Crystal Structures of the Two Isomorphous A-DNA Decamers d(GTACGCGTAC) and d(GGCCGCGGCC)

Taegyun Kim, Taek Hun Kwon, Hyesun Jung, Ja Kang Ku, Muttaiya Sundaralingam,† and Changill Ban*

*Department of Chemistry and Division of Molecular and Life Sciences (BK21), Pohang University of Science and Technology, Pohang, Kyungbuk 790-784, Korea. *E-mail: ciban@postech.ac.kr*
†Department of Chemistry, The Ohio State University, 1060 Carmack Rd., Columbus, Ohio 43210, USA
Received January 9, 2006

To study the effect of sequence on DNA structure, the two decamer crystal structures one alternating, d(GTACGCGTAC), and the other non-alternating, d(GGCCGCGGCC), were solved. Crystals of both decamers belong to the hexagonal space group $P6_122$, with one strand in the asymmetric unit. The unit cell constants of the alternating decamer are a = b = 39.26 Å, c = 77.70 Å. The structure was refined with 1,828 reflections from 8.0 to 2.0 Å resolution to an R value of 21.3% with all DNA atoms and 63 water molecules. The isomorphous non-alternating decamer had unit cell dimensions of a = b = 39.05 Å, c = 82.15 Å. The structure was refined with 2,423 reflections from 8.0 to 2.0 Å resolution to a final R value of 22.2% for all DNA atoms and 65 water molecules. Although the average helical parameters of the decamers are typical of A-DNAs, there are some minor differences between them. The helical twist, rise, x-displacement, inclination and roll alternate in the alternating decamer, but do not in the non-alternating decamer. The backbone conformations in both structures show some differences; the residue G(7) of the alternating decamer is *trans* for α and γ while the *trans* conformations are observed at the residue G(8) of the non-alternating decamer.

Key Words : X-Ray crystallography, A-DNA decamer, Conformational flexibility, Helical parameter

Introduction

The atomic resolution structures of oligonucleotides, particularly those close to or greater than a full turn, provide a wealth of information related to the property of duplex DNA in the cell. However, since the entire periphery of the oligonucleotide duplex is exposed and accessible to the environment, it is argued whether in the crystal DNA conformation of the global and local is affected by the lattice packing forces or the sequence itself or both. A series of studies on the different crystal forms with the same sequence show that the crystal packing mainly forces the local conformational differences.[1-4] In contrast, several other studies show that the local conformation of oligonucleotides can be affected by the sequence itself,[5-7] although the correlation between the sequence and the local conformation is not clear.

In the first alternating A-DNA decamer structure d(GCACGCGTGC) which has previously reported, we showed that its structural feature can be modulated by the sequence alternation.[8] However, verifying the influence of the sequence alternation on the DNA conformation was hard since the decamer was only case crystallized as A-DNA for an alternating sequence. To elucidate the precise influence of the sequence on the DNA conformation, one should study isomorphous structures with different sequences under the same crystal environment. This paper investigated the two decamers d(GTACGCGTAC) and d(GGCCGCGGCC) that were designed and crystallized in the same hexagonal space group $P6_122$ as the previous decamer d(GCACGCGTGC).[8] From the single crystal structures of these decamers, the influence of the alternating and non-alternating sequence on the DNA conformation is revealed.

Experimental Section

(a) Synthesis, crystallization, and data collection. The decamers were synthesized and purified by the previously described protocol.[8] The best crystals of d(GTACGCGTAC) were obtained using the hanging drop vapor diffusion method in the presence of 1 mM of the decamer (single strand concentration), 20 mM of sodium cacodylate buffer (pH 6.0) and 0.5 mM of cobalt hexamine against 45% MPD in the reservoir. Large bi-pyramidal crystals grew at room temperature over four days. The crystals belonged to the hexagonal space group $P6_122$, with unit cell dimensions of a = b = 39.26 Å and c = 77.70 Å. The unit cell indicated that it is isomorphous to other known hexagonal A-DNA decamer structures[1,8,9] with a single strand in the asymmetry unit. The non-alternating decamer d(GGCCGCGGCC) formed several large bi-pyramidal crystals at room temperature in four days under the following conditions: 1 mM of DNA decamer, 40 mM of sodium cacodylate buffer (pH = 7.0) and 0.5 mM of cobalt hexamine against 60% of MPD. The crystals were isomorphous to the alternating decamer d(GTACGCGTAC) with space group $P6_122$ and unit cell dimensions a = b = 39.05 Å and c = 82.15 Å.

For the d(GTACGCGTAC), the largest crystal with dimensions 0.4 × 0.6 × 0.8 mm was mounted in a Lindemann capillary and 2.0 Å resolution x-ray data were collected and processed as described earlier.[8] Of the 16,940 reflections collected, 2,656 were unique, with a $R_{symm}(F)$ = 3.4%. Of

these, 1,828 reflections with $F \geq 5.0\sigma(F)$ in the resolution range of 8.0 to 2.0 Å were used in the refinement. For the d(GGCCGCGGCC), crystal of dimensions 0.3 × 0.4 × 0.6 mm was mounted and 2.0 Å resolution data were collected and processed by the same procedure as above. Of the 19,362 reflections collected, 2,916 reflections were unique, with a $R_{symm}(F)$ of 4.3%. Of these, 2,697 reflections with $F \geq 5.0\sigma(F)$ in the resolution range of 8.0 to 2.0 Å were used in the refinement.

(b) Structure solution and refinement. The atomic coordinates of the isomorphous A-DNA decamer d(GCACGCGTGC)[8] were used as the starting model to solve both the decamer structures. After a rigid body refinement of d(GTACGCGTAC) using the CNS program,[10] the R value dropped to 35.3% for data between 8.0-2.5 Å resolution (RMS = 1.41 Å). The first cycle of positional and isotropic thermal factor refinement reduced the R value to 27.2%. At this point, the correct bases were introduced from omit 3Fo-2Fc and Fo-Fc difference maps, which gave an R value of 24.7%. Further refinement followed by simulated annealing (initial temperature of 4,000 °C) using the 8.0 to 2.0 Å data resulted in an R value of 22.4%. At this stage, all 1828 reflections with $F \geq 5\sigma(F)$ in the resolution range 8.0 to 2.0 Å were included and 3Fo-2Fc omit maps were calculated by omitting one base at a time. Again the model was fitted to the omit maps and 18 water molecules were picked. Picking additional water molecules in further rounds of refinement gave a final R value of 21.3%. The final model contained 63 water molecules and 202 nucleotide atoms in a single strand.

The refinement of the non-alternating decamer d(GGCCGCGGCC) was carried out by the same procedure as above. The final R value was 22.2% for the 2,423 reflections with $F \geq 5\sigma(F)$ in the range of 8.0 to 2.0 Å resolution and the model contained 202 DNA atoms and 65 water molecules. The refinement parameters of the decamers d(GTACGCGTAC) and d(GGCCGCGGCC) are listed in Table 1. The atomic coordinates of both structures will be deposited with the Nucleic Acid Database.[11]

Results and Discussion

Both DNA duplexes (Figure 1) have a narrow and deep major groove and a wide and flat minor groove, with the base pairs inclined to the helix axis, characteristic of A-DNA. In the non-alternating decamer d(GGCCGCGGCC) the minor groove width varies from 9.5 (phosphorous to phosphorous distance less 5.8 Å) to 10.1 Å (average 9.8 Å), being narrow at the middle (9.5 Å) and broad at the termini (10.1 Å). On the contrary, in the alternating decamer d(GTACGCGTAC), it is uniformly broad at both the termini (10.5 Å) and the middle (10.4 Å), with an average value of 10.3 Å. The helical parameters of both structures are shown in Figure 2, and they are in the range of other known A-DNAs. Also, in both decamer structures all the sugar and glycosyl torsion angles adopt the C3'-*endo* pucker and *anti* conformation, respectively.

To substantiate our observation for the sequence effect on the A-DNA conformation, it is necessary to compare the helical parameters of the two isomorphous structures assuming to have identical packing environment. Although the average values of the helical parameters are very similar in both structures, sequence dependent variations were seen

Table 1. Crystal and refinement parameters of d(GTACGCGTAC) and d(GGCCGCGGCC)

	d(GTACGCGTAC)	d(GGCCGCGGCC)
Unit cell dimensions (Å)		
a=b	39.26	39.05
c	77.70	82.15
Unit cell volume (Å^3)	103717.548	108487.6435
Space group	$P6_122$	$P6_122$
Molecule/asymmetry unit (Z-value)	single strand	single strand
Volume per base pair (Å^3)	1,729	1,808
Resolution range (Å)	8.0-2.0	8.0-2.0
Number of reflections ($F \geq 2\sigma(F)$)	1,828	2,423
used in refinement		
Final R-value (%)	21.3	22.2
RMS deviation from ideal geometry		
parameter file used	dna-rna-rep.dna*	dna-rna-rep.dna*
bond lengths (Å)	0.027	0.013
bond angles (°)	3.9	3.6
Final model		
nucleic acid atoms	202	202
water molecules	63	65
Average thermal parameter (Å^2)		
nucleotide	22	26
water molecules	63	67

*Parameter file are calculated from CNS program.

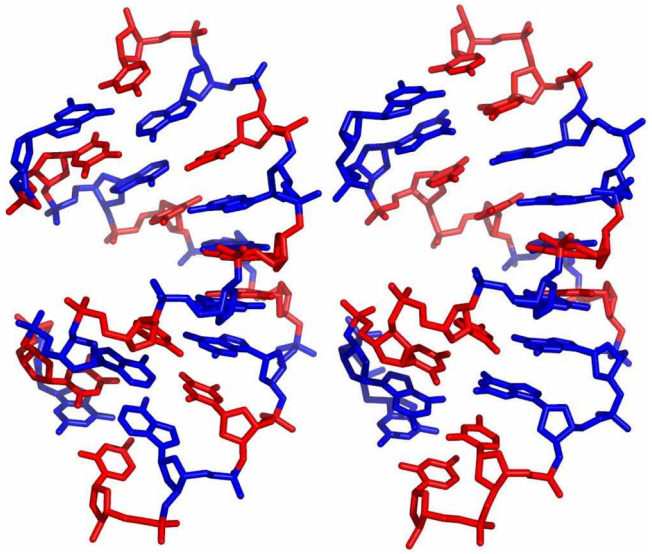

Figure 1. Overall conformations of the A-DNA decamer duplexes d(GTACGCGTAC) (left) and d(GGCCGCGGCC) (right), viewed normal to the dyad in the plane of the page. Two A-DNA decamer duplexes indicated purine and pyrimidine base as blue and red color, respectively.

in some helical parameters (Figure 2). The helical twist, rise, x-displacement, inclination and roll of the alternating decamer structure show an alternation; the twists for the py-pu steps are higher than the pu-py steps. A similar pattern is also found in the rise and inclination of the base pairs. The roll angles also alternate, but the pattern is opposite to those of the x-displacement; the pu-py steps are lower than the py-pu steps found in the previous A-DNA decamer d(GCACG-CGTGC).[8] The roll angles are generally negative and dampened at the termini, while this decamer shows all positive roll angles and the C(4)pG(5) and C(6)pG(7) steps show highly positive roll angles of 18.12°.

Unlike the alternating decamer, the helical parameters of the non-alternating decamer d(GGCCGCGGCC) do not show the alternation, as seen in other isomorphous A-DNA decamers d(GCGGGCCCGC)[1] and d(ACCGGCCGGT).[9] The rise and slide in the central region (C(3)···G(8)) show some alternation with the high values for the py-pu steps and the low values for the pu-py steps. Interestingly, the C(3)pC(4) and G(7)pG(8) steps, which interrupt the alternation of the helical parameters, show the lowest helical twist of 25.1° and the highest negative slide of −1.9 Å.

When the common atoms of the two decamers are superposed, the overall RMS deviation is 0.7 Å (Figure 3). However, the RMS deviation of both structures is bounded at the sugar-phosphate backbones of the 7th (0.7 Å) and the 8th (1.0 Å) residues, indicating that in those regions the backbone conformations of both structures are different. The average torsion angles for the present structure are calculated using the 3DNA program (created and maintained by Dr. Xiang-Jun Lu). In the alternating decamer, the 7th residue

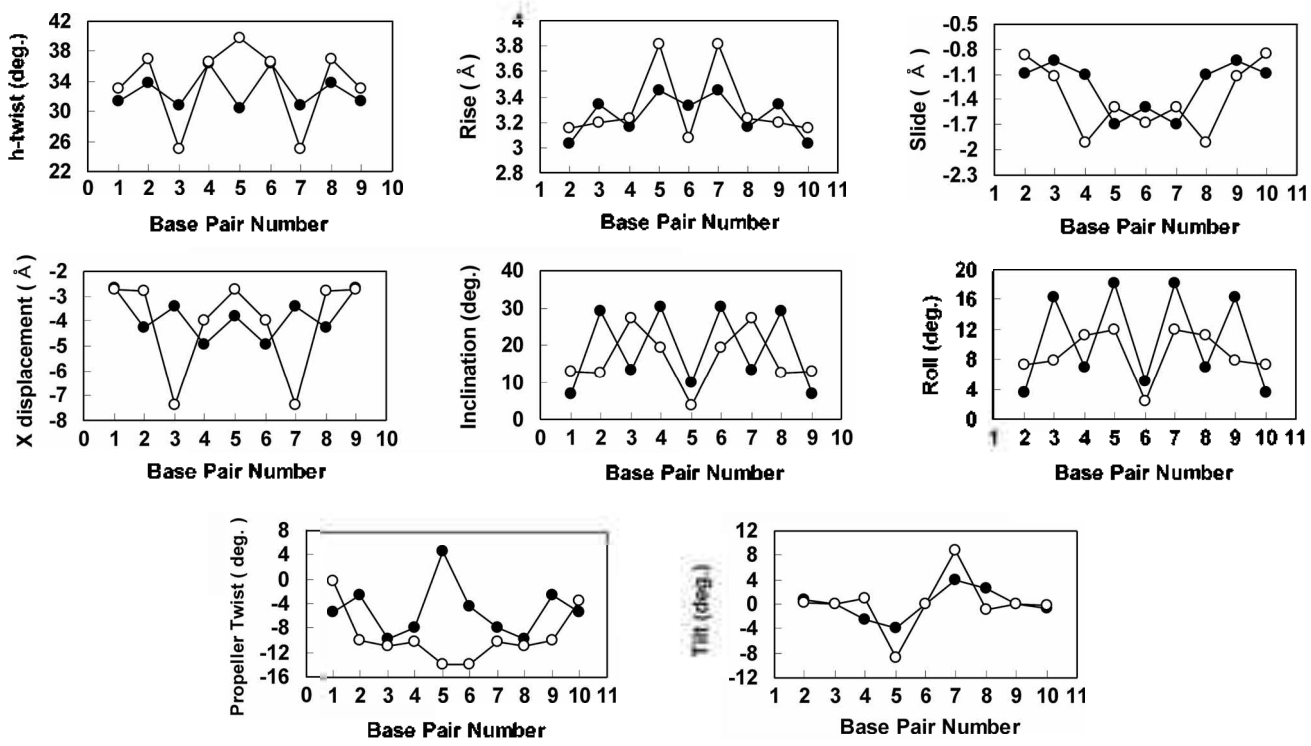

Figure 2. Helical parameter comparisons for the two isomorphous A-DNA decamers d(GTACGCGTAC) (dark circles) and d(GGCCGCGGCC) (open circles). In the alternating decamer, the helical twist and roll show an alternation, but those of the non-alternating decamer are not. The averaged helical parameters of both alternating and non-alternating decamers, helical twist (32.8° *vs.* 33.6°), rise (3.2 Å *vs.* 3.3 Å), slide (−1.2 Å *vs.* −1.3 Å), x-displacement (−3.8 Å *vs.* −4.0 Å), inclination (18.7° *vs.* 16.4°), roll (10.5° *vs.* 8.8°), propeller twist (−5.1° *vs.* −9.3°), and tilt (0.0° *vs.* −1.3°) are in the range seen in other A-DNA structures. The helical parameters were calculated using 3DNA program.

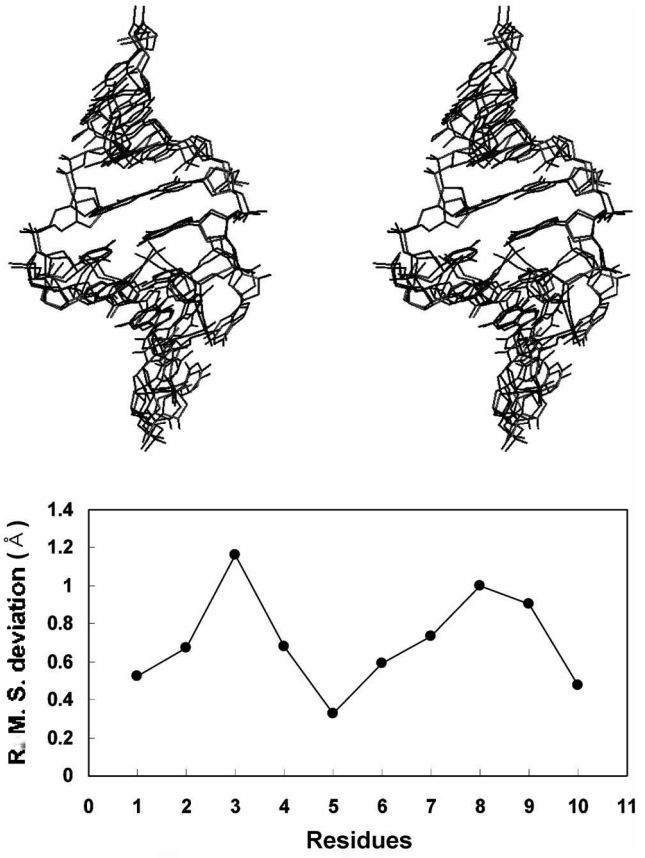

Figure 3. Stereo view of the superposition of the two strands of the decamer duplexes structure (top). The variation in the average rms deviation of each residue upon superposition (bottom). The third residue shows the largest deviation of 1.16 Å.

adopts the nearly extended *trans* conformations for the backbone torsions α (156.8°) and γ (168.8°), as seen in the previous hexagonal A-DNA decamer d(GCACGCGTGC).[8] On the other hand, in the non-alternating decamer, the 8th residue adopts the *trans-trans* conformation with α torsion angle of 139.5° and γ of −165.2°, as found in the other hexagonal A-DNA decamers.[1,9] Note that all hexagonal A-DNA decamers have almost identical intermolecular interactions (see below). Thus, in the hexagonal A-DNA decamer structure, the extended *trans* conformation can be found at either the 7th or 8th residues by the sequence preference (Table 2).

Although the two dyad related duplexes crowd around the center of the reference duplex in the minor groove, a simple common intermolecular interaction is found in both structures. The O4' atom of the 5'-terminal G(1) hydrogen bonds to N2 of the symmetry related base G(7) (Table 3), as previously shown in A-DNA decamer d(GCACGCGTGC) structure.[8] In fact, this intermolecular interaction is conserved in all known hexagonal A-DNA decamer structures.[1,9] Unlike the orthorhombic A-DNA decamer structures, where the two symmetry related molecules are asymmetrically displaced away from the center of the duplex, in the hexagonal A-DNA packing, the symmetrical intermolecular crowding around the center of the duplex could be responsible for the single intermolecular interaction.

Table 2. Total backbone torsion angles for the present d(GTACGCGTAC) and d(GGCCGCGGCC)

torsion angle[§]	α	β	γ	δ	ε	ζ	χ
dna1*							
G	–	–	−37.7	84.9	−161.2	−62.6	−160.2
T	−83.8	177.3	70.6	82.9	−144.8	−81.5	−157.8
A	−48.6	164.6	42.4	82.9	−157.4	−73.2	−148.1
C	−75.5	163.5	70.1	82.4	−147.3	−83.6	−158.4
G	−61.6	171.4	45.8	82.6	−151.2	−73.3	−164.3
C	−40.5	−176.2	22.7	86.5	−173.1	−61.8	−147.7
G	156.8	−166.8	168.8	89.3	−134.5	−88.9	−176.8
T	−24.6	164.7	18.0	85.7	−178.3	−55.8	−148.8
A	−114.9	−169.2	90.2	88.8	−140.8	−87.3	−153.1
C	−40.8	162.4	27.5	86.3	–	–	−123.8
dna2**							
G	–	–	48.5	90.2	−143.8	−80.7	−158.6
G	−52.7	165.0	42.9	84.6	−159.6	−62.8	−156.5
C	−83.3	170.1	71.5	80.3	−147.0	−78.6	−161.7
C	−57.2	167.2	53.4	86.0	−165.0	−77.0	−150.4
G	−88.3	−174.7	60.0	84.0	−155.5	−67.5	−164.8
C	−64.2	−178.6	46.7	80.3	−155.1	−65.0	−158.1
G	−62.3	176.2	54.6	81.9	−169.8	−62.6	−156.8
G	139.5	−177.6	−165.2	94.1	−149.4	−71.2	−172.4
C	−67.3	−176.4	45.1	86.3	−157.0	−71.0	−150.4
C	−51.7	175.7	39.1	94.1	–	–	−137.9

[§]The torsion angles are defined as O3'-P-α-O5'-β-C5'-γ-C4'-δ-C3'-ε-O3'-ζ-P-O5'; *dna1: d(GTACGCGTAC); **dna2: d(GGCCGCGGCC).

In the hexagonal crystal, the wide solvent channel about 20 Å diameter runs down the 6_1 screw axis, which is responsible for the high volume per base pair and the larger amount of water compared with other crystal form. Of the 63 water molecules located in the decamer d(GTACGCGTAC) structure, 44 are directly hydrogen bonded to the DNA and the remaining 8 interact via water bridges (data not shown). 23 water molecules interact with the backbone phosphates and sugar ring oxygen atoms. There are 12 water molecules in the major groove hydrogen bonded to most of the O6/N7 of

Table 3. Intermolecular interactions of the known hexagonal A-DNA decamer structures*

Sequence	Res.	Atom	Dist. (Å)	Atom	Res.	Reference
d(GTACGCGTAC)	G(1)	O4'	3.2	N2	G(7)	this work
d(GGCCGCGGCC)	G(1)	O4'	3.0	N2	G(7)	this work
d(GCACGCGTGC)	G(1)	O4'	3.0	N2	G(7)	8
d(GCGGGCCCGC)	G(1)	O4'	3.0	N2	G(4)	1
	G(1)	N3	3.1	N2	G(5)	
d(ACCGGCCGGT)	A(1)	O4'	3.2	N2	G(4)	9
	A(1)	N3	3.1	N2	G(5)	

*The intermolecular interaction between O4' of the 5'-terminal guanine and N2 of the fourth base pair guanine is conserved in all the hexagonal A-DNA decamer structures.

G, N6/N7 of A and N4 of C. However, the major groove sites of C(10) and both the thymine residues T(2) and T(8) are not hydrogen bonded to water molecules. There are 8 water molecules engaged in the minor groove hydrogen bonding. The hydration pattern of the second decamer is similar to the first; the sugar-phosphate oxygen atoms and the major groove atoms are highly hydrated. Of the 65 water molecules located, 56 are directly hydrogen bonded to the decamer d(GGCCGCGGCC) and 18 interact via water bridges (data not shown). 18 water molecules are engaged in the hydrogen bonding interactions with most of the major groove sites except the G(7) residue while 8 water molecules are hydrogen boned in the minor groove sites.

It is interesting that in both structures the minor groove hydrogen bonding sites are hydrated with almost same ratios compared with the major groove hydrogen bonding sites. 65% (13/20) of the minor groove sites in the alternating decamer structure engaged in the hydrogen bonding with water molecules. This ratio is almost the same with 69% (16/23) of the major groove sites. In the non-alternating decamer structure, 73% (17/23) of the minor groove hydrogen binding sites and 82% (19/23) of the major groove sites are involved in the hydration. This observation therefore clearly demonstrates that the close packing of the symmetry related molecules is not severely blocking the water molecule access into the minor groove. Notice that the hydration ratio difference between both structures may be due to the number of water molecules found in each structure.

Conclusion

The comparison of two isomorphous A-DNA decamer structures of different base sequences show that the overall DNA conformation is very similar while the local conformations show some differences. Both A-DNA decamer structures have the same intermolecular interaction that might be a key factor for the similar overall conformation. However, the differences in the local helical parameters of both structures are mainly influenced by the sequence alternation in one and the non-alternation in the other. In addition, the different position of the extended backbone conformation t, t for the α, γ torsions is probably caused by the sequence difference. This suggests that the inherent conformational flexibility of DNA can be readily perturbed by both base sequence and environment.

Acknowledgment. The authors gratefully acknowledge financial supports from KOSEF through the Center for Integrated Molecular Systems at POSTECH and the Center for Innovative Bio-Physio Sensor Technology (grant number: 02-PJ3-PG6-EV05-0001). This paper is dedicated in memory of our friend and mentor, Dr. Muttaiya Sundaralingam.

References

1. Ramakrishnan, B.; Sundaralingam M. *Biochemistry* **1993**, *32*, 11458.
2. Lipanov, A.; Kopka, M. L.; Kaczor-Grzeskowiak, M.; Quintana, J.; Dickerson, R. E. *Biochemistry* **1993**, *32*, 1373.
3. Jain, S.; Zon, G.; Sundaralingam, M. *Biochemistry* **1991**, *30*, 3567.
4. Shakked, Z.; Guerstein-Guzikevich, G.; Frolow, F.; Rabinovich, D. *Nature* **1988**, *342*, 456.
5. Mooers, B. H. M.; Schroth, G. P.; Baxter, W. W.; Ho, P. S. *J. Mol. Biol.* **1995**, *249*, 772.
6. (a) Dickerson, R. E.; Googsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. USA.* **1994**, *91*, 3579. (b) Bingman, C.; Jain, S.; Zon, G.; Sundaralingam, M. *Nucleic Acids Res.* **1992**, *20*, 6637.
7. Ban, C.; Sundaralingam, M. *Biophysical J.* **1996**, *71*, 1222.
8. Frederick, C. A.; Quigley, G. J.; Teng, M.-K.; Coll, M.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H.-J. *Eur. J. Biochem.* **1989**, *181*, 295.
9. Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Cryst.* **1998**, *D54*, 905.
10. Berman, M. H.; Olson, W. K.; Beveridge, D.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.; Srinvasan, A. R.; Schneider, B. *Biophysical J.* **1992**, *63*, 751.