ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/261569757

Unfolding and Conformational Variations of Thrombin-Binding DNA Aptamers: Synthesis, Circular Dichroism and Molecular Dynamics Simulations

ARTICLE *in* CHEMMEDCHEM · MAY 2014

Impact Factor: 2.97 · DOI: 10.1002/cmdc.201300564

CITATIONS
2

READS
95

8 AUTHORS, INCLUDING:

Han Jin
Peking University
33 PUBLICATIONS **229** CITATIONS

SEE PROFILE

Yi-fu Guan
China Medical University (PRC)
52 PUBLICATIONS **493** CITATIONS

SEE PROFILE

Zhenjun Yang
Peking University
62 PUBLICATIONS **558** CITATIONS

SEE PROFILE

Liangren Zhang
Peking University
146 PUBLICATIONS **1,446** CITATIONS

SEE PROFILE

Available from: Zhenjun Yang
Retrieved on: 13 January 2016

ChemPubSoc Europe

DOI: 10.1002/cmdc.201300564

Unfolding and Conformational Variations of Thrombin-Binding DNA Aptamers: Synthesis, Circular Dichroism and Molecular Dynamics Simulations

Lidan Sun,[a] Hongwei Jin,[a] Xiaoyang Zhao,[b] Zhenming Liu,[a] Yifu Guan,[b] Zhenjun Yang,[a] Liangren Zhang,*[a] and Lihe Zhang[a]

Thrombin-binding DNA aptamer (TBA), with a consensus 15-base sequence: d(GGTTGGTGTGGTTGG), can fold into an antiparallel unimolecular G-quadruplex structure that is necessary for its interaction with thrombin. For the first time, using steered molecular dynamics (SMD) simulations, we have successfully simulated the unfolding process of native TBA G-quadruplex. The unfolding pathway proposed is in agreement with previously reported experimental NMR data. Moreover, the critical intermediate structure in the unfolding pathway, predicted by the NMR results, was identified. The structural characteristics of several TBA oligonucleotides modified with locked nucleoside (LNA) or 2′-*O*-methyl-nucleoside (MNA) at different positions and number were also investigated by CD spectroscopy. An oligonucleotide substituted with either LNA or MNA at position 2 folds into a native-like G-quadruplex, while doubly substituted derivatives of TBA where LNA or MNA is incorporated at positions 11 and 14 are no longer able to form a G-quadruplex. Starting from the same initial intermediate structure, we successfully overcame sampling limitations, and simulated the large conformational variations in structures of native TBA and modified TBAs by classic MD. Analysis of the models showed that inversion of the glycosyl orientation at position 14 contributes significantly to the disruption of G-quadruplex formation in both of the di-substituted modified TBA systems. Our calculations provide a simple and reliable theoretical model that can be used to investigate and predict the effects of the modifications of an oligonucleotide on the resultant G-quadruplex structure. In addition, the computational protocol described can also be applied for other G-quadruplex systems.

Introduction

During the past decade, guanine-rich nucleic acid sequences prone to adopting a variety of stable compact G-quadruplex structures have attracted a tremendous amount of interest for their unique structural features and multiple biological functions.[1,2] According to the strand orientation, G-quadruplexes, consisting of G-quartets and loops, can be classified into parallel or antiparallel structures. G-quartets contain four coplanar guanines that associate via eight Hoogsteen hydrogen bonds and stabilized by specific metal ions, such as K^+ or Na^+, located between two adjacent quartets.[3] Loops connecting the G-quartets, including lateral and diagnose, are formed by several bases located in the intermediate sequence.[4,5] The folding mechanism to form the quadruplex structure and its stability are influenced by factors such as sequence, loop size, Hoogsteen hydrogen bonding, glycosyl orientation, sugar puckering, etc.

The most studied aptamer, thrombin binding DNA aptamer (TBA), is a single oligonucleotide that recognizes its protein target with a consensus 15-base sequence, d(GGTTGGTGTGGTTGG).[6] TBA shows strong anticoagulant properties in vitro and is progressing through preclinical to clinical development.[7] Nuclear magnetic resonance (NMR)[8,9] and X-ray[10] experiments proved that TBA folds into a unimolecular antiparallel G-quadruplex consisting of two guanine quartets connected by two side loops each containing two residues (TT) and one central loop containing three residues (TGT).[11] Even though TBA binds to thrombin protein via its loop regions, it has been confirmed that the TBA G-quadruplex structure is very necessary for binding.[12] To improve the biological or biophysical properties of TBA, many nucleotide modifications have been reported thus far, including the inclusion of a locked nucleic acid (LNA),[13] RNA,[14] 2′-*O*-methyl-RNA (MNA),[15] 8-bromo-guanine,[16] 2′-deoxy-isoguanine,[17] 4-thio-2′-deoxyuridine,[18] an unlocked nucleic acid,[19] 8-amino-2′-deoxyguanosine,[20] or variations that alter the loop size.[21] According to these reported studies, different modifications can promote or disrupt the formation of the G-quadruplex, largely dependent on the position and number of alterations.

It is of high importance to explore the structural factors controlling the structure–activity relationships of TBA. Probing de-

[a] Dr. L. Sun,+ Dr. H. Jin,+ Dr. Z. Liu, Prof. Z. Yang, Prof. L. Zhang, Prof. L. Zhang
State Key Laboratory of Natural and Biomimetic Drugs
School of Pharmaceutical Sciences, Peking University
Beijing 100191 (P. R. China)
E-mail: liangren@bjmu.edu.cn

[b] X. Zhao, Prof. Y. Guan
Key Laboratory of Medical Cell Biology, Ministry of Education
Department of Biochemistry & Molecular Biology
China Medical University
Shenyang 110013, Liaoning (P. R. China)

[+] These authors contributed equally to this work and should be considered as co-first authors.

tails of the G-quadruplex structures can provide lots of useful information for structural studies.[22,23] In this regard, many experimental methods have been utilized to capture and detect these complicated processes, such as single-molecule fluorescence resonance energy transfer (smFRET),[24,25] ultraviolet–visible spectrophotometry (UV-vis),[26] surface plasmon resonance (SPR),[27] NMR,[28] and circular dichroism (CD).[29] Using a hydrogen-bond exchange rate experiment, Mao et al. proposed a possible unfolding pathway in which the terminal guanines are first unzipped.[28] However, detailed information about the structure during the unfolding process and the critical intermediate state cannot be obtained using experimental methods. In recent years, molecular dynamics has been successfully developed and widely applied for the simulation of nucleic acid structures.[30,31] Several studies have reported predicted interactions between thrombin and TBA, modified in the loop regions, by classical molecular dynamics.[12,32]

As mentioned above, a large number of modified derivatives have been synthesized, and these modifications were shown either to promote the folding or disrupt the G-quadruplex. According to previous studies, the G-quadruplex structure is very sensitive to the glycosyl orientation, and opposite glycosyl orientation can lead to total disruption.[33] However, with the time and sampling limitations, the large conformation variations describing the effects of the slight modifications on the substituted structures cannot be obtained by the free molecular dynamics under normal temperature conditions. While many theoretical computational methods[34,35] have been developed and applied to enhance conformation sampling, these methods in turn lead to other problems, such as the need for time-consuming and complicated data analysis. These drawbacks have limited the use of these methods, and the development of a simple model and protocol to solve these problems is required.

Steered molecular dynamics (SMD) simulations[36] are able to give the non-equilibrium force, $f(l)$, as a function of the DNA end-to-end distance (l); similar to single-molecule force experiments.[37] The difficulties encountered when studying G-quadruplex unfolding by classical molecular dynamics can be overcome by using this simulation method. Li et al. have simulated the unfolding process of the human telomeric parallel G-quadruplex and provided two possible folding pathways.[38] However, to the best of our knowledge, experiment data confirming the true folding pathway has yet to be reported. Based on SMD simulations, the probable unfolding pathway can be studied and the critical intermediate structures during the unfolding process also can be predicted.

In this study, atomic SMD simulations were used to predict the unfolding of the antiparallel G-quadruplex of TBA. Compared with classical MD and experimental methods, this technique offers an advantage by enabling large conformational searches and provides detailed structural information to understand the unfolding process. Based on the SMD results, we successfully identified a critical intermediate structure, which was previously predicted by NMR experiments.[28] Starting from the intermediate state, we simulated the refolding process of native TBA by classical MD. In addition, we evaluated the structure characteristics of several TBA oligonucleotides modified with LNA or MNA at different positions and numbers by CD spectroscopy. In order to further corroborate the unfolding pathway and probe the effects of these modifications on the G-quadruplex structure, starting from the intermediate structure, we also simulated the four modified TBA structures by classical MD simulations.

Results and Discussion

Stability of the native TBA G-quadruplex

The stability of the native TBA G-quadruplex structure was first verified by a 4.0 ns free MD simulation. The root mean square deviation (RMSD) plot was calculated with the initial NMR structure[39] as a reference (Figure 1 a). It is worth noting that

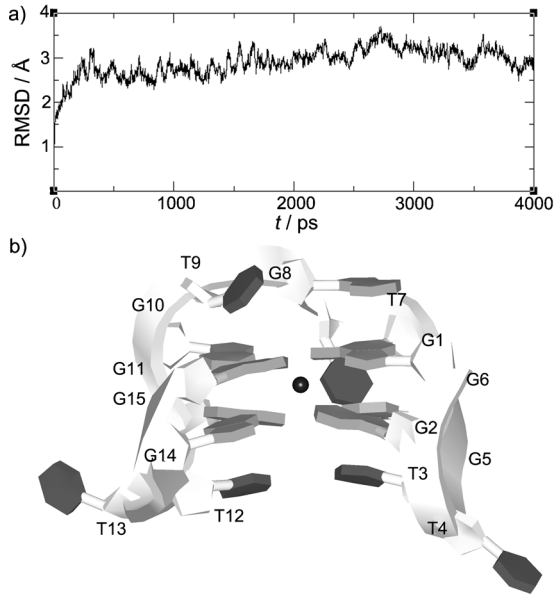


Figure 1. Equilibration dynamics of native TBA. a) The root mean square deviations (RMSDs) of the atoms from their initial positions after preparatory steps during the 4.0 ns simulation of the native TBA. b) The average structure of native TBA extracted from the last 2.0 ns trajectory of the molecular dynamics simulation.

the whole system quickly reached equilibrium in the first 1.0 ns, and the resulting structure was generally stable around an average conformation, reflected by the constant overall RMSD value (3.27 ± 0.31 Å). The average structure, extracted from the last 2.0 ns trajectory, was analyzed (Figure 1 b). From the predicted structure, we can see that the two G-quartets stack well, and the 16 Hoogsteen hydrogen bonds formed by the guanine residues of the G-quartets are also preserved. These calculations confirmed that the previously reported NMR structure[39] can be accurately predicted as a well-defined G-quadruplex by MD simulations.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Unfolding of native TBA

Mao et al. proposed the possible unfolding pathway of native TBA by NMR experiment, but detailed structural information was not described.[28] In addition, Jayapal et al. studied the partially unfolding process of TBA at high temperature using regular MD simulation.[32] Our free MD results confirmed that the reported NMR structure is rather stable. Starting from the NMR structure, the unfolding process of the native TBA G-quadruplex structure was probed by SMD simulation for the first time. In the simulation, the pulling force is applied to the O5′ atom of G1, while the O3′ atom of G15 acts as a reference group. Looking at the force–extension profile over the course of the unfolding process, two distinct peaks with similar magnitude at about 125 ps and 540 ps can be observed, which indicates that the unfolding process undergoes two critical steps (Figure 2 a). The previously reported NMR study[28] also indicated that the unfolding is divided into two steps: first, G1–G15 and G2–G14 open, then G5–G11 and G6–G10 are disrupted. Additionally, in our SMD simulations, we note that the force decreased to very small values and fluctuated between negative and positive values between 250 and 400 ps, indicating that the structure is easily unraveled. This phenomenon was also observed after 600 ps. In most cases, the force is always positive, while some local structures are rather flexible and additional force is unnecessary for variations in the conformation.[40]

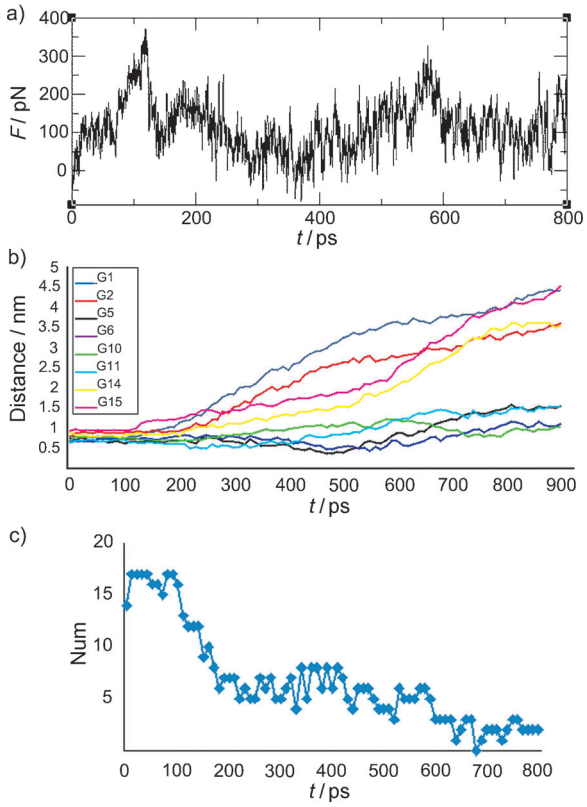


Figure 2. Unfolding of native TBA in aqueous solution. a) Forces as a function of simulation time. b) Distance of guanine residues to the molecular center during stretching. c) The number of hydrogen bonds formed by the two G-quartets of TBA.

In order to further reveal the detailed unfolding process of native TBA, we calculated the distances of the guanine bases to the molecular center during the pulling process. The simulated distances of G1 and G15 to the molecular center increased quickly at first, leading to the unzipping of G1–G15, while the other guanines fluctuated at their equilibrium positions (Figure 2 b). It is worth noting that the first force peak appeared at the same time. Then the G2 and G14 residues started to deviate from their equilibrium positions with the distances increasing, corresponding to the force tough (180 ps) after the first peak. During extension, residues G5, G6, G10 and G11 start to break off until the second force peak. Judging from the relative distances of the adjacent guanines, the local structure formed by G10–G15 and G11–G14 opened last.

Hoogsteen hydrogen bonding is an important structural characteristic of the G-quadruplex, and therefore, the number of hydrogen bonds predicted by SMD simulation can further reveal the detailed unfolding process (Figure 2 c). Initially, 16 hydrogen bonds were observed before the first force peak appeared, suggesting that the structure is a well-defined G-quadruplex. Then, the number of hydrogen bonds rapidly dropped from 16 to 12 at about 125 ps, when the first force peak appeared. The loss of the hydrogen bonds suggests that the G-quadruplex is beginning to undergo large conformational variations. As mentioned above, this variation is mainly contributed to by the breaking of G1–G15. Next, the structure lost another two hydrogen bonds, which corresponds to the deviations of G2–G14. The breaking of G5–G11 and G6–G10 is accompanied by further loss of hydrogen bonds, and the total number decreases to four. In the end, the G-quadruplex unfolded completely.

Now, via analyzing the trajectories of SMD simulations, we have generated a complete picture about the unfolding process of the native TBA G-quadruplex. Figure 3 depicts MD snapshots of the unfolding pathway. During pulling, continuous stretching induces a complete disruption of G1–G15 (125 ps); at about the same time, the first peak in the force curves appears. Then G2–G14 is opened at about 180 ps of the SMD simulation. So far, these observations are in agreement with the previously reported experimental NMR results.[28] In the following 300 ps, G1 and G2 lose interactions with G5 and G6, leading to the unzipping of the T3T4 loop, while the remaining structure remains intact despite the increasing extension. According to the previously reported NMR hydrogen exchange rates experiments,[28] the two TT loops are all open in the final stage of unfolding.

The differences between the observations made in the NMR study and the results of our calculations can be explained by the following reasons: In SMD simulations, the extra pulling force applied to the terminal guanine residue is greater than that experienced in the NMR experiment, leading to a significant overestimation of unfolding forces. In addition, in solution, the terminal structures are more flexible and undergo conformational variation more easily. Furthermore, in the SMD simulations, the force decreases to a very small value and fluctuates between negative and positive values when the terminal guanines (G1 and G2) open. After that, residues G5 and G6

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

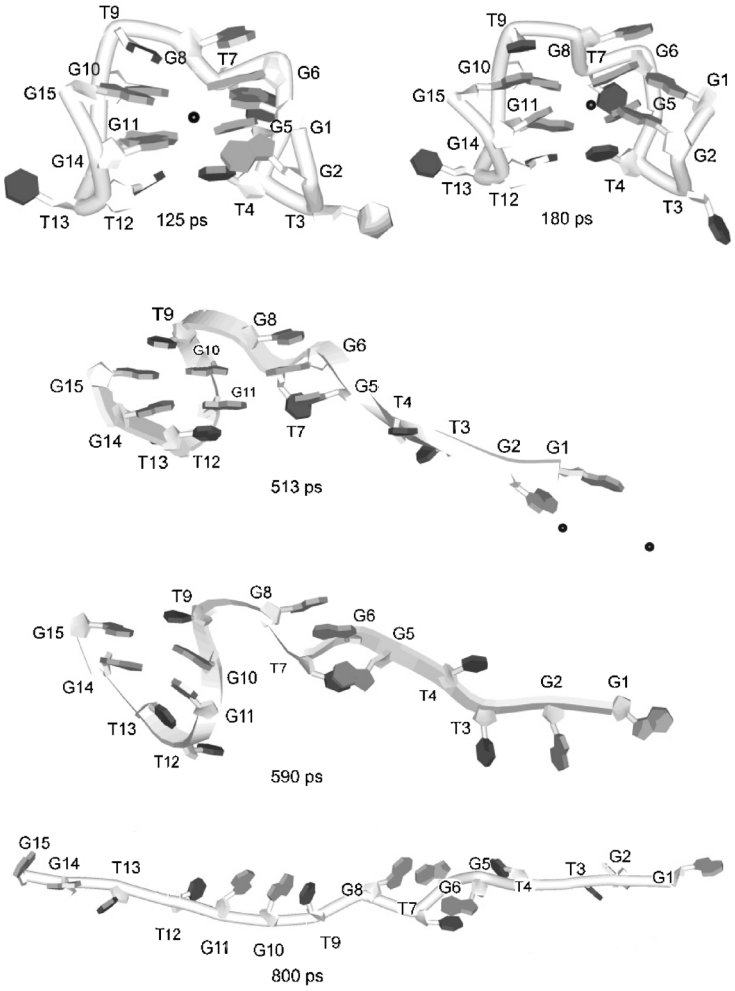


Figure 3. Molecular dynamics snapshots of native TBA demonstrating the unfolding pathway.

start to deviate strongly from their equilibrium positions, leading to the second force peak at 590 ps. A local structure formed by G10–G15 and G11–G14 unfolds, which corresponding to the small force values shown in Figure 2a. The latter results further confirm the experimental observations of the NMR study[28] and prove that the unzipping of T3T4 does not affect the unfolding process. At the end of the simulation, the G-quadruplex structure undergoes a hairpin-to-coil transition. Our SMD simulation results support the unfolding process predicted by the NMR studies and provide meaningful structural information that cannot be obtained by other experiment methods.

Refolding of native TBA

We attempted to simulate refolding of a partially unfolded G-quadruplex by free MD simulation using structure from the SMD simulation as the initial conformation. Theoretically, any initial coordinates taken from the unfolding process can be folded back to the native-like G-quadruplex by MD simulation. However, with the time and sampling limitations, the refolding process can only be observed by starting from the structures corresponding to the first force peak.[35] In this situation, the G-quadruplex is able to avoid the local minima and maintain a partial unfolded structure similar to the G-quadruplex. Combined with our SMD and previously reported NMR results, the starting structure for the refolding of native TBA was taken from the coordinates at 180 ps of the pulling trajectory, which is close to the first force peak when the G1–G15 and G2–G14 unzipped. This critical intermediate structure was predicted in the previous NMR experiment,[28] but the theoretical model was obtained for the first time.

As expected, we successfully simulated the refolding process of native TBA by free MD simulation. Figure 4a shows the RMSD of the structure with respect to the initial structure during free MD simulation. The RMSD curve reached a constant value of about 3.5 Å after 600 ps, which confirmed the stability of the trajectory. The average structure for the last 2.0 ns relative to the whole production run was extracted (Figure 4b). We can see that the partially unfolded structure folded back to its native-like G-quadruplex structure entirely. The four hydrogen bonds reformed between G1–G15 and G2–G14, and the stacking interactions between the two quartets are well preserved. Table 1 shows the hydrogen bond numbers and occupancy of the two quartets during the simulation. For native TBA, the sixteen hydrogen bonds occupancy all exhibited very high values, which further proves the stability of the G-quadruplex. The refolding process also corroborated the intermediate

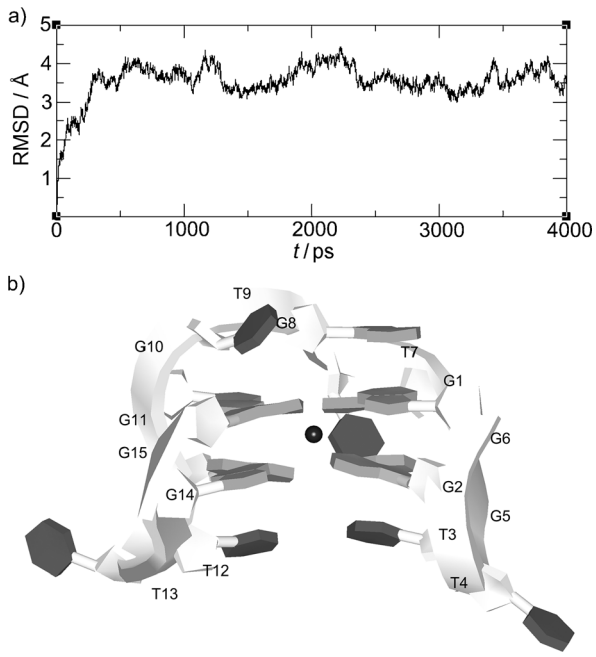


Figure 4. Refolding of native TBA. a) The root mean square deviations (RMSDs) of the atoms from their initial partially unfolded structure during the 4.0 ns simulation. b) The average structure of native TBA over the last 2.0 ns of the molecular dynamics simulation.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1. The hydrogen bond occupancy of G-quartets 1 and 2 during simulations.

		TBA	LNA-2	LNA-11–14	MNA-2	MNA-11–14
G-quartet 1						
15@O6	10@N1	98.8	99.75	–	99.8	99.1
10@N7	6@N2	97.95	99.7	56.37	99.8	22.88
10@O6	6@N1	97.45	99.85	99.55	99.1	23.68
15@N7	10@N2	94.1	99.75	–	99.5	97.2
1@O6	15@N1	88.1	99.75	–	99.3	93.41
1@N7	15@N2	86.5	99.55	–	99.35	78.92
6@N7	1@N2	80.77	99.8	50.42	99.55	–
6@O6	1@N1	80.62	99.85	53.27	99.75	37.06
G-quartet 2						
11@N7	14@N2	99.8	100	–	99.9	99.8
11@O6	14@N1	99.7	99.65	–	99.95	98
5@N7	11@N2	98.8	99.75	97.2	99.65	99.7
2@O6	5@N1	98.5	99.65	77.66	99.4	72.33
2@N7	5@N2	98.2	100	66.02	99.95	82.82
14@O6	2@N1	97.95	99.4	–	99.5	60.74
5@O6	11@N1	96.8	99.45	82.76	99.75	90.91
14@N7	2@N2	96.45	99.7	–	99.7	49.25

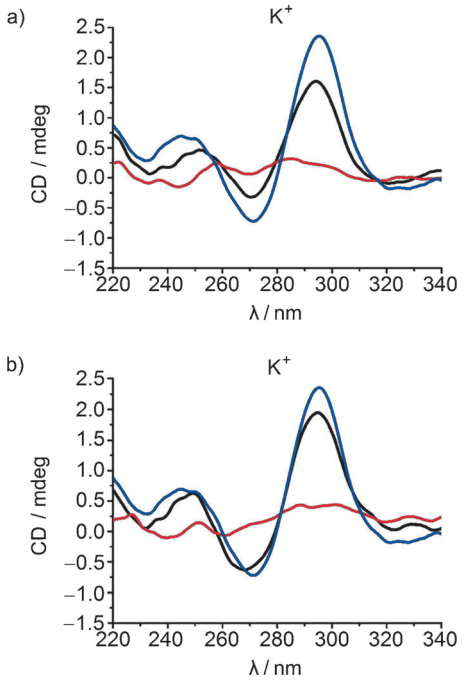


Figure 5. CD spectra of locked nucleoside (LNA) or 2′-*O*-methyl-nucleoside (MNA)-substituted TBAs in 50 mM K^+. a) TBA (black), LNA-2 (blue) and LNA-11–14 (red). b) TBA (black), MNA-2 (blue) and MNA-11–14 (red). Each CD spectrum is an average of at least three scans.

structure and unfolding pathway proposed by SMD simulations. Furthermore, the intermediate structure for the refolding can also be used as a rational model for the study of modified TBA.

Effects of LNA and MNA on the structures of modified TBA

Circular dichroism (CD) spectroscopy

The unmodified TBA in 50 mM K^+ solution shows a characteristics spectrum of an antiparallel G-quadruplex structure: a CD maximum at 295 nm and a CD minimum at 265 nm. Mono-substituted species LNA-2 and MNA-2 show CD spectra almost identical to parent TBA, while the CD spectra of modified TBA species LNA-11–14 and MNA-11–14 suggest a collapse of the antiparallel G-quadruplex structure (Figure 5).

Refolding of LNA-2 and MNA-2

After the simulation, the RMSD curves of LNA-2 (Figure 6a) and MNA-2 (Figure 6b) with respect to the initial structure were also calculated. Both of the RMSD curves reached a similar value—about 3.5 Å after 1.5 ns for LNA-2 and 1.0 ns for MNA-2. This indicates that the two systems reach equilibrium and produce two stable modified structures.

We extracted the average structures for the last 2.0 ns relative to the whole production run (Figure 7). The overall structures, LNA-2 (Figure 7a) and MNA-2 (Figure 7b), are observed to fold back to the native-like G-quadruplex structure entirely. The four hydrogen bonds between G1–G15 and G2–G14 reform during refolding. The stacking interactions between the two quartets are well preserved, with one K^+ ion located in the two G-quartets. For the LNA-2, compared with native TBA, the distance between T4 and T13 increases, and the two hydrogen bonds formed by T4 and T13 become weaker. The conformation of T4 is also different from that observed in native TBA—it deviates slightly from its equilibrium position and the stacking interaction between G2 and T4 become weaker. From a structural point of view, the structural modification decreased the stability of the G-quadruplex, and LNA-2 is the most unstable structure. While the stability is decreased, the binding affinity to thrombin is comparable with native TBA.[41] We also analyzed the hydrogen bonds during the simulations (Table 1). For the two systems, the sixteen hydrogen bonds occupancy all exhibited very high values, which further proves the stability of the G-quadruplexes.

In summary, the average structures and hydrogen bonds analyses of LNA-2 and MNA-2 showed that they can refold to the native-like antiparallel G-quadruplex structures from the stretched structures. These results are in good agreement with the CD spectroscopy results. In addition, the refolding process of the two modified TBA structures corroborate the presence of the intermediate state and further confirmed the folding pathway.

Refolding of di-substituted TBA

Two plots of the RMSD values as a function of the simulation time are shown in Figure 6. For LNA-11–14, the RMSD values kept increasing and did not converge within the simulation timeframe of 4 ns (Figure 6c). For MNA-11–14, the RMSD values increased slowly and finally fluctuated at 4 Å after 3 ns (Figure 6d). These results suggest that some notable structural distortions occur in the two di-substituted structures during the simulations.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

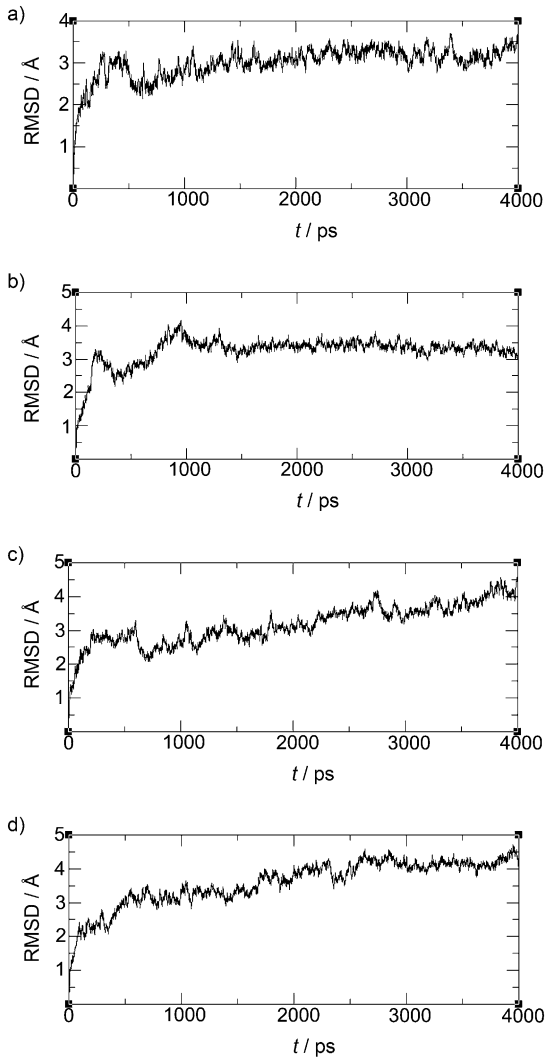


Figure 6. Root mean square deviation (RMSD) curves of the four modified TBAs. a) LNA-2, b) MNA-2, c) LNA-11–14, and d) MNA-11–14.

Looking at the final structure of LNA-11–14 after the simulation (Figure 7 c), we can see that the terminal guanine residues (G1, G2, G14 and G15) completely deviate from the two quartets, whereas a local G-quadruplex structure formed by G5–G11 and G6–G10 is still preserved. During the simulation, the K^+ ion left the modified structure and entered solution. As such, starting from the partially unfolded structure, the LNA-11–14 cannot refold to a native-like G-quadruplex structure. For MNA-11–14 (Figure 7 d), the terminal guanine residues cannot reform stable hydrogen bonds or base stacking interactions, and thus the modified structure cannot fold into a stable G-quadruplex structure. These calculations agree with the CD spectroscopy results. Although the totally unfolded states of LNA-11–14 and MNA-11–14 were not observed within the limited simulation time, the effects of the modifications on the G-quadruplex can be well explained based on the current results.

Table 1 shows the hydrogen bond occupancy of LNA-11–14 and MNA-11–14. For the di-substitution structures, only partial hydrogen bonds are formed, and the hydrogen bond occupancies are all decreased to very low values, especially for the modified guanines at position 14. LNA at position 14 causes the complete loss of the four hydrogen bonds formed with the adjacent guanines, while MNA at position 14 causes the loss of the two hydrogen bonds formed with the native guanine at position 2. In contrast, all of the hydrogen bonds at position 11 are preserved. Modification with LNA or MNA at position 11 is tolerated. Thus, substitution with LNA or MNA at position 14 contributes to the significant disruption of the G-quadruplex.

Glycosyl orientation

Glycosyl orientation plays a critical role in the folding of G-quadruplex,[33] as such, we also studied the glycosyl orientation of the four modified systems. In the native TBA structure, the glycosyl orientation at positions 2 and 11 are *anti* conformations, while the guanine at position 14 is *syn*. The preferred conformation of nucleic analogues LNA and MNA is also *anti*. Since LNA or MNA substitution at either position preserves the original *syn–anti–syn–anti* glycosidic conformation arrangement of the G-tetrads, LNA-2 and MNA-2 can fold into a stable antiparallel G-quadruplex structure. As for the derivatives with substitutions at both positions 11 and 14, LNA-11–14 and MNA-11–14, the glycosyl orientation (*anti*) is inverted with respect to the native structure at position 14, leading to a *syn–anti* glycosidic angle switch of the substituted guanine residues, and the resulting *syn–anti–anti–anti* or *anti–anti–syn–anti* G-tetrad is not the energetically favored configuration of the antiparallel G-quadruplex. Cang et al. indicated that the stability of glycosidic conformational of the antiparallel structure is *syn–anti* > *anti–anti*.[33]

Other modifications, including one or more inverted glycosyl orientations, all led to total disruption.[42] For LNA-11–14, after the simulation, the LNA at position 14 lost all interactions with adjacent guanine residues, which contributed significantly to the disruption in folding, while the LNA at position 11 acted like a native guanine residue. Similar results were also observed in the MNA-11–14 simulation. Starting from the partially unfolded G-quadruplex structure, the disruptive effects of LNA or MNA at position 14 were observed. While the conformation transition, *anti* and *syn*, were not obtained, the effects of the modifications (LNA and MNA) have been well described. In the absence of experiments, this model can be used as an effective tool to predict in a qualitative fashion whether a modified oligonucleotide can fold into a TBA G-quadruplex structure.

Conclusions

Due to the special structural characteristics, the unfolding process and intermediate state of a G-quadruplex are very important for understanding the folding mechanism and structure–activity relationships. Furthermore, with the limitations of sampling and time, the effects of small chemical modifications on a G-quadruplex structure, especially large conformation variations, are very difficult to obtain by classical molecular dynamics (MD) simulations. In this research, for the first time, we simulated the unfolding process of the native thrombin-binding

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

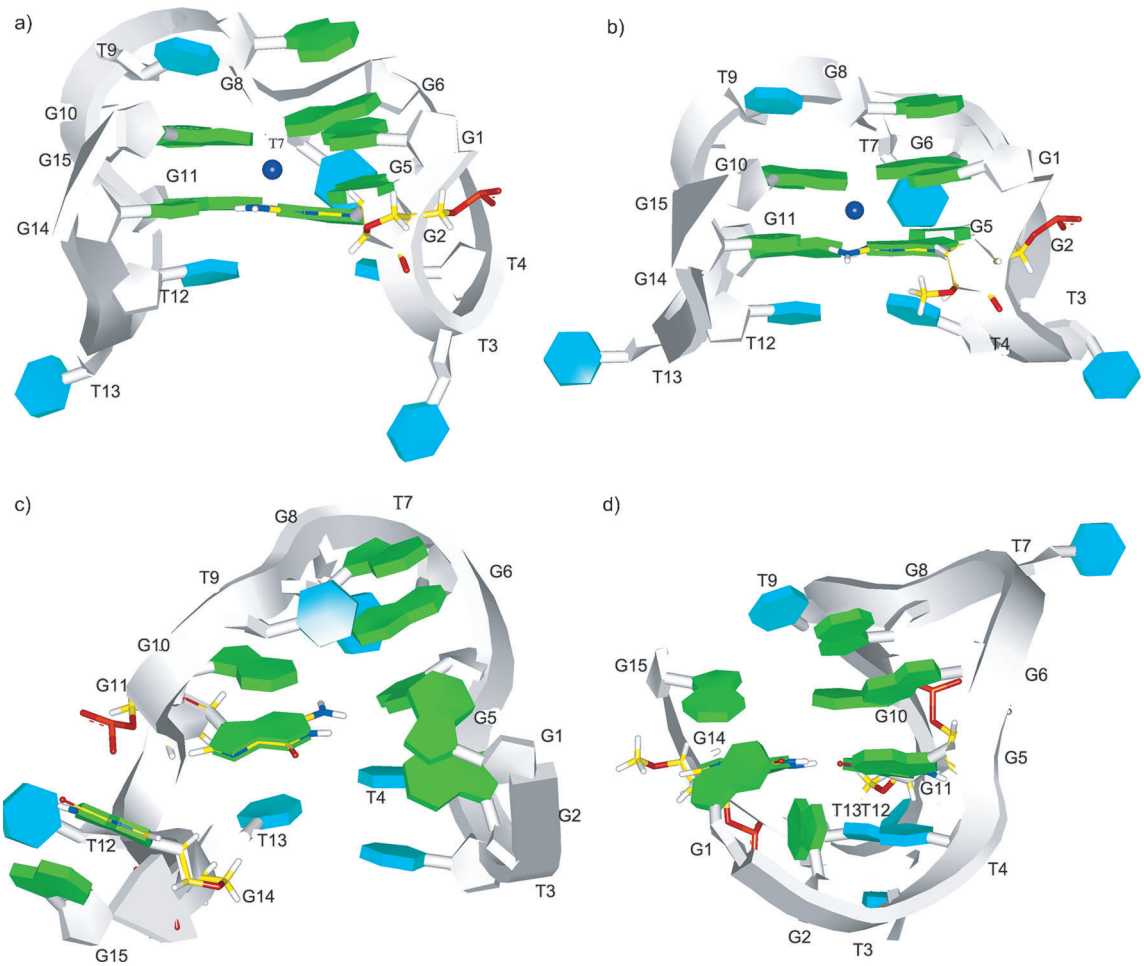


Figure 7. The average structures over the last 2.0 ns the molecular dynamics simulation, with LNA or MNA in stick representation (yellow). a) LNA-2, b) MNA-2, c) LNA-11–14, and d) MNA-11–14.

DNA aptamer (TBA) G-quadruplex structure by theoretical computation and corroborated the existence of a critical intermediate state, which was predicted in previous NMR experiments. Starting from the same intermediate structure, we successfully simulated the large conformational variations in structures of native TBA and modified TBA (LNA-2, MNA-2, LNA-11–14, MNA-11–14).

In the native TBA structure, the glycosyl orientation at positions 2 and 11 are *anti*, while the guanine residue at position 14 is *syn*. The preferential conformation of nucleoside analogues LNA and MNA are *anti*. Therefore, the native TBA and modified TBA derivatives containing LNA-2 and MNA-2 can be refolded into well-defined antiparallel G-quadruplex structures, while LNA-11–14 and MNA-11–14 cannot because of the inverted glycosyl orientations at position 14, which cause major alterations to the hydrogen bonding in these di-substituted derivatives leading to an inability to refold. These conformation variations are agreement with our CD spectroscopy results and further confirm the unfolding pathway as well as intermediate structure predicted by previous NMR studies. Our computations provide a rational model that can be used to predict and investigate the effects of modifications on the TBA G-quadruplex structure. The computational protocol used in this research also can be applied for other G-quadruplex systems.

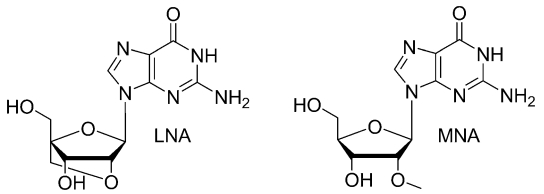
Experimental Section

Construction of the computational models: The NMR structure of antiparallel G-quadruplex TBA (PDB entry 1C35[39]) was used as the starting model for simulations. The TBA and modified structures (Table 2), LNA-2, LNA-11–14, MNA-2 and MNA-11–14, were built using the Discovery Studio 2.5 package.[43] The molecular mechanical parameter of MNA was downloaded from the amber website (http://ambermd.org/). To obtain parameters for LNA, ab initio quantum chemical methods were employed using Gaussian 09.[44] The geometry was fully optimized, and then the electrostatic potentials around atoms of LNA were determined at the HF/6-31G* level of theory. The restrained electrostatic potential (RESP) strategy[45] was used to obtain the partial atomic charges.

Steered molecular dynamics (SMD) simulations: SMD simulations were performed for native TBA using the GROMACS 4.5.3 package[46] and the AMBER03 force field.[47] The temperature ($T = 300$ K) and pressure ($p = 1$ atm) were controlled by langevin thermostats and barostats.[48] The particle-mesh Ewald summation with a 10 Å cutoff was used to compute the long-range electrostatic forces. An

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Chemical structures of locked nucleoside (LNA) or 2′-*O*-methyl-nucleoside (MNA) as well as the sequences of modified TBA used in this study.



Name	Nucleotide sequence[a]
TBA	5′-GGTTGGTGTGGTTGG-3′
LNA-2	5′-GgTTGGTGTGGTTGG-3′
LNA-11–14	5′-GGTTGGTGTGgTTgG-3′
MNA-2	5′-GgTTGGTGTGGTTGG-3′
MNA-11–14	5′-GGTTGGTGTGgTTgG-3′

[a] LNA and MNA are in lower cases, and the position of the modification(s) is given in the name.

integration time step of 2 fs was used. The bonds were constrained using the linear constraint solver (LINCS) algorithm. Conformation of the TBA from NMR experiment served as an initial structure for SMD simulation of 800 ps to characterize the unfolding process. The G-quadruplex was solvated in a cubic box of $6.56\times8.72\times18\,nm^3$; 74 K^+ ions were added to provide electroneutrality. In the simulation, a constant velocity SMD method was used to stretch TBA along two ends of the G-quadruplex. One Hookean spring with force constant ($k_0 = 1000\,kJ\,mol^{-1}\cdot nm^2$) was attached to the O5′ atom of G1 at the end of the quadruplex, and the O3′ atom of G15 was used as a reference. All other atoms were free to move during the simulations. The pulling velocity was 0.01 Å ps^{-1}.

Regular molecular dynamics (MD) simulations: MD simulation was performed with the AMBER 11 molecular simulation package.[49] The AMBER99SB force field was used to describe the molecule. The model was solvated in TIP3P water using an octahedral box, which extended 10 Å away from any solute atom. To neutralize the negative charges of simulated molecules, a K^+ counter ion was added to provide electroneutrality. MD simulations were carried out by using the SANDER module of AMBER 11. The calculations began with 500 steps of steepest descent followed by 500 steps of conjugate gradient minimization with a large constraint of 500 kcal mol^{-1} $Å^{-2}$ on the solute atoms. Then, 1000 steps of steepest descent followed by 1500 steps of conjugate gradient minimization with no restraints on the solute atoms were performed. Subsequently, after 20 ps of MD, during which the temperature was slowly raised from 0 to 300 K with weak (10 kcal mol^{-1} $Å^{-2}$) restraints on the solute, the final unrestrained production simulations of 4.0 ns for the molecules were carried out at constant pressure (1 atm) and temperature (300 K). In the entire simulation, SHAKE was applied to all hydrogen atoms. Periodic boundary conditions with minimum image conventions were applied to calculate the nonbonded interactions. A cutoff of 10 Å was used for the Lennard–Jones interactions.

Preparation of oligonucleotides: The parent TBA (oligonucleotide purification cartridge purified) and LNA- or MNA-substituted TBAs (high-performance liquid chromatography purified) were purchased from TaKaRa Biotech Ltd (Dalian, China). Their stock solutions were prepared with buffer (0.1 mM EDTA, 10 mM sodium phosphate, pH 7.4) at a final concentration of 0.1 mmol L^{-1} and stored at −20 °C. TBAs containing LNA or MNA substitutions at the designated positions were prepared (Table 2). Their solutions were prepared in a buffer of 10 mM Tris-HCl, 0.1 mM EDTA, and 50 mM KCl or 50 mM $CaCl_2$ (pH 7.5) unless specified otherwise.

Circular dichroism (CD) spectroscopy: CD spectra of TBA species were recorded using a spectropolarimeter (JASCO J-810). A 500 μL aliquot of TBA solution in a UV cell of 0.1 cm path length was placed in a thermostable holder. CD spectra were recorded from 200 to 340 nm at room temperature unless specified otherwise. Three scans at a speed of 500 nm min^{-1} were accumulated for every sample. The TBA solution with final concentration of 5 μM used for the CD spectroscopic study was prepared in a buffer (0.1 mM EDTA, 10 mM sodium phosphate, pH 7.4) with either 50 mM KCl or $CaCl_2$.

Acknowledgements

This work was supported by the National S&T Major Project Foundation of China (Grant No. 2011ZX09102-001-17), the National Natural Science Foundation of China (Grant No. 91213302) and the Ministry of Science and Technology of China (Grant No. 2012CB720604).

Keywords: G-quadruplexes · molecular dynamics · nucleosides · simulations · thrombin-binding DNA aptamers

[1] J. L. Huppert, *Philos. Trans. R. Soc. London Ser. A* **2007**, *365*, 2969–2984.
[2] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182–186.
[3] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013–2018.
[4] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
[5] Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149–1171.
[6] L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, J. J. Toole, *Nature* **1992**, *355*, 564–566.
[7] M. Khati, *J. Clin. Pathol.* **2010**, *63*, 480–487.
[8] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, J. Feigon, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745–3749.
[9] K. Y. Wang, S. Mccurdy, R. G. Shea, S. Swaminathan, P. H. Bolton, *Biochemistry* **1993**, *32*, 1899–1904.
[10] I. Russo Krauss, A. Merlino, A. Randazzo, E. Novellino, L. Mazzarella, F. Sica, *Nucleic Acids Res.* **2012**, *40*, 8119–8128.
[11] J. A. Kelly, J. Feigon, T. O. Yeates, *J. Mol. Biol.* **1996**, *256*, 417–422.
[12] B. Pagano, L. Martino, A. Randazzo, C. Giancola, *Biophys. J.* **2008**, *94*, 562–569.
[13] A. Virno, A. Randazzo, C. Giancola, M. Bucci, G. Cirinoc, L. Mayol, *Bioorg. Med. Chem.* **2007**, *15*, 5710–5718.
[14] M. Zaitseva, D. Kaluzhny, A. Shchyolkina, O. Borisova, I. Smirnov, G. Pozmogova, *Biophys. Chem.* **2010**, *146*, 1–6.
[15] B. Sacca, L. Lacroix, J. L. Mergny, *Nucleic Acids Res.* **2005**, *33*, 1182–1192.
[16] S. Goji, J. Matsui, *J. Nucleic Acids* **2011**, *2011*, 316079.
[17] S. R. Nallagatla, B. Heuberger, A. Haque, C. Switzer, *J. Comb. Chem.* **2009**, *11*, 364–369.
[18] S. Mendelboum Raviv, A. Horvath, J. Aradi, Z. Bagoly, F. Fazakas, Z. Batta, L. Muszbek, J. Harsfalvi, *J. Thromb. Haemostasis* **2008**, *6*, 1764–1771.
[19] A. Pasternak, F. J. Hernandez, L. M. Rasmussen, B. Vester, J. Wengel, *Nucleic Acids Res.* **2011**, *39*, 1155–1164.
[20] J. Lopéz de la Osa, C. González, R. Gargallo, M. Rueda, E. Cubero, M. Orozco, A. Avinó, R. Eritja, *ChemBioChem* **2006**, *7*, 46–48.
[21] S. Nagatoishi, N. Isono, K. Tsumoto, N. Sugimoto, *Biochimie* **2011**, *93*, 1231–1238.
[22] R. D. Gray, R. Buscaglia, J. B. Chaires, *J. Am. Chem. Soc.* **2012**, *134*, 16834–16844.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[23] T. Mashimo, H. Yagi, Y. Sannohe, A. Rajendran, H. Sugiyama, *J. Am. Chem. Soc.* **2010**, *132*, 14910 – 14918.
[24] B. Okumus, T. Ha, *G-Quadruplex DNA Methods Protocols* **2010**, *608*, 81 – 96.
[25] P. S. Shirude, S. Balasubramanian, *Biochimie* **2008**, *90*, 1197 – 1206.
[26] C. G. Peng, M. J. Damha, *Nucleic Acids Res.* **2007**, *35*, 4977 – 4988.
[27] Q. J. Tang, X. D. Su, K. P. Loh, *J. Colloid Interface Sci.* **2007**, *315*, 99 – 106.
[28] X. A. Mao, W. H. Gmeiner, *Biophys. Chem.* **2005**, *113*, 155 – 160.
[29] C. M. Olsen, L. A. Marky in *G-Quadruplex DNA: Methods and Protocols*, (Ed.: P. Baumann), Humana Press, Totowa, **2010**, Vol. 608, pp. 147 – 158.
[30] J. Šponer, N. Špačková, *Methods* **2007**, *43*, 278 – 290.
[31] S. Haider, S. Neidle, *G-Quadruplex DNA Methods Protocols* **2010**, *608*, 17 – 37.
[32] P. Jayapal, G. Mayer, A. Heckel, F. Wennmohs, *J. Struct. Biol.* **2009**, *166*, 241 – 250.
[33] X. H. Cang, J. Sponer, T. E. Cheatham, *Nucleic Acids Res.* **2011**, *39*, 4499 – 4512.
[34] T. Takeda, D. K. Klimov, *Biophys. J.* **2009**, *96*, 442 – 452.
[35] S. Kannan, M. Zacharias, *Nucleic Acids Res.* **2011**, *39*, 8271 – 8280.
[36] M. Suan Li, B. Khanh Mai, *Curr. Bioinf.* **2012**, *7*, 342 – 351.
[37] J. Zlatanova, K. van Holde, *Mol. Cell* **2006**, *24*, 317 – 329.
[38] H. Li, E. H. Cao, T. Gisler, *Biochem. Biophys. Res. Commun.* **2009**, *379*, 70 – 75.
[39] V. M. Marathias, P. H. Bolton, *Nucleic Acids Res.* **2000**, *28*, 1969 – 1977.
[40] G. J. Zhao, C. L. Cheng, *Amino Acids* **2012**, *43*, 557 – 565.
[41] L. Bonifacio, F. C. Church, M. B. Jarstfer, *Int. J. Mol. Sci.* **2008**, *9*, 422 – 433.
[42] H. Saneyoshi, S. Mazzini, A. Avino, G. Portella, C. Gonzalez, M. Orozco, V. E. Marquez, R. Eritja, *Nucleic Acids Res.* **2009**, *37*, 5589 – 5601.
[43] Discovery Studio 2.5, Accelrys Inc., San Diego (USA), **2009**.
[44] Gaussian 09 (Revision A.1), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian Inc., Wallingford, CT (USA), **2009**.
[45] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269 – 10280.
[46] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435 – 447.
[47] E. J. Sorin, V. S. Pande, *Biophys. J.* **2005**, *88*, 2472 – 2493.
[48] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. D. Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684 – 3690.
[49] D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, AMBER 11, University of California, San Francisco, USA, **2010**.

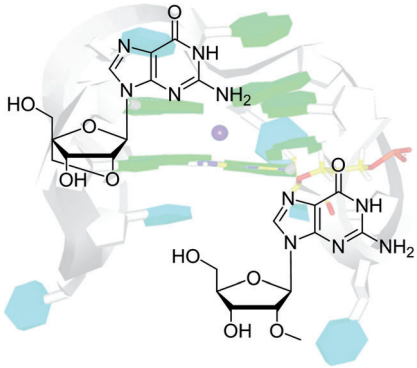
Received: December 29, 2014
Revised: March 14, 2014
Published online on ■■ ■■, 0000

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

L. Sun, H. Jin, X. Zhao, Z. Liu, Y. Guan, Z. Yang, L. Zhang, L. Zhang*

■■ – ■■

Unfolding and Conformational Variations of Thrombin-Binding DNA Aptamers: Synthesis, Circular Dichroism and Molecular Dynamics Simulations



Unraveling the mystery! The unfolding process of thrombin-binding DNA aptamer (TBA) G-quadruplexes and conformational variations of TBA oligonucleotides modified with locked nucleoside or 2′-*O*-methyl-nucleoside were studied by molecular dynamics and circular dichroism spectroscopy. The computational protocol described offers a novel tool for the investigation of other G-quadruplex systems.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim