Structure-Activity Studies of the Binding of Modified Peptide Nucleic Acids (PNAs) to DNA[1]

Birgitte Hyrup,[†] Michael Egholm,[†] Peter E. Nielsen,[‡] Pernilla Wittung,[§] Bengt Nordén,[§] and Ole Buchardt*,[†]

Contribution from the Research Center for Medical Biotechnology, Chemical Laboratory II, The H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark, Department of Biochemistry B, The Panum Institute, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark, and Department of Physical Chemistry, Chalmers University of Technology, S-41296 Gothenburg, Sweden

Received October 21, 1992[⊗]

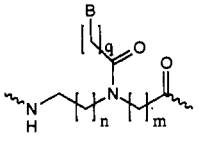
Abstract: Peptide nucleic acid (PNA) oligomers where one of the repeating backbone units is extended with a methylene group to either *N*-(2-aminoethyl)-β-alanine or *N*-(3-aminopropyl)glycine were prepared. Alternatively, the linker to the nucleobase was extended from methylenecarbonyl to ethylenecarbonyl. The thermal stability of the hybrids between these PNA oligomers and complementary DNA oligonucleotides was significantly lower than that of the corresponding complexes involving unmodified PNA. However, the sequence selectivity was retained. Thymidyl decamers with all *N*-(2-aminoethyl)-β-alanine or *N*-(3-aminopropyl)glycine backbones were prepared and shown to be unable to hybridize to the complementary $(dA)_{10}$ oligonucleotides, whereas a PNA decamer containing only ethylenecarbonyl linkers between the nucleobases and the *N*-(2-aminoethyl)glycine backbone showed weak but sequence-specific affinity for complementary DNA. All hybrids involving homopyrimidine PNA oligomers exhibited $(PNA)_2$/DNA stoichiometry, whereas mixed-sequence PNA oligomers formed PNA/DNA duplexes. The preferred binding direction between the modified PNA and DNA in the duplex motifs was antiparallel, as previously reported for complexes involving unmodified PNA.

Introduction

Oligonucleotide analogues are promising candidates as both antisense and antigene drugs, in diagnostics, and as biological tools.[2,3] Analogues with improved affinity and specificity toward complementary oligonucleotides are particularly interesting for such purposes. When DNA analogues are to be used as drugs, a number of issues have to be taken into consideration, e.g., their cellular uptake, biological stability, and the RNase H susceptibility of their hybrids with RNA.[2] Furthermore, ease of synthesis is important for their usefulness.

This has led to the design and synthesis of a wide variety of oligonucleotide analogues.[2,4] The majority of these are only slightly modified relative to natural oligonucleotides, and few attempts have been successful to modify radically the backbone of DNA.

We have recently prepared reagents where the backbone of DNA is replaced by repeating *N*-(2-aminoethyl)glycine units with the nucleobases attached through methylenecarbonyl linkers (**1**, Figure 1). These oligonucleotide analogues, named peptide nucleic acids (PNAs), have retained the hybridization properties of DNA[5-8] and show very high biological stability.[9] They bind to complementary DNA and RNA with surprisingly high affinity, owing partly to their lack of negative charge and presumably to the proper interbase distances, the rigid amido bonds, the high flexibility of the aminoethyl linkers, and eventually intramolecular hydrogen bonding. PNA oligomers containing both purines and pyrimidines form duplexes with complementary DNA, whereas homopyrimidine PNA oligomers bind to complementary DNA with a $(PNA)_2$/DNA stoichiometry probably mediated through Watson–Crick and Hoogsteen base pairing with formation of a triple helix. This triplex formation contributes further to the stability of the homopyrimidine PNA–DNA complexes.



1-4

Figure 1. Schematic drawing of (**1**, $n = 1$, $m = 1$, $q = 1$) the repeating unit in the initial PNA structure, (**2**, $n = 1$, $m = 2$, $q = 1$) the unit with *N*-(2-aminoethyl)-β-alanine backbone (βB), (**3**, $n = 2$, $m = 1$, $q = 1$) the unit with the *N*-(3-aminopropyl)glycine backbone (apgB), and (**4**, $n = 1$, $m = 1$, $q = 2$) the propanoic acid unit with the ethylene carbonyl linker to the nucleobase (paB). B = thyminyl or cytosyl.

PNA was designed by computer model building where a proposed backbone structure was fit in with a $(dT)_{10}(dA)_{10}(dT)_{10}$ triplex in place of the Hoogsteen strand backbone.[10] The number of bonds between each base in PNA corresponds to that in DNA

† The H. C. Ørsted Institute.
‡ The Panum Institute.
§ Chalmers University of Technology.
⊗ Abstract published in *Advance ACS Abstracts*, August 1, 1994.

(1) Abbreviations (standard oligopeptide and oligonucleotide nomenclature is used): apg-, the 3-aminopropylglycine PNA unit; B, nucleobase; Boc, *tert*-butoxycarbonyl; β-, the β-alanine PNA unit; DCC, *N,N'*-dicyclohexylcarbodiimide; DCU, *N,N'*-dicyclohexylurea; DhbtOH, 2,3-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine; DIC, *N,N'*-diisopropylcarbodiimide; DMF, *N,N*-dimethylformamide; DMSO, dimethyl sulfoxide; H-, deprotected terminal amino group; MBHA, methylbenzhydrylamine; $-NH_2$, C-terminal amido group; pa-, the propanoic acid PNA unit; Pfp, pentafluorophenyl; THF, tetrahydrofuran; Z, benzyloxycarbonyl.

(2) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544–584.
(3) Crook, S. T. *Curr. Opin. Biotechnol.* **1992**, *3*, 656–661.
(4) English, U.; Gauss, D. H. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 613–722.
(5) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.
(6) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895–1897.
(7) Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677–9678.
(8) Egholm, M.; Christensen, L.; Behrens, C.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 800–801.
(9) Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchanrd, O.; Sönnichsen, S. H.; Nielsen, P. E. In preparation.
(10) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. In *Antisense Research and Applications*; Crooke, S. T., Lebleu, B., Eds.; CRC Press: Boca Raton, FL, 1992; pp 363–372.

0002-7863/94/1516-7964$04.50/0

Table 1

entry no.	PNA	DNA	T_m/°C[a] X = T (1)	X = C (1)	X = βT (2)	X = βC (2)	X = apgT (3)	X = paT (4)
1	H-T_4(X)T_5-Lys-NH2	d(A_{10})	*72*	53	*59*	45	*61*	*54*
2	-do-	d(A_4CA_5)	62	47	50	36	51	48
3	-do-	d(A_4TA_5)	62	43	50	34	50	47
4	-do-	d(A_4GA_5)	60	*74*	48	*61*	49	46

[a] The melting temperatures of the hybrids were determined as previously described.[7] The solutions were 10 mM in phosphate, 100 mM in NaCl, 0.1 mM in EDTA, and pH was 7.0.

and therefore suggests this to be the optimal length. However, the model building did not allow for direct quantitative comparison between different structures, and in order to compare the properties of the initial PNA with those of closely related compounds, we wanted to extend the backbone[11] or the linker to the nucleobase (**4**, Figure 1) by one methylene group, since this should not much change such properties of PNA as water solubility, achirality, rigidity around the amido bonds, or polarity. The backbone offers two positions for insertion of a methylene group, namely, the 2-aminoethyl part (**3**, Figure 1) and the glycine part (**2**, Figure 1). We here report the synthesis of these modified PNA units, their oligomerization, and the thermal stability of hybrids between the modified PNA oligomers and DNA.

Results and Discussion

Synthesis of Modified PNA Monomers. The syntheses of the thymine- and cytosine-β-alanine analogues, the thymine-(3-aminopropyl)glycine analogue, and the thymine- and cytosine-propanoic acid analogues are outlined in Scheme 1.

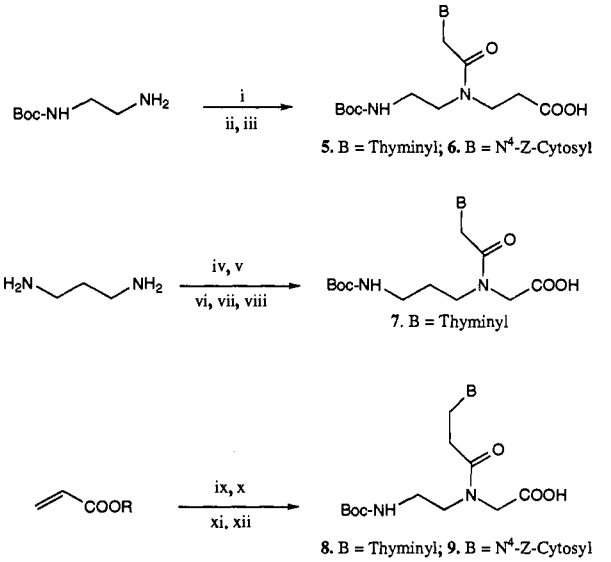
Thermal Stability of Hybrids between Modified PNA Oligomers and DNA. The sequence in the PNA oligomers including modified units was chosen as shown in Tables 1, 2, and 4 for the following reasons: homopyrimidine PNA oligomers were prepared in order to reveal the binding of the modified PNA units in a triple helix mode, whereas mixed-sequence PNA oligomers were expected to show the binding in a duplex mode. Furthermore, we wanted to examine the effect of incorporation of a single modified unit and the binding behavior of oligomers containing only modified units. The single modification was placed in the middle of the oligomers to avoid "end effects". A lysine amide was included at the C-terminus for comparison with the previously reported PNA oligomers. The PNA–DNA binding was examined by T_m measurements (Tables 1–4), which resulted in well-defined melting curves.

(a) Triple Helix Motifs. Single modifications in homopyrimidine motifs were represented by the PNA oligomers H-T_4-(X)T_5-Lys-NH_2 (X = βT (**2**, B = T), βC (**2**, B = C), apgT (**3**, B = T), or paT (**4**, B = T); cf. Figure 1), which were hybridized to either the fully complementary oligodeoxynucleotide or oligodeoxynucleotides with a single mismatch opposite the modified unit. Values for the unmodified PNA oligomers H-T_{10}-Lys-NH_2 and H-T_4CT_5-Lys-NH_2 hybridized to DNA oligomers are included for comparison in Table 1.

Introduction of a single unit with extended backbone in a PNA decamer (Table 1, row 1, X = βT or apgT; row 4, X = βC) caused a decrease in T_m by 13, 11, and 13 °C, respectively, relative to the corresponding unmodified PNA–DNA hybrids. These decreases in T_m are similar to those observed when a single mismatch is introduced in an unmodified PNA–DNA hybrid (Table 1, rows 2–4, X = T; rows 1–3, X = C). However, the sequence specificity is retained, as demonstrated by a further decrease in T_m, when a mismatch is incorporated in the oligodeoxynucleotide opposite the modified unit (Table 1, rows 2–4, X = βT and apgT; rows 1–3, X = βC). This indicates that the modified units in fact recognize the complementary DNA bases. The decrease in stability for the backbone-modified PNA–DNA hybrids, compared to that of the unmodified hybrids, is ascribed to geometric constraints in the PNA and/or a larger loss in entropy upon complex formation. It is interesting to note that the differences in T_m between matched and mismatched *unmodified* $(PNA)_2$/DNA hybrids are comparable to the differences in T_m between matched and mismatched *modified* hybrids (Table

(11) Hyrup, B.; Egholm, M.; Rolland, M.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 518–519.

Scheme 1. Synthesis of PNA Monomers with Extended Backbone or Extended Linker to the Nucleobase[a]



[a] **5**, βT; **6**, βC; **7**, apgT; **8**, paT; **9**, paC. (i) $CH_2CHCOOCH_3$ in CH_3CN, reflux 20 h. (ii) $BCH_2COOPfp$, Et_3N in DMF, 20 h. (iii) Aqueous NaOH, 10 min. (iv) $ClCH_2COOH$. (v) MeOH, HCl. (vi) *p*-NO_2-C_6H_4O-Boc, H_2O/dioxane, pH 10. (vii) BCH_2COOH, DhbtOH, DCC in DMF/CH_2Cl_2. (viii) NaOH, MeOH, 1 h. (ix) Thymine in MeOH, (R = CH_3), catalyst NaOH, reflux 45 h; or N^4-(*Z*)-cytosine in DMF (R = C_2H_5), NaH, 20 h. (x) 2 M NaOH followed by HCl (aqueous). (xi) $BocNHCH_2CH_2NHCH_2COOC_2H_5$ in DMF/CH_2Cl_2, DhbtOH, DCC, 3 h. (xii) 2 M NaOH in MeOH, 1 h (B = thyminyl); or 1 M LiOH in THF, 45 min (B = N^4-(*Z*)-cytosyl).

Table 2. T_m/°C[a] for Hybrids between DNA and PNA Oligomers Consisting Exclusively of Modified Units

entry no.	DNA	PNA H-(paT)$_{10}$-Lys-NH_2	PNA H-(paT)$_4$(paC)-(paT)$_5$-Lys-NH_2
1	d(A_{10})	22	<10
2	d(A_4CA_5)	12	<10
3	d(A_4TA_5)	15	<10
4	d(A_4GA_5)	13	25

[a] See note to Table 1.

Table 3. pH Dependence of T_m/°C[a] Values for PNA/DNA Hybrids

PNA	DNA	pH 5.0	pH 7.0	pH 9.0
H-T_4(βC)T_5-Lys-NH_2	d(A_4GA_5)	67	61	58

[a] See note to Table 1.

Table 4. $T_m \pm 3/°C^a$ for Hybrids between DNA and PNA Oligomers with the Sequence H-GTA-GA(X)-CAC T-Lys-NH_2

entry no.	DNA		PNA X = T	PNA X = βT	PNA X = apgT	PNA X = paT
1	3′-dAGTG-A-TCTAC-5′	parallel	39	28	33	23
2	5′-dAGTG-A-TCTAC-3′	antiparallel	50	40	42	29
3	5′-dAGTG-G-TCTAC-3′		31	22	29	20
4	5′-dAGTG-C-TCTAC-3′			19	24	21
5	H-AGTGATCTAC-LysNH_2	antiparallel	68	60	61	54

[a] The melting temperatures of the hybrids were obtained by CD measurements as described in the text. The CD spectra were recorded on a Jasco 720 spectropolarimeter using a 1-cm quartz cell. The concentration of the oligomers was the same as in the UV measurements,[7] and the solutions were 10 mM in phosphate, 0.1 mM in EDTA, and pH was 7.0.

1, compare X = T with X = βT and X = proT, and X = C with X = βC), and we propose that related changes in conformation take place when these complexes have to accommodate a mismatch.

Results obtained with the PNA oligomer H-T_4(paT)T_5-Lys-NH_2 containing a single unit with extended linker between the backbone and the nucleobase showed the same pattern of hybrid stability as described above (Table 1, rows 1–4, X = paT). The melting temperature is decreased by 18 °C compared to that of the PNA–DNA hybrid involving unmodified PNA, and the T_m decreases further upon incorporation of noncomplementary bases in the DNA opposite the modified unit. The sequence specificity is slightly lower than that of the corresponding reagents containing extended backbone units, as can be seen from the smaller decreases in T_m for mismatched hybrids relative to matched complexes (Table 1, compare X = paT with X = βT or X = apgT).

Considering the relatively large decrease in hybrid stability upon incorporation of a modified unit in the PNA, DNA hybrids with PNA oligomers built up exclusively of these modified units would not be predicted to be particularly stable (unless the decreased stability is due to structural incompatibilities between the (aminoethyl)glycine and the methylene extended backbones). In fact, H-$(\beta T)_{10}$-Lys-NH_2 and H-$(apgT)_{10}$-Lys-NH_2 showed no hypochromicity when mixed with complementary DNA, indicating the absence of stable complexes. Interestingly, the hybrid between H-$(paT)_{10}$-Lys-NH_2 (with extended linker between backbone and all nucleobases) and $(dA)_{10}$ showed a well-defined melting curve, with a T_m of 22 °C (Table 2, row 1). When hybridized to sequences with a single mismatch, the H-$(paT)_{10}$-Lys-NH_2 had a T_m of about 14 °C (Table 2, rows 2–4), indicating that sequence specificity is preserved. H-$(paT)_4$-(paC)$(paT)_5$-Lys-NH_2 hybridized to complementary DNA melts at 25 °C, whereas hybrids with noncomplementary DNA oligomers had T_m values too low to be measured (<10 °C). Since the oligomers with extended linker between the backbone and the nucleobases in principle should allow for base pairing and base stacking as in unmodified PNA, we believe that the fall in stability at least partly can be ascribed to the higher flexibility of the linker to the nucleobases and a less efficient water exclusion from the helix. The oligomers extended in all backbone units did not form stable complexes with complementary DNA, and these results indicate that it is of particular importance to maintain the correct distance between the bases in order to preserve base pairing and allowing a regular backbone geometry, whereas a change in length between the backbone and the nucleobases is more easily accommodated.

In the light of this conclusion, it may seem contradictory that a single backbone-to-base extended paT unit in a PNA oligomer is more decremental to the PNA–DNA complex stability than a single backbone extended βT or apgT unit (Tables 1 and 4). However, in this case, the structural constraints are imposed by the original *N*-(2-aminoethyl)glycine backbone, and in this structural context it is apparently less costly to accommodate an extra methylene group within the backbone than between the backbone and the nucleobase. Simple inspection of a molecular model also supports this notion.

(b) Stoichiometry and pH Dependence of PNA/DNA Hybrids. As expected, the presently reported homopyrimidine PNA oligomers with a single modified unit bind to complementary DNA with a $(PNA)_2$/DNA stoichiometry as determined by UV and CD titrations[7,12] (data not shown), and we propose a triple helix structure for these complexes involving both Watson–Crick and Hoogsteen base pairing. Hoogsteen base pairing requires a protonated cytosine, and consequently the T_m of hybrids containing cytosine should be sensitive to changes in pH. As shown in Table 3, this is the case for the H-$T_4(\beta C)T_5$-Lys-NH_2 oligomer hybridized to its complementary DNA oligomer. At pH 9, T_m is 58 °C, whereas at pH 5 it increases to 67 °C.

(c) Double Helix Motifs. Mixed-sequence PNA oligomers of the form H-GTAGA(X)CACT-Lys-NH_2 (X = βT, apgT, or paT) were prepared and hybridized to either the fully complementary oligodeoxynucleotides or oligodeoxynucleotides with a single mismatch opposite the modified PNA unit. When the UV melting curves were recorded, these were not of well-defined sigmoid shape, and unambiguous T_m values could not be determined. This seemed to be caused by a thermal transition of the PNA oligomers themselves since mixed-sequence PNA decamers typically exhibit T_m values around 41 °C. The melting temperatures of the modified PNA/DNA hybrids lie close to or below this value, giving rise to overlaying transitions in this temperature range. To circumvent this problem, we measured the circular dichroism spectra for the hybrids in the temperature range of 5–70 °C in steps of approximately 5 °C and plotted $\Delta\epsilon$ as a function of temperature at appropriate wavelengths (data not shown). As the PNA oligomers are achiral except in the lysine residue, their CD signals are negligible,[12] and well-defined melting curves could be obtained for the PNA/DNA hybrids. The corresponding T_m values are listed in Table 4, and for comparison, values for the unmodified PNA/DNA hybrid are included. Mixed-sequence PNA oligomers all form 1:1 complexes with DNA as determined by UV titrations, presumably mediated through Watson–Crick hydrogen bonding.

The modified PNA units affect the duplex hybrid stabilities in a manner very similar to that observed in the triplex motifs. The PNA oligomers H-GTAGA(βT)CACT-Lys-NH_2 and H-GTAGA(apgT)CACT-Lys-NH_2, containing a single extended backbone unit, melt at 40 and 42 °C, respectively, when hybridized to the complementary antiparallel (amino-terminus of the PNA facing the 3′-end of the oligonucleotide) DNA target (Table 4, row 2, X = βT or apgT). These T_m values are to be compared to the T_m of 50 °C for the corresponding unmodified PNA/DNA hybrid (Table 4, row 2, X = T). In the case of the PNA oligomer with a single unit extended in the linker between the nucleobase and the backbone, T_m is decreased to 29 °C (Table 4, row 2, X = paT). This is—as also observed in the triplex motifs—a more dramatic decrase than seen for the extended backbones. Comparison of rows 1 and 2 in Table 4 reveals that the antiparallel hybrids are more stable than the parallel hybrids (approximately

(12) Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nordén, B. *J. Am. Chem. Soc.* **1993**, *114*, 6477–6481.

1 °C per base pair). This seems to be a general property of duplex PNA/DNA hybrids.[13]

The effect of incorporating a noncomplementary base in the DNA oligomer opposite the modified PNA unit is shown in rows 3 and 4 (Table 4). In all cases, we see a pronounced decrease in T_m, indicating base-specific recognition of the base at the modified PNA unit in spite of the lowered affinity. Again, the sequence specificity is slightly lower for the PNA oligomer with the extended linker between the nucleobase and the backbone than for the PNA oligomers with extended backbones, as was the case in the triplex motifs.

We have recently demonstrated that two complementary, antiparallel PNA oligomers form a helical duplex.[14] In order to see whether a PNA/PNA helix more easily accommodates the modified PNA units than a PNA/DNA helix does, we hybridized the modified PNA oligomers to the complementary antiparallel PNA oligomer (Table 4, row 5). The unmodified PNA/PNA hybrid melts at 68 °C, whereas a single PNA unit with extended backbone causes T_m to decrease to ca. 60 °C (Table 4, row 5, X = βT or apgT). The extended linker between the backbone and the nucleobase lowers T_m to 54 °C (Table 4, row 5, X = paT), and so it consistently disrupts the helix stability more than the extended backbones. The relative decrease in T_m caused by the modified PNA units is slightly lower in the PNA/PNA hybrids than in the PNA/DNA hybrids. This could indicate that the PNA backbone more easily accommodates a change in the complementary strand than the deoxyribosephosphate backbone of DNA does.

We note that the average decrease in thermal stability per backbone modified base pairing unit is less for PNA_2/DNA triplexes (~14 °C per base triplex, corresponding to 7 °C per base pair) as compared to PNA/DNA duplexes (~11 °C per base pair). Provided that equal destabilization can be attributed to the PNA strand involved in Watson–Crick base pairing regardless of duplex or triplex context, these results indicate that Hoogsteen strand PNA binding is somewhat less sensitive to backbone modifications than Watson–Crick PNA binding to DNA.

On the basis of molecular modeling, it has been proposed that an intramolecular hydrogen bond between the carbonyl oxygen of the methylene carbonyl linker and the aminoethyl NH group of the next backbone unit is important for the proper conformation of the backbone upon formation of PNA–DNA complexes.[15,16] The present results argue against any major importance of such a hydrogen bond since it should be retained in the *N*-(3-aminopropyl)glycine backbone and disrupted in the *N*-(2-aminoethyl)-β-alanine backbone, and we find no major difference in the DNA mimicking properties between PNAs containing these backbone units.

Conclusion

The presently reported results emphasize that the original PNAs to date give the most stable hybrids with DNA and that especially the distance between the nucleobases is of great importance for the hybrid stability. Interestingly, it seems to be of minor importance for the stability in which part of the backbone the extra methylene group is placed. Also, the distance between the backbone and the nucleobases seemes to play a major role in the stability of PNA–DNA hybrids. Not surprisingly, these distances have turned out to be identical to those found in naturally occurring DNA. The ability to regulate the binding affinity between PNA and DNA could be an advantage under certain circumstances.

Experimental Section

General Remarks. NMR spectra were recorded on a JEOL FX 90Q spectrometer, a 250-MHz Bruker instrument, or a Varian 400 with tetramethylsilane as internal standard (except when deuterium oxide was the solvent; then a predetermined reference value was used). For mass spectrometry, a MassLab VG 12-250 quadropole instrument fitted with a VG-FAB source and probe or a tandem mass spectrometer JMS-Hx/Hx110A was used, and melting points were recorded on a Büchi melting point apparatus. Di-*tert*-butyl carbonate, ethylenediamine, thymine, pentafluorophenol, *N,N'*-dicyclohexylcarbodiimide, *N,N'*-diisopropylcarbodiimide, DhbtOH, 1,3-propanediamine, sodium hydride, and chloroacetic acid were all obtained from Aldrich and used without further purification. 4-Nitrophenol (Aldrich) was recrystallized prior to use, and methyl and ethyl acrylate (Aldrich) were distilled. MBHA resin and Boc-(2-Cl-Z)lysine were from Novabiochem. Triethylamine (distilled) and acetonitrile were from Riedel-de Häen. Acetonitrile and *N,N*-dimethylformamide (LabScan, HPLC grade) were dried prior to use over 3-Å molecular sieves, and dioxane (Riedel-de Häen) was passed through basic alumina oxide to remove peroxides. Ether was dried over a sodium/lead alloy. DMSO was from Fluka. All other solvents used were HPLC grade from LabScan. N^4-(*Z*)-Cytosine, N^4-(*Z*)-1-(carboxymethyl)cytosine, and (carboxymethyl)thymine were prepared as described previously.[17] TLC plates (Silicagel 60 F_{254}) were visualized by UV (254 nm) and/or, after heating at 120 °C for 5 min, sprayed with a ninhydrin solution (3 g of ninhydrin in 1000 mL of 1-butanol and 30 mL of acetic acid) and heated again. Silica gel 60 (particle size 0.04–0.063 mm) was used for column chromatography.

***tert*-Butyl 4-Nitrophenylcarbonate.** 4-Nitrophenol (12.75 g; 0.0916 mol) and sodium carbonate (29.14 g; 0.275 mol) were mixed with dioxane (250 mL). Boc anhydride (20.0 g; 0.0916 mol) was transferred to the mixture with dioxane (50 mL), and the solution refluxed for 1 h. Subsequently the suspension was cooled to 0 °C and filtered, and the filtrate was concentrated to 1/3 and poued into water (350 mL) at 0 °C. After being stirred for 0.5 h, the product was collected by filtration, washed with water, and dried over sicapent *in vacuo*. The compound was stored in a freezer. Yield 21.55 g (98%). Mp 71.0–73.0 °C. Anal. for $C_{11}H_{13}NO_5$ found (calcd): C, 55.20 (55.23); H, 5.61 (5.48); N, 5.82 (5.85).

Mono-Boc-ethylenediamine. *tert*-Butyl 4-nitrophenylcarbonate (10.0 g; 0.0418 mol) was dissolved in DMF (50 mL), and the solution added dropwise over a period of 30 min to a mixture of ethylenediamine (27.9 mL; 0.418 mol) and DMF (50 mL). The solution was stirred overnight and evaporated to dryness *in vacuo*, and the resulting oil was dissolved in water (250 mL). pH was adjusted to 3.5 at 0 °C with 4 M hydrochloric acid, after which the solution was filtered and extracted with chloroform (3 × 250 mL). pH was subsequently adjusted to 12 at 0 °C with 2 M sodium hydroxide, and the aqueous solution was extracted with methylene chloride (3 × 300 mL). After treatment with saturated aqueous sodium chloride (250 mL), the methylene chloride solution was dried over magnesium sulfate. The solution was filtered and evaporated to dryness *in vacuo*, resulting in 4.22 g (63%) of the product (oil). ^{1}H NMR (90 MHz; $CDCl_3$): δ 1.44 (s, 9H, *t*-Bu); 2.87 (t, 2H, CH_2NH_2); 3.1 (q, 2H, BocNHCH_2-); 5.62 (b, BocNH-).

***N*-(2-((Boc)amino)ethyl)-β-alanine Methyl Ester Hydrochloride.** Mono-Boc-ethylenediamine (16.28 g; 0.102 mol) was dissolved in acetonitrile (400 mL), and methyl acrylate (91.50 mL; 1.02 mol) was transferred to the mixture with acetonitrile (200 mL). The solution was refluxed overnight under nitrogen in the dark (to avoid polymerization of methyl acrylate). After evaporation of the solution to dryness *in vacuo*, water and ether (200 + 200 mL) was added, the solution was stirred vigorously, and subsequently it was filtered. The aqueous phase was extracted one more time with ether and freeze-dried to yield a yellow solid. Recrystallization from ethyl acetate yielded 13.09 g (46%) of the title compound. Mp 138–140 °C. Anal. for $C_{11}H_{23}N_2O_4Cl$ found (calcd): C, 46.49 (46.72); H, 8.38 (8.20); N, 9.83 (9.91); Cl, 12.45 (12.54). ^{1}H NMR (90 MHz; DMSO-d_6): δ 1.39 (s, 9H, *t*-Bu); 2.81–3.23 (unresolved m, 8H, 2 × -CH_2CH_2-); 3.64 (s, 3H, OCH_3); 7.0 (b, 1H, -CH_2NHCH_2-); 9.33 (b, 1H, BocN*H*-).

(13) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordén, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566–568.

(14) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Nordén, B. *Nature* **1994**, *368*, 561–563.

(15) Almarsson, Ö.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9542–9546.

(16) Almarsson, Ö.; Bruice, T. C.; Kerr, J.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7518–7522.

(17) Dueholm, K.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. Submitted for publication in *J. Org. Chem.*

1-(Carboxymethyl)thymine Pentafluorophenyl Ester. 1-(Carboxymethyl)thymine (5.0 g; 0.027 mol) was dissolved in DMF (20 mL), and methylene chloride (25 mL) was added. Pentafluorophenol (5.0 g; 0.027 mol) was transferred to the mixture with methylene chloride (10 mL). DCC (6.73 g; 0.0326 mol) was added at 0 °C, and after 5 min the ice/water bath was removed. After 3.5 h at room temperature, DCU was removed by filtration and carefully washed successively with methylene chloride (50 mL) and DMF (4 × 50 mL). The filtrate was evaporated to dryness *in vacuo*, yielding a white solid which was recrystallized from 2-propanol. The product was washed three times with petroleum ether and dried over sicapent *in vacuo*. Yield 7.06 g (75%). Mp 192–196 °C. ^{1}H NMR (90 MHz; DMSO-d_6): δ 1.77 (s, 3H, T-CH_3); 4.99 (s, 2H, T-CH_2-); 7.64 (s, 1H, CH in T); 11.51 (s, 1H, NH in T).

***N*-(1-Thyminylacetyl)-*N*-(2-((Boc)amino)ethyl)-β-alanine Methyl Ester.** *N*-(2-((Boc)amino)ethyl)-β-alanine methyl ester hydrochloride (2.0 g; 0.0071 mol) and 1-(carboxymethyl)thymine pentafluorophenyl ester (2.828 g; 0.00812 mol) were dissolved in DMF (50 mL). Triethylamine (1.12 mL; 0.00812 mol) was added, and the mixture was stirred overnight. After addition of methylene chloride (200 mL), the organic phase was extracted with aqueous sodium hydrogen carbonate (3 × 250 mL), half-saturated aqueous potassium hydrogen sulfate (3 × 250 mL), and saturated aqueous sodium chloride (250 mL) and dried over magnesium sulfate. Filtration and evaporation to dryness *in vacuo* resulted in 2.9 g (99%) yield of the title compound (oil). ^{1}H NMR (250 MHz; $CDCl_3$): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.43 (s, 9H, *t*-Bu); 1.88 (s, 3H, T-CH_3); 2.63 (t) and 2.74 (t) (2H, -CH_2CH_2-); 3.25–3.55 (4 × t, 4H, -CH_2CH_2-); 3.65 (2 × t, 2H, -CH_2CH_2-); 3.66 (s) and 3.72 (s) (3H, OCH_3); 4.61 (s) and 4.72 (s) (2H, T-CH_2-); 5.59 (s) and 5.96 (s) (1H, BocNH-); 7.11 (s, 1H, CH in T); 10.33 (s, 1H, NH in T).

***N*-(1-Thyminylacetyl)-*N*-(2-((Boc)amino)ethyl)-β-alanine.** The methyl ester of the title compound (3.0 g; 0.0073 mol) was dissolved in 2 M sodium hydroxide (30 mL). After 15 min, pH was adjusted to 2 at 0 °C with 4 M hydrochloric acid, and the solution was stirred for 2 h. The precipitate was isolated by filtration, washed three times with cold water, and dried over sicapent *in vacuo*. Yield 2.23 g (77%). Mp 170 °C dec. Anal. for $C_{17}H_{26}N_4O_7 \cdot H_2O$, found (calcd): C, 49.49 (49.03); H, 6.31 (6.78); N, 13.84 (13.45). ^{1}H NMR (90 MHz; DMSO-d_6): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.38 (s, 9H, *t*-Bu); 1.76 (s, 3H, T-CH_3); 2.44 and 3.29 (m, 8H, -CH_2-CH_2-); 4.55 (d, 2H, T-CH_2-); 7.3 (s, 1H, CH in T); 11.23 (s, 1H, NH in T). FAB-MS *m/z* 399 (M + 1).

N^4-(*Z*)-1-(Carboxymethyl)cytosine Pentafluorophenyl Ester. N^4-(*Z*)-1-(Carboxymethyl)cytosine (4.00 g; 0.0132 mol) and pentafluorophenol (2.67 g; 0.0145 mol) were mixed with DMF (70 mL) at 0 °C, and DCC (3.27 g; 0.0145 mol) was added. The ice bath was removed after 3 min, and the mixture was stirred for 3 h at room temperature. The precipitated DCU was removed by filtration and washed with DMF, and the filtrate was evaporated to dryness *in vacuo*. The solid residue was treated with methylene chloride (250 mL), stirred vigorously for 15 min, and filtered, and the solution was subsequently washed twice with half-saturated aqueous sodium hydrogen carbonate and once with saturated aqueous sodium chloride, dried over magnesium sulfate, and evaporated to dryness *in vacuo*. The solid residue was recrystallized from 2-propanol (150 mL), and the crystals were washed thoroughly with ether. Yield 3.40 g (55%). Mp 241–245 °C. Anal. for $C_{20}H_{12}N_3F_5O_5$, found (calcd): C, 51.56 (51.18); H, 2.77 (2.58); N, 9.24 (8.95). ^{1}H NMR (90 MHz; $CDCl_3$): δ 4.97 (s, 2H, T-CH_2-); 5.21 (s, 2H, $C_6H_5CH_2$-); 7.31 (d, 1H, C^5H in C); 7.37 (s, 5H, C_6H_5); 7.66 (d, 1H, C^6H in C). FAB-MS *m/z* 470 (M + 1).

***N*-(N^4-(*Z*)-1-Cytosylacetyl)-*N*-(2-((Boc)amino)ethyl)-β-alanine Methyl Ester.** *N*-(2-((Boc)amino)ethyl)-β-alanine methyl ester hydrochloride (2.0 g; 0.0071 mol) and N^4-(*Z*)-1-(carboxymethyl)cytosine pentafluorophenyl ester (3.319 g; 0.0071 mol) were dissolved in DMF (50 mL). Triethylamine (0.99 mL; 0.0071 mol) was added, and the mixture was stirred overnight. After addition of methylene chloride (200 mL), the organic phase was extracted with aqueous sodium hydrogen carbonate (3 × 250 mL), half-saturated aqueous potassium hydrogen sulfate (3 × 250 mL), and saturated aqueous sodium chloride (250 mL) and dried over magnesium sulfate. Filtration and evaporation to dryness *in vacuo* resulted in 3.36 g of a solid compound, which was recrystallized from methanol. Yield 2.42 g (64%). Mp 158–161 °C. Anal. for $C_{25}H_{33}N_5O_8$, found (calcd): C, 55.19 (56.49); H, 6.19 (6.26); N, 12.86 (13.18). ^{1}H NMR (250 MHz; $CDCl_3$): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.41 (s, 9H, *t*-Bu); 2.57 (t) and 2.71 (t) (2H, unresolved -CH_2CH_2-); 3.23–3.58 (m, 6H, unresolved -CH_2CH_2-); 3.63 (s) and 3.68 (s) (3H, OCH_3); 4.82 (s) and 4.87 (s) (2H, CCH_2-); 5.16 (b, 1H, BocN*H*-); 5.21 (s, 2H, $C_6H_5CH_2$-); 7.23–7.75 (unresolved m, 7H). FAB-MS *m/z* 532 (M + 1); 476 (M − *t*-Bu + 1); 432 (M − Boc + 1).

***N*-(N^4-(*Z*)-1-Cytosylacetyl)-*N*-(2-((Boc)amino)ethyl)-β-alanine.** *N*-(N^4-(*Z*)-1-Cytosylacetyl)-*N*-(2-((Boc)amino)ethyl)-β-alanine methyl ester (0.621 g; 0.0012 mol) was dissolved in 2 M sodium hydroxide (8.5 mL) and stirred for 2 h. Subsequently pH was adjusted to 2 at 0 °C with 4 M hydrochloric acid, and the solution was stirred for 2 h. The precipitate was isolated by filtration, washed three times with cold water, and dried over sicapent *in vacuo*. Yield 0.326 g (54%). The white solid was recrystallized from 2-propanol and washed with petroleum ether. Mp 163 °C dec. Anal. for $C_{24}H_{31}N_5O_8 \cdot H_2O$, found (calcd): C, 53.88 (53.82); H, 5.82 (6.21); N, 12.73 (13.08). ^{1}H NMR (250 MHz; $CDCl_3$): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.39 (s, 9H, *t*-Bu); 2.57 (t) and 2.65 (t) (2H, unresolved -CH_2CH_2-); 3.18–3.63 (m, 6H, unresolved -CH_2CH_2-); 4.84 (s) and 4.98 (s) (2H, CCH_2-); 5.19 (s, 2H, $C_6H_5CH_2$-); 5.71 (b, 1H, BocN*H*-); 7.25–7.98 (unresolved m, 7H). FAB-MS *m/z* 518 (M + 1); 462 (M − *t*-Bu + 1); 418 (M − Boc + 1).

***N*-(3-Aminopropyl)glycine Hydrochloride.** Chloroacetic acid (73.05 g; 0.773 mol) was added portionwise over a period of 2 h to 1,3-propanediamine (500 mL; 5.99 mol) in a round-bottomed flask equipped with a mechanical stirrer. The reaction mixture was cooled to about 10 °C during the addition and then left overnight at room temperature. The semisolid obtained by evaporation to dryness *in vacuo* was triturated with DMSO, filtered, washed with ether, and dried over sicapent *in vacuo*. The isolated solid was subsequently dissolved in boiling acetic acid (400 mL), and the title compound was precipitated with ethanol (1500 mL) with stirring at 0 °C for 2 h. The product was filtered off, washed three times with ethanol, and dried over sicapent *in vacuo*. Yield 55.51 g (43%). Mp 203–230 °C. Anal. for $C_5H_{13}N_2O_2Cl$, found (calcd): C, 34.49 (35.60); H, 7.64 (7.72); N, 16.17 (16.62). ^{1}H NMR (90 MHz; D_2O): δ 2.14 (quintet, 2H, -$CH_2CH_2CH_2$-); 2.98–3.22 (2 × t, 4H, -CH_2-CH_2CH_2-); 3.59 (s, 2H, -$NHCH_2COOH$). MS-FAB *m/z* 133 (M + 1); 265 (2M + 1); 398 (3M + 1); 530 (4M + 1).

***N*-(3-Aminopropyl)glycine Methyl Ester Hydrochloride.** *N*-(3-Aminopropyl)glycine hydrochloride (7.149 g; 0.0424 mol) was suspended in methanol (600 mL), and hydrogen chloride gas was bubbled through the mixture at 0 °C during 15 min. After 1 h of reflux, the solution was evaporated to dryness *in vacuo*, and 8.74 g (94%) of the product was isolated as a solid. ^{1}H NMR (250 MHz; D_2O): δ 2.09 (quintet), 2H, -$CH_2CH_2CH_2$-); 3.07 (t, 2H, $ClH_3NCH_2CH_2$-); 3.20 (t, 2H, -CH_2NH_2-$ClCH_2COOCH_3$); 3.80 (s, 3H, OCH_3); 4.05 (s, 2H, -CH_2COOCH_3). ^{13}C NMR: δ 21.8 (-$CH_2CH_2CH_2$-); 34.6 ($ClH_3NCH_2CH_2$-); 42.5 (-CH_2-$NH_2ClCH_2COOCH_3$); 45.5 (-CH_2COOCH_3); 51.7 (OCH_3).

***N*-(3-((Boc)amino)propyl)glycine Methyl Ester.** *N*-(3-Aminopropyl)-glycine methyl ester hydrochloride (1.0 g, 0.0046 mol) was dissolved in water (7.5 mL), and dioxane (7.5 mL) was added. pH was adjusted to 6 with 2 M sodium hydroxide, and *tert*-butyl 4-nitrophenylcarbonate (1.31 g; 0.00548 mol) was transferred to the mixture with dioxane (5 mL). pH was then kept at 10 until the reaction was completed (4 h; indicated by disappearance of the starting material on TLC, eluted with $CHCl_3$, Et_3N, MeOH 7:1:2). After the mixture was cooled to 0 °C, pH was adjusted to 4 using 4 M hydrochloric acid, and the solution was extracted three times with chloroform. Subsequently, pH of the aqueous phase was adjusted to 12 with 2 M sodium hydroxide and extracted three times with ethyl acetate. At last the ethyl acetate phase was washed several times with saturated, aqueous sodium chloride, dried over magnesium sulfate, filtered, and evaporated to dryness *in vacuo*. Yield 0.44 g (39%) oil. ^{1}H NMR (90 MHz; $CDCl_3$): δ 1.4 (s, 9H, *t*-Bu); 1.7 (quintet; 2H, -$CH_2CH_2CH_2$-); 2.7 (t, 2H, -$CH_2NH_2ClCH_2COOCH_3$); 3.2 (q, 2H, BocNHCH_2-); 3.4 (s, 2H, -$NHCH_2COOCH3$); 3.7 (s, 3H, OCH_3). MS (EI) *m/z* 246 (M^+).

***N*-(1-Thyminylacetyl)-*N*-(3-((Boc)amino)propyl)glycine Methyl Ester.** *N*-(3-((Boc)amino)propyl)glycine methyl ester (2.84 g; 0.0115 mol) was dissolved in DMF (35 mL), followed by addition of DhbtOH (2.07 g; 0.0127 mol) and 1-(carboxymethyl)thymine (2.34 g; 0.0127 mol). Subsequently methylene chloride (35 mL) was added, and the mixture was cooled to 0 °C on an ice bath. After addition of DCC (2.85 g; 0.0138 mol), the mixture was stirred at 0 °C for 2 h, followed by 1 h at room temperature. The precipitated DCU was removed by filtration and washed with methylene chloride (25 mL), and a further amount of methylene chloride (150 mL) was added to the filtrate. The organic phase was extracted with half-saturated aqueous sodium hydrogen carbonate (6 × 250 mL), aqueous potassium sulfate (1 vol saturated diluted with 4 vol

water, 3 × 250 mL), saturated aqueous sodium chloride (1 × 250 mL), dried over magnesium sulfate, and evaporated to dryness *in vacuo*. The solid residue was suspended in methylene chloride (35 mL) and stirred for 1 h, and the undissolved DCU was removed by filtration and washed with methylene chloride (25 mL). The filtrate was evaporated to dryness *in vacuo*, and the residue was purified by column chromatography on silica gel, eluting with a mixture of methanol and methylene chloride (gradient from 3 to 7% methanol in methylene chloride). This afforded the title compound as a white solid (3.05 g, 64%). Mp 76–79 °C dec. Anal. for $C_{18}H_{28}N_4O_7$, found (calcd): C, 52.03 (52.42); H, 6.90 (6.84); N, 13.21 (13.58). ^{1}H NMR (90 MHz; $CDCl_3$): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.41 (s) and 1.42 (s) (9H, *t*-Bu); 1.64 (quintet) and 1.86 (quintet) (2H, -$CH_2CH_2CH_2$-); 1.89 (s) and 1.91 (s) (3H, T-CH_3); 3.07 (q) and 3.20 (q) (2H, BocNHCH_2-); 3.45 (t, 2H, -CH_2NCH_2COOCH_3); 3.72 (s) and 3.80 (s) (3H, OCH_3); 4.06 (s) and 4.17 (s) (2H, -CH_2COOCH_3); 4.40 (s) and 4.59 (s) (2H, T-CH_2-); 5.11 (b, 1H, BocN*H*-); 7.04 (s) and 7.05 (s) (1H, CH in T); 8.92 (s) and 8.96 (s) (1H, NH in T). MS-FAB *m/z* 413 (M + 1); 313 (M – Boc + 1).

***N*-(1-Thyminylacetyl)-*N*-(3-((Boc)amino)propyl)glycine.** *N*-(1-Thyminylacetyl)-*N*-(3-((Boc)amino)propyl)glycine methyl ester (3.02 g; 0.00732 mol) was dissolved in methanol (25 mL) and stirred for 1.5 h with 2 M sodium hydroxide (25 mL). The methanol was removed by evaporation *in vacuo*, and pH was adjusted to 2 with 4 M hydrochloric acid at 0 °C. The product was isolated as white crystals by filtration, washed with water (3 × 10 mL), and dried over sicapent *in vacuo*. Yield 2.19 g (75%). Anal. for $C_{17}H_{26}N_4O_7$, H_2O, found (calcd): C, 49.95 (49.03); H, 6.47 (6.29); N, 13.43 (13.45). ^{1}H NMR (90 MHz; DMSO-d_6): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.48 (s, 9H, t-Bu); 1.61 (quintet) and 1.78 (quintet) (2H, -$CH_2CH_2CH_2$-); 1.84 (s, 3H, CH_3 in T); 2.98 (q) and 3.08 (q) (2H, BocNHCH_2-), 3.35 (t) and 3.41 (t) (2H, -CH_2NCH_2COOH); 4.03 (s) and 4.25 (s) (2H, -CH_2COOH); 4.52 (s) and 4.69 (s) (2H, T-CH_2-); 6.79 (t, b) and 6.97 (t, b) (1H, BocN*H*-); 7.38 (s) and 7.49 (s) (1H, CH in T); 11.32 (s) and 11.37 (s) (1H, NH in T). MS-FAB *m/z* 399 (M + 1); 299 (M – Boc + 1).

1-(Carboxyethyl)thymine Methyl Ester. To a suspension of thymine (14.00 g; 0.110 mol) in methanol (200 mL) was added methyl acrylate (39.6 mL; 0.44 mol) and catalytic amounts of sodium hydroxide. The suspension was refluxed in the dark for 45 h and evaporated to dryness *in vacuo*, and the residue was dissolved in methanol (8 mL) under heating. After the solution was cooled on an ice bath, the product was precipitated by addition of ether (20 mL), isolated by filtration, washed with ether (3 × 15 mL) and dried over sicapent *in vacuo*. Yield 11.23 g (48%). Mp 112–119 °C. Anal. for $C_9H_{12}N_2O_4$, found (calcd): C, 51.14 (50.94); H, 5.78 (5.70); N, 11.52 (13.20). ^{1}H NMR (250 MHz; $CDCl_3$): δ 1.85 (s, 3H, T-CH_3); 2.73 (t, 2H, -CH_2COOCH_3); 3.65 (s, 3H, OCH_3); 3.92 (t, 2H, T-CH_2-); 7.16 (s, 1H, CH in T); 9.86 (b, 1H, NH in T). MS-FAB *m/z* 213 (M + 1).

1-(Carboxyethyl)thymine. 1-(Carboxyethyl)thymine methyl ester (5.0 g, 0.024 mol) was suspended in 2 M sodium hydroxide (50 mL) and boiled for 2 h, and pH was subsequently adjusted to 0 with concentrated hydrochloric acid. The solution was extracted with ethyl acetate (10 × 50 mL), and the organic phase was extracted with saturated aqueous sodium chloride, dried over magnesium sulfate, and evaporated to dryness *in vacuo* to give the title compound as a white solid (2.32 g, 50%). Mp 118–121 °C. Anal. for $C_8H_{10}N_2O_4$, found (calcd): C, 48.38 (48.49); H, 5.09 (5.09); N, 13.93 (14.14). ^{1}H NMR (250 MHz; DMSO-d_6): δ 1.81 (s, 3H, T-CH_3); 2.67 (t, 2H, -CH_2COOH); 3.89 (t, 2H, T-CH_2-); 7.57 (s, 1H, CH in T); 11.29 (s, 1H, NH in T); 12.47 (b, 1H, COOH). MS-FAB *m/z* 199 (M + 1).

***N*-(1-Thyminyl-3-propanoyl)-*N*-(2-((Boc)amino)ethyl)glycine Ethyl Ester.** *N*-(2-((Boc)amino)ethyl)glycine ethyl ester (1.0 g; 0.0041 mol) was dissolved in DMF (12 mL), followed by addition of DhbtOH (0.73 g; 0.0045 mol) and 1-(carboxyethyl)thymine (0.89 g; 0.0045 mol). Subsequently, methylene chloride (12 mL) was added, and the mixture was cooled to 0 °C on an ice bath. After addition of DCC (1.01 g; 0.0049 mol), the mixture was stirred at 0 °C for 2 h, followed by 1 h at room temperature. The precipitated DCU was removed by filtration and washed with methylene chloride (25 mL), and a further amount of methylene chloride (50 mL) was added to the filtrate. The organic phase was extracted with aqueous sodium hydrogen carbonate (1 vol saturated diluted with 1 vol water, 6 × 100 mL), aqueous potassium sulfate (1 vol saturated diluted with 4 vol of water, 3 × 100 mL), and saturated aqueous sodium chloride (1 × 100 mL), dried over magnesium sulfate, and evaporated to dryness *in vacuo*. The solid residue was suspended in methylene chloride (15 mL) and stirred for 1 h, and the undissolved DCU was removed by filtration and washed with methylene chloride. The filtrate was evaporated to dryness *in vacuo*, and the residue was purified by column chromatography on silica gel, eluting with a mixture of methanol and methylene chloride (gradient from 1 to 6% methanol in methylene chloride). This afforded the title compound as a white solid (1.02 g, 59%). Anal. for $C_{19}H_{30}N_4O_7$, found (calcd): C, 53.15 (53.51); H, 6.90 (7.09); N, 12.76 (13.13). ^{1}H NMR (250 MHz; $CDCl_3$): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.29 (t, 3H, -$COOCH_2CH_3$); 1.39 (s, 9H, *t*-Bu); 1.87 (s, 3H, T-CH_3); 2.67 (t) and 2.84 (t) (2H, T-CH_2CH_2-); 3.22 (q, 2H, BocNHCH_2-); 3.44 (t) and 3.48 (t) (2H, BocNHCH_2CH_2-); 3.98 (t, 2H, T-CH_2-); 3.99 (s) and 4.04 (s) (2H, -CH_2COOCH_2CH$_3$); 4.17 (q) and 4.19 (q) (2H, -COOCH_2CH$_3$); 5.07 (t, b) and 5.53 (t, b) (1H, BocN*H*-); 7.28 (s) and 7.37 (s) (1H, CH in T); 9.43 (s) and 9.54 (s) (1H, NH in T). MS-FAB *m/z* 427 (M + 1); 327 (M – Boc + 1).

***N*-(1-Thyminyl-3-propanoyl)-*N*-(2-((Boc)amino)ethyl)glycine.** *N*-(1-Thyminyl-3-propanoyl)-*N*-(2-((Boc)amino)ethyl)glycine ethyl ester (0.83 g; 0.00195 mol) was dissolved in methanol (25 mL), and 2 M sodium hydroxide (25 mL) was added. The solution was stirred for 1 h, and then methanol was removed by evaporation *in vacuo*, and pH was adjusted to 2 with 4 M hydrochloric acid at 0 °C. The product was isolated by filtration, washed with ether (3 × 15 mL), and dried over sicapent *in vacuo*. Yield 0.769 g, (99%). Mp 213 °C dec. Analysis for $C_{17}H_{26}N_4O_7$ was not in agreement with the calculated values due to precipitation of the sodium salt of the free acid before acidification. ^{1}H NMR (250 MHz; DMSO-d_6): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.32 (2 × s, 9H, *t*-Bu); 1.71 (s, 3H, T-CH_3); 2.53 (t) and 2.69 (t) (2H, T-CH_2CH_2-); 2.97 (q) and 3.03 (q) (2H, BocNHCH_2-); 3.23 (t) and 3.29 (t) (2H, BocNHCH_2CH_2-); 3.6 (t, 2H, T-CH_2-); 3.90 (s) and 4.05 (s) (2H, -CH_2COOH); 6.77 (t) and 6.86 (t) (1H, BocN*H*-); 7.50 (s, 1H, CH in T); 11.27 (s) and 11.29 (s) (1H, NH in T). MS-FAB *m/z* 399 (M + 1); 343 (M – *t*-Bu + 1); 299 (M – Boc + 1).

N^4-(*Z*)-1-(Carboxyethyl)cytosine. N^4-(*Z*)-Cytosine (8.0 g; 0.033 mol) and sodium hydride (80% in white oil) (0.23 g; 0.0098 mol) were suspended in DMF (90 mL) and stirred at 0 °C under nitrogen for 2 h. Subsequently, ethyl acrylate (10.59 mL; 0.0977 mol) was added dropwise in the dark over a period of 10 min at 0 °C. After 2 h at 0 °C, the reaction was left overnight at room temperature. Evaporation to dryness *in vacuo*, followed by evaporation *in vacuo* from methylene chloride resulted in an yellow solid, which was washed several times with ether until a white solid was obtained. To a suspension of this solid in water (60 mL) was added 2 M sodium hydroxide (60 mL), and after 30 min, pH was adjusted to 2 with 4 M hydrochloric acid. The title compound was isolated by filtration, washed with cold water (3 × 20 mL), and dried over sicapent *in vacuo*. Yield 8.277 g (79%). Mp 183 °C dec. Analysis for $C_{15}H_{15}N_3O_5$ was not in agreement with the calculated values due to precipitation of the sodium salt of the free acid before acidification. ^{1}H NMR (250 MHz; DMSO-d_6): δ 2.69 (t, 2H, -CH_2CH_2COOH); 3.96 (t, 2H, CCH_2CH_2-COOH); 5.19 (s, 2H, $C_6H_5CH_2$-); 6.96 (d, 1H, C^5*H* in C); 7.41 (m, 5H, $C_6H_5CH_2$-); 8.08 (d, 1H, C^6*H* in C). MS-FAB *m/z* 318 (M + 1).

***N*-(N^4-(*Z*)-1-Cytosyl-3-propanoyl)-*N*-(2-((Boc)amino)ethyl)-glycine.** *N*-(2-((Boc)amino)ethyl)glycine ethyl ester (2.0 g; 0.0081 mol) was dissolved in DMF (24 mL), followed by addition of DhbtOH (1.457 g; 0.00893 mol) and N^4-(*Z*)-1-(carboxyethyl)cytosine (2.834 g; 0.00893 mol). Subsequently, methylene chloride (24 mL) was added, and the mixture was cooled to 0 °C on an ice bath. After addition of DCC (2.010 g, 0.00974 mol), the mixture was stirred at 0 °C for 2 h, followed by 1 h at room temperature. The precipitated DCU was removed by filtration and washed with methylene chloride (30 mL), and a further amount of methylene chloride (50 mL) was added to the filtrate. The organic phase was extracted with half-saturated aqueous sodium hydrogen carbonate (6 × 100 mL), aqueous potassium sulfate (1 vol saturated diluted with 4 vol of water, 3 × 100 mL), and saturated aqueous sodium chloride (1 × 100 mL), dried over magnesium sulfate, and evaporated to dryness *in vacuo*. The solid residue was suspended in methylene chloride (25 mL) and stirred for 30 min, and the undissolved DCU was removed by filtration and washed with methylene chloride. The filtrate was evaporated to dryness *in vacuo*, and the residue was purified by column chromatography on silica gel, eluting with a mixture of methanol and methylene chloride (gradient from 5 to 7% methanol in methylene chloride). This afforded the ethyl ester of the title compound as an oil (2.34 g; 0.00429 mol), which was dissolved in tetrahydrofuran (20 mL). 1 M Lithium hydroxide (12.9 mL; 0.013 mol) was added at 0 °C, and the solution was stirred for 45 min. pH was then adjusted to 6 with 1 M hydrochloric acid,

tetrahydrofuran was removed by evaporation *in vacuo*, and pH was further lowered to 2 at 0 °C. The product was isolated by filtration and washed with cold water (3 × 10 mL). It was further purified by dissolving in boiling ethanol (150 mL) and precipitated by dropwise addition of water (300 mL) at 0 °C. After filtration, it was dried over sicapent *in vacuo* and yielded 1.15 g (27% overall yield) of the title compound. Anal. for $C_{24}H_{31}N_5O_8$, found (calcd): C, 55.54 (55.70); H, 6.00 (6.04); N, 13.70 (13.53). ^{1}H NMR (400 MHz; DMSO-d_6): due to limited rotation around the secondary amide, several of the signals were doubled; δ 1.34 (2 × s, 9H, *t*-Bu); 2.68 (t) and 2.79 (t) (2H, CCH_2CH_2-); 2.99 (q) and 3.07 (q) (2H, $BocNHCH_2$-); 3.28 (t) and 3.35 (t) (2H, $BocNHCH_2CH_2$-); 3.95 (t) and 3.96 (t) (2H, CCH_2-); 3.91 (s) and 4.08 (s) (2H, -CH_2COOH); 5.18 (s, 2H, $C_6H_5CH_2$-); 6.72 (t) and 6.83 (t) (1H, BocN*H*-); 6.93 (d, 1H, C^5*H* in C); 7.39 (m, 5H, $C_6H_5CH_2$-); 8.08 (d, 1H, C^6*H* in C); 10.71 (b, 1H, NH on C); 12.6 (b, 1H, -COOH). MS-FAB *m/z* 518 (M + 1).

Oligomerization. PNA oligomers containing the modified units were prepared by standard Merrifield solid-phase peptide synthesis according to previously described protocols.[17] DCC or DIC was used as coupling reagent, and the loading on the MBHA resin was approximately 0.15 mmol/g. All oligomers were purified on reverse-phase HPLC and characterized by FAB mass spectrometry (H-T_4(βT)T_5-Lys-NH_2, *m/z* 2820.71 [M + 1] (calcd 2820.15); H-T_4(βC)T_5-Lys-NH_2, *m/z* 2806.11 [M + 1] (calcd 2806.16); H-(βT)$_{10}$-Lys-NH_2, *m/z* 2946.91 [M + 1] (calcd 2947.30); H-GTAGA(βT)CACT-Lys-NH_2, *m/z* 2868.7 [M + 1] (calcd 2868.2); H-T_4(apgT)T_5-Lys-NH_2, *m/z* 2820.92 [M + 1] (calcd 2820.15); H-(apgT)$_{10}$-Lys-NH_2, *m/z* 2947.17 [M + 1] (calcd 2947.30); H-GTAGA(apgT)CACT-Lys-NH_2, *m/z* 2868.9 [M + 1] (calcd 2868.2); H-T_4(paT)T_5-Lys-NH_2, *m/z* 2821.42 [M + 1] (calcd 2820.15); H-(paT)$_{10}$-Lys-NH_2, *m/z* 2947.29 [M + 1] (calcd 2947.30); H-(paT)$_4$-(paC)(paT)$_5$-Lys-NH_2, *m/z* 2932.57 [M + 1] (calcd 2932.30); H-GTAGA(paT)CACT-Lys-NH_2, *m/z* 2868.3 [M + 1] (calcd 2868.2).

Acknowledgment. This work was supported by ISIS Pharmaceuticals. Anette W. Jørgensen is thanked for recording UV titrations and T_m measurements.