[4] Studies in aqueous media have either registered very low binding constants, or have employed carbohydrate substrates with hydrophobic appendages. For examples, see: Y. Aoyama, Y. Nagai, J. Otsuki, K. Kobayashi, H. Toi, *Angew. Chem.* **1992**, *104*, 785; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 745; A. V. Eliseev, H.-J. Schneider, *J. Am. Chem. Soc.* **1994**, *116*, 6081; K. Kobayashi, Y. Asakawa, Y. Kato, Y. Aoyama, *J. Am. Chem. Soc.* **1992**, *114*, 10307; B.-L. Poh, C. M. Tan, *Tetrahedron* **1993**, *49*, 9581; J. Jimenez-Barbero, E. Junquera, M. Martinpastor, S. Sharma, C. Vicent, S. Penades, *J. Am. Chem. Soc.* **1995**, *117*, 11198; S. A. Staley, B. D. Smith, *Tetrahedron Lett.* **1996**, *37*, 283.

[5] Reports from this laboratory: a) R. P. Bonar-Law, A. P. Davis, B. A. Murray, *Angew. Chem.* **1990**, *102*, 1497; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1407; b) K. M. Bhattarai, R. P. Bonar-Law, A. P. Davis, B. A. Murray, *J. Chem. Soc. Chem. Commun.* **1992**, 752; c) A. P. Davis, S. Menzer, J. J. Walsh, D. J. Williams, *Chem. Commun.* **1996**, 453; d) K. M. Bhattarai, A. P. Davis, J. J. Perry, C. J. Walter, S. Menzer, D. J. Williams, *J. Org. Chem.* **1997**, *62*, 8463. Representative examples from other laboratories: e) Y. Aoyama, Y. Tanaka, S. Sugahara, *J. Am. Chem. Soc.* **1989**, *111*, 5397; f) Y. Kikuchi, Y. Tanaka, S. Sutaro, K. Kobayashi, H. Toi, Y. Aoyama, *J. Am. Chem. Soc.* **1992**, *114*, 10302; g) R. Liu, W. C. Still, *Tetrahedron Lett.* **1993**, *34*, 2573; h) C.-Y. Huang, L. A. Cabell, E. V. Anslyn, *J. Am. Chem. Soc.* **1994**, *116*, 2778. i) M. Inouye, T. Miyake, M. Furusyo, H. Nakazumi, *J. Am. Chem. Soc.* **1995**, *117*, 12416; j) G. Das, A. D. Hamilton, *J. Am. Chem. Soc.* **1994**, *116*, 11139.

[6] a) S. Anderson, U. Neidlein, V. Gramlich, F. Diederich, *Angew. Chem.* **1995**, *107*, 1722; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1596; b) U. Neidlein, F. Diederich, *Chem. Commun.* **1996**, 1493; c) P. B. Savage, S. H. Gellman, *J. Am. Chem. Soc.* **1993**, *115*, 10448; d) R. P. Bonar-Law, J. K. M. Sanders, *J. Am. Chem. Soc.* **1995**, *117*, 259; e) T. Mizutani, T. Murakami, N. Matsumi, T. Kurahashi, H. Ogoshi, *J. Chem. Soc. Chem. Commun.* **1995**, 1257. The receptors in references [6a–c] exploit strong hydrogen bonds to ionic centers, while those in references [6d,e] employ Zn–O interactions.

[7] F. A. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293.

[8] MacroModel V5.5; F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440. Calculations employed the MM2* force field and $CHCl_3$ GB/SA solvation.

[9] **1**: ^{1}H NMR (400 MHz, [D$_6$]DMSO): $\delta = 0.91$ (t, J = 7.0 Hz, 12H, CH_3), 1.38 (m, 16H, $CH_2CH_2CH_3$), 1.76 (quintet, J = 7.0 Hz, 8H, $CH_2CH_2CH_2CH_2$), 4.33, (t, J = 6.5 Hz, 8H, CO_2CH_2), 4.43 (dq, J = 18.0, 5.0 Hz, 16H, CH_2NH), 7.26 (s, 4H, biaryl Ar*H*), 7.81 (s, 8H, biaryl Ar*H*), 8.15 (s, 4H, spacer Ar*H*), 8.46 (s, 8H, spacer Ar*H*), 9.09 (t, J = 5.0 Hz, 4H, NH); ^{13}C NMR (100.6 MHz, DMSO-d$_6$): $\delta = 13.83$ (CH_3), 21.77 (CH_2), 27.60 (CH_2), 27.84 (CH_2), 43.73 (NHCH_2), 65.26 (CO_2CH_2), 124.87 (Ar CH), 129.11 (Ar CH), 130.27 (2 × Ar CH), 130.38 (Ar C), 135.26 (Ar C), 136.70 (Ar C), 139.76 (Ar C), 164.77 (*C*O), 164.80 (*C*O); LSI-MS: *m*/*z*: 1517.7 [*M*H$^+$] (100), 534.9 (20), 460.9 (68).

[10] This unexpected direction of shift may be due to displacement of methanol by the carbohydrate, giving a structure in which not all NH groups are maximally hydrogen-bonded. Alternatively, a conformational change may take place, altering the position of the NH protons in relation to the arene rings.

[11] At first sight it might be thought that, at fast exchange, all eight NH protons would be equivalent. However, the carbohydrate substrate lacks any element of symmetry, and therefore creates eight different NH environments on interaction with the receptor. Coalescence of these eight signals would require fast exchange between eight equivalent binding orientations. The symmetry of **1** allows only four such orientations, and the NH protons are therefore split into two groups of four.

[12] Data from the ^{1}H NMR studies in $CDCl_3/CD_3OH$ (92:8) suggest that the carbohydrate protons experience upfeld shifts of about 0.2–0.3 ppm on complexation. These modest shifts might imply attachment to the exterior of **1**, or might reflect competing shielding/deshielding effects from the arene rings defining the cavity.

[13] ^{1}H NMR titrations analyzed by HOSTEST 5.0. See: a) C. S. Wilcox in *Frontiers in Supramolecular Chemistry and Photochemistry* (Eds.: H.-J. Schneider, H. Dürr), VCH, Weinheim, **1990**, p. 123.

[14] Y. Tanaka, C. Khare, M. Yonezawa, Y. Aoyama, *Tetrahedron Lett.* **1990**, *31*, 6193.

[15] Inverse micelles: N. Greenspoon, E. Wachtel, *J. Am. Chem. Soc.* **1991**, *113*, 7233; T. Kida, D. Furue, A. Masuyama, Y. Nakatsuji, I. Ikeda, *Chem. Lett.* **1996**, 733. "Unimolecular inverse micelles": K. Kobayashi, F. Ikeuchi, S. Inaba, Y. Aoyama, *J. Am. Chem. Soc.* **1992**, *114*, 1105.

[16] F. Franks, P. J. Lillford, G. Robinson, *J. Chem. Soc. Faraday Trans. 1* **1989**, *85*, 2417.

Oligomers with Intercalating Cytosine–Cytosine$^+$ Base Pairs and Peptide Backbone: DNA i-Motif Analogues**

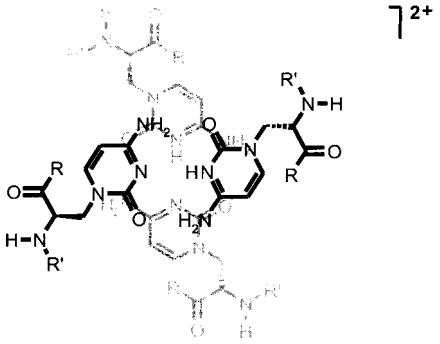
Ulf Diederichsen*

The ends of chromosomes are built up of telomeric DNA with an up to 1000-fold repeating guanine-rich DNA and complementary cytosine-rich sequence.[1] The organization of guanine strands in tetrads based on a guanine donor/acceptor pattern that allows for simultaneous pairing by the Watson–Crick and Hoogsteen sides is a familiar structural motif.[2] Surprisingly, however, the cytosine-rich oligomers are also able to form tetramers since pyrimidine molecules lack a Hoogsteen side they contribute less to the complex stability through stacking interactions. Nevertheless, at pH 4.5–6.5 the semiprotonated C–C$^+$ base pairs are formed in the reverse Watson–Crick mode with three hydrogen bonds building up a parallel double strand; two C–C$^+$ pairing double strands intercalate and stabilize the structure in a four-stranded complex.[3–5] The interaction of consecutive base pairs is provided only by the stacking of the exocyclic carbonyl and amino groups that are oriented with opposed dipoles (Scheme 1). This base-pair arrangement makes a considerable contribution to the stabilization of the pairing complex as indicated by a base-pair distance of only 3.1 Å. DNA tetramers containing C–C$^+$ pairing were called i-motif in reference to intercalating double strands, and were described by Gehring et al. in an NMR structural analysis of the DNA oligomer d(TCCCCC) that forms an i-motif[3] and by Chen et al. with an X-ray structure of d(C_4).[4] In the meantime, the intercalating C–C$^+$ double strands that form C tetrads were established as a structural motif.[5] Unwinding of the double helix is required to facilitate intercalation and leads to a helical twist of only 12.4–16° in the DNA i-motif. As part of

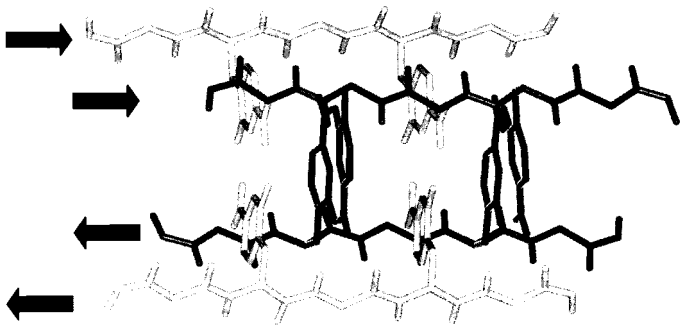
[*] Dr. U. Diederichsen
Institut für Organische Chemie und Biochemie der Technischen Universität München
Lichtenbergstrasse 4, D-85747 Garching (Germany)
Fax: (+49) 89-2891-3210
E-mail: ud@linda.org.chemie.tu-muenchen.de

[**] This work was supported by the Deutsche Forschungsgemeinschaft (Di 542 2-1, Di 542 3-1), the Fonds der Chemischen Industrie, and the Leonhard-Lorenz-Stiftung. I am grateful to Professor H. Kessler, Garching, for generous support.

  1433-7851/98/3716-2273 $ 17.50+.50/0



Scheme 1. Arrangement of consecutive C–C$^+$ base pair planes of an alanyl-PNA tetrad analogous to the DNA i-motif.



Scheme 2. Part of the alanyl-PNA tetrad formed analogous to the DNA i-motif: Oligomers H-(Gly-AlaC)$_4$-Lys-NH$_2$ (**2**) and H-(Gly-AlaC)$_4$-Lys-NH$_2$ (*ent*-**2**) pair in an antiparallel double strand. Two of these double strands stabilize by interpenetration.

our ongoing investigation of alanyl-PNA (PNA = peptide nucleic acids),[6] cytosine–cytosine self-pairing gained our interest because we expected alanyl-PNA to be a suitable i-motif analogue by mimicking backbones that are close to linear.

Alanyl-PNA consists of a regular peptide backbone with an alternating configuration of the alanyl amino acids, which have nucleobases connected in the β position of the side chain. This arrangement leads to the favorable formation of double strands with an extended backbone. The linearity of alanyl-PNA pairing complexes is structurally inherent since the distance of two consecutive side chains is close to the stacking distance of 3.4 Å that is found in DNA. An alternating configuration of the nucleo amino acids is required for orientation of the nucleobases on the same side of the backbone.

To study C–C pairing we synthesized the octamers H-(AlaC-AlaC)$_4$-Lys-NH$_2$ (**1**) and H-(AlaC-AlaC)$_4$-Lys-NH$_2$ (*ent*-**1**; AlaC=β-(cytosine-1-yl)alanine amino acids; amino acids with D configuration are underlined; lysine amide was introduced for solubility reasons).[7] The nucleobase distance hampers intercalation since further extension of the backbone is not possible. In contrast, the oligomers H-(Gly-AlaC)$_4$-Lys-NH$_2$ (**2**) and H-(Gly-AlaC)$_4$-Lys-NH$_2$ (*ent*-**2**) contain a glycine unit as every second amino acid thus providing a spacer that facilitates the intercalation of linear double strands (Scheme 2). Because of the pairing selectivity of alanyl-PNA the oligomers had to be synthesized in both enantiomeric forms: since the reverse Watson–Crick mode offers the only possibility for C–C pairing, antiparallel strand orientation with base pairs formed by heterochiral nucleo amino acids or parallel backbone orientation with homochiral base pairs were required.[6]

We examined the formation of pairing complexes and their stability by temperature-dependent UV spectroscopy by measuring the cooperative destacking of nucleobases as a sigmoidal increase in absorption (hyperchromicity) with raising temperature. The self-pairing tendency of oligomers **1**, *ent*-**1**, **2**, and *ent*-**2** as well as pairing of the equimolar mixtures of the enantiomers **1** + *ent*-**1** or **2** + *ent*-**2** was examined at pH 7 (100 mM NaCl, 10 mM Na_2HPO_4/H_3PO_4) and pH 4.5 (100 mM NaCl, 80 mM tris-acetate).[8] At pH 7 none of the oligomers showed any indication of pairing, which is in agreement with the lack of pyrimidine–pyrimidine recognition in alanyl-PNA.[6] Further, self-pairing of **1** or *ent*-**1** as well as the parallel pairing of enantiomers **1** + *ent*-**1** at pH 4.5 was not observed. Therefore, eight C–C$^+$ base pairs with three hydrogen bonds each are not sufficient for double-strand formation. On the other hand, a stable pairing complex (T_m = 34 °C, 17% hyperchromicity, each 14 μM; Figure 1a) was found when intercalation of the antiparallel C–C$^+$ pairing double strands **2** and *ent*-**2** was facilitated. Complexation was reversible with hysteresis and was also concentration depend-

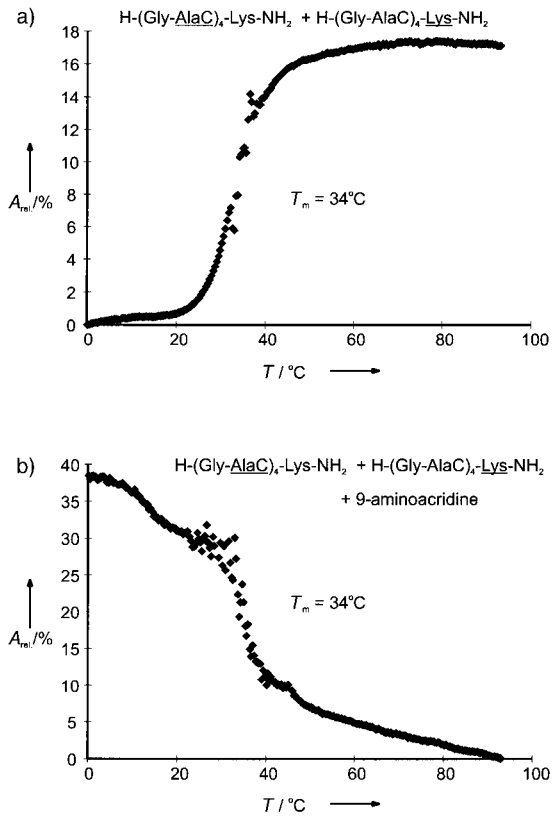


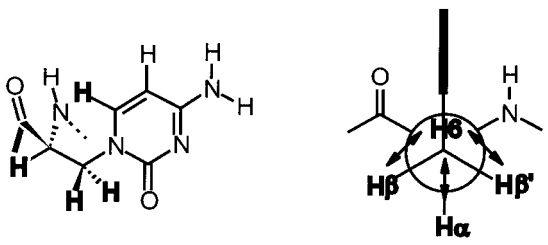
Figure 1. UV melting curves of a) an equimolar mixture of enantiomers **2** and *ent*-**2** (each 14 μM, 100 mM NaCl, 80 mM tris-acetate, pH 4.5, 260 nm) and b) of pairing complex **2** + *ent*-**2** in the presence of 9-aminoacridine (each 15 μM **2**, *ent*-**2**; 150 μM acridine, 100 mM NaCl, 80 mM tris-acetate, pH 4.5, 260 nm).

  1433-7851/98/3716-2274 $ 17.50+.50/0 *Angew. Chem. Int. Ed.* **1998**, *37*, No. 16

ent. Pairing was not observed with oligomer concentrations of lower than 10 μM each. Furthermore, self-pairing of alanyl-PNA **2** or *ent*-**2** was not observed. This is in agreement with parallel double strand orientation being disfavored in the alanyl-PNA series.[9] The alanyl-PNA tetraplex **2** + *ent*-**2** is more stable than the corresponding DNA i-motif. The melting stabilities of DNA were determined under the same conditions as $T_m = 23\,^\circ C$ (30% hyperchromicity, 14 μM) for the tetramer $d(C_4)$ and $T_m = 48\,^\circ C$ (35% hyperchromicity, 11 μM) for hexamer $d(C_6)$ (not shown).

In contrast to the nonoccurrence of the double strand of **1** with eight $C-C^+$ base pairs, the pairing complex **2** + *ent*-**2** was formed with high stability. This indicates—in analogy to the DNA i-motif—an organization of the latter into a tetrad based on intercalation of double strands, with each forming four $C-C^+$ base pairs. The special stacking orientation of the semiprotonated base pairs with opposed dipole orientation of the exocyclic amino and carbonyl functionalities seems to be decisive for the difference in pairing of a $C-C^+$ double strand and a tetramer. In contrast to the DNA i-motif, the alanyl-PNA tetrad is based on antiparallel pairing double strands. By comparison of the overall charge of $C-C^+$ pairing oligomers three cases can be realized: the DNA i-motif has a negative net charge, the self-pairing DNA dimer with an ethylene bridge between positions 3′ and 5′[10] is neutral, whereas alanyl-PNA **2** + *ent*-**2** has an overall positive charge.

Further evidence for **2** + *ent*-**2** organization in tetrads was obtained by gel filtration HPLC (Pharmacia Superdex-Peptide PC 3.2/30). In addition to the main peak with a retention time of 19.0 min a second small peak could be observed at 8.6 min when an aqueous buffer solution (100 mM NaCl, 80 mM tris-acetate, pH 4.5; flowrate 0.1 mL min^{-1}) was used as eluent. The latter signal most likely corresponds to the tetrameric species. In 2D NMR experiments of the **2** + *ent*-**2** complex (500 MHz, D_2O, each 2 mM, 100 mM NaCl, 80 mM tris-acetate) only peaks corresponding to the spin system of one oligomer were observed, as expected for the highly symmetric tetramer. Furthermore, ROESY cross-peaks were found between the proton on C6 (H6) of the nucleobase and the Hβ protons as well as with the Hα proton of the corresponding nucleo amino acid. Qualitatively this indicates a conformation of the nucleo amino acids with the nucleobase oriented antiperiplanar to Hα with the H–C6 pointing towards the backbone (Scheme 3).



Scheme 3. The most likely preferred conformation of the nucleo amino acids in the pairing complex **2** + *ent*-**2**. The observed ROEs (ROE = rotating frame Overhauser enhancement) between H6 and Hα, Hβ, and Hβ′ are indicated by arrows.

The UV melting curve of **2** + *ent*-**2** in the presence of ten equivalents of 9-aminoacridine showed that a degrading of the tetramer occurs, as observed by a hypochromic effect ($T_m =$ 34 °C, about 25% hypochromicity, reversible with hysteresis, each 15 μM alanyl-PNA, 150 μM acridine; Figure 1b). This effect is likely to be caused by the association of the acridine to the single or double strand, in a similar manner as has been described for the DNA dimer d(AC).[11] In contrast, the addition of 9-aminoacridine did not affect the melting curve of DNA $d(C_4)$ under comparable conditions.

Overall, the $C-C^+$ pairing in the alanyl-PNA tetrad represents a DNA i-motif analogue. Because of the inherent linearity of the alanyl-PNA backbone, alanyl-PNA was expected to be a suitable i-motif mimic. This $C-C^+$ pairing is the only exception to pyrimidine–pyrimidine recognition in the alanyl-PNA series. Since the i-motif was not observed for RNA, the sugar moiety and the internucleotide linkage were thought to be essential for formation of the i-motif.[12] The alanyl-PNA analogue indicates that the i-motif is not limited to nucleic acids. Finally, the unusual stacking of base pairs in the i-motif was found to be a prerequisite for the formation of a tetrameric structure.

Received: March 4, 1998 [Z11550IE]
German version: *Angew. Chem.* **1998**, *110*, 2395–2397

Keywords: base pairing · DNA structures · intercalations · peptide nucleic acids

[1] D. Voet, J. G. Voet, *Biochemistry*, 2. ed., Wiley, New York, **1995**.
[2] a) C. Kang, X. Zhang, R. Ratliff, R. Moyzis, A. Rich, *Nature* **1992,** *356*, 126; b) F. W. Smith, J. Feigon, *Nature* **1992**, *356*, 164; c) P. Schultze, F. W. Smith, J. Feigon, *Structure* **1994**, *2*, 221.
[3] K. Gehring, J.-L. Leroy, M. Guéron, *Nature* **1993**, *363*, 561.
[4] L. Chen, L. Cai, X. Zhang, A. Rich, *Biochemistry* **1994**, *33*, 13540.
[5] a) J. L. Leroy, K. Gehring, A. Kettani, M. Guéron, *Biochemistry* **1993**, *32*, 6019; b) J.-L. Leroy, M. Guéron, J.-L. Mergny, C. Hélène, *Nucleic Acids Res.* **1994**, *22*, 1600; c) C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, A. Rich, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11636; d) S. Ahmed, A. Kintanar, E. Henderson, *Nat. Struct. Biol.* **1994**, *1*, 83; e) I. Berger, C. Kang, A. Fredian, R. Ratliff, R. Moyzis, A. Rich, *Struct. Biol.* **1995**, *2*, 416; f) J.-L. Leroy, M. Guéron, *Structure* **1995**, *3*, 101; g) I. Berger, M. Egli, A. Rich, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12116.
[6] a) U. Diederichsen, *Angew. Chem.* **1996**, *108*, 458; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 445; b) U. Diederichsen, *Angew. Chem.* **1997**, *109*, 1966 *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1886; c) U. Diederichsen, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1743; d) U. Diederichsen, *Med. Chem. Lett.* **1998**, *8*, 145.
[7] For synthesis of the nucleo amino acid Boc-AlaC(Z)-OH (Boc = *tert*-butoxycarbonyl, Z = benzyloxycarbonyl) by the nucleophilic ring opening of the Boc-serine lactone with N^4-benzyloxycytosine, see ref. [6b] and P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang, A. Eschenmoser, *Croat Chem. Acta* **1996**, *69*, 535. Oligomerization by solid-phase peptide synthesis (SPPS) was carried out on a 4-methylbenzhydrylamine–polystyrene support loaded with L- or D-lysine(Z)-OH. Activation with *O*-(7-azabenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) provided coupling yields in excess of 95%. All oligomers were characterized by ^{1}H NMR spectroscopy and ESI-MS: **1**: *m/z*: 1587.4 $[MH]^+$, 794.3 $[MH_2^{2+}]$; *ent*-**1**: *m/z*: 794.0 $[MH_2^{2+}]$; **2**: *m/z*: 1094.4 $[MH^+]$, 548.1 $[MH_2^{2+}]$; *ent*-**2**: *m/z*: 1094.2 $[MH^+]$, 548.3 $[MH_2^{2+}]$.
[8] The following temperature program was used to control the heating block for the UV melting curves: 80 °C → −2 °C (20 min) → −2 °C (240 min) → 95 °C (180 min) → −2 °C (180 min) → −2 °C (240 min)

  1433-7851/98/3716-2275 $ 17.50+.50/0

→95 °C (180 min) → – 2 °C (180 min). The temperature was measured directly in solution.
[9] U. Diederichsen, unpublished results. Since interpenetrated C – C^+ base pairs prefer a separation of 3.1 Å and alanyl-PNA with an extended backbone would provide a base-pair distance of 3.6 Å, it is likely that tetrad **2** + *ent*-**2** lowers the distance by inclining the backbone axes or by helication. This would give a preference for the antiparallel backbone orientation; see J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, R. Krishnan, B. Jaun, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 259.
[10] M. Egli, P. Lubini, M. Bolli, M. Dobler, C. Leumann, *J. Am. Chem. Soc.* **1993**, *115*, 5855.
[11] E. Westhof, M. Sundaralingam, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1852.
[12] a) S. Robidoux, M. J. Damha, *J. Biomol. Struct. Dyn.* **1997**, *15*, 529. b) D. Collin, K. Gehring, *J. Am. Chem. Soc.* **1998**, *120*, 4069.

  1433-7851/98/3716-2276 $ 17.50+.50/0