Specific Purine-Purine Base Pairing in Linear Alanyl-Peptide Nucleic Acids

by **Markus F. H. Hoffmann**, **Arndt M. Brückner**, **Thomas Hupp**, **Bernd Engels***, and **Ulf Diederichsen***

Institut für Organische Chemie der Universität Würzburg, Am Hubland, D-97074 Würzburg

Dedicated to *Albert Eschenmoser* on the occasion of his 75th birthday

Purine-purine base pairing with guanine, isoguanine, 2,6-diaminopurine, and xanthine is investigated within the topology of alanyl-PNA. Alanyl-PNA is based on a regular peptide backbone with alternating configuration of the amino acids. The nucleobases are covalently linked as side chains. Their distance in peptides with β-sheet conformation is similar to the favored base-pair stacking distance. Therefore, alanyl-PNA provides self-pairing linear double-strands. The linear double-strand topology does not restrict base-pair size and geometry. The favored base pairs are formed mostly dependent on recognition by H-bonding. The synthesis of the nucleo-amino acids with unnatural nucleobases and their oligomerization is described. Hexamers and a tetramer based on 2,6-diaminopurine-xanthine and guanine-isoguanine base pairs were observed with very high stabilities. For xanthine-xanthine self-pairing, an unusual tridentate reverse *Watson-Crick* pairing mode is indicated, that is only possible with xanthines pairing in different tautomeric forms. To investigate the nature of xanthine-xanthine base pairs in more detail, quantum-chemical calculations were performed. They establish the easier tautomerization of xanthine compared to uracil and indicate that, in the AO basis-set limit, the tridentate pairing mode with mixed tautomers is favored.

1. Introduction. – Unequivocal base-pair complementarity is essential for DNA replication and transcription [1]. Base-pair recognition is based on H-bonding. Therefore, complementarity is given by the order of H-donor and -acceptor positions. As shown for the adenine-thymine base pair (*Fig. 1*), the H-bonding pattern alone is

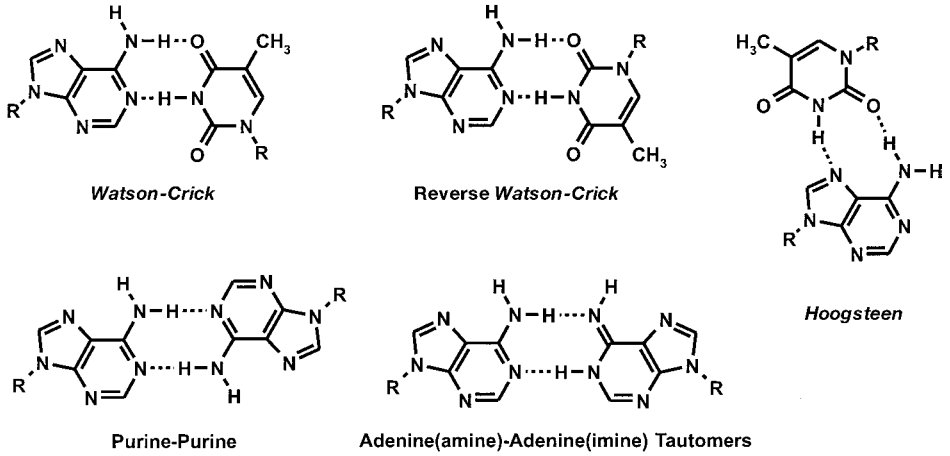


Fig. 1. *Variety of base-pairing in double strands with linear topology*

not at all sufficient for defined recognition: next to the *Watson-Crick* mode known from DNA double strands, there is pairing with the reversed orientation of nucleobases or pairing with the *Hoogsteen* side of purines. Further purine-purine and pyrimidine-pyrimidine pairing is possible, and finally we have to acknowledge pairing of nucleobases with different tautomers. With all these possibilities, it is not obvious how the favored recognition of adenine and thymine looks like. Nature is solving this problem by the helix topology that defines orientation and space of the base pairs and, therefore, restricts the pairing possibilities to the *Watson-Crick* mode of a purine with a pyrimidine [2]. Furthermore, G, C, A, and T are all nucleobases with a clear preference for only one tautomer.

Backbone-modified oligonucleotide analogues might lose the unequivocal complementarity of base pairs because of a change in double-strand topology [3]. Linear double strands are the borderline case with the least restrictions. There is no limit in base-pair size and orientation. Therefore, all purine-purine, purine-pyrimidine, and pyrimidine-pyrimidine combinations are allowed with all pairing modes. With base-pair planes being orthogonal to the backbone, even antiparallel and parallel strand orientation should be as likely.

With alanyl-peptide nucleic acids (alanyl-PNA), we introduced such a linear pairing system (*Fig. 2*) [4]: nucleo-amino acids based on alanyl units with covalently bound nucleobases on the side chains form a regular peptide backbone with alternating configuration of the amino acids. Because the distance between consecutive side chains of peptides with β-sheet conformation is very similar to the favored stacking distance, alanyl-PNA double strands are favored with linear topology. The orientation of nucleobases is well-defined, and the double strand has high rigidity.

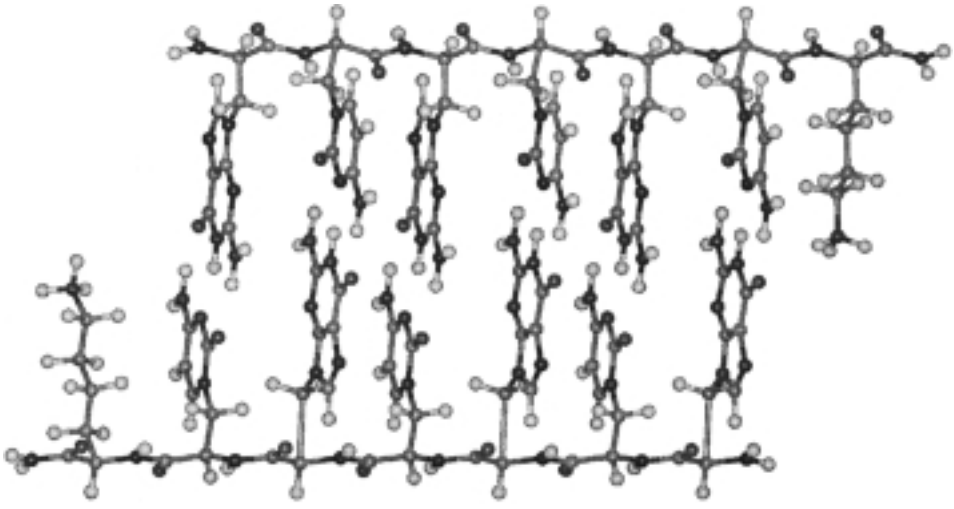


Fig. 2. *Model of a linear alanyl-PNA double strand.* The regular peptide strand has alternating configuration. Nucleobases are attached to the alanyl side chains. Double-strand formation is based on H-bonding (here tridentate *Watson-Crick* pairing between guanine and cytosine), base stacking, and solvation. The linearity of the double strand results from the distance of consecutive side chains with a peptide backbone in β-sheet conformation (3.6 Å) that determines the base-pair distance close to the ideal stacking distance (3.4 Å).

As expected, in alanyl-PNA all purine-purine and purine-pyrimidine base-pair combinations were found, and the preferred pairing mode differs from base pair to base pair [5]. The stabilities for hexameric double strands are in a similar order ($T_m = 21–32°$), because they are based on twodentate base pairs. Only the hexamer based on tridentate G-C pairs shows a much higher stability ($T_m = 58°$). This indicates that clear base-pair complementarity also might be reached in linear double strands if only base pairs with three H-bonds are considered. Further, the donor/acceptor complementarity of nucleobases should be unequivocal, and, for the overall double-strand geometry, nucleobase pairs should be isosteric within the same double strand. Therefore, we examined the purine-purine base pairing of 2,6-diaminopurine (D) with xanthine (X), and of isoguanine (I) with guanine (G) within the alanyl-PNA double strand (*Fig. 3*) (for a detailed study of G-I and X-D base pairs in homo-DNA, see [6]). Tridentate H-bonding in general is only possible over the *Watson-Crick* site. The base pairs D-X and I-G in the *Watson-Crick* mode are both isosteric and should lead to highest stability in a double strand. A tridentate reverse *Watson-Crick* mode is possible for D-X but not for I-G. Other base-pair combinations between D, X, I, and G are expected with only two H-bonds and, therefore, should not interfere with the desired D-X, I-G pairs.

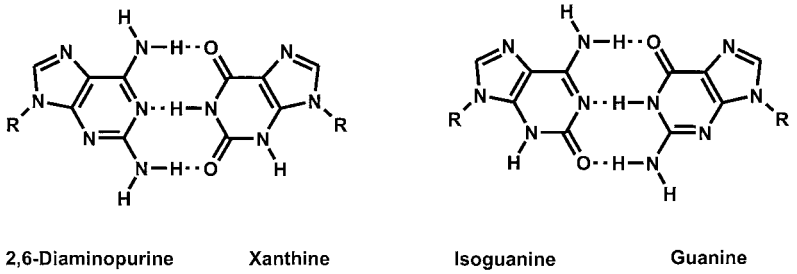
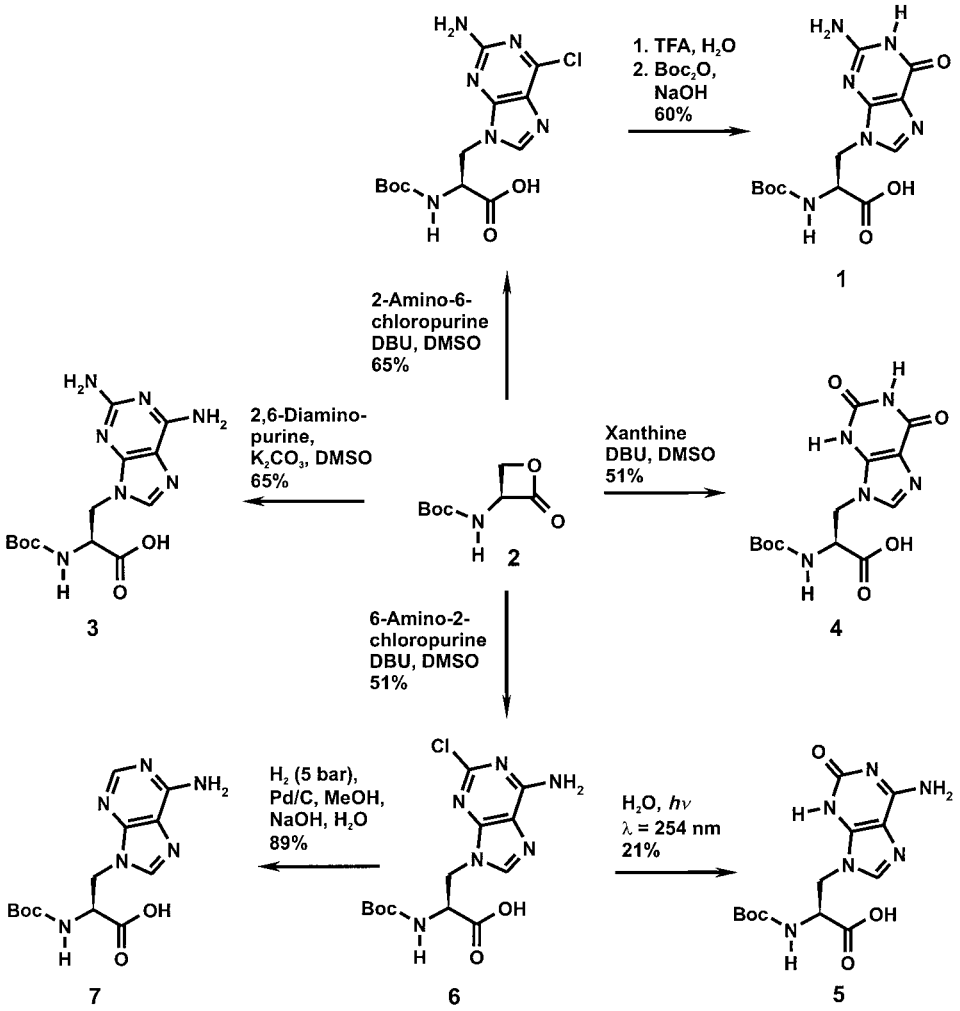


Fig. 3. *Isomorphous tridentate 2,6-diaminopurine-xanthine and isoguanine-guanine base pairs*

2. Synthesis of Purinyl-Nucleo-Amino Acids and Their Oligomerization. – The first stereoselective synthesis of alanyl nucleo-amino acids was described for the adeninyl and thyminyl derivatives by nucleophilic ring opening of Boc-serinelactone [7][8]. In a similar manner, the guaninyl nucleo-amino acid **1** was prepared starting from 2-amino-6-chloropurine and Boc-serinelactone **2**, followed by nucleobase transformation (*Scheme*) [4e]. Under similar conditions, also the 2,6-diaminopurine derivative **3** and the xanthine nucleo-amino acid **4** were prepared. Next to the desired *N*(9)-nucleo-amino acids, *N*(7)-regioisomers were formed as minor side-products that had been separated by column chromatography. The synthesis of the isoguaninylalanine **5** was performed photochemically (256 nm) [9] in aqueous solution from 6-amino-2-chloropurinyl derivative **6**. Nucleo-amino acid **6** was accessible from ring opening of Boc-serinelactone **2** with 6-amino-2-chloropurine which was prepared following a procedure of *Brown* and *Weliky* [10]. The *N*(9)-connectivity was established by hydrogenolytic dehalogenation [11] of amino acid **6**, providing the known adeninyl derivative **7** [8].

Scheme. *Synthesis of Purinyl Nucleo-Amino Acids by Ring Opening of Boc-Serinelactone*



The optical purity of all nucleo-amino acids was determined by Boc-deprotection, coupling with Boc-L-Ala-OSu, and HPLC analysis of the diastereoisomers. At least 94% of the desired enantiomer was found.

Oligomerization of the nucleo-amino acids was provided by solid-phase peptide synthesis on a MBHA-polystyrene (MBHA = 4-methylbenzhydrylamine) resin loaded with *N*-α-Boc-ω-[(benzyloxy)carbonyl]-protected L- or D-lysine. The lysine amide at the C-terminal end of the PNA oligomers was incorporated to increase the solubility of the oligomer. Coupling was accomplished with 5 equiv. of the respective nucleo-amino acid and activation with *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) over 60–120 min (dependent on the nucleo-amino acid) in a coupling yield $>95\%$. After deprotection and cleavage from the solid support with

2,2,2-trifluoroacetic acid (TFA)/trifluoromethanesulfonic acid/*m*-cresol 8:1:1, the oligomers were precipitated with Et_2O and purified by HPLC on a *RP-C18* column. Oligomers were characterized by ESI-MS.

3. Base Pairing between 2,6-Diaminopurine and Xanthine. – The interaction of 2,6-diaminopurine and xanthine was investigated with hexamers H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH_2 (**8**) and H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH_2 (*ent*-**8**) (AlaD = β-(2,6-diaminopurin-9-yl)alanine, AlaX = β-(xanthin-9-yl)alanine; D-configured nucleo-amino acids are underlined). The sequential order of nucleobases determines the D-X pairing with antiparallel strand orientation. In analogy to G-C base pairing in alanyl-PNA [4b], the *Watson-Crick* pairing mode requires pairing with homochiral nucleo-amino acids, whereas the reverse *Watson-Crick* mode is possible with heterochiral base pairs (*Fig. 4*). Therefore, double-strand formation with six D-X base pairs in the reverse *Watson-Crick* mode can be observed with self-pairing oligomers, whereas the *Watson-Crick* double strand requires pairing of enantiomeric strands.

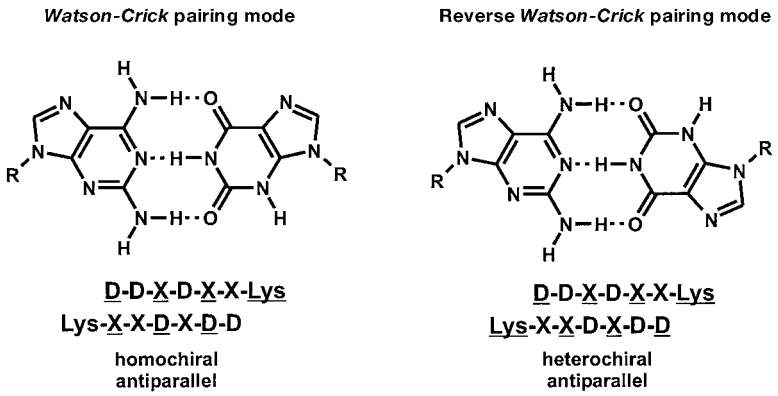


Fig. 4. *Two possible tridentate pairing modes for 2,6-diaminopurine-xanthine.* Double-strand formation in the reverse *Watson-Crick* mode requires self-pairing alanyl-PNA oligomers. The *Watson-Crick* pairing is only possible with enantiomeric alanyl-PNAs.

The double-strand stability of alanyl-PNA was analyzed with temperature-dependent UV spectroscopy in 0.01M Na_2HPO_4/H_3PO_4, pH 7, 0.1M NaCl. As known from DNA chemistry, the cooperative destacking of nucleobases during separation of double strands can be observed as sigmoidal increase of the absorption [12]. The temperature where only half of the oligomers exist as double strand (T_m) is given as the value for the double-strand stability. For self-pairing of H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH_2 (**8**), a stability of $T_m = 54°$ (5 μM, 15% hyperchromicity) was found. The enantiomer *ent*-**8** provided a stability of $T_m = 57°$ (7 μM, 14% hyperchromicity). The equimolar mixture of enantiomers **8** and *ent*-**8** only gave a T_m of 52° (3 μM each, 10% hyperchromicity). The pairing of enantiomers is not more stable than the self-pairing complexes, and the drop in stability nicely shows the concentration dependence of double-strand formation. Therefore, D-X base pairing in alanyl-PNA is favored in the reverse *Watson-Crick* mode. The self-pairing double strand of hexamer **8** is C_2-

symmetrical. This also is indicated by CD spectra at temperatures between 5 and 55°, which do not show any *Cotton* effect.

Compared to the *Watson-Crick* pairing G-C hexamers ($T_m = 58°$, 6 μM each), D-X pairing is slightly less stable probably due to lower stacking contributions. The dipole orientation of neighboring xanthines as well as 2,6-diaminopurines is very similar within the linear double strand of alanyl-PNA. Nevertheless, D-X base pairing in the reverse *Watson-Crick* mode still fits to the concept of base-pair selectivity based on high double-strand stability by tridentate base pairing. The D-X donor/acceptor patterns of the *Watson-Crick* and the reverse *Watson-Crick* mode are identical. Therefore, it is likely that the D-X *Watson-Crick* pairing contribution is not much lower in stability. In oligomers containing all G, I, D, and X nucleobases, a common base-pair geometry should be possible with G-I and D-X *Watson-Crick* base pairs.

4. Guanine-Isoguanine Base Pairing. – The synthesis of oligomer H-(AlaI-AlaI-AlaG-AlaI-AlaG-AlaI-AlaG)-Lys-NH_2 (AlaG = β-(guanin-9-yl)alanine, AlaI = β-(isoguanin-9-yl)alanine) failed probably due to aggregation phenomena on the solid support. Therefore, first evidence for G-I pairing comes from tetramer H-(AlaI-AlaG)$_2$-Lys-NH_2 (**9**). The G-I self-pairing double strand with reverse *Watson-Crick* pairing mode (*Fig. 5*) was found to be quite stable ($T_m = 35°$, 8 μM, 14% hyperchromicity). The G-C tetramer with identical H-bonding pattern had a lower stability with $T_m = 12°$ (12 μM, 8% hyperchromicity). This indicates a higher stacking contribution within the all-purine oligomer.

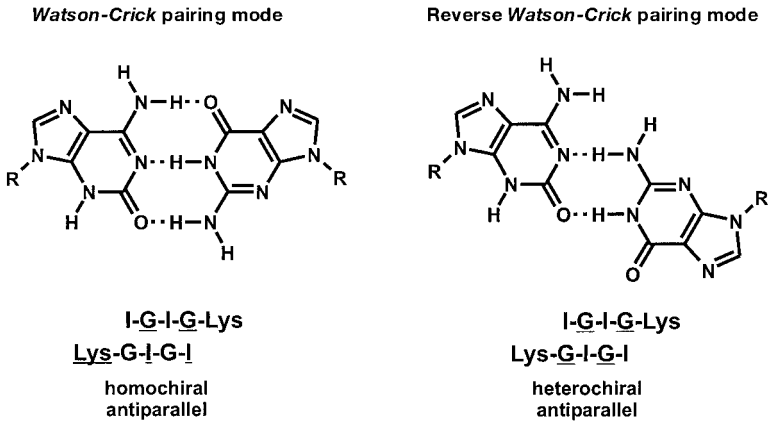


Fig. 5. *Pairing modes for isoguanine-guanine.* Only *Watson-Crick* pairing (pairing of enantiomeric oligomers) results in a tridentate stabilization.

5. Xanthine-Xanthine Base Pairing. – As part of our studies of purine-purine recognition, we also investigated the xanthine-xanthine base pairing. A twodentate base pairing was expected either in the *Watson-Crick* mode (homochiral pairing of enantiomeric oligomers) or in the reverse *Watson-Crick* mode (heterochiral, self-pairing), respectively (*Fig. 6*). Therefore, the high stability of the self-pairing double strand H-(AlaX-AlaX-AlaX-AlaX-AlaX-AlaX)-Lys-NH_2 (**10**; $T_m = 48°$, 3 μM, 20% hyperchromicity) was surprising. The equimolar mixture of xanthine hexamers **10** and

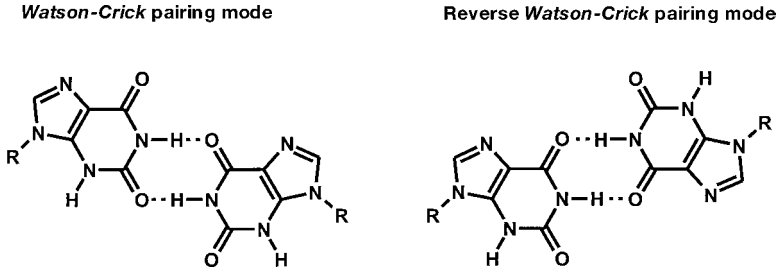


Fig. 6. *Xanthine-xanthine base pairs with two* H-bonds

ent-**10** was less stable than the self-pairing complexes ($T_m = 42°$, 1.5 μM each, 10% hyperchromicity), indicating that the reverse *Watson-Crick* geometry is favored.

Compared with other stabilities known for hexamer double strands in alanyl-PNA ($T_m = 21-32°$ for twodentate base pairing), the xanthine hexamer looks more like a double strand based on tridentate nucleobase pairs. Tridentate xanthine-xanthine base pairing indeed is possible with xanthine nucleobases in different tautomeric forms (*Fig. 7*). The most stable xanthine tautomer is the diketo-xanthine but with the 6-enol-2-keto-3*H*-, 6-enol-2-keto-1*H*-, 2-enol-6-keto-1*H*-, 2-enol-6-keto-3*H*-, and the dienol-xanthine tautomers exist that indeed can form tridentate X-X base pairs (*Fig. 8*).

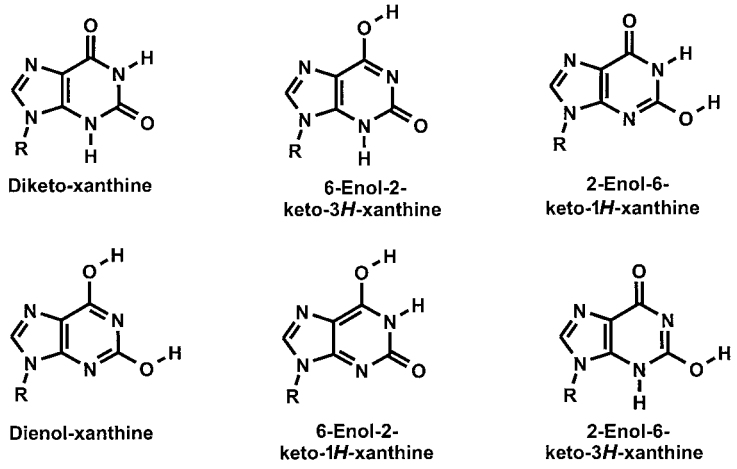


Fig. 7. *Xanthine tautomers*

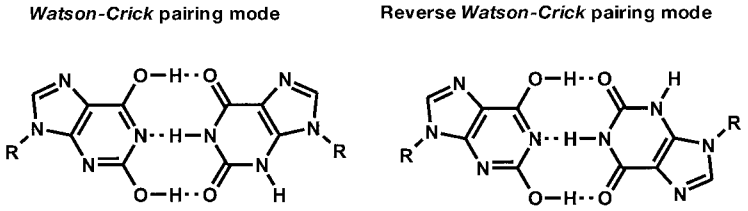


Fig. 8. *Xanthine-xanthine base pairs with mixed tautomers*

Double-strand formation with the reverse *Watson-Crick* pairing mode using mixed diketo- and dienol-xanthine base pairs would perfectly fit our experimental results: this X-X pairing complex with tridentate H-bonding is isosteric with the D-X in the self-pairing oligomer H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH_2 (**8**) and, therefore, also has C_2 symmetry. This interpretation is in agreement with the CD spectra of the xanthine hexamer at various temperatures. No *Cotton* effect was found as it was the case for the D-X oligomer **8**. Further, 2,6-diaminopurine is the isosteric nucleobase of xanthine in the tautomeric dienol form. The base pairs D-X and X-X were found in the same pairing mode and with an identical nucleobase orientation in their respective hexamer double strands. Both double strands are isosteric. The lower stability of the X-X hexamer is probably due to the energy that is needed to provide the dienol tautomer of xanthine.

6. Calculations of Xanthine Base Pairing Using Different Tautomers. – The surprising stability of the xanthine hexamer could be explained by a tridentate base-pairing mode involving tautomeric forms. The involvement of tautomers in the pairing of nucleobases was often discussed but could never be proved [3][13–15]. The question arises if and why the relative stability of the various pairing modes of xanthine differ considerably from other nucleobases. To answer this question, we computed the tautomeric forms of xanthine (*Fig. 7*) and the stabilities of different X-X base pairs. Similar computations were performed for the smaller uracil (for the tautomeric forms of uracil studied before, see [16]) to examine the particularity of xanthine regarding tautomerization.

7.1. *Uracil Tautomers as a Reference for Xanthine Studies. Table 1* contains the relative energies of selected tautomeric forms of uracil with respect to the most stable diketo form as a function of various theoretical approaches. It shows that the energy gap of 2-enol-4-keto-3*H*-uracil **B** to the most stable conformer **A** (*Fig. 9*) is *ca.* 11–13 kcal/mol. However, the conformer **B′** with the ability to form a tridentate base pair is already by *ca.* 20 kcal/mol higher in energy than the most stable diketo form **A**. *Table 1* also shows that, in comparison to the most accurate CCSD(T) approach, the density-functional theory (B3LYP functional) overestimates the energy gap between the various tautomeric forms. One important example is the diol form **C**. The energy gaps computed with the more elaborate MP2 approach are very similar to those obtained with the CCSD(T) approach.

7.2. *Xanthine Tautomers.* In *Table 2*, the energy gaps between the most stable diketo form of xanthine and the other tautomeric forms are summarized (*Fig. 7*). Tautomers with the possibility for tridentate base pairing are 2-enol-6-keto-1*H*-xanthine, 6-enol-2-keto-3*H*-xanthine, 2-enol-6-keto-3*H*-xanthine, and dienol-xanthine. While 2-enol-6-keto-1*H*-xanthine can build up tridentate base pairs with 6-enol-2-keto-3*H*-xanthine and 2-enol-6-keto-3*H*-xanthine, the dienol-xanthine is able to form a tridentate base pair with the diketo form. *Table 2* shows that 2-enol-6-keto-1*H*-xanthine **B**, which represents the most stable enol-keto-tautomeric form, is only by *ca.* 2–4 kcal/mol less stable than the diketo form **A**. However, similar to uracil, the conformer **B′** of 2-enol-6-keto-1*H*-xanthine (*Fig. 9*) that has the ability for tridentate base pairs is much higher in energy (10–12 kcal/mol). Since 2-enol-6-keto-1*H*-xanthine can form tridentate base pairs only with the even less stable 6-enol-2-keto-3*H*-xanthine (12–14 kcal/mol) or 2-

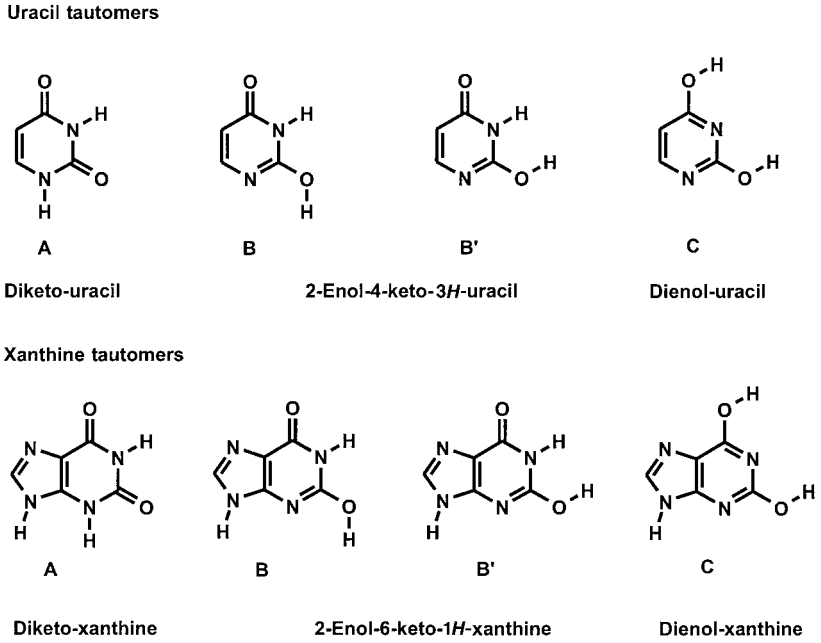


Fig. 9. *Calculated uracil and xanthine tautomers.* In case of the enol-keto tautomers, the conformers **B** and **B′** also differ in energy.

Table 1. *Computed Relative Energies of Uracil Tautomers.* All energies are given with respect to the diketo form. (All values are given in kcal/mol.)

Tautomer	Basis set	B3LYP	MP2 [a])	CCSD(T) [a])
2-Enol-4-keto-3*H*	6-31G(d)	12.4 [d])/20.8 [c])	13.1 [b])	12.4 [b])
	6-31++G(d,p)	11.1 [b])/19.4 [c])	10.5 [b])	10.2 [b])
Dienol	6-31G(d)	15.3 [d])/16.5 [e])	14.4 [d])	14.6 [d])
	6-31++G(d,p)	12.8 [d])/14.0 [e])	10.9 [d])	10.8 [d])

[a]) The geometrical parameters were taken from the corresponding B3LYP calculation.
[b]) and [c]) refer to conformers **B** and **B′**, respectively, in *Fig. 9*.
[d]) The more stable conformer of dienol-uracil can be obtained by rotating the O–H bond of the 2-enol group of conformer **C** by 180°.
[e]) Refers to conformer **C** of dienol-uracil.

enol-6-keto-3*H*-xanthine (21–22 kcal/mol), the resulting base pairs would be too high in energy to represent stable pairing modes.

Applying the 6-31G++G(d,p) AO basis set in combination with the B3LYP functional, we predict that the energy gap between the dienol and the diketo form is *ca.* 7 kcal/mol. If the more sophisticated MP2 approach is used, the gap decreases to 5.8 kcal/mol. The destabilization of the dienol form with respect to the diketo form is of the same magnitude as an approximate bond energy of a H-bond (5–6 kcal/mol)

Table 2. *Computed Relative Energies of Xanthine Tautomers.* All energies are given with respect to the diketo form (all values are given in kcal/mol).

Tautomer	Basis set	B3LYP	MP2
2-Enol-6-keto-1*H*	6-31G(d)	4.9 [a])/12.4 [b])	4.2 [a])
	6-31++G(d,p)	3.4 [a])/10.7 [b])	–
Dienol	6-31G(d)	9.7 [c])/9.7 [d])	9.6 [c])/–
	6-31++G(d,p)	6.9 [c])/7.1 [d])	5.6 [c])/5.8 [d])
	TZVPP	–/–	4.4 [c])/4.6 [d])
6-Enol-2-keto-3*H*	6-31G(d)	14.0	–
	6-31++G(d,p)	12.4	–
2-Enol-6-keto-3*H*	6-31G(d)	23.0	–
	6-31++G(d,p)	20.9	–

[a]) and [b]) refer to conformers **B** and **B′**, respectively, in *Fig. 9*.
[c]) The more stable conformer of dienol-xanthine can be obtained by rotating the O–H bond of the 2-enol group of conformer **C** by 180°.
[d]) Refers to conformer **C** of dienol-xanthine.

[14]. Therefore, the tridentate base pair of the diketo and the dienol form (*Fig. 8*) is very similar in energy compared to the twodentate base pair of the diketo-xanthine tautomer (*Fig. 6*).

7.3. *Xanthine-Xanthine Base Pairs.* In *Table 3*, the dimerization energies for these two forms are summarized. Further, relative energies of the various base pairs are listed in comparison to the reverse *Watson-Crick* self-pairing of diketo-xanthine. If the 6-31G(d) AO basis set in combination with the B3LYP functional is employed, the dimerization energy of the twodentate base pair is about two thirds of the tridentate base pair, reflecting similar strength of all the different H-bonds involved. Despite the higher dimerization energy, the tridentate base pair is found to be *ca.* 3 kcal/mol higher in energy than the twodentate base pair. Employing the larger 6-31++G(d,p) AO basis set, the energy difference decreases to less than 1 kcal/mol. If the MP2 approach is used in combination with the same AO basis, the energy difference is only 0.5 kcal/mol. The increasing stability of the tridentate base-pairing mode results, because the energy gap between the diketo and the dienol form of xanthine decreases if the basis set is

Table 3. *Comparison between the Twodentate Base-Pairing Mode of the Diketo-Diketo Xanthine Base Pair with the Tridentate Base-Pairing Mode of the Diketo-Dienol Xanthine Base Pair* (all values are given in kcal/mol)

Pair	Pairing mode	E_{dim}		ΔE (B3LYP) [a])		MP2 [a])
		6-31G(d)	6-31++G(d,p)	6-31G(d)	6-31++G(d,p)	6-31++G(d,p)
Diketo-Diketo	reverse *WC*	–10.0	–9.7	0.0	0.0	0.0
Diketo-Diketo	*WC*	–9.3	–9.1	0.7	0.6	0.3
Diketo-Dienol	reverse *WC*	–16.2	–16.0	2.8	0.8	0.5
Diketo-Dienol	*WC*	–16.4	–16.1	2.6	0.7	0.5

[a]) Relative energy with respect to the diketo-diketo xanthine base pair in the *Watson-Crick* pairing mode. Positive numbers indicate that the diketo-diketo reverse *Watson-Crick* xanthine base pair is lower in energy.

enlarged or if the MP2 method is employed (*Table 2*). To estimate the AO basis-set limit, one has to take into account that the energy difference between the dienol-xanthine tautomer and the diketo-xanthine decreases from 5.8 kcal/mol to 4.6 kcal/mol if the even more flexible TZVPP basis set is used (*Table 2*). Therefore, we expect that MP2 computations near the AO basis-set limit will predict that the tridentate base pair of diketo-xanthine with dienol-xanthine is more stable than the twodentate base pair of the diketo-xanthine self-pairing. Additional computations employing larger AO basis sets are under way.

7. Conclusions. – Double strands with linear topology have no size and geometry restriction for base pairing. Therefore, recognition of two large purines can be observed with different pairing modes. Within the linear double strands of alanyl-PNA, twodentate purine-purine base pairing was known, *i.e.*, between A-A ($T_m = 21°$ for a hexamer) and A-G ($T_m = 32°$ for a hexamer). Unequivocal base-pair complementarity in linear systems needs tridentate purine-purine pairing. The 2,6-diaminopurine-xanthine base pair has the required stability ($T_m = 52°$ for a hexamer). The self-pairing of guanine and isoguanine already had an enormous stability ($T_m = 35°$ for a tetramer). Further, xanthine-xanthine pairing turned out to be too stable to be based on twodentate H-bonding. Base pairs using two different tautomers to form tridentate base pairs fit best to the experimental results. This was confirmed by computational studies that indicate that, in the basis-set limit, the tridentate base pairing of the diketo-xanthine with the dienol tautomer is more stable than the twodentate self-pairing of diketo-xanthine. The computations clearly prove the marked difference between the xanthine system and uracil. Furthermore, our calculations show that this difference is due to the energy of the dienol form of xanthine, which is more stable than the equivalent form of uracil.

Support by the *Deutsche Forschungsgemeinschaft* (Di 542/2-2) and the *Fonds der Chemischen Industrie* is gratefully acknowledged.

Experimental Part

General. All reagents were of anal. grade and used as supplied. Solvents were of the highest grade available. HPLC: *Pharmacia Äkta* basic with *YMC-Pack ODS, RP-C18*, 250 × 20 mm, 5 μm, 120 Å for preparation and 250 × 4.6 mm, 5 μm, 120 Å for anal. samples. The oligomer concentration was calculated by taking the extinction coefficient at 90° as being the sum of the extinction coefficients of the nucleo-amino acids. The optical purity was determined by HPLC of the amino-acid dimers obtained with Boc-(*S*)-Ala-OSu. M.p.: *Büchi SMP-20* apparatus. Optical rotation: *Perkin-Elmer Polarimeter 241 MC*. UV Melting curves: *Perkin-Elmer* UV/VIS *Lambda 10* with *Peltier* temp. programmer *PTP 1*. CD Spectra: *Jasco J-715* with *Peltier*-type temp. control system *PTC-348W*. NMR Spectra: *Bruker AC 250*, *Bruker DMX 400*, or *Bruker DMX 500* spectrometer. ESI-MS: *TSQ 700 Finnigan* spectrometer.

*(*S*)-*N*-[(*tert*-Butoxy)carbonyl]-β-(2,6-diaminopurin-9-yl)alanine* (**3**). A suspension of 2,6-diaminopurine (3.47 g, 22.7 mmol), Boc-L-serinelactone **2** (3.30 g, 17.3 mmol), and K_2CO_3 (2.39 g, 17.3 mmol) in DMSO (55 ml) was stirred at r.t. for 2 h. After addition of AcOH (0.99 ml, 17.3 mmol), the solvent was removed. Purification by chromatography (silica gel; AcOEt/MeOH 8:2; after removal of 2,6-diaminopurine AcOEt/MeOH/AcOH 80:20:5), repeated co-evaporation with toluene, and freeze drying from H_2O yielded **3** (3.82 g, 65%). White powder. M.p. 239° (dec.). ee > 94% (determined by HPLC of the dimer prepared with Boc-(*S*)-Ala-OSu on *RP C18*; 0–25% MeCN in 30 min, t_R (*like*) 24.9 min; t_R (*unlike*) 26.2 min). TLC (i-PrOH/H_2O/AcOH 5:2:1, sat. with NaCl): R_f 0.60. $[\alpha]_{589}^{23} = -25.8$ ($c = 0.3$, MeOH). UV (0.01M Na_2HPO_4/H_3PO_4, pH 7, 0.1M NaCl): 260, 286. IR: 3362*s*, 3210*s*, 2978*m*, 2973*m*, 2950*w*, 1598*s*, 1413*m*, 1165*m*, 1026*w*, 791*w*. ^{1}H-NMR ((D$_6$)DMSO): 1.27 (*s*, Boc); 3.77–3.95 (*m*, H–C(α), 1 H–C(β)); 4.41 (*d*, $J = 12$ Hz, 1 H–C(β)); 5.65 (*s*, NH_2);

5.95 (*m*, Boc N*H*); 6.51 (*s*, NH_2); 7.49 (*s*, H–C(8)). ^{13}C-NMR ((D$_6$)DMSO): 28.2 (Boc); 45.5 (C(β)); 55.9 (C(α)); 77.9 (Boc); 113.0 (C(5)); 138.1 (C(8)); 152.2 (C=O (Boc)); 155.2 (C(4)); 156.1 (C(2)); 160.2 (C(6)); 172.0 (COOH). ESI-MS: 338.0 (100, $[M+H]^+$).

*(*R*)-N-[(*tert*-Butoxy)carbonyl]-β-(2,6-diaminopurin-9-yl)alanine* (*ent*-**3**). The synthesis was accomplished according to the procedure for **3**.

*(*S*)-N-[(*tert*-Butoxy)carbonyl]-β-(xanthin-9-yl)alanine* (**4**). A suspension of xanthine (2.50 g, 16.3 mmol) and Boc-L-serinelactone **2** (2.03 g, 10.8 mmol) in DMSO (30 ml) was cooled to 10° and, after 15 min, 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU; 2.2 ml) was added. After 2.5 h, the mixture was quenched with AcOH (1.10 ml, 19.2 mmol) and concentrated. Purification by chromatography (silica gel; AcOEt/MeOH/H_2O 80:14:6; after removal of xanthine AcOEt/MeOH/H_2O/AcOH 80:14:6:6), repeated co-evaporation with toluene, and freeze drying from H_2O yielded **4** (1.75 g, 51%). White powder. M.p. 212° (dec.). ee >94% (determined by HPLC of the dimer prepared with Boc-(*S*)-Ala-OSu on *RP C18*; 0–25% MeCN in 30 min, t_R (*like*) 23.1 min; t_R (*unlike*) 24.3 min). TLC (i-PrOH/H_2O/AcOH 5:2:1, sat. with NaCl): R_f 0.44. $[\alpha]_{589}^{23} = -71.9$ ($c = 0.17$, MeOH). UV (0.01M Na_2HPO_4/H_3PO_4, pH 7, 0.1M NaCl): 275. IR: 3386*m*, 2980*m*, 2879*m*, 2806*m*, 1699*s*, 1687*s*, 1596*m*, 1391*m*. ^{1}H-NMR ((D$_6$)DMSO): 1.28 (*s*, Boc); 4.17–4.43 (*m*, H–C(α), 1 H–C(β)); 4.65 (*m*, 1 H–C(β)); 6.73 (*m*, BocNH); 7.75 (*s*, H–C(8)); 8.25 (*s*, NH of xanthine); 8.76 (*s*, NH of xanthine). ^{13}C-NMR ((D$_6$)DMSO): 28.2 (Boc); 47.7 (C(β)); 54.5 (C(α)); 78.2 (Boc); 106.8 (C(5)); 144.0 (C(8)); 149.9 (C(4)); 151.6 (C(2)); 154.5 (C=O (Boc)); 156.0 (C(6)); 171.8 (COOH). ESI-MS: 340.1 (100, $[M+H]^+$).

*(*R*)-N-[(*tert*-Butoxy)carbonyl]-β-(xanthin-9-yl)alanine* (*ent*-**4**). The synthesis was accomplished according to the procedure for **4**. The anal. data of *ent*-**4** are identical except for $[\alpha]_{589}^{23} = +50.0$ ($c = 0.26$, MeOH).

*(*S*)-β-(6-Amino-2-chloropurin-9-yl)-N-[(*tert*-butoxy)carbonyl]alanine* (**6**). DBU (743 µl, 4.97 mmol) was added to a suspension of 6-amino-2-chloropurine (702 mg, 4.14 mmol) [8] in DMSO (20 ml). Within 30 min, a soln. of Boc-L-serinelactone **2** (930 mg, 4.97 mmol) in DMSO (10 ml) was added. The mixture was stirred for 3 h before the reaction was terminated with AcOH (284 µl, 4.97 mmol). The solvent was removed, and MeOH (10 ml) was added to the mixture. After centrifugation from 6-amino-2-chloropurine, the mixture was concentrated and purified by chromatography (silica gel; AcOEt/MeOH 4:1; after removal of 6-amino-2-chloropurine AcOEt/MeOH/AcOH 8:2:1). After repeated co-evaporation with toluene, **6** (746 mg, 51%) was isolated as a colorless solid. TLC ($CHCl_3$/MeOH/H_2O/AcOH 70:30:3:0.3): R_f 0.27. IR: 3338*s*, 3199*m*, 1654*s*, 1597*s*, 1314*m*, 1250*m*, 1165*m*. ^{1}H-NMR ((D$_6$)DMSO): 1.06 (*s*, 2 H, Boc); 1.26 (*s*, 7 H, Boc); 4.07–4.22 (*m*, H–C(α), 1 H–C(β)); 4.53 (*m*, 1 H–C(β)); 6.48 (*m*, BocNH); 7.62 (*s*, NH_2); 7.95 (*s*, H–C(8)). ^{13}C-NMR ((D$_6$)DMSO): 28.2 (Boc), 45.2 (C(β)); 54.8 (C(α)); 78.1 (Boc); 117.7 (C(5)); 142.0 (C(8)); 151.2 (C(4)); 152.9 (C(2)); 155.0 (C=O (Boc)), 156.8 (C(6)); 171.3 (COOH). ESI-MS: 357 (22, $[M+H]^+$).

*(*R*)-β-(6-Amino-2-chloropurin-9-yl)-N-[(*tert*-butoxy)carbonyl]alanine* (*ent*-**6**). The synthesis was accomplished according to the procedure for **6**.

*(*S*)-N-[(*tert*-Butoxy)carbonyl]-β-(isoguanin-9-yl)alanine* (**5**). Compound **6** (256 mg, 718 µmol) was dissolved in H_2O (200 ml). The soln. was irradiated at 256 nm (15 W) for 13 h. The mixture was filtered, and the solvent was removed. Purification by HPLC on *RP C18* silica gel (5–18% MeCN in 20 min, t_R 13 min) yielded **5** (52 mg, 21%). Colorless solid. TLC (AcOEt/MeOH/H_2O/AcOH 6:2:2:1, sat. with NaCl): R_f 0.33. ee 96% (determined by HPLC of the dimer prepared with Boc-(*S*)-Ala-OSu on *RP C18*; 7–13% MeCN in 30 min, t_R (*like*) 13.9 min; t_R (*unlike*) 14.6 min). $[\alpha]_{589}^{27} = -50.7$ ($c = 0.26$, H_2O). UV (MeOH): 253, 295. IR: 3368*s*, 3109*m*, 2980*m*, 1686*s*, 1618*s*, 1400*m*, 1386*w*, 1166*m*. ^{1}H-NMR ((D$_6$)DMSO): 1.14 (*s*, 2 H, Boc); 1.30 (*s*, 7 H, Boc); 4.02 (*m*, 1 H–C(β)); 4.16 (*m*, H–C(α)); 4.30 (*m*, 1 H–C(β)); 6.37 (*s*, NH, purine); 6.88 (*s*, BocNH); 7.47 (*s*, NH_2); 7.55 (*s*, H–C(8)). ^{13}C-NMR ((D$_6$)DMSO): 27.9 (Boc); 28.2 (Boc); 43.3 (C(β)); 53.0 (C(α)); 78.6 (Boc); 109.0 (C(5)); 139.8 (C(8)); 152.1 (C(2)); 155.4 (C(6)); 155.9 (C=O (Boc)); 171.6 (COOH). ESI-MS: 338 (70, $[M+H]^+$).

*(*R*)-N-[(*tert*-Butoxy)carbonyl]-β-(isoguanin-9-yl)alanine* (*ent*-**5**). The synthesis was accomplished according to the procedure for **5**.

*(*S*)-β-(Adenin-9-yl)-N-[(*tert*-Butoxy)carbonyl]alanine* (**7**). Compound **6** (52.0 mg, 146 µmol) was dissolved in MeOH (20 ml) and NaOH (2 ml of 1M soln.) was added. In an autoclave, Pd/C (20 mg, 5% Pd) and H_2 (5 bar) were added. The mixture was stirred for 16 h, filtered over *Celite*, and the solvent was removed. Purification by chromatography on *RP C18* (silica gel; H_2O/MeOH 9:1) provided, after freeze drying, **7** (42 mg, 89%) as a colorless solid. The anal. data of **7** correspond with the *N*(9)-regioisomer obtained by nucleophilic opening of Boc-L-serinelactone **2** with adenine [8].

General Procedure for SPPS of Alanyl-PNA. Oligomerization was performed as a solid-phase peptide synthesis on a 4-methylbenzhydrylamine(MBHA)-polystyrene resin (50 mg, 15.45 mmol) loaded with (*S*)- or (*R*)-Boc-lysine(Z)-OH (0.309 mmol/g, Z = (benzyloxy)carbonyl) in a small column. For each coupling step, an

excess of 5 equiv. *N*-Boc-protected nucleo-amino acid (77.25 mmol) was used and activated by *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU; 26.40 mg, 69.53 mmol) and EtN (i-Pr)$_2$ (26.3 µl, 154.5 mmol) in DMF (600 µl). After swelling the resin for 20 min, the following procedure was repeated for every nucleo-amino acid unit: *1*) deprotection twice, for 3 min with CF_3COOH (TFA)/*m*-cresol 95:5 (2 ml); *2*) washing five times each with CH_2Cl_2/DMF 1:1 (2 ml) and pyridine (2 ml); *3*) coupling step, 60–120 min gently moving the reaction column; *4*) washing three times each with CH_2Cl_2/DMF 1:1 (2 ml), DMF/piperidine 95:5 (2 ml), and CH_2Cl_2/DMF 1:1 (2 ml). Finally, the alanyl-PNA was washed twice with TFA (2 ml) and cleaved from the solid support within 1 h using 1.6 ml TFA/trifluoromethanesulfonic acid/*m*-cresol 8:1:1. The dark brown soln. was concentrated to 400 µl, and the alanyl-PNA precipitated with Et_2O (5 ml) as a white solid. The PNA was separated using a centrifuge, followed by HPLC (*RP C18*). The yield of each coupling step was estimated from HPLC to be higher than 95%.

H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH$_2$ (**8**): HPLC: 21.2 min, *RP C18*, gradient: 0–20% MeCN in 30 min. ESI-MS: 1466.5 (100, $[M+H]^+$).

H-(AlaD-AlaD-AlaX-AlaD-AlaX-AlaX)-Lys-NH$_2$ (*ent*-**8**): HPLC: 21.2 min, *RP C18*, gradient: 0–20% MeCN in 30 min. ESI-MS: 1466.5 (100, $[M+H]^+$).

H-(AlaI-AlaG-AlaI-AlaG)-Lys-NH$_2$ (**9**): HPLC: 29.7 min, *RP C18*, gradient: 2.5–7.5% MeCN in 30 min. ESI-MS: 513.9 (72, $[M+2H]^{++}$).

H-(AlaX-AlaX-AlaX-AlaX-AlaX-AlaX)-Lys-NH$_2$ (**10**): HPLC: 18.0 min, *RP C18*, gradient: 0–30% MeCN in 30 min. ESI-MS: 1472.3 (100, $[M+H]^+$).

H-(AlaX-AlaX-AlaX-AlaX-AlaX-AlaX)-Lys-NH$_2$ (*ent*-**10**): HPLC: 18.1 min, *RP C18*, gradient: 0–30% MeCN in 30 min. ESI-MS: 1472.3 (100, $[M+H]^+$).

General Procedure for UV Melting Curves and CD Spectroscopy. The oligomers (6–12 µM) were dissolved in a Na_2HPO_4/H_3PO_4 buffer (pH 7.0, 0.01M) containing NaCl (0.1M) and placed in a UV cell (10 mm). For UV melting curves, the following temperature program was used to control the heating block: 80° → –2° (60 min) → –2° (240 min) → 90° (180 min) → –2° (180 min) → –2° (240 min) → 90° (180 min) → –2° (180 min). The temp. for UV spectra was measured directly in the soln., and for CD spectra the temp. of the heating block was determined.

Computational Methods. In the present work, density-functional theory (DFT) and *Møller-Plesset* perturbation theory [17] up to the second order (MP2) are used for the description of the various base-pairing modes. The DFT computations were performed with the B3LYP functional [18][19]. For the refinement of the geometries, an anal. gradient optimization was employed. Most of the geometry optimizations were performed with the DFT approach, because bond lengths obtained from MP2 approach differ less than 0.01 Å from the DFT values. The minima were characterized by frequency calculations. DFT Computations were performed with the *Gaussian 98* program package [20]. For the MP2 calculations, the *Gaussian 98* [20] and the *Turbomole* program package [21] were used. For MP2 computations, which were performed in combination with the TZVPP basis set, the RIMP2 ansatz [22][23] was employed.

The accuracy of DFT and MP2 was checked by CCSD(T) computations, because computed energy differences between tautomers are strongly influenced by the theoretical approach [24][25][26]. To study the influence of the AO basis set on the computed energies, various basis sets such as 6-31G(d) or 6-311G(d,p) [27][28] were used. The influence of diffuse functions was also tested (6-31++G(d,p) [29]). In addition, we employed the TZVPP basis set which represents a (11s6p2d1f) AO basis in a [5s3p2d1f] contraction [30].

To correct the basis-set superposition error (BSSE) [17], the *Boys-Berndardi* counterpoise correction (CP) [31] was employed.

REFERENCES

[1] W. Saenger, 'Principles of Nucleic Acid Structure', Springer, New York, 1994.
[2] S. Neidle, 'Nucleic Acid Structure', Oxford University Press, Oxford, 1999.
[3] 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.
[4] a) U. Diederichsen, *Angew. Chem., Int. Ed.* **1996**, *35*, 445; b) U. Diederichsen, *Angew. Chem., Int. Ed.* **1997**, *36*, 1886; c) U. Diederichsen, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1743; d) U. Diederichsen, D. Weicherding, *Synlett* **1999**, *S1*, 917; e) U. Diederichsen, D. Weicherding, *ChemBioChem*, submitted.
[5] U. Diederichsen in 'Bioorganic Chemistry – Highlights and New Aspects', Eds. U. Diederichsen, T. Lindhorst, L. A. Wessjohann, B. Westermann, Wiley-VCH, Weinheim, 1999, pp. 255–261.

[6] K. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1998**, *81*, 375.
[7] a) L. D. Arnold, T. H. Kalantar, J. C. Vederas, *J. Am. Chem. Soc*. **1985**, *107*, 7105; b) L. D. Arnold, R. G. May, J. C. Vederas, *J. Am. Chem. Soc.* **1988**, *110*, 2237.
[8] a) P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang, A. Eschenmoser, *Croatica Chim. Acta* **1996**, *69*, 535; b) P. Lohse, *'Synthese und Eigenschaften von oligomeren Nukleodipeptamidinium-Salzen'*, Dissertation, ETH 9937, Zürich, 1992.
[9] a) V. Nair, D. A. Young, *J. Org. Chem.* **1985**, *50*, 406; b) Z. Kazimierczuk, R. Mertens, W. Kawczynski, F. Seela, *Helv. Chim. Acta* **1991**, *74*, 1742; c) F. Seela, R. Mertens, Z. Kazimierczuk, *Helv. Chim. Acta* **1992**, *75*, 2298.
[10] G. B. Brown, V. S. Weliky, *J. Org. Chem*. **1958**, *23*, 125.
[11] R. Volpini, E. Camaioni, S. Vittori, L. Barboni, C. Lambertucci, G. Cristalli, *Helv. Chim. Acta* **1998**, *81*, 145.
[12] D. Pörschke, *Biopolymers* **1971**, *10*, 1989.
[13] P. Hobza, J. Sponer, *Chem. Rev.* **1999**, *99*, 3247 and refs. cit. therein.
[14] J. R. Sambrano, A. R. de Souza, J. J. Queralt, J. Andrés, *Chem. Phys. Lett.* **2000**, 317, 437.
[15] J. Florián, J. Leszczynski, *J. Am Chem. Soc*. **1996**, *118*, 3010.
[16] a) J. Leszczynski, *J. Phys. Chem.* **1992**, *96*, 1649; b) J. W. Boughton, P. Palay, *Int. J. Quant. Chem*. **1993**, *47*, 49; c) D. A. Estrin, L. Paglieri, G. Corougiu, *J. Phys. Chem*. **1994**, *98*, 5653; d) M. Orozco, B. Hernandez, F. J. Lueque, *J. Phys. Chem*. **1998**, *102*, 5228.
[17] F. Jensen, 'Introduction to Computational Chemistry', Wiley, Chichester, England, 1999.
[18] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[19] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.
[20] Gaussian 98 (Revision A.7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, A. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc. Pittsburgh, PA, 1998.
[21] TURBOMOLE, R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Haase, M. Häser, H. Horn, C. Huber, U. Huniar, M. Kattannek, C. Kölmel, M. Kollwitz, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, Quant. Chem. Group, University of Karlsruhe, Germany, since 1988.
[22] K. Eikhorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *242*, 652.
[23] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Theor. Chem. Acc*. **1998**, *294*, 143.
[24] I. R. Gould, I. H. Hillier, N. A. Bourton, R. J. Hall, *J. Mol. Spectr.* **1995**, *331*, 147.
[25] R. Kobayashi, *J. Phys. Chem. A* **1998**, *102*, 10813.
[26] G. Fogarasi, *J. Mol. Struct*. **1997**, *413–414*, 271.
[27] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1997**, *56*, 2257.
[28] R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 650.
[29] M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265.
[30] R. Ahlrichs, A. Schäfer, C. Huber, *J. Chem. Phys.* **1994**, *100*, 5829.
[31] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 49.

Received August 21, 2000