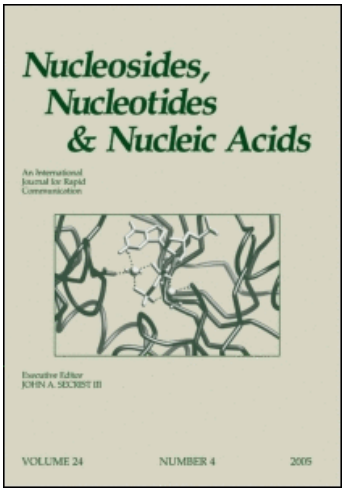

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Crystal Structures of Four A-DNA Decamers that Contain C5-Propyne-Cytosines

Yi-Gui. Gao[a]; Andrew H. J. Wang[a]
[a] Department of Cell & Structural Biology, University of Illinois at Urbana-Champaign, Urbana, IL

To cite this Article Gao, Yi-Gui. and Wang, Andrew H. J.(1998) 'Crystal Structures of Four A-DNA Decamers that Contain C5-Propyne-Cytosines', Nucleosides, Nucleotides and Nucleic Acids, 17: 9, 2093 — 2102
To link to this Article: DOI: 10.1080/07328319808004752
URL: http://dx.doi.org/10.1080/07328319808004752

NUCLEOSIDES & NUCLEOTIDES, 17(9-11), 2093-2102 (1998)

CRYSTAL STRUCTURES OF FOUR A-DNA DECAMERS THAT CONTAIN C5-PROPYNE-CYTOSINES

Yi-Gui Gao and Andrew H.-J. Wang*

Department of Cell & Structural Biology
University of Illinois at Urbana-Champaign, Urbana, IL 61801

ABSTRACT

Nucleic acids incorporating C5-propynepyrimidines have significantly-increased stability which is valuable for antisense applications. Crystal structures of four variants of the d(ACCGGCCGGT) decamer, with each cytosine individually replaced by a C5-propynecytosine (abbreviated **Y**), have been analyzed at high resolution in an attempt to understand the structural basis of the stabilization effect by this novel chemical modification. The five independent **NpY** dinucleotide steps in these four structures show that the hydrophobic propyne group is extensively stacked with the 5'-side base. A model of the poly(purine):poly(C5-propynepyrimidine) A-DNA duplex shows that the deep major groove is substantially filled up by the C5-propyne groups from the poly(C5-propynepyrimidine) strand, displacing many water molecules. These observations may explain the enhanced stability of the duplex incorporating C5-propynepyrimidines.

INTRODUCTION

Modifications on various positions of nucleic acid bases have profound effects on the structure and stability of nucleic acids. For example, methylation or bromination at the C5-position of cytosine facilitates the B- to Z-DNA formation for the alternating $(dC\text{-}dG)_n$ sequence.[1] More recently we showed that methylation at the C8-position of guanine has an even more potent effect on the B-Z equilibrium.[2] It has been suggested that the stability of the Z-DNA structure of poly-d(m^5C-G) may be partly related to the enhanced hydrophobicity due to the clustering of methyl groups on the surface of Z-DNA helix.[3] Likewise the methylation at the C5-position of cytosine appears to stabilize the A-DNA structure.[4] Therefore it is of interest to know the effect of other hydrophobic modifications on the structure-stability relation of nucleic acids.

One of the novel hydrophobic modifications involves the introduction of a propyne substituent at the C5-position of pyrimidine (**Figure 1**). It has been shown that oligonucleotides containing C5-propyne-pyrimidines hybridize to their complementary target sequences with a substantial increase in their T_m values.[5] Therefore those C5-propyne-pyrimidine-containing oligonucleotides may serve as excellent antisense agents for a variety of applications.[6]

In order to probe the structural bases of the effect of the propyne substituent at the C5-position of pyrimidine on DNA, we synthesized a series of DNA decamers relating to d(ACCGGCCGGT) in which all four cytosines have been individually mutated to C5-propyne-cytosines (abbreviated as **Y**). This sequence was chosen because the A-DNA structure of the unmodified d(ACCGGCCGGT) decamer has been determined by us before.[7] It should be noted that the A-DNA conformation is very similar to that of the RNA duplex, except for the additional 2'-hydroxyl groups bordering the two edges of the wide minor groove in RNA. The similarity between A-DNA and RNA makes it relevant to study the **Y**-containing duplexes in the A-DNA conformation because the structural understanding could be extrapolated to the RNA duplexes which are formed during sense-antisense nucleic acid hybridization.

In this paper we address these issues by solving the crystal structures of four **Y**-containing decamer oligonucleotides in two crystal forms (either in the $P6_122$ or the $P2_12_12_1$ space group) and compare their conformations.

MATERIALS AND METHODS

All four Y-containing decamers, d(A**Y**CGGCCGGT) (decamer **I**), d(AC**Y**GGCCGGT) (decamer **II**), d(ACCGG**Y**CGGT) (decamer **III**), and d(ACCGGC**Y**GGT) (decamer **IV**) were synthesized using automated DNA synthesizers and purified by Sepharose S100 gel filtration column chromatography. The **Y** phosphoramidite was purchased from Glen Research. Preparations of the various crystal forms of A-DNA listed in **Table 1** follow the procedures described previously.[8] Typically, the crystallization solution contained 4 mM DNA duplex (2.5 μl), 100 mM sodium cacodylate buffer (pH 6.5) (10 μl), 30% 2-methyl-2,4-pentanediol (2-MPD) (5 μl), plus varying amounts of the counter ions (2.5 μl). For the molecules **I**, **II** and **IV**, crystallization solutions also contained 2 mM of spermine tetrachloride. For molecule **III**, crystallization solution contained 2 mM of $[Co(NH_3)_6]^{3+}$ ion. All solutions were equilibrated with 50% 2-MPD.

Suitable crystals were chosen and mounted individually in a thin-walled glass capillary and sealed with a droplet of the crystallization mother liquor for data collection. The diffraction data sets were collected at ambient room temperature (~25 ^oC) to a resolution of 1.9 Å as indicated in **Table 1**. We used a Rigaku R-AXIS IIc image

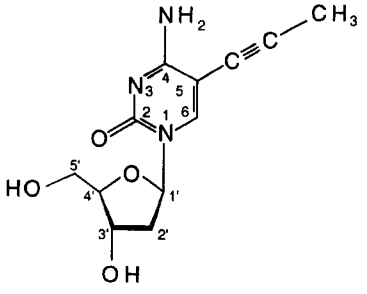

Figure 1. Chemical structures of C5-propynedeoxycytidine.

Table 1. Crystal and refinement data of four A-DNA d(ACCGGCCGGT) decamers (**I-IV**) containing C5-propynecytosines.

	I	II	III	IV[a]
a (Å)	24.85	38.84	39.25	39.65
b (Å)	45.12	38.84	39.25	39.65
c (Å)	47.90	79.01	78.06	79.12
Space group	$P2_12_12_1$	$P6_122$	$P6_122$	$P6_122$
R_{merge} (%)[b]	5.9	6.4	5.7	7.4
Completeness to 2 Å (%)	70.6	91.7	76.4	95.0
# reflections (>2.0 $\sigma(F_o)$)	3,110	2,208	2,642	2,883
R-factor	0.186	0.195	0.184	0.196
R_{free} (10% data)	0.286	0.312	0.279	0.310
r.m.s.d. bond distance (Å)[c]	0.013	0.017	0.020	0.022
r.m.s.d. bond angle (deg)[c]	3.36	3.60	3.53	3.54
No. of waters	54	51	58 + 2 Co	60

a: The four decamers are **I**: d(A**Y**CGGCCGGT), **II**: d(AC**Y**GGCCGGT), **III**: d(ACCGG**Y**CGGT), and **IV**: d(ACCGGC**Y**GGT).
b: $R_{merge}=\Sigma|I-\langle I\rangle|/\Sigma\langle I\rangle$.
c: Root-mean-square deviation from ideal bond length and bond angle, resulted from the refinement.

plate system mounted on an RU-200 rotating-anode x-ray generator at a power of 50 KV and 80 mA for data collection. The fine-focused x-ray beam (0.3 mm) was collimated with a graphite monochromator to produce the CuK_α radiation (1.5418 Å). The image plate to crystal distance was 90 mm. Forty-five and sixty frames of oscillation data (2^o /10 min) were collected to cover a range of 90^o and 120^o of the ϕ angle for the $P2_12_12_1$ form and the $P6_122$ form, respectively. After data processing using the Molecular Structure Corporation (Woodlands, TX) software package, each dataset produced a 2-4 fold of redundancy in the observed reflections for both crystal forms. The R_{merge} (for all $F_o>2.0\ \sigma(F_o)$ reflections) values of the datasets range from 5.7 to 7.4% (**Table 1**). Care was taken to account for intensity saturation on the image plate for strong reflections. Some frames were recollected with a faster oscillation speed (2^o/2 min) in order to recover those strong (saturated) reflections.

The atomic coordinates from the canonical crystal structures were used as the starting models for the refinement using the simulated annealing (SA) procedure of the X-PLOR package.[9] For the $P6_122$ and the $P2_12_12_1$ crystals, the coordinates from the Mg^{2+}-form of d(ACCGGCCGGT)[7] and the Mg^{2+}-form of r(GCG)d(TATACGC),[10] respectively, were used with the sequence of the molecule suitably changed. The force field parameters for DNA nucleotides were from toph11.dna and param11x.dna files in X-PLOR. Those for the propyne group were obtained from comparison with related structures in the Cambridge Crystal Database. Each cycle of the SA refinement consisted of heating at 3000 °K and then slowly cooling by 10° decrements to 300° K over 1.8 ps, followed by 120 steps of conjugated gradient minimization. After many cycles of refinement with all available data, the R-factor in general reached ~30% at 2.0 Å resolution. Water molecules and metal ions were then located from subsequent Fourier ($2|F_o|-|F_c|$) maps and added to the refinement. At this stage, the heating temperature was 1000 °K in order to keep the water molecules from moving to random positions due to high temperature. The final crystallographic R-factor was near 19%. The crystallographic data along with the refinement results are summarized in **Table 1**. The final atomic coordinates of the structure have been deposited at the Brookhaven Protein Databank.

RESULTS AND DISCUSSION

Conformations of A-DNA decamers

In this study, we analyzed four new structures of **Y**-containing DNA decamer molecules which were crystallized in two crystal lattices ($P2_12_12_1$ and $P6_122$). Inspection of their global conformational parameter indicates that they all adopt conformation very close to the canonical A-DNA (**Figure 2**). The duplexes have the

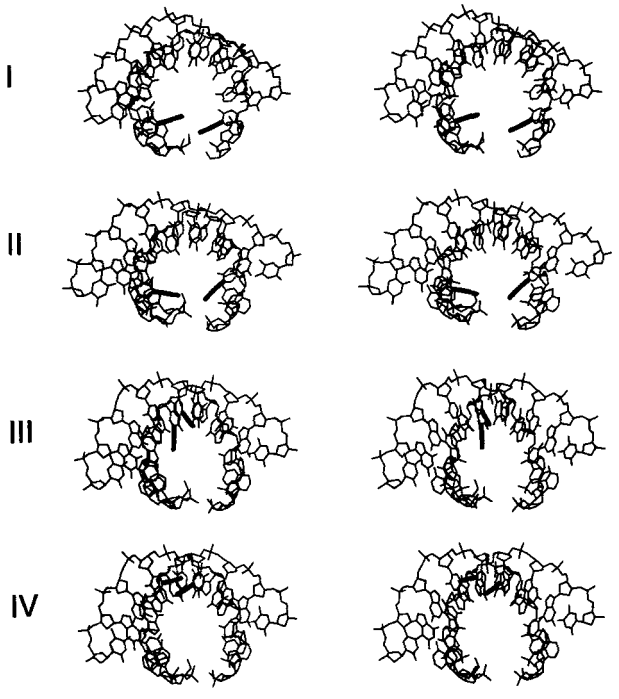

Figure 2. Molecular structures of the four C5-propynecytosine-containing A-DNA decamer helixes crystallized in two space groups.

characteristic wide and shallow minor groove and the deep and narrow major groove. The averaged base pair inclination angle is in the range of 15^{o}-17^{o} and the averaged x-displacement of the base pairs is about -4.0 Å. All sugars have the N-type pucker (C3'-endo). Nevertheless, some small variations exist between the two crystal forms. The average rise per residue values in the $P6_122$ and $P2_12_12_1$ forms are 2.48 Å and 2.71 Å, respectively, whereas the x-displacements for the two respective crystal forms are -4.00 Å and -4.90 Å. As observed before,[10] we note that molecules crystallized in the same space group have a more similar conformation, whereas the same molecule crystallized in different space groups has different (local) conformations. For example, superposition of the structures of d(ACCGGCCGGT) (native), molecules **II**, **III** and **IV** in the $P6_122$ form have show r.m.s.d. (using common atoms only) ranging from 0.36 to 1.06 Å. On the other hand molecule **I** (in $P2_12_12_1$) has r.m.s.d. (using common atoms only) ranging from 1.14 to 1.51 Å with the molecules in $P6_122$ space group.

Conformations of C5-propynecytosine nucleotides

The C5-propynecytosine nucleotide is located in different sequence locations of the decamers. However the conformation of those **Y** nucleotides are similar, i.e., N-type pucker (pseudorotation angle 16° and amplitude 35°), *anti* glycosyl angle (χ=201°). The r.m.s.d. among the four **Y** nucleotides is 0.26 Å.

The linear propyne substituent is very hydrophobic and bulky. Its presence in the deep major groove of A-DNA results in increased hydrophobicity of the major groove. A large volume located between the N4 amino and the 5'-phosphate groups is now occupied by the propyne groups, displacing some water molecules which are normally hydrogen bonded to N4 and phosphate. In fact, the terminal methyl group of the propyne substituent is making van der Waals contacts with the phosphate group from both its own and the preceding (5'-side) phosphate groups. For example, the terminal methyl group of the propyne substituent of **Y**3 in decamer **II** is 3.42 Å to O1P of **Y**3 and 3.29 Å to O1P of C2 nucleotides. The corresponding distances in decamers **III** and **IV** are 3.89 Å and 3.51 Å for the methyl of **Y**6 to O1P of **Y**6 and G5, and 3.26 Å and 2.99 Å for the methyl of Y7 to O1P of **Y**7 and C5, respectively. **Figure 3** shows the environment of the propyne group using decamer **III** as an example. It is of interest to note that a conserved bridging water molecule is normally found located between two adjacent phosphates in the A-DNA.[12] The insertion of the propyne group removes the bridging water molecule from this location.

The interaction of the propyne group with the neighboring base pairs is somewhat dependent on the sequence context as shown in the base-base stacking patterns of the five independent 5'-Np**Y** steps of the decamers (**Figure 4**). In the 5'-Pu-p-**Y** steps, the propyne group is stacked extensively with the imidazole ring of the purine base. In the 5'-Py-p-**Y** steps, the propyne group is stacked extensively with the pyrimidine base. Such hydrophobic stacking interactions no doubt contribute to the stability of the duplex. In contrast there is no stacking interaction between the propyne group with the base on the 3' side.

We have used the structural information regarding the propyne groups in the deep groove of A-DNA to probe the conformation of poly(purine)-poly(C5-propynepyrimidine) duplex. A model of such a duplex has been constructed (**Figure 5**). It can be seen that the hydrophobic propyne groups from the poly(C5-propynepyrimidine) strand occupy the deep groove nearly completely. In some sense this is analogous to the hydrophobic core of proteins which provide the driving forces for the folding and stability of proteins. It is likely that the strong hydrophobic interactions derived from the contiguous propyne groups of the poly(C5-propynepyrimidine) strand provide the same role in the folding and stability of RNA duplexes.

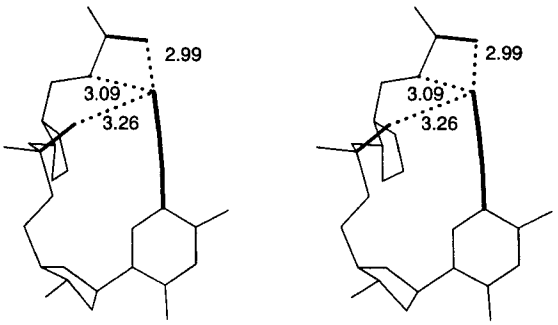

Figure 3. Detailed structure surrounding the propyne group using the **Y7** nucleotide of decamer **III** as an example.

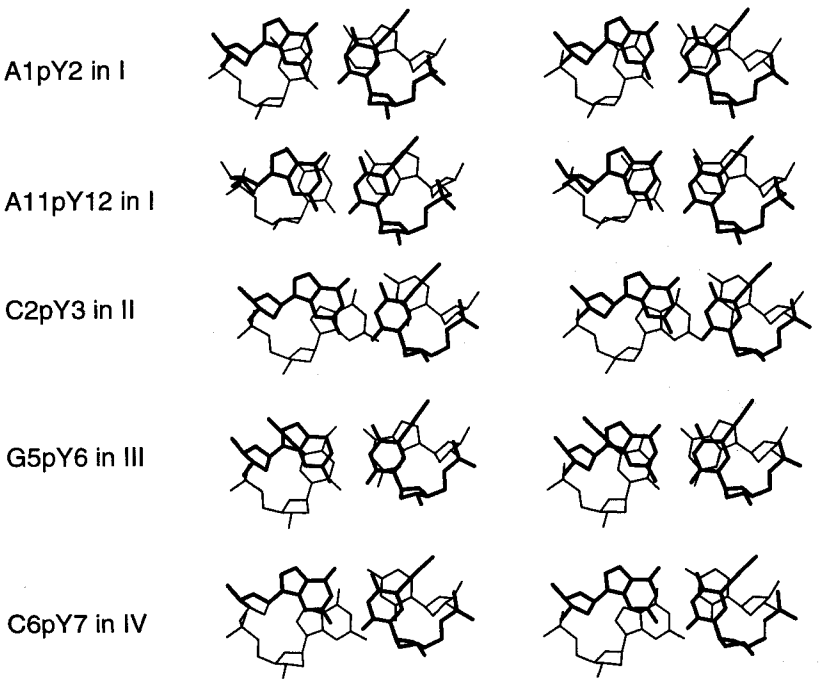

Figure 4. Comparisons of the base-base stacking interactions of the five independent Np**Y** steps in four crystal structures.

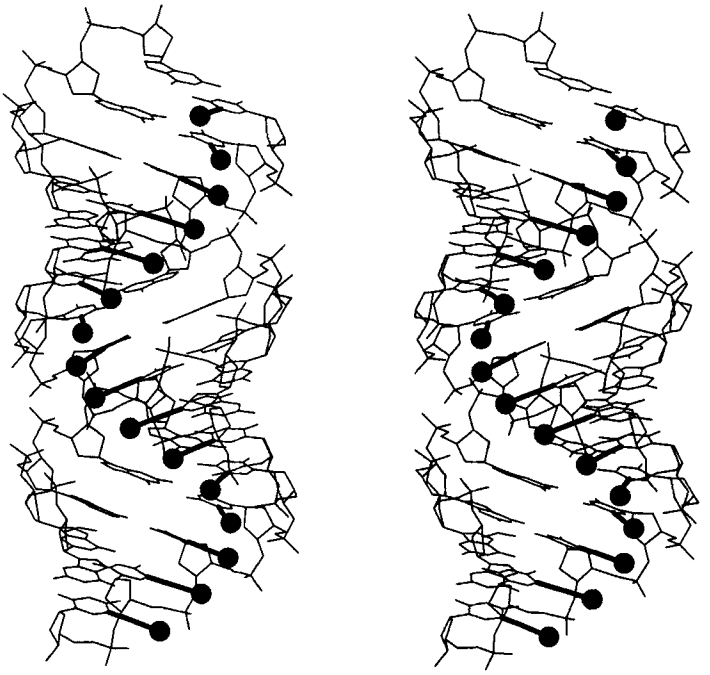

Figure 5. A model of 1.5 turns of the poly(deoxyguanosine)-poly(C5-propynedeoxycytidine) A-DNA duplex.

Binding of $[Co(NH_3)_6]^{3+}$ to A-DNA

The crystallization of a particular Y-containing sequence of d(ACCGGCCGGT) into one of the two different crystal lattices depending on the choice of metal ions and the position of Y in the sequence. $[Co(NH_3)_6]^{3+}$ ions produced the hexagonal form for d(ACCGGYCGGT).

It is interesting to mention that $[Co(NH_3)_6]^{3+}$ is extremely potent in promoting, at sub-micromolar concentration, the B- to Z-DNA transition for poly-d(m^5C-G).[1,13] It is equally effective in promoting the formation of the DNA four-way junction (Holliday junction).[14] More recently, it has been shown that the activity of a hammerhead ribozyme (whose crystal structure has been determined recently[15] is strongly dependent on the Mg^{2+} concentration and $[Co(NH_3)_6]^{3+}$ can substitute Mg^{2+} effectively.[16] Therefore $[Co(NH_3)_6]^{3+}$ is a novel ion that is useful in the study of the structure-function of nucleic acids.[17]

In the native d(ACCGGCCGGT)-$P6_122$ form, three independent $[Co(NH_3)_6]^{3+}$ ions are located.[11] One of them, the $[Co(NH_3)_6]^{3+}$ ion located at the center of the deep

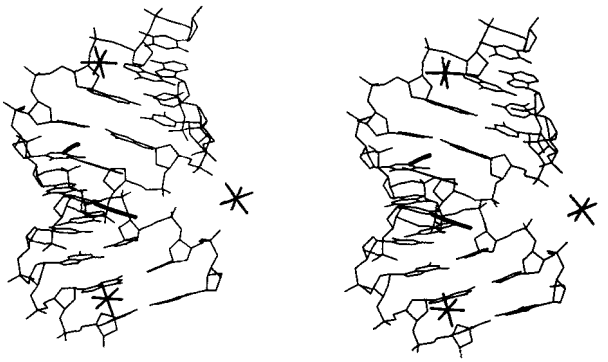

Figure 6. Interactions between $[Co(NH_3)_6]^{3+}$ ions and A-DNA in the crystal structure of the $P6_122$ form of d(ACCGGYCGGT). Two different binding modes are found, one bridges two phosphate groups across the major groove and the other adheres to the edges of two adjacent guanines (G8 and G9).

major groove near the G4-G5 bases of a duplex, is absent in the d(ACCGGYCGGT) (decamer **II**) crystal due to the bulky propyne group of **Y6** (**Figure 6**). In the two binding modes shown in **Figure 6**, a single $[Co(NH_3)_6]^{3+}$ ion is bound to the edges of two adjacent intra-stranded guanines through hydrogen bonds to the N7 and O6 sites of guanines. Such interactions are similar to those found in Z-DNA with $[Co(NH_3)_6]^{3+}$ bridging a guanine base (at N7/O6 sites) and the 3'-phosphate group by hydrogen bonds.[18] Thus $[Co(NH_3)_6]^{3+}$ and related ions (spermine and spermidine) are able to stabilize alternative DNA conformations by providing a very effective hydrogen bonding network and a high positive charge density (+3) for the neutralization of negatively-charged phosphates in DNA/RNA.

ACKNOWLEDGMENTS

This work was supported by NIH grant GM-41612 (A. H.-J. W). We thank Dr. H. Robinson for his assistance.

REFERENCES

1. Behe, M.; Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. USA.* **78**, 1619-1623.
2. Sugiyama, H.; Kawai, K.; Matsunaga, A.; Fujimoto, K.; Saito, I.; Robinson, H.; Wang, A. H.-J. (1996) *Nucleic Acids Res.* **24**, 1272-1278, and references therein.

3. Fujii, S.; Wang, A. H.-J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. (1982) *Nucleic Acids Res.* **10**, 7879-7892.

4. Frederick, C. A.; Saal, D.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H.-J.; Rich, A. (1987) *Biopolymers* **26**, S145-160.

5. Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. (1992) *Tetrahedron Lett.* **33**, 5307-5310.

6. Wagner, R. W.; Matteucci, M. D.; Lewis, J. G.; Gutierrez, A. J.; Moulds, C.; Froehler, B. C. (1993) *Science* **260**, 1510-1513. Flanagan, W. M.; Su, L. L.; Wagner, R. W. (1996) *Nature Biotechnology* **14**, 1139-1145.

7. 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. (1989) *Eur. J. Biochem.* **181**, 295-307.

8. Wang, A. H.-J.; Y.-G. Gao. (1990) *Methods* **1**, 91-99.

9. Brünger, A. (1993) X-PLOR (version 3.1), The Howard Hughes Medical Institute and Yale University, New Haven, CT.

10. Wang, A. H.-J.; Fujii, S.; van Boom, J. H.; van der Marel, G. A.; van Boeckel, S. A. A.; Rich, A. (1981) *Nature* **299**, 601-604.

11. Gao, Y.; Robinson, H.; van Boom, J. H.; Wang, A. H.-J. (1995) *Biophys. J.* **69**, 559-568, and references therein.

12. Saenger, W.; Hunter, W. N.; Kennard, O. (1986) *Nature* **324**, 385-388.

13. Peck, L. J.; Nordheim, A.; Rich, A.; Wang, J. C. (1982) *Proc. Natl. Acad. Sci. USA.* **79**, 4560-4564.

14. Duckett, D. R.; Murchie, A. I. H.; Lilley, D. M. J. (1990) *EMBO J.* **9**, 583-590.

15. Pley, H. W.; Flaherty, K. M.; McKay, D. B. (1994) *Nature* **372**, 68-74.

16. Bassi, G. S., Mollegaard, N.-E.; Murchie, A. I. H.; von Kitzing, E.; Lilley, D. M. J. (1995) *Nature Struct. Biol.* **2**, 45-55.

17. Robinson, H.; Wang, A. H.-J. (1996) *Nucleic Acids Res.* **24**, 676-682.

18. Gessner, R. G.; Quigley, G. J.; Wang, A. H.-J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. (1985) *Biochemistry* **24**, 237-250.