ARTICLE

OBC
www.rsc.org/obc

Side chain homologation of alanyl peptide nucleic acids: pairing selectivity and stacking†

Ulf Diederichsen,*[a] Daniel Weicherding[a,b] and Nicola Diezemann[a]

[a] *Institut für Organische und Biomolekulare Chemie, Tammannstr. 2, D-37077, Göttingen, Germany. E-mail: udieder@gwdg.de; Fax: +49 551 392944; Tel: +49 551 393221*

[b] *Clondiag Chip Technologies GmbH, Löbstedter Str. 103-105, D-07749, Jena, Germany. E-mail: daniel@clondiag.com*

Received 28th July 2004, Accepted 13th January 2005
First published as an Advance Article on the web 14th February 2005

Alanyl peptide nucleic acids (alanyl-PNAs) are oligomers based on a regular peptide backbone with alternating configuration of the amino acids. All side chains are modified by covalently linked nucleobases. Alanyl-PNAs form very rigid, well defined, and linear double strands based on hydrogen bonding of complementary strands, stacking, and solvation. Side chain homology was examined by comparing a methylene linker (alanyl-PNA) with an ethylene linker (homoalanyl-PNA), a trimethylene linker (norvalyl-PNA), and PNA sequences with mixed linker length between nucleobase and backbone. Side chain homology in combination with a linear double strand topology turned out to be valuable in order to selectively manipulate pairing selectivity (pairing mode) and base pair stacking.

Introduction

Pairing selectivity, stability, and functionality of oligonucleotides is determined not only by the nucleobases as recognition units but also by the constitution and conformation of the backbone.[1,2] The sugar phosphodiester backbone of DNA induces the double helix topology which is crucial for the integrity of the genetic code.[3] It defines the distance of anomeric centers and, therefore, the space and geometry of the base pairs in DNA. Only the Watson–Crick pairing mode is allowed for purine–pyrimidine base pairs. In contrast, pairing combinations and pairing modes in general are not restricted by a linear double strand topology.[4] Furthermore, linear double strands are lacking dynamic effects like helicalization or unwinding. Therefore, linear duplexes are simplified model systems with the potential to study all kinds of recognition interactions with the base stack like base pairing of canonical and artificial nucleobases, various pairing modes, stacking interactions, and intercalation.

In this regard, alanyl peptide nucleic acids (alanyl-PNAs) were investigated extensively as model system for a DNA base stack with linear topology.[5] Alanyl-PNA is based on a regular peptide backbone composed of alanyl units with alternating configuration. The nucleobases are covalently linked to alanyl side chains in the β-position (Fig. 1).[6,7] These oligomers are able to form linear double strands based on hydrogen bond recognition and stacking of the nucleobases. Simple models as well as experimental results suggest that an alanyl-PNA duplex is formed with collinear strands and orthogonal base pairs.[8–10] The distance of two base pairs (3.6 Å) results from the distance of consecutive side chains in peptides with β-strand conformation. Since this distance is close to stacking in DNA (3.4 Å) alanyl-PNA double strands are linear, rigid, and well defined. As expected for linear double strands, base pair formation is not restricted regarding its size (purine–purine and purine–pyrimidine) and pairing mode (Watson–Crick and Hoogsteen). Alanyl-PNA turned out to be a valuable model for examining small structural changes because of low topological restrictions in combination with the rigidity of the double strand.

Homologation is a common way to modify biooligomers in respect to altering their binding affinity, enzymatic stability or even the backbone topology.[11–13] Aminoethylglycine-PNA is able to recognize DNA and RNA adapting their helix topology. Homologation leads to significant destabilization of the oligonucleotide-PNA double strand due to discordance with the DNA or RNA helix topology.[14,15] Herein we investigated the influence of side chain homologation of alanyl-PNA on pairing selectivity and stacking. Modifications of these systems with linear topology allow conclusions to be drawn on the interdependence between recognition and insertion of methylene groups. Alanyl, homoalanyl, and norvalyl nucleo amino acids were prepared and oligomerized in order to investigate the pairing properties of the respective peptide nucleic acids.

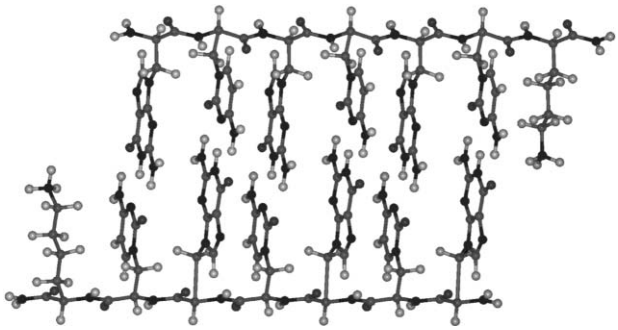


Fig. 1 Presentation of an alanyl-PNA double strand as a rigid and structurally well defined model system for a DNA nucleobase stack with linear topology.

Results and discussion

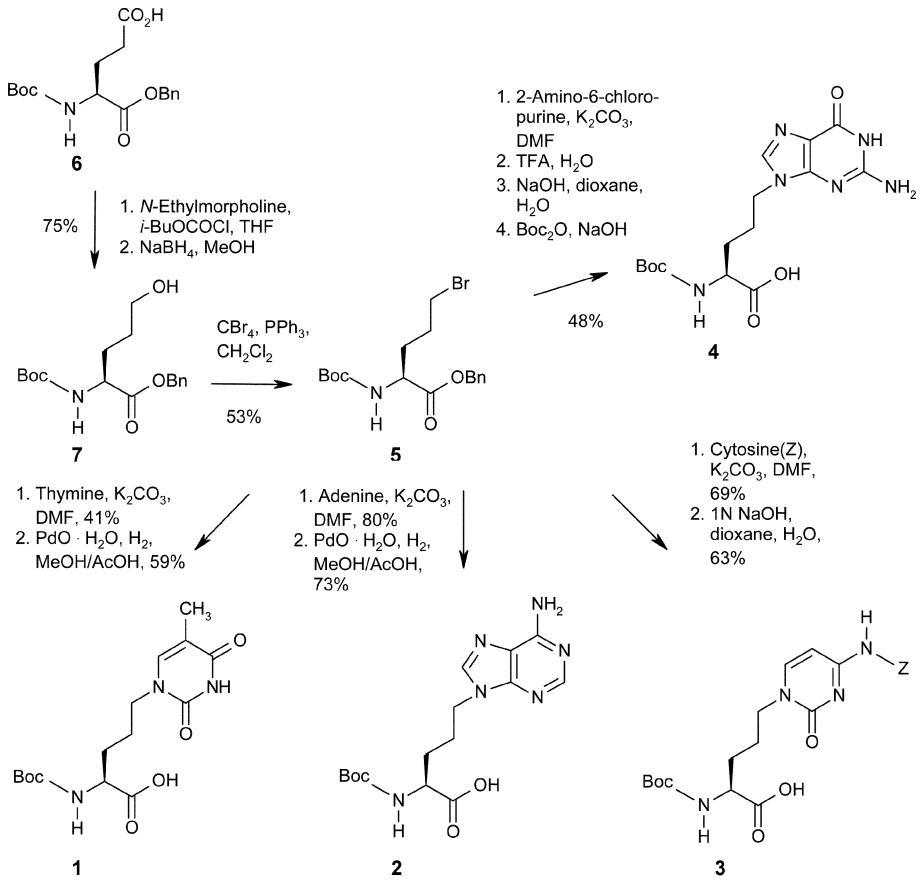
Nucleo amino acids as building blocks of the respective oligomers had to be prepared in high enantiomeric purity. A series of alanyl, homoalanyl, norvalyl and mixed PNA hexamers was obtained by manual solid phase peptide synthesis following Boc-strategy. Finally, the comparison of double strand stabilities of various PNA oligomers enabled us to draw conclusions on base pairing selectivities and base pair stacking dependent on side chain homologation.

Synthesis of nucleo amino acids and their oligomerization

Alanyl nucleo amino acids were prepared by nucleophilic ring opening of butoxycarbonyl (Boc) protected serine lactone with the respective nucleobase.[16–18] For the synthesis of homoalanyl

† Electronic supplementary information (ESI) available: synthesis details and spectroscopic characterisation. See http://www.rsc.org/suppdata/ob/b4/b411545g/

DOI: 10.1039/b411545g



Scheme 1 Synthesis of the norvalyl nucleo amino acids.

nucleo amino acids, substitution of Boc-protected γ-bromo homoalanyl benzyl ester with the nucleobases was used chiefly following procedures of Taddei *et al.*[19–22] Synthetic details for both series of nucleo amino acids are provided as electronic supplementary information (ESI).†

The norvalyl nucleo amino acids **1**–**4** also were obtained by substitution of a primary bromide with the respective nucleobase (Scheme 1). Bromide **5** was available by side chain reduction of the benzylester of Boc-protected glutamic acid **6** followed by Appel reaction[23] of the generated primary alcohol **7**. Nucleophilic substitution was successful with high yields using thymine, adenine, Z-protected cytosine or 2-amino-6-chloropurine without having a notable problem with the generation of undesired *N*3 or *N*7 regioisomers. The *N*9-alkylated guanine was generated under TFA conditions with the need of Boc-reprotection. Finally, ester saponification was provided either hydrogenolytically or with NaOH.

Oligomerization of the alanyl, homoalanyl, or norvalyl nucleo amino acids was provided by solid phase peptide synthesis on a MBHA-polystyrene (MBHA = 4-methylbenzhydrylamine) resin loaded with *N*-α-Boc-ω-benzyloxycarbonyl protected L- or D-lysine. The lysine amide at the C-terminal end of the PNAs was incorporated to increase the solubility of the oligomers. Coupling was done with 5 equivalents of the respective nucleo amino acid and activation with *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) and 1-hydroxy-7-azabenzotriazol (HOAt) over 30–90 min (dependent on the nucleo amino acid) exceeding a yield of 95% for each amino acid coupling step. After deprotection and cleavage from the solid support with TFA–trifluoromethanesulfonic acid–*m*-cresol 8 : 1 : 1, the oligomers were precipitated with diethyl ether and purified by HPLC on a RP-C18 column. All oligomers prepared within this study (Table 1) were characterized using ESI-MS; for a few examples ^{1}H-NMR spectroscopy was applied. The configurational integrity was indicated by mirror imaged CD spectra for oligomers that were synthesized in both enantiomeric forms.

Table 1 Synthesized alanyl, homoalanyl, and norvalyl PNA oligomers

Oligomer[a]	No.
H-(*AlaG*-AlaG-*AlaC*-AlaG-*AlaC*-AlaC)-*Lys*-NH_2	**8**
H-(AlaG-*AlaG*-AlaC-*AlaG*-AlaC-*AlaC*)-Lys-NH_2	ent-**8**
H-(*AlaG*-**HalG**-*AlaC*-**HalG**-*AlaC*-**HalC**)-*Lys*-NH_2	**9**
H-(AlaG-***HalG***-AlaC-***HalG***-AlaC-***HalC***)-Lys-NH_2	ent-**9**
H-(***HalG***-**HalG**-***HalC***-**HalG**-***HalC***-**HalC**)-*Lys*-NH_2	**10**
H-(**HalG**-***HalG***-**HalC**-***HalG***-**HalC**-***HalC***)-Lys-NH_2	ent-**10**
H-(*AlaG*-NvaG-*AlaC*-NvaG-*AlaC*-NvaC)-*Lys*-NH_2	**11**
H-(AlaG-*NvaG*-AlaC-*NvaG*-AlaC-*NvaC*)-Lys-NH_2	ent-**11**
H-(*NvaG*-NvaG-*NvaC*-NvaG-*NvaC*-NvaC)-*Lys*-NH_2	**12**
H-(NvaG-*NvaG*-NvaC-*NvaG*-NvaC-*NvaC*)-Lys-NH_2	ent-**12**
H-(*AlaA*-AlaA-*AlaT*-AlaA-*AlaT*-AlaT)-*Lys*-NH_2	**13**
H-(AlaA-*AlaA*-AlaT-*AlaA*-AlaT-*AlaT*)-Lys-NH_2	ent-**13**
H-(***HalA***-**HalA**-*AlaT*-**HalA**-*AlaT*-AlaT)-*Lys*-NH_2	**14**
H-(***HalA***-AlaT-***HalA***-AlaT-***HalA***-AlaT)-*Lys*-NH_2	**15**
H-(*NvaA*-NvaA-*NvaT*-NvaA-*NvaT*-NvaT)-*Lys*-NH_2	**16**
H-(NvaA-*NvaA*-NvaT-*NvaA*-NvaT-*NvaT*)-Lys-NH_2	ent-**16**

[a] AlaG = β-(9-guaninyl)alanine, AlaC = β-(1-cytosinyl)alanine, AlaA = β-(9-adeninyl)alanine, AlaT = β-(1-thyminyl)alanine, HalG = γ-(9-guaninyl)homoalanine, HalC = γ-(1-cytosinyl)homoalanine, HalA = γ-(9-adeninyl)homoalanine, NvaG = δ-(9-guaninyl)norvaline, NvaC = δ-(1-cytosinyl)norvaline, NvaA = δ-(9-adeninyl)norvaline, NvaT = δ-(1-thyminyl)norvaline; D-configured nucleo amino acids are in italics.

The stability of double strands was determined in aqueous solution (0.1 M NaCl, 0.01 M Na_2HPO_4/H_3PO_4, pH 7.0) by temperature dependent UV spectroscopy. Since the duplex separation into single strands is a cooperative process destacking can be detected by an increase of absorption. The temperature at the turning point of the sigmoidal curve indicates the stability of the complex.

Side chain homology

The peptide backbone of alanyl-PNA duplexes adopts a β-sheet conformation forced by double strand formation (Fig. 1).[6] Therefore, consecutive side chains with alternating configuration have a distance of 3.6 Å, which is close to the favored stacking distance of nucleobase pairs. Because the separation of amino acid side chains equals the preferred distance for π–π interactions, the double strand should not lower the base pair distance by helicalization or inclination. Therefore, alanyl-PNA double strands are linear by design and provide a perfect topology in order to study the influence of side chain homology. Within this idealized model, the side chain length should not affect the structural inherent linearity but has consequences for pairing selectivity and stacking orientation.

Effects on pairing selectivity

The pairing selectivity is defined by the pattern of proton donor and acceptor positions given by the kind of nucleobase and their orientation. The nucleobase orientation is dependent on the constitution and conformation of the backbone and the side chain linker. Therefore, the pairing selectivity is strongly influenced by the length of the linker between the backbone and the nucleobase (Fig. 2). In PNAs with linear topology the side chains of consecutive nucleo amino acids have a *gauche* like orientation with respect to the backbone. With all bonds of the side chains being staggered this 60° angle between successive side chains results in a mirror image orientation of neighboring alanyl-PNA nucleobases. The Hoogsteen sites of purines are alternately oriented up and down. For identical nucleobases the order of donor and acceptor positions at the Watson–Crick site is reversed from one nucleobase to the other. In an idealized homoalanyl-PNA strand, the analysis of the nucleobase orientation gives the same result as described for alanyl-PNA except for an inverted orientation of each individual nucleobase, whereas for norvalyl-PNA the situation should be exactly the same as described for alanyl-PNA.

In contrast, for PNAs built of alternating alanyl and homoalanyl nucleo amino acids a different pairing selectivity is expected since all Hoogsteen sites are oriented alike. Therefore, all nucleobases are oriented with a comparable order of donor and acceptor positions at the Watson–Crick site. The alanyl/homoalanyl alternated PNA should have the closest similarity to DNA double strands where all Hoogsteen sites are oriented towards the major groove. Furthermore, the alternation of longer homoalanyl and shorter alanyl nucleo amino acids contributes to the pairing selectivity since effective double strand formation requires pairing of an alanyl with a homoalanyl unit for geometrical reasons.

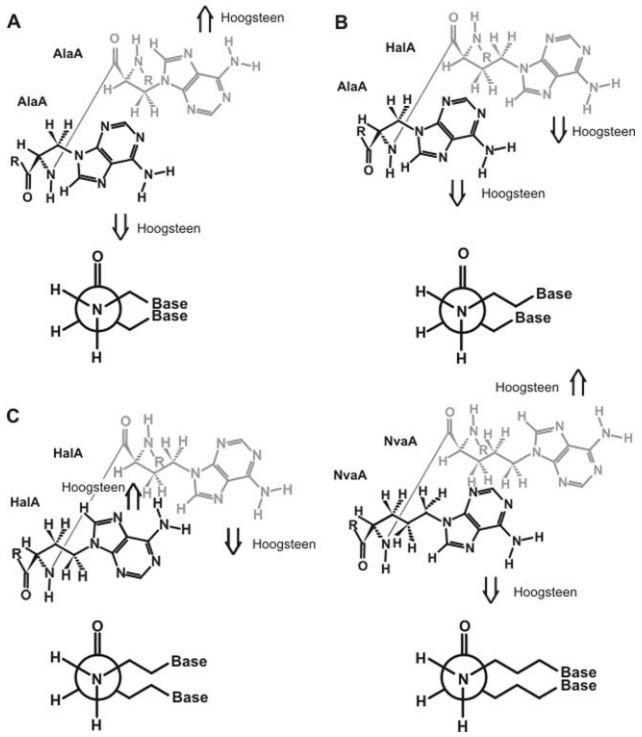


Fig. 2 Pairing selectivity influenced by side chain homology: **A.** in alanyl-PNA neighboring purines are oriented in opposite directions; **B.** alternating alanyl/homoalanyl-PNA oligomers have uniformly oriented purines and therefore are the best representation of DNA; **C.** the orientation of purines in homoalanyl and norvalyl-PNA is like in alanyl-PNA.

Guanine–cytosine pairing selectivity

Experimental evidence for the influence of side chain homology on pairing selectivity was investigated with G–C pairing oligomers. In alanyl-PNA the G–C pairing is possible at pH 7.0 either in the more stable Watson–Crick pairing mode forming three hydrogen bonds or with the reverse Watson–Crick mode based on only two H-bonding interactions (Fig. 3). The selfpairing complex of the alanyl-PNA H-(*AlaG*-AlaG-*AlaC*-AlaG-*AlaC*-AlaC)-*Lys*-NH_2 (**8**) is formed with six G–C base pairs in an antiparallel strand orientation. This antiparallel strand orientation is given by the sequential order of guanine and cytosine. Previously it was shown that in the selfpairing complex of oligomer **8** the G–C base pairs are formed in the reverse Watson–Crick mode with nucleo amino acids of opposite configuration (heterochiral pairing).[7] It was also shown that pairing of enantiomeric alanyl-PNA oligomers **8** and ent-**8** leads to base pairs formed with homochiral nucleo amino acids.[7] The G–C pairing of the enantiomers **8** and ent-**8** requires the Watson–Crick pairing mode.

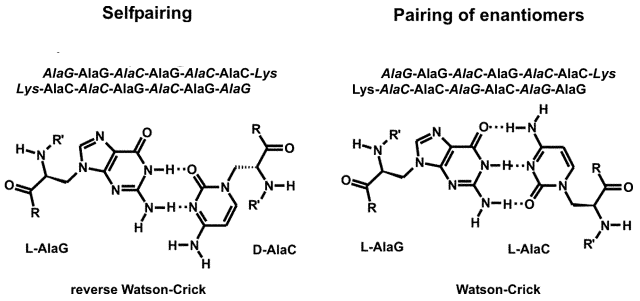


Fig. 3 In alanyl-PNA antiparallel selfpairing (heterochiral base pairs) is possible in the reverse Watson–Crick mode, whereas pairing of enantiomers (homochiral base pairs) requires the Watson–Crick mode.

Indeed, higher stability was found by temperature dependent UV spectroscopy for the double strand containing enantiomeric oligomers **8** and ent-**8** in the Watson–Crick mode (T_m = 58 °C, 13% hyperchromicity (*H*), 6 μM each). In contrast, the selfpairing complex of oligomer **8** based on reverse Watson–Crick recognition with two hydrogen bonds had only a stability of T_m = 40 °C (16% *H*, 12 μM).

For the alanyl/homoalanyl alternating PNA oligomers H-(*AlaG*-HalG-*AlaC*-HalG-*AlaC*-HalC)-*Lys*-NH_2 (**9**) and ent-**9** the situation was reversed which would be in accordance with the altered Watson–Crick/reverse Watson–Crick selectivity. The stability for the selfpairing complex of **9** of T_m = 52 °C (22% *H*, 12 μM) was not exceeded by the stability of an equimolar mixture of enantiomers **9** and ent-**9**. This indicates a higher stability of the Watson–Crick pairing double strands preventing the reverse Watson–Crick mode from forming. The selectivity for homoalanyl-PNAs H-(*HalG*-HalG-*HalC*-HalG-*HalC*-HalC)-*Lys*-NH_2 (**10**) and ent-**10** turned out to be comparable to alanyl-PNA since Watson–Crick pairing of enantiomers **10** and ent-**10** (T_m = 37 °C, 25% H, 6 μM each) was more

Table 2 Double strand formation dependent on side chain length

PNA	Oligomers	Pairing mode	T_m[a]
alanyl	**8**	reverse WC	40 °C
alanyl	**8** + ent-**8**	WC	58 °C
alanyl/homoalanyl	**9**	WC	52 °C
alanyl/homoalanyl	**9** + ent-**9**	reverse WC	not formed
homoalanyl	**10**	reverse WC	33 °C
homoalanyl	**10** + ent-**10**	WC	37 °C
alanyl/norvalyl	**11**	reverse WC	18 °C
alanyl/norvalyl	**11** + ent-**11**	WC	16 °C
norvalyl	**12**	reverse WC	13 °C
norvalyl	**12** + ent-**12**	WC	15 °C

[a] All double strand stabilities were determined for hexamers with the sequence GGCGCC at an overall concentration of 12 μM in 0.1 M NaCl, 0.01 M Na_2HPO_4/H_3PO_4, pH 7.0.

Table 3 Double strand formation of A–T pairing oligomers

PNA	Oligomers	Pairing mode	T_m[a]
alanyl	**13**	rev. Hoogsteen, rev. WC	25 °C
alanyl	**13** + ent-**13**	WC, Hoogsteen	not formed
alanyl/homoalanyl	**14**	WC	22 °C
alanyl/homoalanyl	**15**	WC	19 °C
norvalyl	**16**	rev. Hoogsteen, rev. WC	11 °C
norvalyl	**16** + ent-**16**	WC, Hoogsteen	13 °C

[a] All double strand stabilities were determined for A–T pairing hexamers at an overall concentration of 6 or 12 μM in 0.1 M NaCl, 0.01 M Na_2HPO_4/H_3PO_4, pH 7.0.

stable than selfpairing of **10** (T_m = 33 °C, 28% *H*, 12 μM). Furthermore, norvalyl/alanyl alternated sequences H-(*AlaG*-NvaG-*AlaC*-NvaG-*AlaC*-NvaC)-*Lys*-NH_2 (**11**, T_m = 18 °C, 135% *H*, 12 μM) and ent-**11** as well as the norvalyl oligomers H-(*NvaG*-NvaG-*NvaC*-NvaG-*NvaC*-NvaC)-*Lys*-NH_2 (**12**, T_m = 13 °C, 90% *H*, 12 μM)) and ent-**12** were investigated expecting the same base pair selectivity as in alanyl-PNA (Table 2). Surprisingly, the equimolar mixture of enantiomeric oligomers **11** and ent-**11** (T_m = 16 °C, 85% *H*, 6 μM each) and **12** and ent-**12** (T_m = 15 °C, 160% *H*, 6 μM each) turned out to be within the same order of stability as the selfpairing complexes. Therefore, we can conclude that with the longer norvalyl linker the base pair selectivity is no longer determined by the idealized PNA geometry. The conformational flexibility of the side chains seems sufficient to allow all kinds of base pair orientations.

Despite the differences based on bi- or tri-dentate pairing modes, for each step of homologation from alanyl-PNA (T_m = 58 °C) to homoalanyl-PNA (T_m = 37 °C) and norvalyl-PNA hexamers (T_m = 15 °C) a drop in stability was observed. This effect is likely to be attributed to the higher flexibility of the longer side chains but partly also to different stacking patterns that might be induced by homology as will be discussed later.

Adenine–thymine Watson–Crick selectivity

In principle the influence of side chain homology on the selectivity of A–T pairing in alanyl and homoalanyl-PNA is the same as described for G–C pairing oligomers.[24] The only difference arises from the possibility for A–T base pairs to perform Hoogsteen or reverse Hoogsteen pairing besides the Watson–Crick and reverse Watson–Crick mode all with comparable stability. In alanyl-PNA with antiparallel strand orientation, homochiral base pairing (enantiomeric oligomers) favors the Watson–Crick or Hoogsteen mode. The selfpairing with heterochiral base pairs is only possible in the reverse Watson–Crick or reverse Hoogsteen mode. For alanyl/homoalanyl alternated sequences these pairing preferences are reversed.

In accordance with these topological requirements, double strand formation of alanyl-PNA H-(*AlaA*-AlaA-*AlaT*-AlaA-*AlaT*-AlaT)-*Lys*-NH_2 (**13**, T_m = 25 °C, 14% *H*, 6 μM) was observed preferentially as a selfpairing complex compared to the equimolar complex of alanyl-PNAs **13** and ent-**13** (Table 3). In the alanyl-PNA series the reversed Hoogsteen mode seems favored for A–T.[25] In the mixed alanyl/homoalanyl-PNA series for oligomer H-(*HalA*-AlaT-*HalA*-AlaT-*HalA*-AlaT)-*Lys*-NH_2 (**15**, T_m = 19 °C, 5% *H*, 6 μM) also selfpairing was observed as the preferred complex. Norvalyl-PNA oligomer H-(*NvaA*-NvaT-*NvaA*-NvaT-*NvaA*-NvaT)-*Lys*-NH_2 (**16**, T_m = 11 °C, 6% *H*, 12 μM) provided a slight preference for pairing with its enantiomer ent-**16** (T_m = 13 °C, 8% *H*, 6 μM each). The overall drop in stability induced by side chain homologation was not as significant for A–T paired oligomers as was observed for G–C paired PNAs. The similar quality of the Watson–Crick and the Hoogsteen pairing modes in A–T base pairs is likely to result in more alternatives for better stabilization also of the homologized oligomers.

So far only mixed oligomers with an alternation of alanyl and homoalanyl nucleo amino acids were investigated. Therefore, oligomer H-(*HalA*-HalA-*AlaT*-HalA-*AlaT*-AlaT)-*Lys*-NH_2 (**14**) was prepared in order to determine the influence of nucleo amino acids with different side chain lengths being not in sequential alternation. As shown in Fig. 4 there should be a difference in pairing possibilities for oligomers **14** and **15**, indeed. The Watson–Crick sites of both oligomers **14** and **15** are all oriented towards the same direction (in Fig. 4 towards the reader) whereas the orientation of the Hoogsteen sites is dependent on the sequence: in oligomer **15** with alternation of the alanyl and homoalanyl amino acids they are all oriented alike and can be used for hybridization with a complementary strand. Oligomer **14** has two neighboring adeninyl homoalanines, however only one of them could possibly participate in pairing with the complementary strand. This clearly favors the Watson–Crick pairing since it is not affected by the missing alternation of alanyl and homoalanyl amino acids. Therefore, we can conclude that the stability of oligomer **14** (T_m = 22 °C, 12% *H*, 6 μM) is based on six A–T Watson–Crick pairs. Furthermore, the circular dichroism spectra of the selfpairing complexes of oligomers **14** and **15** are very similar (not shown) indicating that both alanyl/homoalanyl oligomers prefer the Watson–Crick pairing mode. The CD spectra of the double strand with the comparable alanyl-PNA sequence H-(*AlaA*-AlaA-*AlaT*-AlaA-*AlaT*-AlaT)-*Lys*-NH_2 (**13**) forming the reverse Hoogsteen mode differ from the spectra of the higher homologues **14** and **15**.

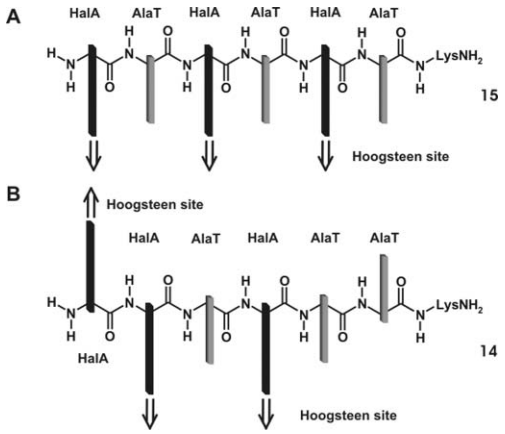


Fig. 4 The sequential order of alanyl and homoalanyl nucleo amino acids determines the pairing mode: **A**. alanyl/homoalanyl alternating PNA **15** allows pairing over the Watson–Crick or Hoogsteen site; **B**. two succeeding homoalanyl nucleo amino acids in oligomer **14** strongly favour pairing over the Watson–Crick site.

Effects on base stacking and solvation

As indicated so far, specific recognition by hydrogen bond formation has an influence on the PNA double strand stabilities but in addition stacking of the aromatic base pairs and solvation need to be considered. Furthermore, the PNAs with longer side chains are less preorganized entropically lowering the double strand stability. In case of oligonucleotide helices it is quite difficult to differentiate between these contributions because base pair orientations are topologically restricted and oligonucleotide double strands are highly dynamic. In contrast, the linear PNA double strands **8**–**12** provide base stacks with the same kind and the same amount of hydrogen bonds. They only differ in the respective base pair orientation and the length of the side chain linker. Within the Watson–Crick pairing oligomers we already attributed the drop in stability from alanyl-PNA (T_m = 58 °C, 13% *H*), alanyl/homoalanyl-PNA (T_m = 52 °C, 22% *H*), homoalanyl-PNA (T_m = 37 °C, 25% *H*), alanyl/norvalyl-PNA (T_m = 16 °C, 85% *H*), and norvalyl-PNA (T_m = 15 °C, 160% *H*) to stacking, solvation effects and loss of preorganization. The influence of side chain homology on base pair stacking in linear double strands was derived in an idealized form from Maruzen models.[26] The stacking of neighboring base pairs is indicated in Fig. 5 assuming a well staggered linker conformation. It is remarkable that by the inclusion of an additional methylene group in the linker not just the pairing sites can be altered but also the whole orientation of the nucleobase regarding the backbone gets reversed. In the case of alanyl-PNA, stacking of neighboring nucleobases in the G–C Watson–Crick pairing mode is allowed with opposite orientation. In mixed alanyl/homoalanyl PNA succeeding nucleobases are oriented alike to give a strong resemblance to the stacking of G–C-pairs in B-DNA[27,28] where all Hoogsteen sites have the same orientation pointing towards the major groove. The nucleobases in alanyl/homoalanyl PNA have a similar orientation of dipoles and intersecting base pairs.

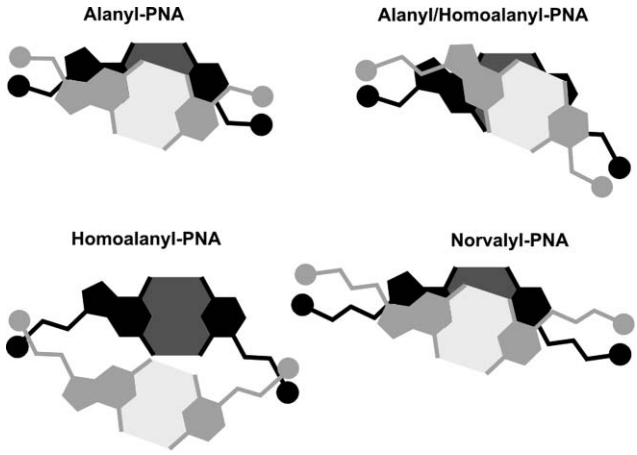


Fig. 5 Base pair stacking in linear PNA is dependent on the length of the side chain linker as shown for the G–C Watson–Crick mode in an idealized presentation.

Alanyl and alanyl/homoalanyl-PNA both provide a good overlap of consecutive nucleobase pairs. In contrast, homoalanyl-PNA suffers from an opposite outward orientation of both base pairs due to the additional methylene groups. Homoalanyl-PNA is likely to overcome the idealized situation pictured in Fig. 5 by changing the side chain conformation. Nevertheless, it seems impossible to compensate the low stacking contribution which therefore would be structurally inherent. The idealized stacking of norvalyl-PNA is identical to alanyl-PNA. The enlarged linker of each base pair allows conformational reorganization that also might be forced by the solvent polarity. As pointed out, this conformational flexibility might even allow for a change of pairing modes. The high hyperchromicities determined for the norvalyl containing PNAs also point to an optimized base pair orientation made possible by higher side chain flexibility. From the results of norvalyl containing PNA we need to emphasize that the double strand rigidity in the systems with longer side chain linker is not sufficient for determination of stacking orientations any more.

The dependence of PNA base stacking on the linker length clearly needs to be considered next to hydrogen bonding, solvation and preorganization. Side chain homology offers the possibility to specifically design base stacks with various nucleobase orientations. Previously we described another type of base stack with all the nucleobase pairs being completely oriented alike. This system was obtained by backbone homology using β-homoalanyl-PNA.[29,30]

Conclusions

The linearity and rigidity of PNAs based on a regular peptide backbone offer a valuable scaffold to observe the influence of the linker length between backbone and nucleobase on pairing selectivity and base pair stacking. It was shown that side chain homology affects the pairing selectivity by altering the orientation of Hoogsteen sites as well as donor/acceptor positions at the Watson–Crick site. Furthermore, three different stacking situations can be set up dependent on the use of alanyl or homoalanyl nucleo amino acids. Norvalyl-PNA with a three atom linker between the nucleobase and the backbone also provided double strands. Nevertheless, the predictability of pairing mode and stacking orientation were lost because of the higher flexibility of the side chain.

Overall, the different types of PNA are valuable model systems and tools for the investigation of base stack mediated processes like intercalation, electron transfer, and interactions with the base stack. For isosteres of natural biooligomers these model studies indicate that the extension of a side chain linker by a methylene group does affect the functional unit regarding facial orientation as well as the order of donor/acceptor positions. Modifications of biomolecules aiming functional similarity should better be isosteres or homologized by a two atom unit and not by one methylene group only.

Experimental

General

All reagents were of analytical grade and used as supplied. Solvents were of the highest grade available. NMR spectra were recorded with a Bruker AC 250, Bruker DMX 400 or Bruker DMX 500 spectrometer. ESI mass spectra were measured with a TSQ 700 Finnigan spectrometer and high-resolution mass spectra with a Bruker Apex IV FT-ICR MS instrument. For optical rotation a Perkin-Elmer Polarimeter 241 MC, for elemental analysis a Heraeus EA 415-0, and for melting point determination a Büchi SMP-20 apparatus were used. HPLC was done on Pharmacia Äkta basic using YMC-Pack ODS, RP-C18 with a linear gradient of A (0.1% TFA in water) to B (0.1% TFA in acetonitrile–water 9 : 1) or to B′ (0.1% TFA in acetonitrile). Oligomers were purified using 250 × 20 mm, 5 µm, 120 Å for preparation and 250 × 4.6 mm, 5 µm, 120 Å for analytical samples. UV melting curves were measured with a Perkin-Elmer UV/Vis Lambda 10 with a Peltier Temperature Programmer PTP 1 or a JASCO V-550 UV/Vis spectrometer equipped with a JASCO ETC-505S/ETC-505T temperature controller. CD spectra were recorded with a Jasco J-715 or Jasco J-810 with a Peltier Type Temperature Control System PTC-348W. The oligomer concentration was calculated by taking the extinction coefficient at 80 °C as being the sum of the extinction coefficients of the nucleo amino acids. The optical purity was determined by HPLC analysis of the amino acid dimers obtained with Boc-(*S*)-Ala-OSu or Boc-(*R*)-Phe-OSu.

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-hydroxy-norvaline benzyl ester 7. (*S*)-*N*-Boc-Glutamic acid benzyl ester (**6**) (7.00 g, 20.8 mmol) was dissolved in anhydrous THF (100 mL). *N*-Ethylmorpholine (2.39 mL, 22.8 mmol) and chloroformic acid isobutyl ester (2.98 mL, 22.82 mmol) was added at −10 °C and stirred for 10 min at −10 °C. After adding sodium borohydride (2.35 g, 62.2 mmol) MeOH (400 mL) was added within 30 min. The mixture was stirred for 15 min at −10 °C and 30 min at room temperature. After neutralization with 1 N HCl (42 mL) and evaporation the product was dissolved in AcOEt (150 mL). The organic layer was washed with 1 N HCl (150 mL), H_2O (150 mL), 6% aq $KHCO_3$ (150 mL) and brine (150 mL). It was dried over Na_2SO_4, filtered and evaporated. Purification by column chromatography on silica (hexane–AcOEt 1 : 1) gave amino acid **7** (5.03 g, 75%) as a colorless oil. R_f [hexane–AcOEt 1 : 1] = 0.31; $[\alpha]^{20}_D$ −2.4 (*c* 1.0 in MeOH); δ_H (200 MHz; $CDCl_3$; Me_4Si) 1.43 (9 H, s, Boc), 1.51–2.01 (4 H, m, H-β, H-γ), 3.63 (2 H, m, H-δ), 4.39 (1 H, m, H-α), 5.10–5.25 (3 H, m, N*H*Boc, CH_2Ph), 7.37 (5 H, m, Ph); δ_C (50.3 MHz; $CDCl_3$; Me_4Si) 28.2, 28.3 (C_{prim}–Boc), 29.3, 53.2 (C-α), 61.9 (C-δ), 67.0 (*C*H_2Ph), 80.0 (C_{qu}–Boc), 128.3 (C_{tert}–Ph), 128.4 (C_{tert}–Ph), 128.6 (C_{tert}–Ph), 135.4 (C_{qu}–Ph), 155.5 (*C*ONH), 172.7 (*C*OBn); λ_{max}(MeOH)/nm 258; ESI-MS *m*/*z*: 346.4 [M + Na]$^+$; HRMS calcd for $C_{17}H_{25}NO_5$ [M + H]$^+$ 324.1805, found 324.1807.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-hydroxy-norvaline benzyl ester ent-7. The synthesis followed the procedure for enantiomer **7**. The analytical data of ent-**7** and **7** are identical except for $[\alpha]^{20}_D$ +2.3 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-bromo-norvaline benzyl ester 5. (*S*)-*N*-Boc-δ-Hydroxy-norvaline benzyl ester (**7**) (5.03 g, 15.6 mmol) was dissolved in anhydrous CH_2Cl_2 (100 mL). At −5 °C and under exclusion of light carbon tetrabromide (10.3 g, 31.1 mmol) was added. Triphenylphosphine (8.16 g, 31.1 mmol) was dissolved in anhydrous CH_2Cl_2 (50 mL), cooled at −5 °C and added to the cold reaction mixture. After stirring for 2 h at room temperature the solvent was evaporated. The residue was purified by column chromatography on silica (hexane–AcOEt 9 : 1) to yield amino acid **5** (3.18 g, 53%) as a yellow oil. R_f [hexane–AcOEt 9 : 1] = 0.21; $[\alpha]^{20}_D$ −20.5 (*c* 1.6 in MeOH); δ_H (300 MHz; $CDCl_3$; Me_4Si) 1.43 (9 H, s, Boc), 1.74–2.03 (4 H, m, H-β, H-γ), 3.38 (2 H, m, H-δ), 4.38 (1 H, m, H-α), 5.06 (1 H, m, N*H*Boc), 5.18 (2 H, m, CH_2Ph), 7.38 (5 H, m, Ph); δ_C (50.3 MHz; $CDCl_3$; Me_4Si) 28.3 (C_{prim}–Boc), 28.5, 31.4, 32.8, 52.8 (C-α), 67.2 (*C*H_2Ph), 80.0 (C_{qu}–Boc), 128.3 (C_{tert}–Ph), 128.5 (C_{tert}–Ph), 128.6 (C_{tert}–Ph), 135.2 (C_{qu}–Ph), 155.3 (*C*ONH), 172.2 (*C*OBn); λ_{max}(MeOH)/nm 258; ESI-MS *m*/*z*: 408.0 [M + Na]$^+$; HRMS calcd for $C_{17}H_{24}BrNO_4$ [M + Na]$^+$ 408.0781, found 408.0777.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-bromo-norvaline benzyl ester ent-5. The synthesis followed the procedure for enantiomer **5**. The analytical data of ent-**5** and **5** are identical except for $[\alpha]^{20}_D$ +15.6 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(1-thyminyl)norvaline benzyl ester. To a solution of (*S*)-*N*-Boc-δ-bromo-norvaline benzyl ester (**5**) (1.20 g, 3.11 mmol) in anhydrous DMF (110 mL) dry K_2CO_3 (644 mg, 4.66 mmol), thymine (588 mg, 4.66 mmol) and tetrabutylammonium iodide (115 mg, 310 μmol) were added. After 3 d AcOH (267 μL, 4.66 mmol) was added and evaporated. Purification by column chromatography on silica (hexane–AcOEt 2 : 3) gave (*S*)-*N*-Boc-δ-(1-thyminyl)-norvaline benzyl ester (550 mg, 41%) as a white solid. R_f [hexane–AcOEt 2 : 3] = 0.26; $[\alpha]^{20}_D$ −16.5 (*c* 1.0 in MeOH); δ_H (300 MHz; $CDCl_3$; Me_4Si) 1.44 (9 H, s, Boc), 1.59–1.91 (4 H, m, H-β, H-γ), 1.89 (3 H, s, CH_3–T), 3.71 (2 H, m, H-δ), 4.38 (1 H, m, H-α), 5.10–5.25 (3 H, m, N*H*Boc, CH_2Ph), 6.98 (1 H, br s, H-6), 7.36 (5 H, m, Ph), 8.43 (1 H, br s, N*H*–T); δ_C (50.3 MHz; $CDCl_3$; Me_4Si) 12.2 (*C*H_3–T), 24.8, 28.2 (C_{prim}–Boc), 29.8, 47.5 (C-δ), 52.5 (C-α), 67.2 (*C*H_2Ph), 80.1 (C_{qu}–Boc), 128.2–128.6 (C_{tert}–Ph), 135.1 (C_{qu}–Ph), 140.3 (C-6), 150.9 (C-2), 155.5 (*C*ONH), 164.4 (C-4), 172.1 (*C*OBn); λ_{max}(MeOH)/nm 269; ESI-MS *m*/*z*: 454.3 [M + Na]$^+$, 885.1 [2 M + Na]$^+$; HRMS calcd for $C_{22}H_{29}N_3O_6$ [M + H]$^+$ 432.2129, found 432.2126.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(1-thyminyl)norvaline benzyl ester. The synthesis followed the procedure for its enantiomer. The analytical data were identical except for $[\alpha]^{20}_D$ +15.3 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(1-thyminyl)norvaline 1. PdO·H_2O (250 mg, 1.78 mmol) was added to a solution of (*S*)-*N*-Boc-δ-(1-thyminyl)norvaline benzyl ester (500 mg, 1.16 mmol) in MeOH (10 mL) and AcOH (250 μL). After saturation with hydrogen, the reaction mixture was stirred for 2 h. The precipitated nucleo amino acid was dissolved in MeOH–AcOH (9 : 1) and separated from the palladium oxide. After coevaporation with toluene the residue was purified by RP silica column chromatography (H_2O to H_2O–MeOH 3 : 1) to yield nucleo amino acid **1** (230 mg, 59%, >96% e.e.) as a white solid (HPLC for H-(*R*)-Phe-(*S*)-NvaT-OH: t_R = 21.4 min, gradient 13–21% B in 30 min). R_f [MeOH–AcOEt 1 : 4] = 0.48; $[\alpha]^{20}_D$ +5.4 (*c* 1.0 in MeOH); δ_H (200 MHz; CD_3OD; Me_4Si) 1.43 (9 H, s, Boc), 1.54–1.80 (4 H, m, H-β, H-γ), 1.86 (3 H, s, CH_3–T), 3.75 (2 H, m, H-δ), 4.00 (1 H, m, H-α), 7.45 (1 H, s, H-6); δ_C (50.3 MHz; CD_3OD; Me_4Si) 12.3 (*C*H_3–T), 26.4, 28.8 (C_{prim}–Boc), 30.8, 47.7–50.3 (C-δ, *C*D_3OD), 55.9 (C-α), 80.3 (C_{qu}–Boc), 111.5 (C-5), 143.3 (C-6), 153.4 158.5 (C-2, *C*ONH), 166.8 (C-4); λ_{max}(MeOH)/nm 271 (ε/dm^3 mol^{-1} cm^{-1} 8 157); ESI-MS *m*/*z*: 386.4 [M − H + 2 Na]$^+$; HRMS calcd for $C_{15}H_{23}N_3O_6$ [M + H]$^+$ 364.1479, found 364.1481.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(1-thyminyl)norvaline ent-1. The synthesis followed the procedure for enantiomer **1**. The analytical data of ent-**1** and **1** are identical except for $[\alpha]^{20}_D$ −6.3 (*c* 0.5 in MeOH) and HPLC for H-(*R*)-Phe-(*R*)-NvaT-OH: t_R = 20.6 min, gradient 13–21% B in 30 min.

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(9-adeninyl)norvaline benzyl ester. To a solution of (*S*)-*N*-Boc-δ-bromo-norvaline benzyl ester **5** (1.10 g, 2.85 mmol) in anhydrous DMF (150 mL) dry K_2CO_3 (590 mg, 4.27 mmol), adenine (577 mg, 4.27 mmol) and tetrabutylammonium iodide (108 mg, 290 μmol) were added. After 3 d AcOH (245 μL, 4.27 mmol) was added and evaporated. Purification by silica column chromatography (MeOH–AcOEt 1 : 9) yielded the (*S*)-*N*-Boc-δ-(9-adeninyl)-norvaline benzyl ester (998 mg, 80%) as a white solid. R_f [MeOH–AcOEt 1 : 9] = 0.39; $[\alpha]^{20}_D$ −10.7 (*c* 1.0 in MeOH); δ_H (300 MHz; $CDCl_3$; Me_4Si) 1.43 (9 H, s, Boc), 1.60–2.00 (4 H, m, H-β, H-γ), 4.19 (2 H, m, H-δ), 4.42 (1 H, m, H-α), 5.15 (2 H, m, CH_2Ph), 5.33 (1 H, d, ^3J(H,H) = 8.4 Hz, N*H*-Boc), 5.71 (2 H, br s, H/D-exchangeable, NH_2), 7.34 (5 H, m, H–Ph), 7.72 (1 H, s, H-2, H-8), 8.34 (1 H, s, H-2, H-8); δ_C (50.3 MHz; $CDCl_3$; Me_4Si) 26.1, 28.2 (C_{prim}–Boc), 29.7, 43.0 (C-δ), 52.7 (C-α), 67.1 (*C*H_2Ph), 80.0 (C_{qu}–Boc), 119.4 (C-5), 128.3 (C_{tert}–Ph), 128.4 (C_{tert}–Ph), 128.5 (C_{tert}–Ph), 135.1 (C_{qu}–Ph), 140.2 (C-8), 149.9 (C-4), 152.8 (C-2), 155.4, 155.5, 172.1 (*C*OBn); λ_{max}(MeOH)/nm 261; ESI-MS *m*/*z*: 463.3 [M + Na]$^+$, 903.1 [2 M + Na]$^+$; HRMS calcd for $C_{22}H_{28}N_6O_4$ [M + H]$^+$ 441.2245, found 441.2241.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(9-adeninyl)norvaline benzyl ester. The synthesis followed the procedure for its enantiomer. The analytical data are identical except for $[\alpha]^{20}_D$ +11.0 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(9-adeninyl)norvaline 2. PdO·H_2O (300 mg, 2.67 mmol) was added to a solution of (*S*)-*N*-Boc-δ-(9-adeninyl)norvaline benzyl ester (600 mg, 1.36 mmol) in MeOH (10 mL) and AcOH (300 μL). After saturation with hydrogen, the reaction mixture was stirred for 2 h. The precipitated nucleo amino acid was dissolved in MeOH–AcOH (9 : 1) and separated from the palladium oxide. After coevaporation with toluene the residue was purified by RP silica column chromatography (H_2O to H_2O–MeOH 3 : 1) to yield

nucleo amino acid **2** (348 mg, 73%, >95% e.e.) as a white solid (HPLC for H-(*R*)-Phe-(*S*)-NvaA-OH: t_R = 24.3, gradient 8–19% B in 30 min). R_f [MeOH–AcOEt 1 : 4] = 0.52; $[\alpha]^{20}_D$ +3.9 (*c* 0.9 in MeOH); δ_H (300 MHz; [D$_6$]DMSO; Me$_4$Si) 1.35 (9 H, s, Boc), 1.59 (2 H, m, H-β, H-γ), 1.84 (2 H, m, H-β, H-γ), 3.90 (1 H, m, H-α), 4.13 (2 H, t, ^3J(H,H) = 6.6 Hz, H-δ), 6.96 (1 H, d, ^3J(H,H) = 7.8 Hz, N*H*-Boc), 7.11 (2 H, s, NH_2), 8.09 (1 H, s, H-2, H-8), 8.12 (1 H, s, H-2, H-8); δ_C (75.5 MHz; [D$_6$]DMSO; Me$_4$Si) 26.3, 28.1 (C(*C*H$_3$)$_3$), 42.5, 53.1 (C-α), 77.9 (C-δ, C$_{prim}$–Boc), 81.9 (C-δ, C$_{qu}$–Boc), 118.7 (C-5), 140.7 (C-8), 149.5 (C-4), 152.2 (C-2), 155.4, 155.9, 173.7 (*C*OOH); λ_{max}(MeOH)/nm 260 (ε/dm^3 mol^{-1} cm^{-1} 16 407); ESI-MS *m*/*z*: 351.4 [M + H]$^+$, 373.3 [M + Na]$^+$; HRMS calcd for $C_{15}H_{22}N_6O_4$ [M + H]$^+$ 351.1772, found 351.1775.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(9-adeninyl)norvaline ent-2. The synthesis followed the procedure for enantiomer **2**. The analytical data of ent-**2** and **2** are identical except for $[\alpha]^{20}_D$ −11.5 (*c* 0.8 in MeOH) and HPLC for H-(*R*)-Phe-(*R*)-NvaA-OH: t_R = 24.5 min, gradient 8–19% B in 30 min.

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(*N*4-benzyloxycarbonyl-1-cytosinyl)norvaline benzyl ester. To a solution of (*S*)-*N*-Boc-δ-bromo-norvaline benzyl ester (**5**) (1.00 g, 2.59 mmol) in anhydrous DMF (150 mL) dry K_2CO_3 (537 mg, 3.88 mmol), *N*4-benzyloxycarbonyl cytosine (952 mg, 3.88 mmol) and tetrabutylammonium iodide (96.2 mg, 260 μmol) were added. After 3 d AcOH (222 μL, 3.88 mmol) was added and evaporated. Purification by column chromatography on silica (AcOEt) gave the (*S*)-*N*-Boc-δ-(*N*4-benzyloxycarbonyl-1-cytosinyl)-norvaline benzyl ester (959 mg, 69%) as a white solid. R_f [AcOEt] = 0.45; $[\alpha]^{20}_D$ +30.4 (*c* 1.0 in MeOH); δ_H (300 MHz; CDCl$_3$; Me$_4$Si) 1.43 (9 H, s, Boc), 1.64–1.91 (4 H, m, H-β, H-γ), 3.86 (2 H, m, H-δ), 4.39 (1 H, m, H-α), 5.11–5.23 (6 H, m, C*H*$_2$Ph, N*H*-Boc), 7.14 (1 H, d, ^3J(H,H) = 10.8 Hz, H-5), 7.32–7.40 (10 H, m, H–Ph), 7.51 (1 H, d, ^3J(H,H) = 10.8 Hz, H-6); δ_C (50.3 MHz; CDCl$_3$; Me$_4$Si) 24.7, 28.2 (C$_{prim}$–Boc), 29.8, 49.8 (C-δ), 52.5 (C-α), 67.2, 67.8, 80.1 (C$_{qu}$–Boc), 94.8 (C-5), 128.3–128.6 (C$_{tert}$–Ph), 135.0 (C$_{qu}$–Ph), 135.1 (C$_{qu}$–Ph), 148.5 (C-6), 152.3, 155.5, 162.1, 172.1 (*C*OBn); λ_{max}(MeOH)/nm 245, 295; ESI-MS *m*/*z*: 573.4 [M + Na]$^+$, 1123.3 [2 M + Na]$^+$; HRMS calcd for $C_{29}H_{34}N_4O_7$ [M + H]$^+$ 551.2500, found 551.2503.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(*N*-4-benzyloxycarbonyl-1-cytosinyl)norvaline benzyl ester. The synthesis followed the procedure for the enantiomer. The analytical data were identical except for $[\alpha]^{20}_D$ −27.3 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(*N*4-benzyloxycarbonyl-1-cytosinyl)norvaline 3. To a solution of (*S*)-*N*-Boc-δ-(*N*4-benzyloxycarbonyl-1-cytosinyl)norvaline benzyl ester (400 mg, 730 μmol) in dioxane–H$_2$O (60 mL, 3 : 2) 1 N sodium hydroxide (940 μL) was added. After stirring for 1 d at room temperature 1 N HCl was added until pH 6 was obtained. After coevaporation with toluene the residue was purified by column chromatography on RP silica (H$_2$O to H$_2$O–MeOH 3 : 1) yielding nucleo amino acid **3** (211 mg, 63%, >97% e.e.) as a white solid (HPLC for H-(*R*)-Phe-(*S*)-NvaC-OH: t_R = 14.9 min, gradient 25–40% B in 20 min). R_f [MeOH] = 0.53; $[\alpha]^{20}_D$ +39.2 (*c* 1.0 in MeOH); δ_H (300 MHz; [D$_6$]DMSO; Me$_4$Si) 1.34 (9 H, s, Boc), 1.42–1.67 (4 H, m, H-β, H-γ), 3.57 (1 H, m, H-α), 3.74 (2 H, m, H-δ), 5.16 (2 H, s, C*H*$_2$Ph), 5.95 (1 H, m, N*H*-Boc), 6.92 (1 H, d, ^3J(H,H) = 7.2 Hz, H-5), 7.37 (5 H, m, H–Ph), 8.02 (1 H, d, ^3J(H,H) = 7.2 Hz, H-6), 10.66 (1 H, br s, N*H*–C); δ_C (50.3 MHz; CD$_3$OH; Me$_4$Si) 26.1, 28.8 (C$_{prim}$–Boc), 31.2, 51.4 (C-δ), 56.5 (C-α), 68.5 (*C*H$_2$Ph), 80.0 (C$_{qu}$–Boc), 96.8 (C-5), 129.3 (C$_{tert}$–Ph), 129.4 (C$_{tert}$–Ph), 129.6 (C$_{tert}$–Ph), 137.1 (C$_{qu}$–Ph), 150.7 (C-6), 154.4, 157.6, 158.4, 164.6, 179.1 (*C*OOH); λ_{max}(MeOH)/nm 241, 296 (ε/dm^3 mol^{-1} cm^{-1} 14 179); ESI-MS *m*/*z*: 505.4 [*M* − H + 2 Na]$^+$; HRMS calcd for $C_{22}H_{28}N_4O_7$ [M + H]$^+$ 461.2031, found 461.2033.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(*N*4-benzyloxycarbonyl-1-cytosinyl)norvaline ent-3. The synthesis followed the procedure for enantiomer **3**. The analytical data of ent-**3** and **3** are identical except for $[\alpha]^{20}_D$ −8.7 (*c* 1.0 in MeOH) and HPLC for H-(*R*)-Phe-(*R*)-NvaC-OH: t_R = 13.4 min, gradient 25–40% B in 20 min.

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(2-amino-6-chloro-9-purinyl)-norvaline benzyl ester. To a solution of (*S*)-*N*-Boc-δ-bromo-norvaline benzyl ester (**5**) (1.50 g, 3.88 mmol) in anh. DMF (150 mL) dry K_2CO_3 (805 mg, 5.82 mmol), 2-amino-6-chloropurine (988 mg, 5.82 mmol) and tetrabutylammonium iodide (144 mg, 390 μmol) were added. After 2 d AcOH (334 μL, 5.82 mmol) was added and evaporated. Purification by column chromatography on silica (hexane–AcOEt 1 : 2) gave (*S*)-*N*-Boc-δ-(2-amino-6-chloro-9-purinyl)-norvaline benzyl ester (1.60 g, 87%) as a white solid. R_f [hexane–AcOEt 1 : 2] = 0.29; $[\alpha]^{20}_D$ −14.9 (*c* 1.4 in MeOH); δ_H (300 MHz; CDCl$_3$; Me$_4$Si) 1.44 (9 H, s, Boc), 1.52–1.94 (4 H, m, H-β, H-γ), 4.09 (2 H, m, H-δ), 4.47 (1 H, m, H-α), 5.08–5.23 (4 H, m, C*H*$_2$Ph, NH_2), 5.38 (1 H, d, ^3J(H,H) = 12.0 Hz, N*H*), 7.32 (m, 5 H; H–Ph), 7.69 (s, 1 H; H-8); δ_C (50.3 MHz; CDCl$_3$; Me$_4$Si) 25.9, 28.2 (C$_{prim}$–Boc), 29.6, 42.8 (C-δ), 53.6 (C-α), 67.2 (*C*H$_2$Ph), 80.2 (C$_{qu}$–Boc), 125.0 (C-5), 128.3 (C$_{tert}$–Ph), 128.6 (C$_{tert}$–Ph), 128.6 (C$_{tert}$–Ph), 135.0 (C$_{qu}$–Ph), 142.2 (C-8), 151.3, 153.7, 155.4, 159.1, 172.1 (*C*OBn); λ_{max}(MeOH)/nm 248, 311; ESI-MS *m*/*z*: 497.3 [M + Na]$^+$, 970.9 [2 M + Na]$^+$; HRMS calcd for $C_{22}H_{27}ClN_6O_2$ [M + H]$^+$ 475.1855, found 475.1857.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(2-amino-6-chloro-9-purinyl)-norvaline benzyl ester. The synthesis followed the procedure for the enantiomer. The analytical data were identical except for $[\alpha]^{20}_D$ +20.5 (*c* 1.0 in MeOH).

(*S*)-δ-(9-Guaninyl)norvaline benzyl ester. (*S*)-*N*-Boc-δ-(2-Amino-6-chloro-9-purinyl)norvaline benzyl ester (900 mg, 1.89 mmol) was dissolved in TFA–H$_2$O (8 mL, 3 : 1) and stirred for 2 d at room temperature. The solvent mixture was coevaporated with toluene. After drying *in vacuo* a white solid was obtained in quantitative yield. R_f [CHCl$_3$–MeOH–H$_2$O–AcOH 70 : 30 : 3 : 0.3] = 0.25; $[\alpha]^{20}_D$ −0.2 (*c* 0.9 in MeOH); δ_H (300 MHz; CDCl$_3$; Me$_4$Si) 1.94–2.08 (4 H, m, H-β, H-γ), 4.21–4.35 (3 H, m, H-δ, H-α), 5.27 (2 H, q, ^3J(H,H) = 12.0 Hz, C*H*$_2$Ph), 7.36 (5 H, m, H–Ph), 9.05 (1 H, s, H-8); δ_C (50.3 MHz; CDCl$_3$; Me$_4$Si) 26.0, 28.3, 45.4 (C-δ), 53.3 (C-α), 69.2 (*C*H$_2$Ph), 108.7 (C-5), 128.3 (C-8), 129.7 (C$_{tert}$–Ph), 129.8 (C$_{tert}$–Ph), 129.8 (C$_{tert}$–Ph), 136.3 (C$_{qu}$–Ph), 151.6, 154.9, 157.3, 170.0 (*C*OBn); λ_{max}(MeOH)/nm 257; ESI-MS *m*/*z*: 357.4 [M + H]$^+$, 713.3 [2 M + H]$^+$, 735.2 [2 M + Na]$^+$; HRMS calcd for $C_{17}H_{20}N_6O_3$ [M + H]$^+$ 357.1670, found 357.1674.

(*R*)-δ-(9-Guaninyl)norvaline benzyl ester. The synthesis followed the procedure for the enantiomer. The analytical data were identical except for $[\alpha]^{20}_D$ + 1.3 (*c* 1.0 in MeOH).

(*S*)-*N*-*tert*-Butoxycarbonyl-δ-(9-guaninyl)norvaline 4. (*S*)-δ-(9-Guaninyl)norvaline benzyl ester (874 mg, 3.28 mmol) was dissolved in H$_2$O–1 N NaOH–dioxane (25 mL, 1 : 1 : 2, pH 9) and stirred for 4 h. At 0 °C di-*tert*-butyldicarbonate (787 mg, 3.61 mmol) was added, stirred for 45 min at 0 °C and 3 d at room temperature keeping the pH slightly higher than 9. Finally, the reaction mixture was acidified to pH 6 and evaporated. The residue was purified by column chromatography on RP silica (AcOEt–MeOH 4 : 1, 1% AcOH) to provide (*S*)-*N*-Boc-δ-(9-guaninyl)norvaline **4** (633 mg, 55%, >98% e.e.) as a white solid (HPLC for H-(*R*)-Phe-(*S*)-NvaG-OH: t_R = 11.7 min, gradient 12–20% B in 20 min). R_f [CHCl$_3$–MeOH–H$_2$O–AcOH 70 : 30 : 3 : 0.3] = 0.20; $[\alpha]^{20}_D$ +22.9 (*c* 1.0 in MeOH); δ_H (300 MHz; [D$_6$]DMSO; Me$_4$Si) 1.34 (9 H, s, Boc), 1.94–1.72 (4 H, m, H-β, H-γ), 3.72 (1 H, br s, H-α), 3.88 (2 H, m, H-δ), 6.11 (1 H, m, N*H*-Boc), 6.64 (2 H, br s, NH_2), 7.61 (1 H, s, H-8), 11.0 (1 H, br s, N*H*-guanine); δ_C (75.5 MHz; [D$_6$]DMSO; Me$_4$Si)

26.0, 28.1 (C_{prim}–Boc), 29.9, 42.5 (C-δ), 54.4 (C-α), 77.5 (C_{qu}–Boc), 116.5 (C-5), 137.2 (C-8), 151.1, 153.7, 154.8, 157.0, 175.7 (*C*OOH); λ_{max}(MeOH)/nm 255 (ε/dm^3 mol^{-1} cm^{-1} 12 106); ESI-MS *m*/*z*: 389.5 $[M + Na]^+$; HRMS calcd for $C_{15}H_{22}N_6O_5$ $[M + H]^+$ 367.1724, found 367.1725.

(*R*)-*N*-*tert*-Butoxycarbonyl-δ-(9-guaninyl)norvaline ent-4. The synthesis followed the procedure for enantiomer **4**. The analytical data of ent-**4** and **4** were identical except for $[\alpha]^{20}_D$ −17.8 (*c* 1.0 in MeOH) and HPLC for H-(*R*)-Phe-(*R*)-NvaG-OH: t_R = 12.1 min, gradient 12–20% B in 20 min.

General method for SPPS of alanyl, homoalanyl and norvalyl-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*)-lysine(Z)-OH (0.309 mmol g^{-1}, Z = benzyloxycarbonyl) in a small column. For each coupling step an excess of four equivalents *N*-*tert*-butoxycarbonyl protected nucleo amino acid (77.3 mmol) was used and activated by *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU, 26.4 mg, 69.5 mmol), 5 equivalents 1-hydroxy-7-azabenzotriazol (HOAt), and *N*,*N*-diisopropylethylamine (26.3 μL, 155 mmol) in DMF (600 μL). After swelling of the resin for 20 min the following procedure was repeated for every nucleo amino acid unit: (1) deprotection twice, for 3 min with 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, 30–90 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 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 solution was concentrated to 400 μL, and the alanyl, homoalanyl or norvalyl-PNA precipitated with diethyl ether (5 mL) as a white solid. The PNA was separated using a centrifuge, followed by purification with HPLC (RP-C18). The yield of each coupling step was estimated from HPLC to be higher than 95%.

H-(*AlaG*-AlaG-*AlaC*-AlaG-*AlaC*-AlaC)-*Lys*-NH$_2$ 8. HPLC (gradient: 5–15% B′ in 30 min): t_R = 15.3 min; ESI-MS *m*/*z*: 674.0 $[M + 2H]^{2+}$.

H-(AlaG-*AlaG*-AlaC-*AlaG*-AlaC-*AlaC*)-Lys-NH$_2$ ent-8. HPLC (gradient: 5–15% B′ in 30 min): t_R = 17.1 min; ESI-MS *m*/*z*: 674.1 $[M + 2H]^{2+}$.

H-(*AlaG*-HalG-*AlaC*-HalG-*AlaC*-HalC)-*Lys*-NH$_2$ 9. HPLC (gradient: 5–15% B′ in 30 min): t_R = 20.4 min; ESI-MS *m*/*z*: 1388.4 $[M + H]^+$.

H-(AlaG-*HalG*-AlaC-*HalG*-AlaC-*HalC*)-Lys-NH$_2$ ent-9. HPLC (gradient: 5–15% B′ in 30 min): t_R = 23.7 min; δ_H (500 MHz; D_2O; Me_4Si) 1.29 (2 H, m, H-δ Lys), 1.56 (3 H, m, H-β, H-γ Lys), 1.69 (1 H, m, H-β Lys), 1.98 (2 H, m, H-β HalC), 2.05– 2.30 (4 H, m, H-β HalC, H-β HalG), 2.88 (2 H, t, ^3J(H,H) = 7 Hz, H-ε Lys), 3.62 (1 H, m, H-γ HalG), 3.76 (3 H, m, H-γ HalG, H-γ HalC), 3.90–4.05 (4 H, m, H-β AlaC, H-γ HalG), 4.06–4.20 (2 H, m, H-α Lys, H-α HalG), 4.17 (2 H, m, H-β AlaC), 4.30 (2 H, m, H-α HalG, H-α HalC), 4.47–4.63 (3 H, m, H-α AlaG, H-α AlaC), 5.94 (1 H, d, H-5 AlaC), 5.99 (1 H, d, ^3J(H,H) = 8 Hz, H-5 AlaC), 6.03 (1 H, d, H-5 HalC), 7.53 (1 H, d, ^3J(H,H) = 8 Hz, H-6 AlaC), 7.60 (1 H, d, ^3J(H,H) = 8 Hz, H-6 AlaC), 7.65 (1 H, s, H-8 AlaG), 7.68 (1 H, d, ^3J(H,H) = 8 Hz, H-6 HalC), 7.72 (1 H, s, H-8 AlaG), 7.82 (1 H, s, H-8 HalG). P. E. COSY, TOCSY, ROESY spectra were used for the assignment; ESI-MS *m*/*z*: 1388.4 $[M + H]^+$, 695.1 $[M + 2H]^{2+}$.

H-(*HalG*-HalG-*HalC*-HalG-*HalC*-HalC)-*Lys*-NH$_2$ 10. HPLC (gradient: 5–15% B′ in 30 min): t_R = 21.7 min; ESI-MS *m*/*z*: 716.2 $[M + 2H]^{2+}$.

H-(HalG-*HalG*-HalC-*HalG*-HalC-*HalC*)-Lys-NH$_2$ ent-10. HPLC (gradient: 5–15% B′ in 30 min): t_R = 19.7 min; ESI-MS *m*/*z*: 1430.6 $[M + H]^+$, 716.3 $[M + 2H]^{2+}$.

H-(*AlaG*-NvaG-*AlaC*-NvaG-*AlaC*-NvaC)-*Lys*-NH$_2$ 11. HPLC (gradient: 7–14% B in 30 min): t_R = 19.8 min; ESI-MS *m*/*z*: 715.8 $[M + 2 H]^{2+}$; HRMS calcd for $C_{57}H_{75}N_{33}O_{13}$ $[M + 2H]^{2+}$ 715.8184, found 715.8185.

H-(AlaG-*NvaG*-AlaC-*NvaG*-AlaC-*NvaC*)-Lys-NH$_2$ ent-11. HPLC (gradient: 7–14% B in 30 min): t_R = 20.0 min; ESI-MS *m*/*z*: 716.2 $[M + 2 H]^{2+}$.

H-(*NvaG*-NvaG-*NvaC*-NvaG-*NvaC*-NvaC)-*Lys*-NH$_2$ 12. HPLC (gradient: 10–12% B in 30 min): t_R = 13.4 min; ESI-MS *m*/*z*: 758.4 $[M + 2 H]^{2+}$, 1514.6 $[M + H]^+$; HRMS calcd for $C_{63}H_{87}N_{33}O_{13}$ $[M + 3H]^{3+}$ 505.5793, found 505.5789.

H-(NvaG-*NvaG*-NvaC-*NvaG*-NvaC-*NvaC*)-Lys-NH$_2$ ent-12. HPLC (gradient: 10–12% B in 30 min): t_R = 14.0 min; ESI-MS *m*/*z*: 758.4 $[M + 2 H]^{2+}$, 1514.6 $[M + H]^+$.

H-(*AlaA*-AlaA-*AlaT*-AlaA-*AlaT*-AlaT)-*Lys*-NH$_2$ 13. HPLC 21.0 min, gradient: 5–30% B′ in 30 min; ESI-MS *m*/*z*: 1343.4 $[M + H]^+$, 672.6 $[M + 2H]^{2+}$.

H-(AlaA-*AlaA*-AlaT-*AlaA*-AlaT-*AlaT*)-Lys-NH$_2$ ent-13. HPLC 21.0 min, gradient: 5–30% B′ in 30 min; ESI-MS *m*/*z*: 1343.3 $[M + H]^+$, 672.5 $[M + 2H]^{2+}$.

H-(*HalA*-*NvaC*-HalA-*AlaT*-HalA-*AlaT*-AlaT)-*Lys*-NH$_2$ 14. HPLC 15.6 min, gradient: 10–20% B′ in 30 min; λ = 260 nm. δ_H (250 MHz; D_2O; Me_4Si) 1.19–1.25 (2 H, m, H-γ Lys), 1.50–1.66 (11 H, m, thymine, H-δ Lys), 1.88–2.30 (6 H, m, H-β), 2.35–2.50 (2 H, m, H-β Lys), 2.75–2.90 (2 H, t, H-ε Lys), 3.90–4.20 (12 H, m, H-γ HalA, H-γ HalT, H-α HalA, H-α HalT), 4.50–4.60 (2 H, m, H-α Lys), 7.10 (1 H, s, thymine), 7.16 (1 H, s, thymine), 7.25 (1 H, s, thymine), 7.92–8.11 (6 H, m, adenine); ESI-MS *m*/*z*: 694.2 $[M + 2H]^{2+}$, 1385.7 $[M + H]^+$.

H-(*HalA*-AlaT-*HalA*-AlaT-*HalA*-AlaT)-*Lys*-NH$_2$ 15. HPLC (gradient:5–20% B′ in 30 min) t_R = 11.1 min; λ = 260 nm; δ_H (250 MHz; D_2O; Me_4Si) 1.25–1.45 (2 H, m, Lys), 1.50–1.85 (10 H, m, H-β, Lys), 2.05–2.50 (6 H, m, H-β HalA), 2.90 (2 H, m, H-ε Lys), 4.05–4.50 (13 H, m, H-α, H-γ HalA), 7.25 (1 H, s, AlaT), 7.35 (2 H, s, AlaT), 8.10–8.40 (6 H, m, HalA); MALDI-TOF-MS *m*/*z*: 1385.6 $[M + H]^+$.

H-(*NvaA*-NvaA-*NvaT*-NvaA-*NvaT*-NvaT)-*Lys*-NH$_2$ 16. HPLC (gradient: 12.5–14.0% B in 30 min): t_R = 27.5 min; ESI-MS *m*/*z*: 756.7 $[M + 2 H]^{2+}$, 1511.5 $[M + H]^+$; HRMS calcd for $C_{66}H_{90}N_{30}O_{13}$ $[M + 3H]^{3+}$ 504.5841, found 504.5845.

H-(NvaA-*NvaA*-NvaT-*NvaA*-NvaT-*NvaT*)-Lys-NH$_2$ ent-16. HPLC (gradient: 12.5–14% B in 30 min): t_R = 27.9 min; ESI-MS *m*/*z*: 756.7 $[M + 2 H]^{2+}$, 1511.6 $[M + H]^+$.

General method 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, 0.01 M) containing NaCl (0.1 M) and placed in a UV cell (10 mm). For UV melting curves the following temperature program was used to control the heating block: 80 °C → −2 °C (60 min) → −2 °C (180 min) → 90 °C (180 min) → −2 °C (180 min) → −2 °C (180 min) → 90 °C (180 min) → −2 °C (180 min).

Acknowledgements

Support by the Deutsche Forschungsgemeinschaft (SFB 416 and GRK 782) and the Fonds der Chemischen Industrie is gratefully acknowledged.

References

1 *Nucleic Acid Structure*, ed. S. Neidle, Oxford University Press, Oxford, 1999.
2 C. Branden and J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York, 1999.
3 J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, R. Krishnan, B. Jaun, C. Leumann and A. Eschenmoser, *Helv. Chim. Acta*, 1993, **76**, 259.
4 A. Eschenmoser and M. Dobler, *Helv. Chim. Acta*, 1992, **75**, 218.
5 U. Diederichsen, in *Bioorganic Chemistry—Highlights and New Aspects*, ed. U. Diederichsen, T. Lindhorst, L. A. Wessjohann and B. Westermann, Wiley-VCH, Weinheim, 1999, pp. 255–261.
6 (*a*) U. Diederichsen, *Angew. Chem.*, 1996, **108**, 458; (*b*) U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 445.
7 (*a*) U. Diederichsen, *Angew. Chem.*, 1997, **109**, 1966; (*b*) U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 1886.
8 U. Diederichsen, *Bioorg. Med. Chem. Lett.*, 1997, **7**, 1743.
9 (*a*) U. Diederichsen, *Angew. Chem.*, 1998, **110**, 2395; (*b*) U. Diederichsen, *Angew. Chem., Int. Ed.*, 1998, **37**, 2273.
10 U. Diederichsen and D. Weicherding, *SYNLETT*, 1999, **S1**, 917.
11 (*a*) J. Gante, *Angew. Chem.*, 1994, **106**, 1780; (*b*) J. Gante, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 1699.
12 R. Declercq, A. van Aerschot, R. J. Read, P. Herdewijn and L. van Meervelt, *J. Am. Chem. Soc.*, 2002, **124**, 928.
13 M. Böhringer, H.-J. Roth, J. Hunziker, M. Göbel, R. Krishnen, A. Giger, B. Schweizer, J. Schreiber, C. Leumann and A. Eschenmoser, *Helv. Chim. Acta*, 1992, **75**, 1416.
14 B. Hyrup, M. Egholm, M. Rolland, R. H. Berg, P. E. Nielsen and O. Buchardt, *J. Chem. Soc., Chem. Commun.*, 1993, 518.
15 B. Hyrup, M. Egholm, P. E. Nielsen, P. Wittung, B. Norden and O. Buchardt, *J. Am. Chem. Soc.*, 1994, **116**, 7964.
16 P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang and A. Eschenmoser, *Croat. Chim. Acta*, 1996, **69**, 535.
17 L. D. Arnold, T. H. Kalantar and J. C. Vederas, *J. Am. Chem. Soc.*, 1985, **107**, 7105.
18 L. D. Arnold, R. G. May and J. C. Vederas, *J. Am. Chem. Soc.*, 1988, **110**, 2237.
19 A. Lenzi, G. Reginato and M. Taddei, *Tetrahedron Lett.*, 1995, **36**, 1713.
20 A. Lenzi, G. Reginato, M. Taddei and E. Trifilieff, *Tetrahedron Lett.*, 1995, **36**, 1717.
21 P. Ciapetti, A. Mann, A. Schoenfelder, M. Taddei, E. Trifilieff, I. Canet and J. L. Canet, *Lett. Peptide Sci.*, 1997, **4**, 341.
22 P. Ciapetti, F. Soccolini and M. Taddei, *Tetrahedron*, 1997, **53**, 1167.
23 (*a*) R. Appel, *Angew. Chem.*, 1975, **87**, 863; (*b*) R. Appel, *Angew. Chem., Int. Ed. Engl.*, 1975, **14**, 801.
24 U. Diederichsen and H. W. Schmitt, *Tetrahedron Lett.*, 1996, **37**, 475.
25 The preference of A–T pairing alanyl-PNA for the reverse Watson–Crick or reverse Hoogsteen mode was investigated by 2,6-diaminopurine/thymine alanyl-PNA hexamers. The reverse Watson–Crick mode of diaminopurine/thymine is formed with three hydrogen bonds whereas the reverse Hoogsteen mode is only bidentate. Since diaminopurine/thymine alanyl-PNA hexamers provided comparable stabilities to the A/T alanyl-PNA the reverse Hoogsteen mode is likely to be formed in both cases.
26 Maruzen HGS Molecular Structure Model, Tokyo, Japan.
27 R. E. Dickerson, in *Nucleic Acid Structure*, ed. S. Neidle, Oxford University Press, Oxford, 1999, pp. 145–197.
28 U. Heinemann and C. Alings, *J. Mol. Biol.*, 1989, **210**, 369.
29 U. Diederichsen and H. W. Schmitt, *Angew. Chem.*, 1998, **110**, 312.
30 (*a*) A. M. Brückner, P. Chakraborty, S. H. Gellman and U. Diederichsen, *Angew. Chem.*, 2003, **115**, 4532; (*b*) A. M. Brückner, P. Chakraborty, S. H. Gellman and U. Diederichsen, *Angew. Chem., Int. Ed.*, 2003, **42**, 4395.