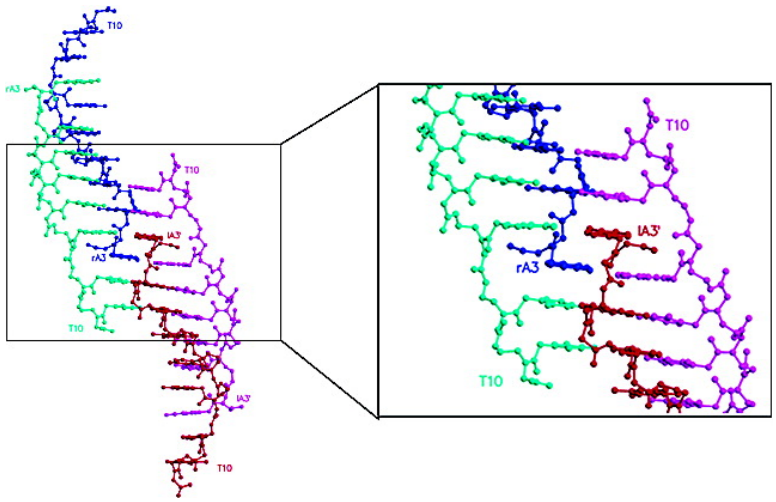
Article

Crystal Structure of a Partly Self-Complementary Peptide Nucleic Acid (PNA) Oligomer Showing a Duplex–Triplex Network

Britt Petersson, Bettina Bryde Nielsen, Hanne Rasmussen, Ingrid Kjller Larsen, Michael Gajhede, Peter E. Nielsen, and Jette Sandholm Kastrup

J. Am. Chem. Soc., **2005**, 127 (5), 1424-1430• DOI: 10.1021/ja0458726 • Publication Date (Web): 13 January 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 6 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 01/13/2005

Crystal Structure of a Partly Self-Complementary Peptide Nucleic Acid (PNA) Oligomer Showing a Duplex−Triplex Network†

Britt Petersson,‡ Bettina Bryde Nielsen,‡ Hanne Rasmussen,‡,|| Ingrid Kjøller Larsen,‡ Michael Gajhede,‡ Peter E. Nielsen,*,§ and Jette Sandholm Kastrup*,‡

Contribution from the Biostructural Research, Department of Medicinal Chemistry, The Danish University of Pharmaceutical Sciences, Universitetsparken 2, DK-2100 Copenhagen, Denmark, and Department of Medical Biochemistry and Genetics, The Panum Institute, University of Copenhagen, Blegdamsvej 3c, DK-2200 Copenhagen, Denmark

Received July 9, 2004; E-mail: pen@imbg.ku.dk; jsk@dfuni.dk

Abstract: The X-ray structure of a partly self-complementary peptide nucleic acid (PNA) decamer (H-GTAGATCACT-L-Lys-NH_2) to 2.60 Å resolution is reported. The structure is mainly controlled by the canonical Watson−Crick base pairs formed by the self-complementary stretch of four bases in the middle of the decamer ($G_4A_5T_6C_7$). One right- and one left-handed Watson−Crick duplex are formed. The two PNA units C_9T_{10} change helical handedness, so that each PNA strand contains both a right- and a left-handed section. The changed handedness in C_9T_{10} allows formation of Hoogsteen hydrogen bonding between C_9T_{10} and G_4A_5 of a PNA strand in an adjacent Watson−Crick double helix of the same handedness. Thereby, a PNA−PNA−PNA triplex is formed. The PNA unit A_3 forms a noncanonical base pair with A_8 in a symmetry-related strand of opposite handedness; the base pair is of the A−A reverse Hoogsteen type. The structural diversity of this PNA demonstrates how the PNA backbone is able to adapt to structures governed by the stacking and hydrogen-bonding interactions between the nucleobases. The crystal structure further shows how PNA oligomers containing limited sequence complementarity may form complex hydrogen-bonding networks.

Introduction

Peptide nucleic acid (PNA) is a synthetic and achiral DNA mimic in which the phosphoribose backbone of DNA has been substituted with a pseudopeptide consisting of *N*-(2-aminoethyl)-glycine units with a methylene carbonyl linker connecting to the nucleobase[1] (see Figure 1). The stability of PNA in biological fluids[2] and its unique chemical and structural properties[3] have attracted attention for possible applications in molecular biology,[4] in diagnostics,[4−6] and within drug discovery.[7,8]

† Atomic coordinates and structure factors have been deposited in PDB (ID Code 1XJ9) and NDB (ID Code UP0006).

‡ The Danish University of Pharmaceutical Sciences.

§ University of Copenhagen.

|| Current address: Protein Biophysics and Delivery, Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd, Denmark.

(1) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500.
(2) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Berg, R. H.; Kim, S. K.; Nordén, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566−568.
(3) Eriksson, M.; Nielsen, P. E. *Q. Rev. Biophys.* **1996**, *29*, 369−394.
(4) (a) Ray, A.; Nordén, B. *FASEB J.* **2000**, *14*, 1041−1060. (b) Nielsen, P. E. *Curr. Opin. Biotechnol*. **2001**, *12*, 16−20. (c) Pellestor, F.; Paulasova, P. *Int. J. Mol. Med.* **2004**, *13*, 521−525.
(5) Radwanska, M.; Magez, S.; Perry-O'Keefe, H.; Stender, H.; Coull, J.; Sternberg, J. M.; Büscher, P.; Hylding-Nielsen, J. J. *J. Clin. Microbiol.* **2002**, *40*, 4295−4297.
(6) Igloi, G. L. *Expert Rev. Mol. Diagn.* **2003**, *3*, 17−26.

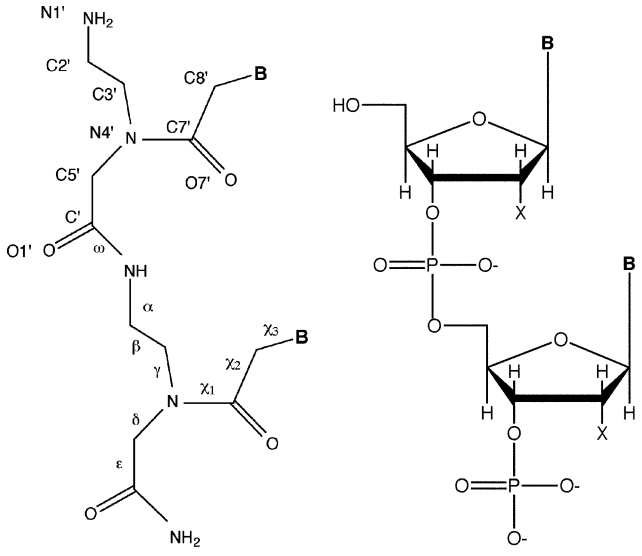


Figure 1. Chemical formulas of PNA (left) and DNA/RNA (right). "**B**" denotes the nucleobases, and "X" denotes H in DNA and OH in RNA.

PNA has proven a very potent structural mimic of natural nucleic acids, being capable of forming homoduplexes,[9−12] heteroduplexes,[13−15] and heterotriplexes[16] as demonstrated by X-ray crystallography and NMR structure determinations. These

10.1021/ja0458726 CCC: $30.25

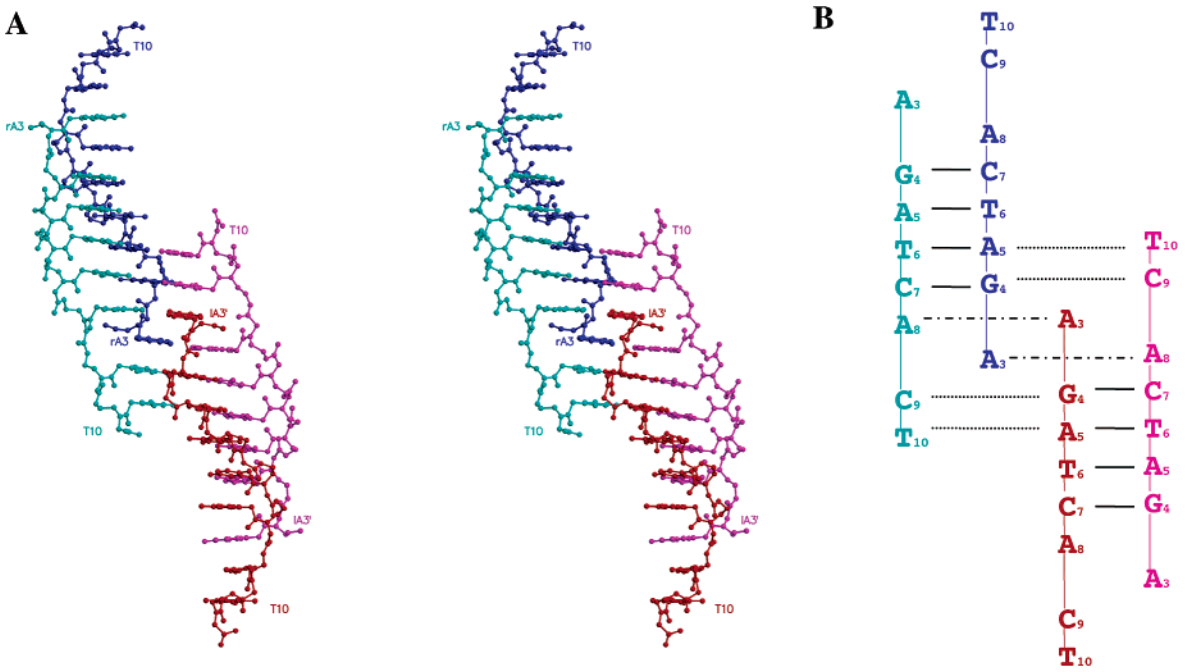


Figure 2. (A) Structure of the PNA complex with one predominantly right-handed strand (blue) and one predominantly left-handed strand (red) in the asymmetric unit of the crystal (shown in stereo). The two other strands (right-handed in cyan and left-handed in magenta) are generated by crystallographic symmetry. The first two PNA units could not be located in the electron density map. (B) Schematic drawing of the observed structure of the PNA complex (color coding as in (A)). In the right-handed strands, A_3-A_8 are right-handed and C_9 and T_{10} are left-handed. In the left-handed strands, A_3-A_8 are left-handed and C_9 and T_{10} are right-handed. Watson−Crick base pairing is represented by solid lines, Hoogsteen base pairing by dotted lines, and A−A reverse Hoogsteen base pairing by dot−dash lines.

structures show that PNA prefers a P-form helix, which is different from both the A- and B-forms of RNA and DNA. Evidence has also been presented that PNA oligomers may form triplexes,[17] hairpins,[18] and quadruplexes.[19] However, more complex tertiary structures as found for instance in RNA has not been described for PNA. Most PNA oligomers show fairly sharp thermal transitions with quite strong hyperchromicity indicating an ordered stacked structure of the single strand, and thereby a propensity for self-organization and aggregation not seen with oligonucleotides. This is particularly pronounced for PNAs with some self-complementarity, such as the well-studied PNA H-GTAGATCACT-L-Lys-NH_2.[20] Intrigued by this, we set out to investigate whether the PNA could be crystallized and if so to determine its structure by X-ray crystallography. This was indeed possible, and we now report the three-dimensional structure of this "single-stranded" PNA at 2.6 Å resolution. The structure reveals a hitherto unseen diversity of structural motifs (including duplex, Hoogsteen-type triplex, backbone chirality shift, and a novel adenine−adenine base pairing−intercalation structure) adopted by the PNA in order to maximize base-stacking and base-pairing interactions. Thus, an intricate network is built in the crystal, demonstrating the potential of PNA to form tertiary structure as well as the high flexibility of the backbone.

Results

The X-ray structure of the partly self-complementary PNA (H-GTAGATCACT-L-Lys-NH_2) has been determined to 2.60 Å resolution. Two PNA strands are observed in the asymmetric unit of the crystal: one strand is predominantly right-handed and one is predominantly left-handed. These strands form double helices with their symmetry-related partners as shown in Figure 2. This leads to a continuous stacking of the nucleobases throughout the crystal with alternating right- and left-handed helices. The positions of the eight C-terminal PNA units in both PNA strands were located in the electron density maps. However, the position of the L-lysine residue on the C-terminal PNA unit could not be interpreted.

The overall structures of both the right- and left-handed double helices are mainly controlled by the self-complementary region of four bases ($G_4A_5T_6C_7$) in the middle of the PNA capable of forming canonical Watson−Crick base pairing (see Table 1 and Figures 2 and 3A). The antiparallel double helix has a conformation corresponding to the P-form helix observed for PNA.[9,16] The average helical parameters are given in Table 2. The twist, rise, and tilt are very similar for strands of both

(7) Elayadi, A. N.; Corey, D. R. *Curr. Opin. Investig. Drugs* **2001**, *2*, 558−561.
(8) Koppelhus, U.; Nielsen, P. E. *Adv. Drug. Delivery Rev.* **2003**, *55*, 267−280.
(9) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98−101.
(10) Haaima, G.; Rasmussen, H.; Schmidt, G.; Jensen, D. K.; Kastrup, J. S.; Stafshede, P. W.; Norden, B.; Buckhardt, O.; Nielsen, P. E. *New J. Chem.* **1999**, *23*, 833−840.
(11) Eldrup, A. B.; Nielsen, B. B.; Haaima, G.; Rasmussen, H.; Kastrup, J. S.; Christensen, C.; Nielsen, P. E. *Eur. J. Org. Chem.* **2001**, 1781−1790.
(12) Rasmussen, H.; Liljefors, T.; Petersson, B.; Nielsen, P. E.; Kastrup, J. S. *J. Biomol. Struct. Dyn.* **2004**, *21*, 495−502.
(13) Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12021−12026.
(14) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777−780.
(15) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410−413.
(16) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838−1841.
(17) Wittung, P.; Nielsen, P.; Nordén, B. *J. Am. Chem. Soc.* **1997**, *119*, 3189−3190.
(18) Armitage, B.; Ly, D.; Koch, T.; Frydenlund, H.; Ørum, H.; Schuster, G. B. *Biochemistry* **1998**, *37*, 9417−9425.
(19) Krishnan-Ghosh, Y.; Stephens, E.; Balusubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944−5945.
(20) Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Nordén, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544−5552.

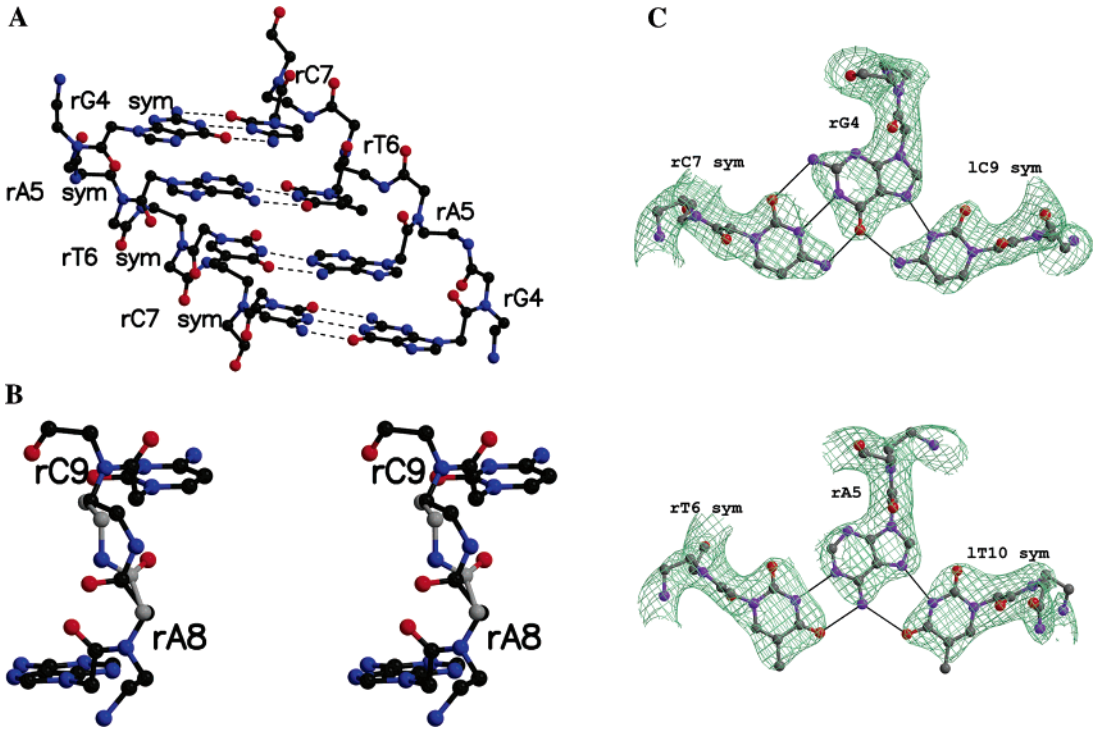


Figure 3. Details of structural features of partly self-complementary PNA. (A) The Watson−Crick base pairing of the right-handed duplex. Hydrogen bonds between the nucleobases are shown as black dotted lines. (B) The change of helical handedness from right-handed (rA_8) to left-handed (rC_9) PNA (shown in stereo). Two conformations of the PNA backbone are shown with light and dark gray carbon atoms. (C) The base triplets C = G·C(+) (upper) and T-A·T (lower) of the right-handed triplex. The final $2F_o-F_c$ electron density map is shown (green; contoured at 1.0 σ). Hydrogen bonds between the nucleobases are shown as black lines. The atoms are colored according to type: C, gray; N, blue; O, red.

Table 1. Hydrogen-Bonding Interactions between Nucleobases in the Predominantly Right-Handed Strand of Partly Self-Complementary PNA[a]

PNA unit	H-bonding partner	H-bonding classification
rA_3[b]	lA_8 sym[c]	A·A trans Watson−Crick/Hoogsteen (parallel)
rA_3'	rA_8 sym	A·A C6-NH_2−N1 single hydrogen bond
rG_4	rC_7 sym	G=C cis Watson−Crick/Watson−Crick
rG_4	lC_9 sym	G·C(+) cis Watson−Crick/Hoogsteen
rA_5	rT_6 sym	A-T cis Watson−Crick/Watson−Crick
rA_5	lT_{10} sym	A·T cis Watson−Crick/Hoogsteen
rT_6	rA_5 sym	T-A cis Watson−Crick/Watson−Crick
rC_7	rG_4 sym	C=G cis Watson−Crick/Watson−Crick
rA_8	lA_3' sym	A·A trans Watson−Crick/Hoogsteen (parallel)
rA_8	rA_3' sym	A·A C6-NH_2−N1 single hydrogen bond
rC_9	lG_4 sym*	C(+)·G cis Watson−Crick/Hoogsteen
rT_{10}	lA_5 sym*	T·A cis Watson−Crick/Hoogsteen

[a] The interactions are classified according to the Leontis−Westhof nomenclature.[22] The bases in the predominantly left-handed strand show corresponding hydrogen-bonding interactions. [b] rA_3 and rA_3' denote the two different conformations of A_3 in the predominantly right-handed strand, and lA_3 and lA_3' denote the conformations in the predominantly left-handed strand. [c] Sym and sym* indicate two strands generated by different crystallographic symmetry operations.

handedness; however, deviations are observed in X-displacement and inclination. The X-displacements mostly resemble that of A-RNA, whereas the rise is identical to that of B-DNA. A smaller twist is seen in PNA compared to B-DNA and A-RNA. In PNA, this leads to a more extended double helix with 18−20 bases per turn compared to 10−11 in B-DNA and A-RNA. The base-pair inclination is much larger than that observed in the first reported PNA−PNA hexamer,[9] but is similar to the inclination of corresponding backbone[10] and nucleobase[11] modified hexamers.

Interestingly, the two last PNA units in the PNA (C_9T_{10}) change helical handedness (Figures 2 and 3B). The backbone torsional angle values of individual PNA units are listed in Table 3. The change in helical handedness is apparent from the different signs of the torsional angles of the PNA units, which is especially evident from the torsional angle χ_3. The changed handedness in C_9T_{10} allows formation of Hoogsteen hydrogen bonding with the Watson−Crick base pairs (G_4A_5) of a third PNA strand of the same handedness, and thereby leads to triplex formation (Table 1 and Figures 2 and 3C). The hydrogen-bonding interactions of the nucleobases are similar to those of DNA and RNA homo- or heterotriplexes.[16,21]

The PNA unit A_3 is observed in two distinct conformations (rA_3 and rA_3', and lA_3 and lA_3', for the predominantly right-

Table 2. Helical Parameters (Average Values) of PNA, B-DNA, and A-RNA Double Helices

	X-displacement (Å)	inclination (deg)	twist (deg)	rise (Å)	base tilt (deg)	bases per turn
PNA decamer (right)[a]	−3.5	−26.7	19.1	3.4	0.0	19
PNA decamer (left)[a]	7.9	−8.9	−18.0	3.3	0.0	20
PNA hexamer[b]	−8.3	0.3	19.8	3.2	1.0	18
B-DNA[c]	0.0	1.5	36.0	3.4	0.0	10
A-RNA[d]	−5.3	15.8	32.7	2.8	0.0	11

[a] Average values are given for the self-complementary region of four bases ($G_4A_5T_6C_7$) in the middle of the PNA decamer capable of forming canonical Watson−Crick base pairing. [b] PNA−PNA hexamer.[9] [c] Ideal B-form DNA helix.[21] [d] Ideal A-form RNA helix.[21]

Table 3. Characterization of Conformations of Partly Self-Complementary PNA

PNA unit	α[a]	β	γ	δ	ϵ	ω	χ_1	χ_2	χ_3	distance between bases[b]	parallel stacking characterization[c]
rA_3		136	−141	110	−7	163	48	135	116	6.5	with rC_9 sym, partly with rA_8 sym and lA_3'
rA_3'		138	−93	−113	34	−178	−1	146	−146	4.5	partly with rG_4, lC_9 sym, lA_8 sym, and lA_3
rG_4	−105 (128)	133 (−124)	81 (−86)	93	−14	178	−3	−173	69	4.9	with rA_5 and rA_8 sym
rA_5	−84	53	65	76	180	−177	1	−173	87	4.8	with rT_6 and rG_4, partly with rC_7 sym
rT_6	−95	51	71	84	158	−167	5	−163	88	4.8	with rC_7 and rA_5
rC_7	−94	59	58 (76)	101 (75)	162 (−15)	−174 (−174)	3	−169	96	4.3	with rA_8 and rT_6
rA_8	−126 (57)	98 (82)	103 (64)	−141 (−179)	12 (−99)	−175 (144)	−31 (−38)	−159	103	7.8	with rC_7, partly with rA_3 sym, lA_3 sym*, and rG_4 sym
rC_9	−82 (84)	−110 (−140)	−93 (−87)	−69 (−60)	2	161	29 (19)	152	−91	4.5	with rT_{10}, rA_3 sym, and lA_3 sym*
rT_{10}	−54	−72	−72	−85	−80		17	149	−83		with rC_9
lA_3'		−131	128	−97	−13	153	−29	−139	−126	6.2	with lC_9 sym, partly with lA_8 sym and rA_3
lA_3		119	112	83	1	−151	83	156	114	4.7	with rC_9 sym, partly with lG_4, rA_8 sym, and rA_3'
lG_4	82 (−107)	−112 (138)	−49 (112)	−108	−153	153	−7	−168	−89	4.8	with lA_5 and lA_8 sym
lA_5	117	−78	−73	−112	−160	148	8	−175	−94	4.9	with lT_6 and lG_4
lT_6	118	−69	−68	−92	176	168	−4	178	−79	4.7	with lC_7 and lA_5
lC_7	108	−52	−75 (−85)	−100 (−81)	−165 (10)	162 (175)	−3	167	−87	4.2	with lA_8 and lT_6
lA_8	105 (−56)	−57 (−70)	−75 (−41)	−106 (−152)	37 (−114)	−158 (142)	−7 (9)	173	−106	8.1	with lC_7, partly with lA_3' sym, rA_3' sym*, and lG_4 sym
lC_9	−119 (62)	167 (152)	85 (108)	100 (111)	146	−168	18 (7)	−175	76	4.2	with lT_{10}, lA_3' sym, and rA_3' sym*
lT_{10}	−94	76	72	−149	−134		0	−174	90		with lC_9

[a] Torsion angle (deg) definitions: α = C′−N1′−C2′−C3′; β = N1′−C2′−C3′−N4′; γ = C2′−C3′−N4′−C5′; δ = C3′−N4′−C5′−C′; ϵ = N4′−C5′−C′−O1′; ω = C5′−C′−N1′−C2′; χ_1 = C3′−N4′−C7′−C8′; χ_2 = N4′−C7′−C8′−N1/N9; χ_3 = C7′−C8′−N1/N9−C2/C4 (Figure 1). Numbers in parentheses refer to the torsion angles for the second conformation about the amide bond. [b] The distance (Å) given is between N1 (pyrimidine bases)/N9 (purine bases) and N1 (pyrimidine bases)/N9 (purine bases) of the next PNA unit. [c] Sym and sym* indicate two strands generated by different crystallographic symmetry operations.

and left-handed strands, respectively) pointing in separate directions (see Figure 4A,C). In essence, these are equivalent to backbone conformations of opposite handedness (Table 3). The two conformations of A_3 give rise to two different base-pairing interactions of A_3 with A_8 of a symmetry-related strand (Table 1 and Figure 4B). rA_3 and lA_3' are engaged in reverse Hoogsteen hydrogen bonding[21] to the symmetry-related A_8 of opposite handedness (Figures 2 and 4A,B), by interstrand intercalation of A_3 between A_8 and C_9 of the corresponding duplex strand. The reverse Hoogsteen hydrogen bonding is in the family of the trans Watson−Crick/Hoogsteen base pair according to the annotation of Leontis−Westhof.[22] The rA_3' and lA_3 are interacting through a single hydrogen bond (C6-NH_2 to N1) to A_8 of the corresponding "duplex" strand.

Thirty-three water molecules have been located in the structure of the PNA. Most of these are found to occupy similar hydration sites in the structure as observed in other crystal structures of PNA.[9−13,16] The water molecules bridging the N1′ or O1′ in the backbone and the O2 (pyrimidine) or N3 (purine) of the bases in the "minor groove" are found only in the PNA residues forming Watson−Crick base pairing ($G_4A_5T_6C_7$). Water molecules are also seen, forming hydrogen bonds to the carbonyl oxygen (O7′) in the linker of the PNA skeleton. These water molecules are distributed equally among all substructures, and are not favored in, e.g., the Watson−Crick duplex.

To examine whether the complex structure observed in the crystal, in particular the triplex part, might also be present or play a role in solution, thermal denaturation studies were undertaken. The existence of a well-defined bimolecular or higher molecular complex in solution was clearly indicated by a T_m concentration dependency of PNA (Figure 5A). Formation of Hoogsteen base pairing between a protonated C_9 and G_4 is stabilized by low pH. Thus, contributions of this interaction to the overall stability of the complex should be reflected in the pH dependency of the T_m. However, T_m did not increase at pH below 6.0 (Figure 5B); in fact, T_m decreased. Therefore, it is concluded that in solution the Watson−Crick duplex is not stabilized by triplex formation.

Discussion

The ability of polynucleotides to form secondary and tertiary structures is pivotal for the biological function of structural and catalytic RNAs (e.g., ribosomal RNA, tRNA, and ribozymes), and is also a prerequisite for generation of polynucleotide aptamers. The nucleic acid mimicking properties of PNA in terms of duplex, triplex, and hairpin formation is well recognized, but higher order structures as demonstrated here have not been seen previously. The present results illustrate the structural diversity of PNA interactions and emphasize the general importance of base stacking to guide hydrogen bonding and vice versa.

Thus, the crystal structure reveals several highly surprising features relating to nucleic acid recognition in general and to PNA properties in particular, such as interstrand intercalation, triplex formation, and backbone chirality shift. The structure is clearly controlled by the Watson−Crick complementary tetramer in the middle of the sequence forming an antiparallel duplex

(21) *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: New York, 1999.
(22) Leontis, N. B.; Westhof, E. *RNA* **2001**, *7*, 499−512.

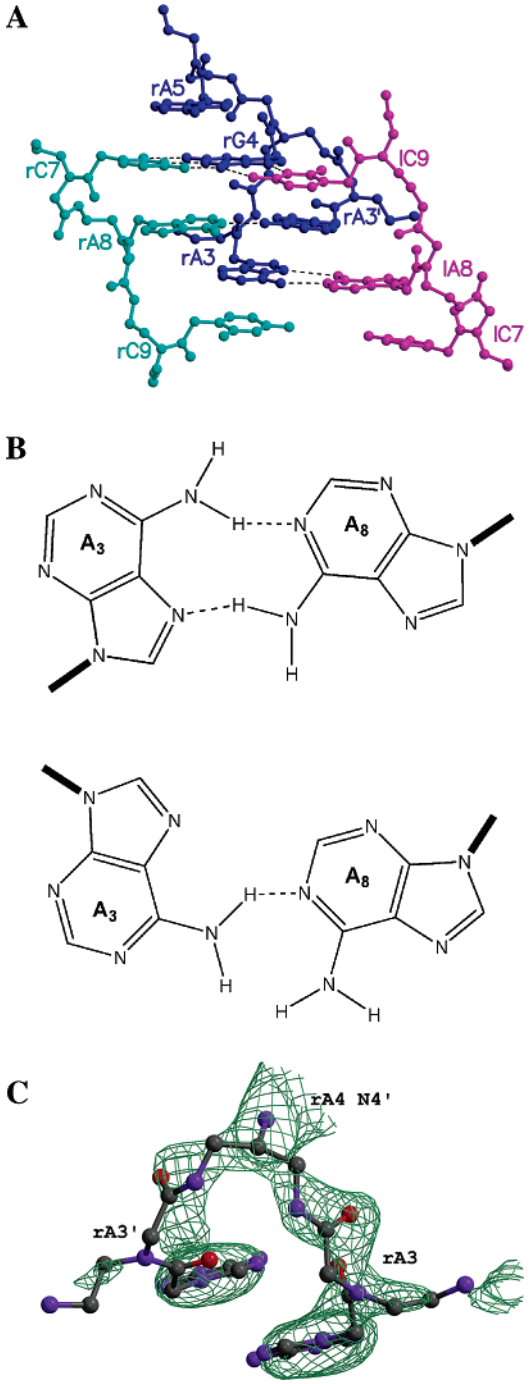


Figure 4. (A) Closeup view of the two conformations of A_3 in the predominantly right-handed strand (rA_3 and rA_3': blue) and their interactions with A_8 from two different symmetry-related strands. In the two conformations of A_3, this PNA unit is located at very different positions giving rise to nonstandard A−A base pairing. rA_3 forms reverse Hoogsteen base pairing with lA_8 of a left-handed symmetry-related molecule (magenta). In the rA_3' conformation, a hydrogen bond is formed between C6-NH_2 and N1 of rA_8 of a right-handed symmetry-related molecule (cyan). The same interactions are formed in the predominantly left-handed strand (not shown). Hydrogen bonds between the nucleobases are shown as black dotted lines. (B) Schematic representations of the reverse Hoogsteen base pairing (upper) and the C6-NH_2 to N1 hydrogen bonding (lower). (C) Omit electron density map (green; contoured at 1 σ) calculated from a partly refined model where the PNA units A_3 in the right- and left-handed strands were omitted. The fully refined structure of the partly self-complementary PNA is shown and color coded according to atom type.

with "dangling tails" of three nucleobases at each end (Figure 2). Two such duplexes are then making head-to-tail interactions

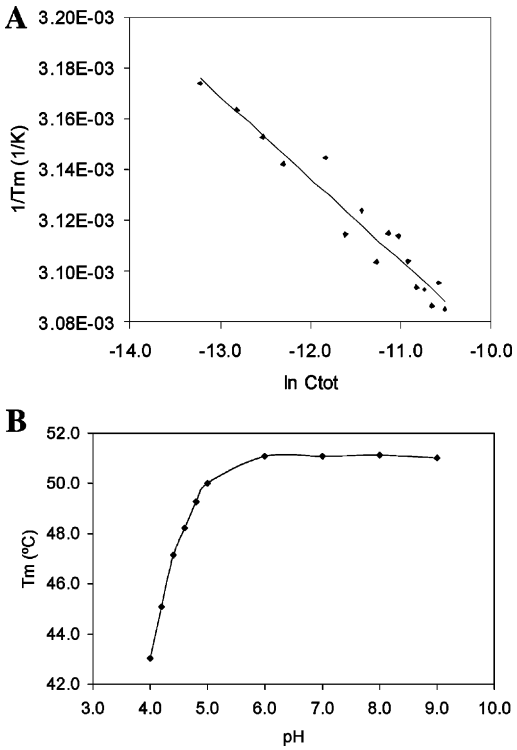


Figure 5. UV melting experiments. (A) T_m's dependency on concentration of the partly self-complementary PNA represented in a $1/T_m$ versus ln C_{tot} plot, where C_{tot} is the total concentration of PNA in the samples. The linearity of this plot shows that a complex between two or more strands of the PNA is formed. (B) T_m's dependency on pH represented in a T_m versus pH plot. T_m increases with pH until pH 6.0, implying that triplex or higher order structure is not stabilizing the Watson−Crick duplex.

in a very unique way in which A_3 from the one strand is switched to the other strand of the duplex by intercalation between A_8 and C_9 (Table 3 and Figures 2 and 4A). This arrangement is stabilized by reversed Hoogsteen base pairing between A_3 in duplex 1 and A_8 in duplex 2, and further positions C_9 and T_{10} of one tail to form regular triplex Hoogsteen bonding to G_4 and A_5 of the other duplex (Figure 3C). Indeed, this is the first reported three-dimensional structure of a PNA−PNA−PNA triple helix and confirms the results from circular dichroism measurements showing the formation of PNA triplexes in solution.[17] With this arrangement a continuous, multimeric structure is formed in the crystal. In an alternative arrangement giving a similar overall structure, A_3 stacks continuously in the PNA strand with G_4 and pairs with A_8 in the duplex (Table 3 and Figures 2 and 4). This base pair is, however, only stabilized by a single hydrogen bond (between N1 and C6-NH_2). The two structures are related by interchanging the positions of rA_3 and lA_3'. The present structure shows that PNA is capable of forming a supramolecular polymer in the crystal. In that respect, this higher order structure is comparable to what have been observed for RNA and DNA, e.g., for the short DNA fragment dATAT where polymers are formed in the crystal by Watson−Crick hydrogen bonding.[23]

Overall, the structure emphasizes how the oligomers by "molecular puzzle" self-assemble to ascertain continuous base stacking as well as to optimize base pairing. This principle is naturally fully analogous to that controlling the folding of

(23) Viswamitra, M. A.; Kennard, O.; Jones, P. G.; Sheldrick, G. M.; Salisbury, S.; Favello, L.; Shakked, Z. *Nature* **1978**, *273*, 687−688.

biologically functional RNA molecules in ribozymes[24,25] and artificially developed aptamers.[26,27] The structure also shows that PNA oligomers, because of the high flexibility and noncharged nature of their backbone, exhibit a large structural versatility, and future studies will reveal how the repertoire of structures adoptable by PNA compares to that of natural nucleic acids. The present examples of structural diversity may also provide, at least part of, the explanation why some PNA oligomers have a tendency for aggregation in solution, by formation of interstrand networks through non-Watson−Crick nucleobase interactions.

A large variety of "nonstandard" nucleobase interactions are known from the "RNA structural world". Indeed, the trans Watson−Crick/Hoogsteen base pair between A_3 in one strand and A_8 from a symmetry-related strand in the PNA (Figures 2 and 4) observed in the present PNA structure has previously been observed in a structure of an RNA duplex with an internal loop[28] and in the structure of the yeast initiator tRNA.[29] In the 30S and 50S ribosomal subunits, an A−A trans Watson−Crick/Hoogsteen base pair is present at several locations in both double and triple helices.[30,31] In the present structure, A_3 and A_8 are in two parallel strands of opposite handedness, which of course is different from the situation in RNA molecules.[28−31] The large variety of interactions observed in the present structure gives rise to a series of different conformations of the PNA backbone (Table 3). The impressive structural diversity displayed by RNA likewise is associated with distinct variations in backbone conformations,[32,33] now amounting to 32 conformational classes of RNA. However, the backbone conformations of PNA do not naturally group into any of the 32 conformational classes of RNA due to the fundamental differences between the two backbones.

Most interestingly, the two C-terminal PNA units (C_9T_{10}) within the PNA strand must change handedness (Table 3 and Figures 2 and 3B,C) to interact with the adjacent "duplex" of opposite handedness. Induction of preferred handedness (chirality) to PNA strands or duplexes by terminally attached chiral ligands such as amino acids has previously been reported.[34,35] This conversion has been suggested to arise from the site of the stereogenic center and to migrate through the stack of bases.[34,35] The chiral induction was observed to be a relatively slow process, and one mechanism considered for this was chirality shift within the backbone. The present PNA structure clearly demonstrates that such a shift is indeed structurally feasible. Of note, a model of double-stranded polynucleotides has previously been proposed, in which right- and left-handed DNA sequences alternate along the chain.[36,37]

It has previously been suggested—contrary to chemical and structural intuition—that the PNA backbone possesses a low degree of flexibility.[13,38] The present crystal structure demonstrates a very high adaptability of the backbone of PNA. It is possible for a single PNA strand to not only change local handedness, but also to adopt an extended conformation that allows for intercalation of a nucleobase from another strand. The high adaptability of the PNA backbone allows PNA to form several types of "standard" hydrogen-bonding schemes seen for DNA and RNA. We speculate that the possibility of changes in local handedness might create an even larger and yet undiscovered repertoire of nucleobase stacking and hydrogen-bonding interactions in higher order PNA structures than found in RNA. These results would thus inspire the search for PNA "foldamers" and aptamers, and for PNAzymes using backbone functionalized PNA.[39] Also, the design of PNA nanostructures may gain inspiration from the present findings.

Experimental Section

Crystallization and Data Collection. The PNAs (H-GTAGAT-CACT-L-Lys-NH_2 and H-GTAGA-(5-BrU)-CACT-L-Lys-NH_2) were synthesized as described previously.[40] The hanging-drop vapor diffusion method was used to crystallize the PNA. Crystals of H-GTAGAT-CACT-L-Lys-NH_2 grew in a drop of 1 μL of a 2.5−5 mg/mL aqueous solution of the PNA mixed with a drop of 1 μL of reservoir solution containing 0.1 M $MgCl_2$, 10% poly(ethylene glycol) 8000, and 10% glycerol, and equilibrated against 0.5 mL of reservoir solution at 20 °C. A native data set was collected at 293 K on one crystal (size 150 × 40 × 10 μm) using synchrotron radiation, $\lambda = 1.0000$ Å (ESRF beamline BM14, Grenoble) and a MAR image plate detector. The crystals belong to space group *C*2 with cell dimensions $a = 55.57$ Å, $b = 41.91$ Å, $c = 30.14$ Å, and $\beta = 117.65°$.

Crystals of a 5-bromouracil (5-BrU) derivative were obtained by the sitting-drop vapor diffusion method using a drop consisting of 3 μL of 1.9 mg/mL aqueous solution of the 5-BrU derivative and 2 μL of reservoir solution containing 20% 2-propanol, 0.05 M $CaCl_2$, and 0.1 M sodium acetate at pH 4.8. The drops were equilibrated against 0.5 mL of reservoir solution. After repeated macroseeding, the crystals grew to a maximum size of 350 × 50 × 10 μm. The derivative crystals belong to space group *C*2 with the unit cell dimensions $a = 55.86$ Å, $b = 42.09$ Å, $c = 29.96$ Å, and $\beta = 120.14°$. The crystals were transferred through a reservoir solution containing additional 30% glycerol as a cryoprotectant before data collection. A MAD data set was collected at four different wavelengths from one crystal at 110 K using synchrotron radiation (EMBL beamline BW7A, Hamburg) and a MAR CCD detector.

Both native and derivative data were processed using the programs DENZO and SCALEPACK[41] and the CCP4 suite of programs.[42] The data statistics are shown in Table 4.

X-ray Structure Determination and Structure Refinement. Two bromine atom sites were located with the program SOLVE, version 2.00,[43] using data between 20.0 and 2.6 Å (figure of merit (FOM) =

(24) Steitz, T. A.; Moore, P. B. *Trends Biochem. Sci.* **2003**, *28*, 411−418.
(25) Lilley, D. M. *Trends Biochem. Sci.* **2003**, *28*, 495−501.
(26) Brody, E. N.; Gold, L. *J. Biotechnol.* **2000**, *74*, 5−13.
(27) Burgstaller, P.; Jenne, A.; Blind, M. *Curr. Opin. Drug Discov. Dev.* **2002**, *5*, 690−700.
(28) Baeyens, K. J.; De Bondt, H. L.; Pardi, A.; Holbrook, S. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12851−12855.
(29) Basavappa, R.; Sigler, P. B. *EMBO J.* **1991**, *10*, 3105−3111.
(30) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Nature* **2000**, *407*, 340−348.
(31) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905−920.
(32) Schneider, B.; Moravek, Z.; Berman, H. M. *Nucleic Acids Res.* **2004**, *32*, 1666−1677.
(33) Murray, L. J. W.; Arendall, W. B., III; Richardson, D. C.; Richardson, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13904−13909.
(34) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Nordén, B. *Nature* **1994**, *368*, 561−563.
(35) Wittung, P.; Eriksson, M.; Lyng, R.; Nielsen, P. E.; Nordén, B. *J. Am. Chem. Soc.* **1995**, *117*, 10167−10173.
(36) Rodley, G. A.; Scobie, R. S.; Bates, R. H. T.; Lewitt, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2959−2963.
(37) Millane, R. P.; Rodley, G. A. *Nucleic Acids Res.* **1981**, *9*, 1765−1773.
(38) Sen, S.; Nilsson, L. *J. Am. Chem. Soc.* **1998**, *120*, 619−631.
(39) (a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1939−1942. (b) Püschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707−4710.
(40) Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J. *J. Peptide Sci.* **1995**, *3*, 175−183.
(41) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.
(42) Collaborative Computational Project, Number 4. *Acta Crystallogr., Sect. D* **1994**, *D50*, 760−763.
(43) Terwilliger, T. C.; Berendzen, J. *Acta Crystallogr., Sect. D* **1999**, *D55*, 849−861. www.solve.lanl.gov.

Table 4. Statistics on Diffraction Data and Refinement[a]

Diffraction Data

	native	5-BrU λ_1	5-BrU λ_2	5-BrU λ_3	5-BrU λ_4
wavelength, λ (Å)	1.0000	0.9177	0.9185	0.9110	0.9218
resolution (Å)	30.0−2.6 (2.69−2.60)	20.0−2.4 (2.49−2.40)	20.0−2.4 (2.49−2.40)	20.0−2.4 (2.49−2.40)	20.0−2.4 (2.49−2.40)
unique reflections	1839	2443	2445	2437	2425
completeness (%)	98.6 (95.6)	99.6 (98.8)	99.6 (97.9)	99.5 (97.5)	98.7 (89.8)
multiplicity	3.0	4.7	4.7	4.7	3.6
$R_{sym}(I)$[b] (%)	4.5 (18.3)	5.7 (31.5)	5.8 (33.1)	6.2 (37.0)	5.5 (33.0)
$I/\sigma(I)$	22.4 (5.3)	15.5 (3.2)	15.6 (2.9)	15.1 (2.8)	14.5 (2.3)

Refinement

	native
resolution (Å)	27.0−2.60
R-value/R_{free}-value[c] (%)	25.5/27.7
total number of non-hydrogen atoms	380
total number of water molecules	33
average B-values of PNA (Å^2)	42
average B-values of water molecules (Å^2)	50
rms[d] deviation of bond lengths (Å)	0.009
rms deviation of bond angles (deg)	1.3

[a] Numbers in parentheses refer to data for the outer bin. [b] R_{sym} = agreement between symmetry-related reflections. [c] R_{free}-value = cross-validation R-value for test set of reflections (10%) omitted during the refinement. [d] rms= root-mean-square.

0.44 for all reflections). Refinement of the positions of the two bromine atoms and density modifications were performed using RESOLVE, version 2.00[44] (FOM = 0.52 for all reflections). The calculated electron density was easily interpreted, and two PNA strands, one predominantly right-handed and one predominantly left-handed, were located. Eight out of ten PNA units were traced in the electron density using the program O.[45] Following model building, one refinement cycle with CNS[46] using the MAD data to 2.40 Å resolution was performed. Then, interchanging cycles of model building and refinements using the native data to 2.60 Å resolution were performed with the programs O[45] and CNS,[46] respectively. In refinements, the maximum-likelihood method, individual B-factors, and bulk solvent corrections were applied. The two first PNA units and the L-Lys residue in each PNA strand could not be interpreted in the electron density maps. Water molecules were added gradually.

The two conformations of the amide bond between the PNA units C_7A_8 and A_8C_9 in each strand (in one conformation, the carbonyl oxygen points toward the C-terminus, and in the other conformation the carbonyl oxygen points into the minor groove) were refined using a constant occupancy of 0.5. Similarly, the PNA unit A_3 of both the right- and left-handed strands were observed in two conformations (apparently right- and left-handed conformations, denoted rA_3, rA_3', lA_3, and lA_3'). Initially, only the rA_3 and lA_3' conformations were modeled, but various omit maps clearly indicated the presence of an additional conformation even at this intermediate resolution of 2.6 Å. However, the electron density is more well-defined for rA_3 in the right-handed strand and for lA_3' in the left-handed strand than for rA_3' and lA_3 (Figure 4C). The two conformations were refined with equal occupancy of 0.5. By the inclusion of these double conformations in the refinement, both the conventional R-value and R_{free} decreased significantly. For statistics on refinement, see Table 4. The program Curves[47] was used to study the helix conformation of the Watson−Crick duplexes of the PNA. The noncanonical base pair database, NCIR,[48] was used in the analysis of the A−A base pairs. Figures 2A, 3A,B, and 4A were made with the programs Molscript[49] and Raster3D,[50] and Figures 3C and 4C were made with the programs Bobscript[51] and Raster3D.[50]

UV Thermal Melting Experiments. The sample solutions were prepared with buffers containing 10 mM sodium phosphate, 100 mM NaCl, and 0.1 mM EDTA. In the study of concentration dependency, pH was 7.0 and the concentration of the partly self-complementary PNA was varied from 0.9 to 27 μM. In the study of pH dependency, pH was varied between 4.0 and 9.0 and the concentration of PNA was 25.3 μM. The solutions were heated to 90.0 °C and then annealed to form the complex prior to the measurements. Absorbance versus temperature profiles were obtained at 260 nm on a Cary 300 BIO UV−Visible Spectrophotometer, interfaced with a Zitech computer. Quartz cells with 1.0 cm path length were used to hold the samples. The temperature of the sample was increased continuously from 2.0 to 90.0 °C at a rate of 1.0 °C/min using an ethylene glycol water-circulating bath fitted with a thermoregulator (Cary temperature controller), and the absorbance of the sample was measured every 0.5 °C. Reproducible melting curves were obtained for each solution. The UV melting curves were analyzed as described previously.[52]

Acknowledgment. We thank Jolanta Jørgensen and Brian Rosenberg for technical assistance. This work was supported by grants from The Dansync Center for Synchrotron Radiation, PharmaBiotec, The Lundbeck Foundation, Apotekerfonden of 1991, and The Danish Natural Science Research Council. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and thank Andrew Thompson for assistance in using beamline BM14. We acknowledge access to the EMBL BW7A beamline at the DORIS storage ring, DESY, Hamburg, and support from the European Community—Research Infrastructure Action under the FP6 "Structuring the European Research Area Programme" Contract No. RII3/CT/2004/5060008.

JA0458726

(44) (a) Terwilliger, T. C. *Acta Crystallogr., Sect. D* **1999**, *D55*, 1863−1871. (b) Terwilliger, T. C. *Acta Crystallogr., Sect. D* **2000**, *D56*, 965−972. (c) Terwilliger, T. C. *Acta Crystallogr., Sect. D* **2001**, *D57*, 1755−1762. (d) Terwilliger, T. C. *Acta Crystallogr., Sect. D* **2001**, *D57*, 1763−1775.
(45) Jones, T. A.; Zou, J.-Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr., Sect. A* **1991**, *A47*, 110−119.
(46) Brünger, A. T. *Acta Crystallogr., Sect. D* **1998**, *D54*, 905−921.
(47) Lavery, R.; Sklenar, H. J. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(48) Nagaswamy, U.; Voss, N.; Zhang, Z.; Fox, G. E. *Nucleic Acid Res.* **2000**, *28*, 375−376.
(49) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(50) Merritt, E. A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505−524.
(51) Esnouf, R. M. *Acta Crystallogr., Sect. D* **1999**, *D55*, 938−940.
(52) Ratilainen, T.; Nordén, B. In *Peptide Nucleic Acids: Methods and Protocols*; Nielsen, P. E., Ed.; Methods in Molecular Biology 208; Humana Press Inc.: Totowa; NJ, 2002; pp 59−88.