Comparison of Models of Thrombin-Binding 15-mer DNA Aptamer by Molecular Dynamics Simulation

R. V. Reshetnikov^{1*}, A. V. Golovin¹, and A. M. Kopylov²

¹Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 11991 Moscow, Russia; fax: (495) 939-4195; E-mail: r.reshetnikov@gmail.com
²Faculty of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia; fax: (495) 939-8846; E-mail: kopylov.alex@gmail.com

Received December 18, 2009

Abstract—Two models of 15-mer thrombin-binding DNA aptamer (15TGT) were comparatively analyzed by molecular dynamics simulation using the GROMACS software package. The two original models of 15TGT were obtained by NMR and X-ray analyses. The models significantly differ in the topology of loops and the direction of oligodeoxyribonucleotide chain. The evolution of the two structures in parm99 force fields and parmbsc0 optimized for nucleic acids was analyzed in our adaptation of GROMACS architecture. It is shown that the best system for description of the 15TGT structure is the model obtained by X-ray analysis in the parmbsc0 force field.

DOI: 10.1134/S0006297910080109

Key words: aptamer, molecular dynamics simulation, thrombin

Aptamers are synthetic oligonucleotides specific to targets varying from small molecules to whole cells [1, 2]. The aptamers are obtained by the SELEX (Systematic Evolution of Ligands by Exponential Enrichment) method [3]. One of the first successful targets of the aptamers was the serine protease thrombin, which plays a key role in blood clotting [4-6].

There are a number of DNA aptamers to thrombin, but the structure is known only for the 15-mer aptamer 5'-GGTTGGTGTGGTTGG-3' called 15TGT. The structure of this aptamer was studied in a complex with thrombin by X-ray analysis [7]. Using NMR, the structure of the aptamer was obtained in free state [8] and in a complex with thrombin [9] as well as in complexes with various ions [10, 11].

In all cases, 15TGT has so-called G-quadruplex structure. Eight guanines form two plane G-quartet (G-quadruplex core) attached together by three lateral-type loops: a central trinucleotide with TGT sequence (TGT-loop) and two lateral dinucleotide loops with TT sequence (TT-loop) (Fig. 1).

The models of 15TGT structure obtained by X-ray analysis and NMR differ significantly in configuration of

loops and direction of the polynucleotide chain. In addition, there is a difference between the aptamer-thrombin interaction: in the X-ray model the protein interacts with the TGT-loop of the oligonucleotide, while in the NMR model the interaction is mediated by the TT-loop. Kelly et al. have shown that the model of 15TGT obtained by NMR and aligned with the X-ray model has better R-factor than the X-ray structure itself [12]. In addition, the data for aptamer-thrombin interactions in solution are controversial for the two models. Thus, Ikebukuro et al. demonstrated that G-for-C substitution in the TGT-loop leads to significant decrease in thrombin-inhibiting activity of the aptamer [13], which might indicate breakage of probable DNA-protein contact. At the same time, mutations in the TT-loops influence the binding efficiency of thrombin [14, 15].

Thus, there is currently no full model of the aptamer structure. Comparative analysis *in silico* of evolution of both variants of the aptamer structure by molecular dynamics (MD) will reveal possible structure of the aptamer—protein complex that is immensely important for rational design of thrombin-specific aptamers.

There are several works checking the correctness of the MD method for description of G-quadruplex structures [16-20]. It has been shown that the G-quadruplex core is well modeled by MD and has rigid structure, while

^{*} To whom correspondence should be addressed.

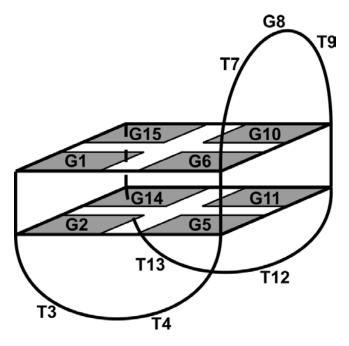


Fig. 1. Scheme of 15TGT structure. The polynucleotide chain is shown by the line. Nucleotide residues are designated by letters and numbers.

the G-quadruplex loops are unsuitable objects for MD modeling. There are three types of G-quadruplex loops: diagonal, propeller (with opposite direction of the chain), and lateral [21]. Sponer's team has shown that the parm94 and parm99 force fields incorrectly predict the topology of diagonal and propeller loops of the structures (G4 T4 G4)₂ and AGGG(TTAGGG)₃, respectively [16-18]. However, Pagano et al. have shown that the lateral loops of 15TGT are correctly described by the parm98 force field in short simulations [22].

In the year 2007, the parmbsc0 force field optimized for simulation of nucleic acids by the parm99 force field was elaborated [23]. This force field accurately describes the topology of canonical DNA structures in a context of very long time of simulation and correctly describes many noncanonical DNA structures [24]. The parameters of

the force field enable reliable prediction of the structure of diagonal TTA-loops of the G-quadruplexes [17].

There are two works on MD of 15TGT and its derivatives. In the study of Pagano et al. [22], it was shown that the final averaged structure is almost the same as the original NMR structure of 15TGT, but the behavior of the aptamer structure was not analyzed in detail. In the study of Jayapal et al. [25], it was shown that the NMR model of 15TGT with disrupted topology regains its initial structure in 2 nsec of MD in force field OPLS-AA. The records of this force field for nucleic acids were presented at http://rnp-group.genebee.msu.su/3d/ff.htm, but the parameters of the torsion angles of deoxyribose appeared to be non-optimal and to require further development (personal communication of Maik Goette).

In the present work, static and dynamic properties of two 15TGT models obtained by MD with long-term trajectories using the parameters of the parm99 and parmbsc0 force fields are described. The X-ray structure has been shown to have abnormal fluctuations of the G-quadruplex core in both these force fields. The topology of the final structure significantly differs from the original. The NMR structure has rigid G-quadruplex core, and the geometry of the loops of the final structure corresponds to the NMR data. These facts indicate incorrectness of the solution of the X-ray structure. The NMR structure in the parmbsc0 force field was shown to be the optimal description system of the 15TGT structure.

MATERIALS AND METHODS

The X-ray structure of 15TGT was taken from the structure of human α -thrombin—aptamer complex (PDB ID 1hut [7]). The NMR structure of 15TGT was taken from PDB file 148d, the eighth frame [8].

The GROMACS 4.0 software package was used for MD simulation and the analysis of trajectories [26-28]. The simulations were carried out in two force fields: AMBER-99 ϕ , which is parm99 [29] modified and ported to GROMACS by Sorin and Pande [30]; parmbsc0 [23] ported to the GROMACS by the authors of the present

Table 1. Simulation parameters

Aptamer model	Force field	Trajectory length, nsec	System content, atoms			
			DNA	water	Na ⁺	K ⁺
X-Ray 15TGT	parm 99	150	488	22 516	13	1
X-Ray 15TGT	parmbsc0	150	488	27 344	13	1
NMR 15TGT	parm 99	150	488	22 524	13	1
NMR 15TGT	parmbsc0	150	488	22 524	13	1

work. The simulations in explicit solvent were carried out at 300 K under the control of velocity rescaling thermostat [31] at constant pressure and using the PME (particle mesh Ewald) method [32] to take into account the electrostatic interactions. The aptamers were put into the center of a triclinic cell with distance to the borders 15 Å. The cell was filled by TIP4P water [33]; negative charge of the system was compensated by sodium ions. To stabilize the structure, potassium ions were added to the center of the G-quadruplex. Total content of the systems is given in Table 1.

RESULTS

Initial models of spatial structure of 15TGT. There are two sources of structural data for 15TGT: NMR [8] and X-ray analysis [7]. The NMR structure of 15TGT is characterized by stacking of nitrogen-containing bases G8 and T9 with the upper G-quartet of G-quadruplex and T4–T13 interaction stacked with lower G-quartet. The X-ray structure has different topology of the loops: only G8 is stacked with the upper G-quartet but not T4 and T13; although they are parallel to the lower G-quartet, they do not interact with each other. The most significant difference between the two structures is reverse direction of the polynucleotide chain.

The observed contradiction occurs due to the fact that the aptamer molecule is asymmetric but having local symmetric elements such as G-quartets. Electron density maps exactly characterize the G-quartets, but ambiguity appears after the analysis of the loops.

Final models of spatial structure of 15TGT. Both of the 15TGT models are characterized by stable trajectories of the MD in both the parm99 and parmbsc0 force fields. The NMR model retained all structural properties in both the force fields: there is G8 and T9 stacking with the upper G-quartet, and the T4–T13 pair stacked with the lower G-quartet (Fig. 2, a and b).

In the X-ray model in the parm99 force field, the conformation of lateral TT-loops has changed resulting in formation of new stacking of T4–T13, while stacking of G8 with the upper G-quartet is retained. Stacking of T4–T13 was also found in the NMR model of the aptamer complex with Sr²⁺ (PDB ID 1rde) [10]. In the parmbsc0 force field, the X-ray model has significantly changed topology (RMSD relative to the initial structure is 3.63 Å); the final conformation has defective geometry of the G-quadruplex core and does not have the T4–T13 pair (Fig. 2, c and d).

Dynamics of the 15TGT models. To describe the dynamics of the15TGT structure, the motions of the nitrogen atoms of heterocyclic bases of DNA were followed since the motions of nitrogen-containing bases are crucial for the overall motion of the molecule. For the analysis of each structure, the last 40 nsec of the trajectory of MD were taken into account.

According to the graph of mean-square fluctuations (Fig. 3), the heterocyclic bases of the G-quadruplex are characterized by minimal deviations from the basic structure and there is no significant difference in dynamics of the G-quartets between the 15TGT models. The following regularities in dynamics of nitrogen-containing bases of the loops were revealed. First, it is not surprising that their mobility is much higher than that of heterocyclic bases of the G-quadruplex. Second, mobility of the loops is higher in the parmbsc0 than in the parm99 force field. The average amplitude of mean-square fluctuations of the nitrogen-containing bases of the loops of the X-ray structure is lower than that of the NMR structure. A conclusion that the X-ray structure is better stabilized than the NMR structure would be trivial, but this suggestion was proved to be incorrect.

Motion parameters of the G-quadruplex: assessment of planarity and rigidity. In the present work we attempted to elaborate the parameters adequately describing the motions of the G-quadruplex. For this goal, two types of motion are marked out: "in plane" — fluctuations of the guanine in the G-quartet plane, and "from plane" — tendency of the G-quartet to form tepee structure.

For the first type of motions, the distance between the centers of mass of heterocyclic bases of guanine of the G-quartets was used as a parameter (Fig. 4a). The second type of the motions was described by the distance between mass centers of the two quartets, which have O6 and N9 atoms as vertices, respectively (Fig. 5, a and b). As in case of construction of the graph of root-mean-squared fluctuations, the last 40 nsec of the MD trajectories were used for the analysis.

"In plane" motions. To construct a graph of the distances between the centers of mass of nitrogen-containing bases of the G-quartets, a maximum of the four distances was taken at each point in time. Detailed statistics of the maximal distances are shown in Table 2.

According to Fig. 4, the least rigid structure of the upper G-quartet is the 15TGT X-ray model in the parm99 force field since its fluctuations deviate from the average most often and far, which is much connected with change in distance between G1 and G6 (Table 2). The fluctuations of the other distances of the upper G-quartet are in the range of 0.5 Å from the average. The rigidity of the lower G-quartet is similar, but the amplitudes of fluctuations of the G2–G14 and G11–G5 pairs are greater than that of the others (Table 2).

"From plane" motions. The X-ray model of 15TGT in the parmbsc0 force field demonstrates no planarity of the upper and lower G-quartets. The G-quadruplex of the model acquires sharp tepee structure (see Fig. 2d). In the parm99 force field, motions of the G-quartets of the X-ray structure almost do not differ from that of the NMR structure with "from plane" fluctuation amplitude in the range of 0.6 Å (Fig. 5); stable structure is observed despite the fact that the lower G-quartet of the X-ray

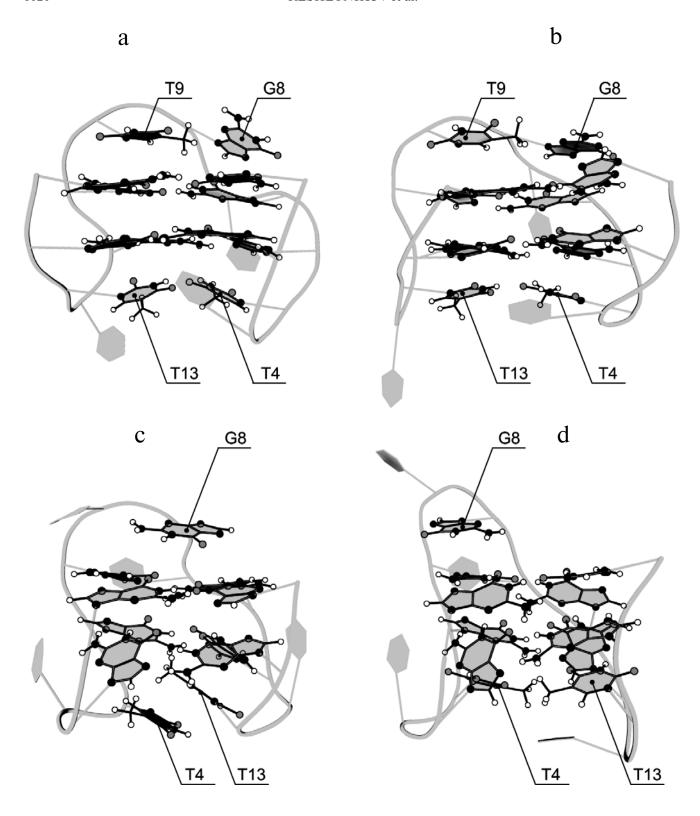


Fig. 2. Static models of spatial structure of 15TGT after MD simulation. a, b) NMR model of 15TGT in the parm99 and parmbsc0 force fields, respectively. The models resemble to each other as well as the initial structure of PDB 148d [8]. They are characterized by stacking of G8 and T9 with the upper G-quartet and T4—T13 pair interactions in stacking with the lower G-quartet. c, d) X-Ray model of 15TGT in the parm99 and parmbsc0 force fields, respectively. In the parm99 force field T4—T13 stacking defective in planarity of the lower G-quartet appears. In the parmbsc0 force field the G-quadruplex core has acquired tepee structure. T4 and T13 do not interact with each other.

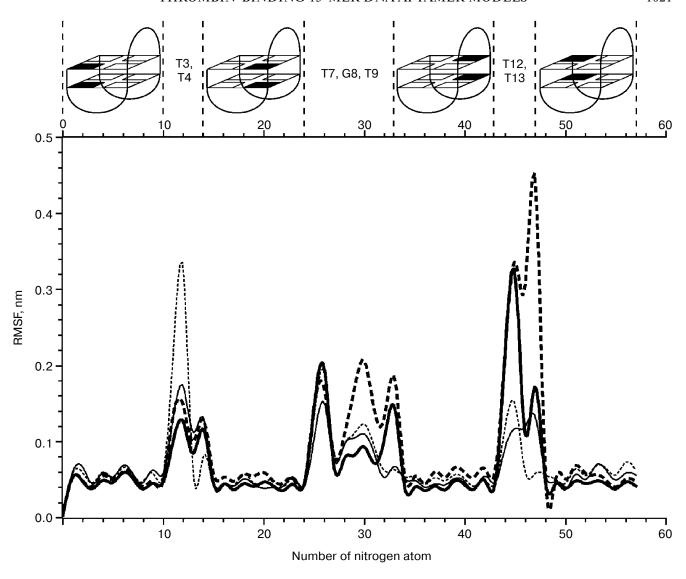


Fig. 3. A graph of root-mean-squared fluctuations (RMSF) of the 15TGT nitrogen atoms. The panel in the top shows the correspondence between the structural elements and the numbered atoms. The atoms of G-quartet bases are marked by a scheme of 15TGT with corresponding black-marked bases. The thin solid line is the X-ray model in the parm99 force field; the thin dashed line is the X-ray model in the parm99 force field; the thick dashed line is the NMR model in the parmbsc0 force field (similar for Figs. 4 and 5).

structure is not planar due to the T4–T13 stacking (see Fig. 2c).

Amplitudes of fluctuations of the upper G-quartet (0-0.8 Å) are lower than those of the lower G-quartet (0-1.2 Å) (Figs. 5c and 5d, respectively). This fact can be explained by stabilization of the upper G-quartet by stacking with nitrogen-containing bases of the TGT-loop located above the G-quartet and limiting "from plane" motions.

DISCUSSION

Currently there are two different variants of 15TGT structure, those obtained by X-ray and NMR analysis, to

which two variants of thrombin complex models correspond. In consideration of the current experimental data, giving preference to one of the models is impossible. In the present work the evolution of spatial structure of the two models of thrombin-specific aptamer is comparatively analyzed by the computational simulation method of molecular dynamics.

The parameters of the AMBER parm94/99 force fields are most widely used for molecular dynamics simulation of nucleic acids due to the ability to describe DNA and RNA structures correctly and in short simulation times (up to 10 nsec). However, in multi-nanosecond simulations the presence of irreversible α/γ transition was shown, which causes the breakage of sugar-phosphate

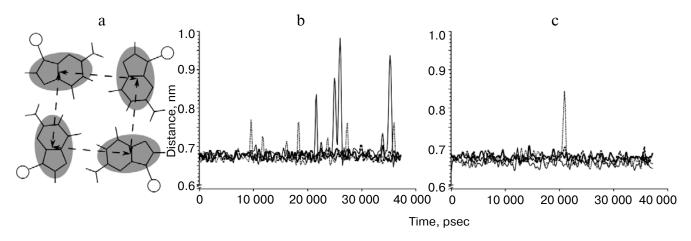


Fig. 4. Motions of the G-quadruplexes "in plane". a) Scheme of measured distances between mass centers of G-quartets bases. Conditional distribution of the masses of guanine heterocyclic bases is shown by gray ovals; for (b) and (c), the maximum of the four distances for each time frame is shown on the ordinate. b) Fluctuations of the upper G-quartet "in plane" are characterized by average value of measured distances of 0.67 nm and low amplitude (in the range of 0.05 nm for the NMR model in both force fields). Periodic leaps of the amplitude characterize insufficient stabilization of the G-quadruplex core of the X-ray model by heterocyclic bases of the TGT-loop. c) Fluctuations of the lower G-quartet "in plane" are also characterized by average value of measured distances of 0.67 nm and low amplitude (in the range of 0.05 nm for the NMR model in both force fields and for the X-ray model in the parm99 force field). A single amplitude leap corresponding to the X-ray structure is observed in the parmbsc0 force field.

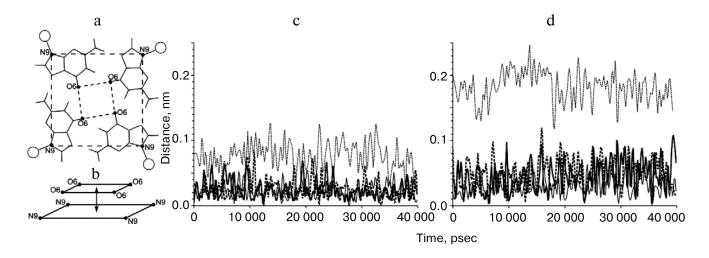


Fig. 5. Motions of the G-quadruplexes "from plane". a) Scheme of the relative position of the two quartets: the outer quartet formed by the N9 atoms and the inner quartet formed by the O6 atoms. b) Scheme of formation of tepee structure: the inner quartet leaves the plane of the G-quartet and planarity of the G-quartet breaks. For (c) and (d), the distance between the two centers of mass of the inner and outer quartets is shown on the ordinate. c) Fluctuations of the upper G-quartet "from plane" are characterized by average value of the measured distance of 0.02 nm and an amplitude 0.05-0.08 nm for the NMR model in both force fields and for the X-ray model in the parm99 force field. For the X-ray model in the parmbsc0 force field, the average is 0.08 nm and the amplitude is 0.11 nm. d) Fluctuations of the lower G-quartet "from plane" are characterized by average value of the measured distance of 0.04 nm and amplitude 0.06-0.12 nm for the NMR model in both force fields and for the X-ray model in the parm99 force field. For the X-ray model in the parmbsc0 force field, the average is 0.18 nm and the amplitude is 0.12 nm.

backbone of DNA [34]. Re-parameterization of the α/γ torsion angles leads to the appearance of the novel version of the parm99 force field called parmbsc0, which can reproduce structures of canonical and noncanonical DNA in very short time ranges (up to microseconds) [23, 24].

For the study of MD of nucleic acids, the GRO-MACS software package is used due to its advantage in

rate and same functionality compared with analogs [26]. To perform the MD of aptamers, the parmbsc0 force field was ported to the GROMACS architecture. In the present work, two models of the 15TGT aptamer based on NMR and X-ray analysis data were studied. All simulations were performed in the parm99 and parmbsc0 force fields.

Distance	Name	X-Ray mod	el of 15TGT	NMR model of 15TGT	
Distance		parm 99	parmbsc0	parm 99	parmbsc0
Upper G-quartet	G1-G6	42.7	28.6	26.7	40.5
	G1-G15	17.2	33.4	16.6	19.5
	G10-G6	13.5	20.5	38.1	15
	G10-G15	26.6	17.5	18.6	25
Lower G-quartet	G2-G5	10.5	0	13.3	19.8
	G2-G14	45.7	55.9	30	18.75
	G11-G5	21.3	42.6	33.2	39
	G11-G14	22.5	1.5	23.5	22.5

Table 2. Distribution of maximal distances between guanines of G-quartets, %

As a result of MD simulation, the NMR model of 15TGT did not significantly change in either of the force fields. The X-ray model of 15TGT changed its initial conformation in both force fields and acquired topology close to the NMR model in a complex with Sr²⁺ in the parm99 force field [10], while in the parmbsc0 force field the geometry of G-quadruplex acquired tepee structure not typical for G-quartet (see Fig. 2).

The fact that the final structure of the X-ray model after 150 nsec of MD simulation differs from the initial one could be considered as an artifact appearing due to incorrect parameterization of the force fields. However, this is disproved by the behavior of the NMR structure very well conforming to the spectroscopic data. In addition, significant mobility of the G2-G14 and G11-G5 base pairs "in plane" (see Table 2) well correlates with an earlier suggested mechanism of thermal denaturation of 15TGT, according to which misfolding of the aptamer structure starts with breakage of hydrogen bonds between the G2-G14 and G11-G5 nucleotide pairs of the lower quartet [35]. The parm99 and parmbsc0 force fields correctly reproduce the structure of the G-quadruplex core [16-18, 23]. However, G-quartets of the X-ray structure clearly behave differently in different force fields (Figs. 2 and 5). Redundant stabilization of the X-ray model in the parm99 force field is likely to be due to non-optimal parameters of several key torsion angles, viz α (O3'-P-O5'-C5') and γ (C3'-C4'-C5'-O5').

As a result of the analysis, it is shown that the most correct spatial structure of the DNA-aptamer 15TGT is described by the NMR model (PDB ID 148d) in the parmbsc0 force field. Despite the fact that the loops of the G-quadruplex structures are very difficult objects for modern force fields [16-18, 23], the topology of the central TGT-loop and the lateral TT-loops of the final structure are very close to the data of NMR spectroscopy.

There are now many variants of the structures of DNA-aptamers to thrombin, which can be structurally based on 15TGT and the properties of which differ from the initial properties for 15TGT [13]. Detailed descrip-

tion of the initial structure enables rational design of antithrombin aptamers for creating novel anticoagulant drugs.

Development of novel computational methods for study of dynamic behaviour of G-quadruplexes extends opportunities for study of these unusual DNA structures that is important not only for study of the structure of G-quadruplex aptamers, but for many other DNA structures, for instance, telomere tails and artificial ionophores.

The authors are deeply obliged to Academician of the Russian Academy of Sciences A. A. Bogdanov for productive discussion of the results and valuable recommendations. The authors appreciate the authorities of the Laboratory of Parallel Informational Technologies of the Research Computing Center of Lomonosov Moscow State University, which has the SKIF MSU "Chebyshev" supercomputer, on which all the computations were performed.

This work was supported by grants from the Ministry of Science (02.512.11.2242/1-08) and Russian Foundation for Basic Research (08-04-01244-a).

REFERENCES

- Nimjee, S. M., Rusconi, C. P., and Sullenger, B. A. (2005) Annu. Rev. Med., 56, 555-583.
- Shamah, S. M., Healy, J. M., and Cload, S. T. (2008) Acc. Chem. Res., 41, 130-138.
- 3. Tuerk, C., and Gold, L. (1990) Science, 249, 505-510.
- Bock, L. C., Griffin, J. L., Latham, C. A., Vermaas, E. H., and Toole, J. J. (1992) *Nature*, 355, 564-566.
- Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A., and Feigon, J. (1993) *Proc. Natl. Acad. Sci. USA*, 90, 3745-3749.
- Tasset, D. M., Kubik, M. F., and Steiner, W. (1997) J. Mol. Biol., 272, 688-698.
- Padmanabhan, K., Padmanabhan, K. P., Ferrara, J. D., Sadler, J. E., and Tulinsky, A. (1993) *J. Biol. Chem.*, 268, 17651-17654.

- Schultze, P., Macaya, R. F., and Feigon, J. (1994) J. Mol. Biol., 235, 1532-1547.
- 9. Padmanabhan, K., and Tulinsky, A. (1996) *Acta Crystallogr. Biol. Crystallogr.*, **52**, 272-282.
- Mao, X., Marky, L. A., and Gmeiner, W. H. (2004) J. Biomol. Struct. Dyn., 22, 25-33.
- Marathias, V. M., and Bolton, P. H. (2000) Nucleic Acids Res., 28, 1969-1977.
- Kelly, J. A., Feigon, J., and Yeates, T. O. (1996) J. Mol. Biol., 256, 417-422.
- 13. Ikebukuro, K., Okumura, Y., Sumikura, K., and Karube, I. (2005) *Nucleic Acids Res.*, **33**, 108.
- 14. Hecke, A., and Mayer, G. (2005) *J. Am. Chem. Soc.*, **127**, 822-823.
- Mendelboum, R. A., Horvath, J., Aradi, Z., Bagoly, F., Fazakas, Z., et al. (2008) *J. Thromb. Haemost.*, 6, 1764-1771.
- Fadrna, E., Spackova, N., Stefl, R. J., Koca, T. E., Cheatham,
 T. E., and Sponer, J. (2004) *Biophys. J.*, 87, 227-242.
- Fadrna, N., Spackova, J., Sarzynska, J., Koca, M., Orozco, M., Cheatham, T. E., et al. (2009) J. Chem. Theory Comput., 5, 2514-2530.
- 18. Sponer, J., and Spackova, N. (2007) *Methods*, **43**, 278-290.
- Hazel, P., Parkinson, G. N., and Neidle, S. (2006) Nucleic Acids Res., 34, 2117-2127.
- Haider, S., Parkinson, G. N., and Neidle, S. (2008) Biophys. J., 95, 296-311.
- 21. Neidle, S. (2009) Curr. Opin. Struct. Biol., 19, 239-250.

- Pagano, B., Martino, L., Randazzo, A., and Giancola, C. (2008) *Biophys. J.*, 94, 562-569.
- Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham,
 T. E., et al. (2007) *Biophys. J.*, 92, 3817-3829.
- 24. Perez, A., Luque, F. J., and Orozco, M. (2007) *J. Am. Chem. Soc.*, **129**, 14739-14745.
- 25. Jayapal, P., Mayer, G., Heckel, A., and Wennmohs, F. (2009) *J. Struct. Biol.*, **166**, 241-250.
- Lindahl, E., Hess, B., and van der Spoel, D. (2001) J. Mol. Mod., 7, 306-317.
- Van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005) *J. Comput. Chem.*, 26, 1701-1718.
- 28. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) J. Chem. Theory Comput., 4, 435-447.
- Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., et al. (1995) J. Am. Chem. Soc., 117, 5179-5197.
- 30. Sorin E. J., and Pande, V. S. (2005) *Biophys. J.*, **88**, 2472-2493.
- 31. Bussi, G., Donadio, D., and Parrinello, M. (2007) *J. Chem. Phys.*, **126**, 014101-014107.
- 32. Darden, T. D., York, D., and Pedersen, L. (1993) *J. Chem. Phys.*, **98**, 10089-10092.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.*, 79, 926-935.
- 34. Orozco, M., Noy, A., and Perez, A. (2008) *Curr. Opin. Struct. Biol.*, **18**, 185-193.
- 35. Mao, X., and Gmeiner, W. H. (2005) *Biophys. Chem.*, **113**, 155-160.