

Theoretical study of peptide nucleic acid with (2'R,4'R)-Prolyl-(1S,2S)-2-aminocyclobutanecarboxylic acid backbone binding to DNA and self-pairing

Amphawan Maitarad¹, Nattawee Poomsuk¹, Chittima Laohpongspaisan¹, Khatcharin Siriwong^{1,*}

¹ Materials Chemistry Research Center, Department of Chemistry and Center for Innovation in Chemistry, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand

*Corresponding Author: skhatcha@kku.ac.th

Abstract

The binding ability of pyrrolidinyl peptide nucleic acid (PNA) binding to its complementary DNA and self-pairing has been studied experimentally. However, a detailed understanding of the binding property is still unclear due to the lack of their crystallographic data. In this work, structural and energetic properties of PNA-DNA and PNA-PNA duplexes were studied using molecular dynamics (MD) simulations and quantum calculations. The studied pyrrolidinyl PNA backbone was (2'R,4'R)-prolyl-(1S,2S)-2-aminocyclobutanecarboxylic acid. MD simulations of three different forms (A-, B- and P-form) were performed in order to investigate the probable duplex conformation. As the results, PNA-PNA duplex exhibited the structural feature between A- and P-type conformations, while PNA-DNA double helix clearly showed the characteristic of B-form. In addition, quantum calculations revealed that the interaction of PNA-DNA duplex was larger than that of PNA-PNA duplex, indicating higher intrinsic stability of PNA-DNA compared to PNA-PNA double strand. This research may lead to a design of PNA for further applications.

Keywords: Peptide nucleic acid; Molecular dynamics simulation; Quantum calculation

1. Introduction

The characterization of deoxyribonucleic acid (DNA) duplex in terms of structure and stability was completely studied in the last decade of the twentieth century. As DNA is rapidly digested by nuclease enzyme, its applications as molecular tools in molecular biology and biotechnology are limited. The biological stability is therefore the aim of chemical modification of oligonucleotide. In addition, the study of binding affinity and specificity is still retained [1–3].

Peptide nucleic acid (PNA) is a DNA analog in which the sugar-phosphate backbone of DNA is replaced by poly(*N*-aminoethylglycine) with the nucleobases attached through a methylenecarbonyl linkage at the glycine nitrogen, namely aegPNA (see Figure 1), which was first introduced by Nielsen and co-workers [4]. PNA is very specific interactions with DNA, RNA and self-pairing. Moreover, it is also more chemically and biochemically stable than natural nucleic acids. Therefore, its applications, especially in genetic therapeutic agent and nanomaterials, have been widely studied [5–9].

In recent years, Vilaivan and co-workers have synthesized a new series of pyrrolidinyl PNA backbones based on the development of the selectivity and binding affinity with its complementary sequences [10–15]. It is interesting that pyrrolidinyl PNA with (2'R,4'R)-prolyl-(1S,2S)-2-aminocyclobutanecarboxylic acid backbone, namely acbcPNA (Figure 1), can form PNA self-hybrid with thermal stability (T_m) of 45.7 °C. On the contrary, its binding affinity with DNA revealed more stable with T_m value of 77.7 °C [14]. Unfortunately, to date their three-dimensional structures are not successfully verified by NMR or crystallographic data. Thus, an understanding of the structural and binding properties is unclear.

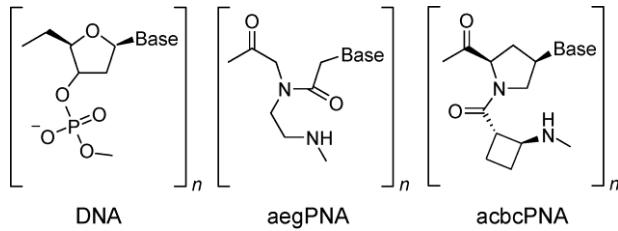


Figure 1 The monomeric units of nucleic acids.

In this study, a theoretical approach based on molecular dynamics (MD) simulation and quantum calculation was used to fulfill the understanding of structural and energetic properties of PNA-PNA and PNA-DNA duplexes with acbcPNA backbone. To the best of our knowledge, this is the first report for MD simulation and interaction energy of acbcPNA system.

2. Materials and Methods

2.1 MD simulations

MD simulations of 10-mer acbcPNA with base sequence of *N*-GCGAATTCGC-*C* binding to its complementary DNA strand and self-pairing in antiparallel direction were performed in an aqueous solution. For PNA-PNA system, the simulations were started with A-, B- and P-type conformations in order to investigate the probable duplex conformation. According to a previous study [16,17], only a typical B-form was used as starting structure for PNA-DNA duplex. All starting structures were generated as follows. First, four types of acbcPNA monomeric unit (adenine (A), cytosine (C), guanine (G) and thymine (T)) were prepared by one-to-one mapping of acbcPNA atoms onto DNA atoms of the corresponding nucleotide. Next, DNA-DNA duplexes with three different conformations were constructed using 3DNA program [18] based on the helical parameters listed in Table 1 (other parameters were set to zero). The DNA-DNA duplexes were then modified to PNA-PNA and PNA-DNA hybrids by replacing DNA units with the prepared PNA units based on the coordinates of the nucleobase atoms. Finally, the backbone of PNA units was linked to each other and then re-optimized with a molecular mechanics method. Since acbcPNA is not standard residues, its force field parameters are not available. They were therefore prepared by using the restrained electrostatic potential (RESP) method as described elsewhere [17].

Table 1 Helical parameters of A-, B- and P-type duplexes.

	Slide (\AA)	Rise (\AA)	Twist ($^{\circ}$)
A-form ^a	0.0	2.6	32.7
B-form ^a	0.0	3.4	36.0
P-form ^b	-2.5	3.3	20.0

^a The data were taken from Ref. [1].

^b The data were taken from [18].

All MD simulations were performed in the isothermal-isobaric ensemble ($T = 300 \text{ K}$ and $P = 1 \text{ atm}$) using AMBER12 program [20] with standard conditions, i.e. Parm10 force field combined with the prepared force field parameters, TIP3P water box extended by 10 \AA in each direction from the duplex [21], time step of 2 fs, periodic boundary conditions, cutoff of 9 \AA for nonbonded interactions, SHAKE algorithm employed to constrain all hydrogen bonds [22], and particle-mesh Ewald algorithm for treatment of electrostatic interactions [23]. For PNA-DNA duplex, the system was neutralized with 9 sodium counterions. The simulation protocols for heating and equilibration steps followed a previous work [17]. The unrestrained MD simulations were performed for 10 ns and the coordinates were stored every 1 ps, yielding 10000 MD structures.

2.2 Interaction energy

The MD trajectory was employed in order to investigate the interaction between two strands of duplex. The interaction energy (ΔE) defined as the energy difference between duplex and its single strands is expressed as $\Delta E = E_{\text{duplex}} - (E_{\text{strand 1}} + E_{\text{strand 2}})$ where E_{duplex} , $E_{\text{strand 1}}$ and $E_{\text{strand 2}}$ are the energies of duplex, strand 1 and strand 2, respectively. These energies were calculated using Gaussian 09 program [24] with B3LYP/6-31G(d,p) method.

3. Results and Discussion

3.1 MD trajectory analysis

The root mean squared deviations (RMSDs) were used to evaluate the conformational stability of the simulated system. The RMSDs of MD structures were analyzed by taking into account all duplex atoms with respect to their starting structure.

For PNA-DNA duplex, the RMSD rapidly increased to about 2 Å, and then remained stable for the rest of the simulation (Figure 2a). The average RMSD was 2.1 Å. When 5 MD structures sampled every 2 ns were superimposed with the starting B-form structure, it was found that the simulated structures were well-defined duplex as shown in Figure 2b. These results indicated that the obtained MD trajectory was stable in geometrical term, and the sampled configurational space corresponded to a double helix.

For PNA-PNA system, the RMSDs of all trajectories with respect to their starting conformations showed a sharp increase to 3 Å in the first 200 ps as shown in Figure 3a. The RMSD of the A-trajectory was rather stable after 2 ns with highly average value of about 5.2 Å. For the B-trajectory, it exhibited a divergence from the starting B-type duplex. The RMSD fluctuated up to about 8 Å during 4–5 ns, although the average RMSD (~4.8 Å) was smaller than that of the A-trajectory. Among all MD trajectories, the simulation started with P-form seemed to provide the most stable structure. Its RMSD fluctuated around 4 Å throughout the simulation time with the average value of 4.2 Å. The superimposition of selected MD structures with their corresponding starting structures revealed that the MD structures started with P-form showed a characteristic of double helix, whereas the MD structures started with A- and B-type duplexes did not express a feature of double helix (Figure 3b). This implied that the probable conformation of PNA-PNA duplex was a canonical P-form. Hence, for further investigation, only P-trajectory was employed.

To investigate the structural feature of the simulated structures in term of base pair orientation, the helical parameters were evaluated using 3DNA program [18] based on the last 8000 MD structures. Two terminal base pairs were excluded to avoid the end effect. As listed in Table 2, the base-pair parameters (shear, stretch, stagger, opening and buckle) of both PNA-PNA (P-trajectory) and PNA-DNA duplexes were slightly deviated from zero, whereas the propeller-twist was deviated about -13° . This indicated that the base pairs of simulated systems were almost in planarity.

Considering base-step parameters, shift and tilt were fluctuated around zero, while roll parameter was deviated about 13° and 6° for PNA-PNA and PNA-DNA duplexes, respectively. The obtained values of these parameters were normally found in common duplexes [26]. However, it is worth noting that only slide, rise and twist are the major parameters used to describe the duplex conformations. As the results, PNA-DNA duplex obviously exhibited B-form characteristic (see Tables 1 and 2) which was similar to the conformation of PNA with (2'R,4'R)-prolyl-(1S,2S)-2-aminocyclopentanecarboxylic acid backbone binding to DNA as reported previously [17]. For PNA-PNA duplex, the slide and rise values revealed a P-form characteristic, whereas the twist parameter exhibited the conformation between A-form and P-form. The average MD structures of both duplexes were illustrated in Figure 4.

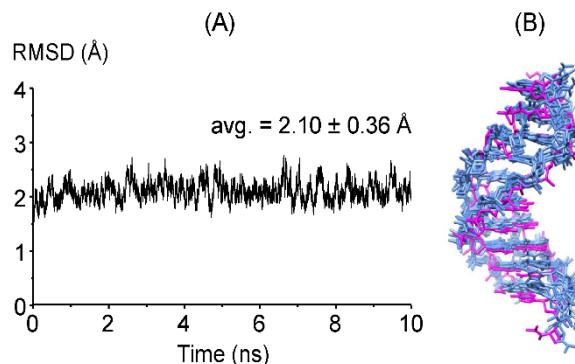


Figure 2 (A) RMSD plot of PNA-DNA duplex and (B) superimposition of 5 MD derived structures (in indigo) with their starting structure (in pink). The image of molecular structures was produced by UCSF Chimera program [25].

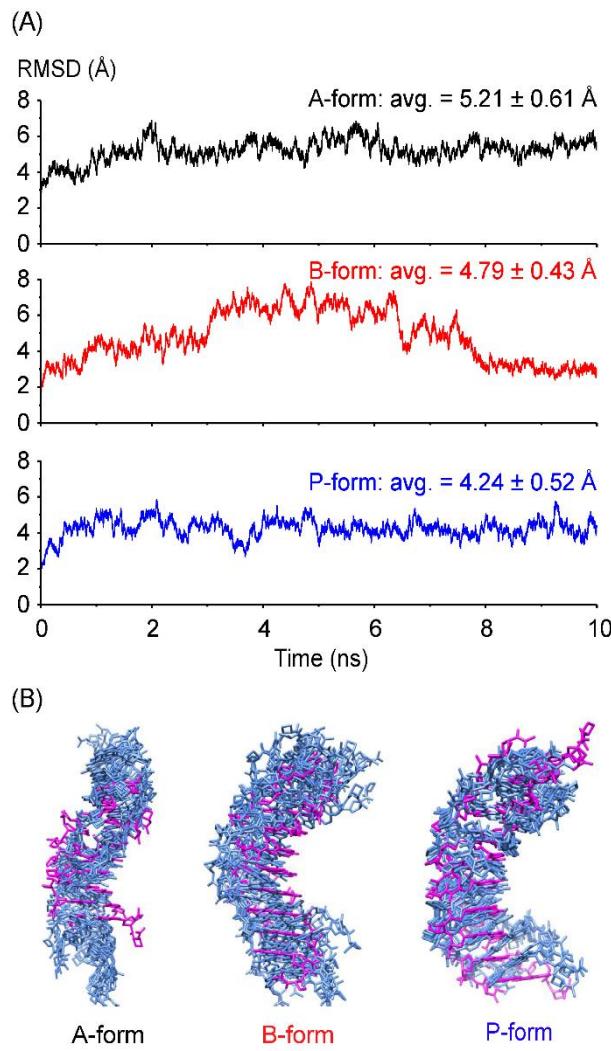


Figure 3 (A) RMSD plots of PNA-PNA duplexes of various forms and (B) superimposition of 5 MD derived structures (in indigo) with their starting structure (in pink) for each form.

Table 2 The average helical parameters and their standard deviations (in parentheses).

Parameters	PNA-PNA	PNA-DNA
Shear (Å)	0.0 (0.3)	0.1 (0.3)
Stretch (Å)	-0.1 (0.1)	0.0 (0.1)
Stagger (Å)	-0.1 (0.2)	-0.1 (0.4)
Buckle (°)	0.3 (6.8)	-1.7 (9.3)
Prop-Tw (°)	-13.8 (5.5)	-13.4 (7.2)
Opening (°)	-1.1 (1.6)	2.7 (4.0)
Shift (Å)	0.0 (0.3)	-0.2 (0.4)
Slide (Å)	-1.1 (2.2)	-0.7 (0.5)
Rise (Å)	3.4 (1.3)	3.4 (0.3)
Tilt (°)	0.2 (1.7)	2.3 (4.7)
Roll (°)	12.6 (8.0)	5.9 (4.9)
Twist (°)	27.5 (7.8)	36.1 (3.9)

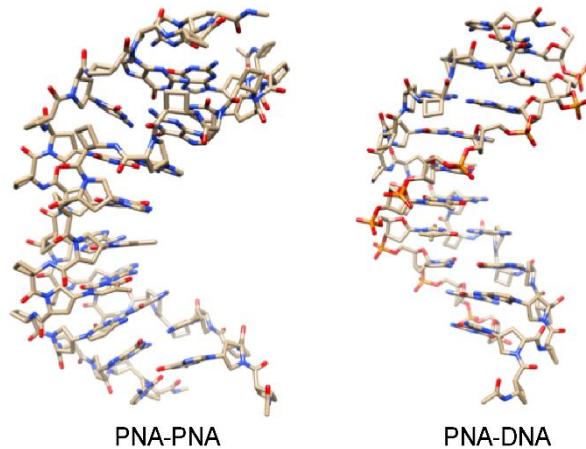


Figure 4 Average structures of (A) PNA-PNA, and (B) PNA-DNA duplexes.

3.2 Interaction energy

For each duplex, the interaction energy of 16 MD structures was calculated using B3LYP/6-31G(d,p) method. These structures were sampled every 500 structures from the last 8 ns MD trajectory. It was found that PNA-DNA system yielded higher interaction compared to PNA-PNA duplex, in agreement with the experimental thermal stability (T_m) (Table 3). Actually, the interaction energy cannot be considered as the binding affinity for comparing with the T_m of duplex. Instead, the binding free energy, which is the combination of interaction energy, solvation free energy and entropy terms, should be used. Nevertheless, the larger negative value of the interaction energy may imply that PNA-DNA is higher intrinsically stable than PNA-PNA double strand.

Table 3 The interaction energy (ΔE) calculated with B3LYP/6-31G(d,p) method.

Systems	ΔE (kcal/mol)	T_m (°C) ^a
PNA-PNA	-203.3 ± 4.0	45.7
PNA-DNA	-211.0 ± 3.4	77.7

^a The experimental T_m values were taken from Ref. [14].

3.3 Intra-strand base stacking

The intra-strand base stacking of adjacent bases was evaluated based on the average MD structure and two terminal base pairs were not taken into account to avoid the end effect. The results showed that almost intra-strand adjacent bases of PNA-PNA duplex were found to overlap, except A₅-T₆ and A₁₅-T₁₆ pairs, while all pairs of PNA-DNA duplex exhibited the base stacking. This suggested that the base pairs of PNA-DNA duplex were better orientation to retain the hydrophobic interaction, i.e. π - π stacking, between adjacent bases than those of PNA-PNA duplex. This may affect the interaction of duplex, leading to the lower interaction energy of PNA-PNA duplex as described above.

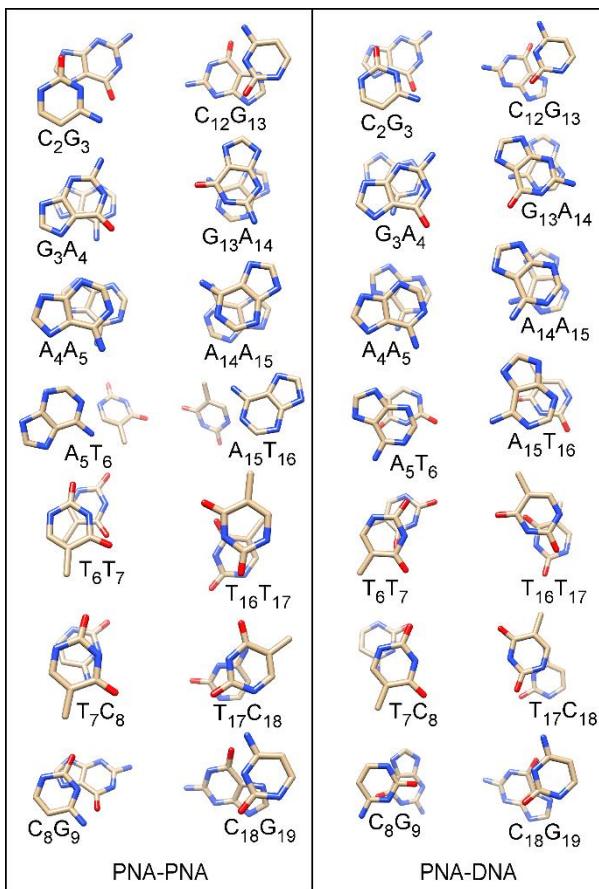


Figure 5 Intra-strand base stacking of PNA-PNA and PNA-DNA duplexes.

4. Conclusions

In this work, the molecular modeling approach based on MD simulation and quantum calculation was used to fulfill the understanding of structural and energetic properties of PNA-PNA and PNA-DNA duplexes, where the PNA backbone was ($2'R,4'R$)-prolyl-(1*S*,2*S*)-2-aminocyclobutanecarboxylic acid. MD simulations revealed that the PNA-DNA duplex provided stable trajectory in geometrical term and the simulated structures clearly showed the characteristic of B-form, while the PNA-PNA duplex exhibited the conformational feature between A- and P-type duplexes. In addition, the orientation of base stacking in PNA-DNA was better than that in PNA-PNA duplex which could affect the hydrophobic interaction between adjacent bases. It was found that the interaction energy of PNA-DNA duplex was larger negative than that of PNA-PNA duplex indicating higher intrinsic stability of PNA-DNA compared to PNA-PNA double strand. Such finding may lead to a design of PNA for further applications.

5. Acknowledgments

This work was financially supported by Materials Chemistry Research Center, Khon Kaen University, and Uttaradit Rajabhat University. National e-Science Infrastructure Consortium (URL: <http://www.e-science.in.th>) is acknowledged for computing resources.

6. References

- [1] Blackburn, G.M., Gait, M.J., Loakes, D., Williams, D.M., 2006. Nucleic acids in chemistry and biology, 3rd ed. RSC Publishing, Cambridge.
- [2] Dahm, R., 2005. Friedrich Miescher and the discovery of DNA. *Developmental Biology* 278, 274–288.
- [3] Dahm, R., 2008. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Human Genetics* 122, 565–581.
- [4] Nielsen, P.E., Egholm, M., Berg, R.H., Buchardt, O., 1991. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* 254, 1497–1500.

- [5] Egholm, M., Buchardt, O., Nielsen, P.E., Berg, R.H., 1992. Peptide nucleic acids (PNA): oligonucleotide analogs with an achiral peptide backbone. *Journal of the American Chemical Society* 114, 1895–1897.
- [6] Hyrup, B., Nielsen, P.E., 1996. Peptide nucleic acids (PNA): synthesis, properties and potential applications. *Bioorganic & Medicinal Chemistry* 4, 5–23.
- [7] Ray, A., Nordén, B., 2000. Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. *The FASEB Journal* 14, 1041–1060.
- [8] Porcheddu, A., Giacomelli, G., 2005. Peptide nucleic acids (PNAs), a chemical overview. *Current Medicinal Chemistry* 12, 2561–2599.
- [9] Gambari, R., 2014. Peptide nucleic acids: a review on recent patents and technology transfer. *Expert Opinion on Therapeutic Patents* 24, 267–294.
- [10] Vilaivan, T., Suparpprom, C., Harnyuttanakorn, P., Lowe, G., 2001. Synthesis and properties of novel pyrrolidinyl PNA carrying β -amino acid spacers. *Tetrahedron Letters* 42, 5533–5536.
- [11] Vilaivan, T., Srisuwannaket, C., 2006. Hybridization of pyrrolidinyl peptide nucleic acids and DNA: selectivity, base-pairing specificity, and direction of binding. *Organic Letters* 8, 1897–1900.
- [12] Taechalertpaisarn, J., Sriwarom, P., Boonlua, C., Yotapan, N., Vilaivan, C., Vilaivan, T., 2010. DNA-, RNA- and self-pairing properties of a pyrrolidinyl peptide nucleic acid with a (2'R,4'S)-prolyl-(1S,2S)-2-aminocyclopentanecarboxylic acid backbone. *Tetrahedron Letters* 51, 5822–5826.
- [13] Vilaivan, C., Srisuwannaket, C., Ananthanawat, C., Suparpprom, C., Kawakami, J., Yamaguchi, Y., Vilaivan, T., 2011. Pyrrolidinyl peptide nucleic acid with α/β -peptide backbone: a conformationally constrained PNA with unusual hybridization properties. *Artificial DNA, PNA & XNA* 2, 50–59.
- [14] Mansawat, W., Vilaivan, C., Balázs, Á., Aitken, D.J., Vilaivan, T., 2012. Pyrrolidinyl peptide nucleic acid homologues: effect of ring size on hybridization properties. *Organic Letters* 14, 1440–1443.
- [15] Vilaivan, T., 2015. Pyrrolidinyl PNA with alpha/beta-dipeptide backbone: from development to applications. *Accounts of Chemical Research* 48, 1645–1656.
- [16] Siriwong, K., Chuichay, P., Saen-oon, S., Suparpprom, C., Vilaivan, T., Hannongbua, S., 2008. Insight into why pyrrolidinyl peptide nucleic acid binding to DNA is more stable than the DNA-DNA duplex. *Biochemical and Biophysical Research Communications* 372, 765–771.
- [17] Poomsuk, N., Siriwong, K., 2013. Structural properties and stability of PNA with (2'R,4'R)- and (2'R,4'S)-prolyl-(1S,2S)-2-aminocyclopentanecarboxylic acid backbone binding to DNA: a molecular dynamics simulation study. *Chemical Physics Letters* 588, 237–241.
- [18] Rasmussen, H., Kastrup, J.S., Nielsen, J.N., Nielsen, J.M., Nielsen, P.E., 1997. Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å resolution. *Nature Structural & Molecular Biology* 4, 98–101.
- [19] Lu, X.J., Olson, W.K., 2003. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research* 31, 5108–5121.
- [20] AMBER 12, 2012. University of California, San Francisco.
- [21] Jorgensen, W., 1981. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *Journal of the American Chemical Society* 103, 335–340.
- [22] Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C., 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *Journal of Computational Physics* 23, 327–341.
- [23] Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *Journal of Chemical Physics* 98, 10089–10092.
- [24] Gaussian 09 Revision B.01, 2009. Gaussian, Wallingford CT.
- [25] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605–1612.
- [26] Saenger, W., 1984. *Principles of nucleic acid structure*, Springer-Verlag, New York.