S3 The 3D model of GO18-T-d with G-quadruplex structure

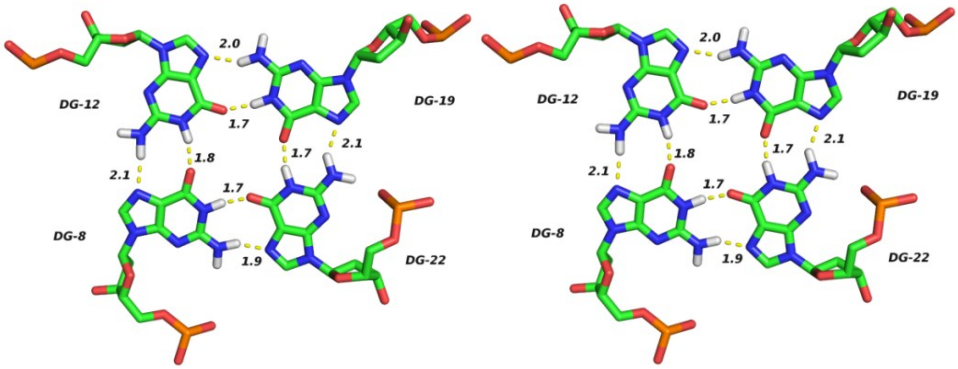


Fig. S4 The G-quadruplex structure of GO18-T-d and interaction

5. Binding site of the aptamer GO18-T-d

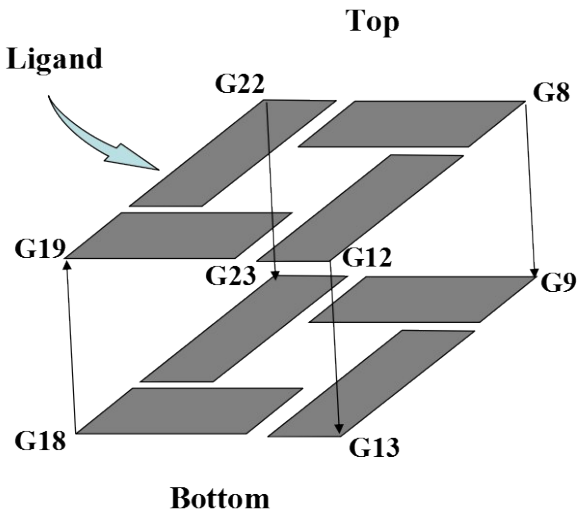


Fig. S5 Binding site of GO18-T-d with G-quadruplex structure

6. Docking results between aptamer G-quadruplex structure and ligands

Table S2. Docking results between aptamer GO18-T-d and ligands

Ligands	lowest binding energy kcal/mol	mean binding energy kcal/mol
GTX1/4	-8.56	-8.37
GTX2/3	-6.09	-5.96
STX	-5.77	-5.67
neoSTX	-5.58	-5.43

Tab S3. Descriptive statistics of total energy between GO18-T-d and GTX1/4

Data	Ntotal	Mean	Standard Deviation	Minimum	Median	Maximum
	247	-97848.78984	240.44601	-98531.724	-97830.022	-97166.936

References

1. J. Concepcion, K. Witte, C. Wartchow, S. Choo, D. F. Yao, H. Persson, J. Wei, P. Li, B. Heidecker, W. L. Ma, R. Varma, L. S. Zhao, D. Perillat, G. Carricato, M. Recknor, K. Du, H. Ho, T. Ellis, J. Gamez, M. Howes, J. Phi-Wilson, S. Lockard, R. Zuk and H. Tan, *Combinatorial chemistry & high throughput screening*, 2009, **12**, 791-800.
2. H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan and B. Schneider, *Biophysical journal*, 1992, **63**, 751-759.
3. H. M. Berman, T. N. Bhat, P. E. Bourne, Z. Feng, G. Gilliland, H. Weissig and J. Westbrook, *Nature structural biology*, 2000, **7 Suppl**, 957-959.
4. J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones and D. Yang, *Nucleic acids research*, 2007, **35**, 2440-2450.
5. B. I. Kankia and L. A. Marky, *Journal of the American Chemical Society*, 2001, **123**, 10799-10804.
6. E. Krieger, G. Koraimann and G. Vriend, *Proteins*, 2002, **47**, 393-402.
7. E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein and G. Vriend, *Proteins*, 2004, **57**, 678-683.
8. P. K. Athyala, J. R. Kanwar, M. Alameen, R. K. Kanwar, S. Krishnakumar, J. Watson, U. Vetrivel and J. Narayanan, *Biochemical and biophysical research communications*, 2015, DOI: 10.1016/j.bbrc.2015.11.109.
9. A. Kettani, S. Bouaziz, W. Wang, R. A. Jones and D. J. Patel, *Nature structural biology*, 1997, **4**, 382-389.
10. G. N. Parkinson, F. Cuenca and S. Neidle, *Journal of molecular biology*, 2008, **381**, 1145-1156.
11. A. T. Phan, V. Kuryavyi, H. Y. Gaw and D. J. Patel, *Nature chemical biology*, 2005, **1**, 167-173.
12. C. Bazzicalupi, M. Ferraroni, A. R. Bilia, F. Scheggi and P. Gratteri, *Nucleic acids research*, 2013, **41**, 632-638.
13. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *Journal of computational chemistry*, 2009, **30**, 2785-2791.
14. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *Journal of computational chemistry*, 2004, **25**, 1605-1612.
15. D. Seeliger and B. L. de Groot, *Journal of computer-aided molecular design*, 2010, **24**, 417-422.
16. T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1993, **98**, 10089.
17. D. S. Cerutti, R. E. Duke, T. A. Darden and T. P. Lybrand, *Journal of chemical theory and computation*, 2009, **5**, 2322.