PNA-Directed Triple-Helix Formation by N^7-Xanthine

Robert H. E. Hudson,* Mykhaylo Goncharenko, Andrew P. Wallman, Filip Wojciechowski

Department of Chemistry, The University of Western Ontario, London, Ontario N6A 5B7, Canada
Fax +1(519)6613022; E-mail: rhhudson@uwo.ca

Received 16 August 2004

Abstract: We report the first example of alkylation of underivatized xanthine with chloroacetic acid to yield a separable mixture of N^7- and N^9-(methylenecarboxyl)xanthine and its conversion to a peptide nucleic acid monomer compatible with Fmoc-based oligomerization chemistry. Additionally, we have simultaneously prepared the N^7- and N^9-PNA monomers of guanine by alkylation of 2-*N*-isobutyrylguanine which were subsequently separated. Molecular modeling of the nucleobase base triplets indicates that N^7-xanthine and N^7-guanine form isomorphous triplets with adenine and guanine, respectively. We also show that polyamides containing N^7-xanthine are compatible with triple-helix formation.

Key words: xanthine, peptide nucleic acid, isomorphous, triplex

PNA is a synthetic mimic of DNA in which the sugar-phosphate backbone of the natural nucleic acid has been replaced with a polyamide backbone but maintains the natural nucleobases.[1] Possessing an uncharged polyamide backbone, PNA avidly binds to complementary nucleic acids and may form PNA:DNA:PNA triple helices dependent on the sequence context.[2] In order to favor triplex formation for amenable sequences, bisPNAs that contain two PNA domains covalently connected are commonly used.[3] Thus, one domain is responsible for Watson–Crick recognition[4] of a polypurine sequence, while the other is involved in Hoogsteen binding.[5]

Various modified nucleobases have been incorporated into both DNA and PNA oligomers to investigate and optimize recognition of natural nucleic acids by triplex formation. The most studied triplex, the pyrimidine motif, has the requirement for protonation of cytosine residues in the Hoogsteen strand. Consequently, such triple-stranded complexes are most stable in an environment that is generally below physiological pH. In order to overcome the pH dependence of triplex stability, nucleobase analogs of protonated cytosine such as N^7-guanine have been utilized in DNA chemistry[6] and more recently in PNA.[7] As well, the synthetic pyrimidine pseudoisocytosine, which is also a protonated cytosine analog, has been successfully utilized in PNA oligomers.[8]

The use of the synthetically more accessible N^7-guanine in conjunction with thymine in the Hoogsteen strand is not optimal because the N^7-G*G:C and T*A:T base triplets are not isomorphous. This will presumably cause a distortion in the backbone of the third strand polymer because the distance and orientation of the site of attachment of the N^7- or N^1-methylenecarbonyl group is different relative to the purine target (Figure 1).

Based on examination of molecular models, we reasoned that N^7-linked xanthine would be an excellent candidate for recognition of adenine and is isomorphous with N^7-guanine (Figure 1, c).[9] In addition, the purine base may deliver a stabilizing effect based on stacking interactions within the Hoogsteen strand.

Due to the number of potentially nucleophilic sites of xanthine, direct alkylation usually shows poor regioselectively. Regioselective monoalkylation has been described when protecting groups were used,[10,11] via trimethylsilylated intermediates,[12] or by use of a (tri-*n*-butylphosphine)cobalt(III) complexes.[13] As well, there are limited examples of regioselective alkylation to produce the N^9-isomer.[14] Finally, the reported reference to PNA containing N^9-xanthine should rather indicate hypoxanthine (inosine).[15]

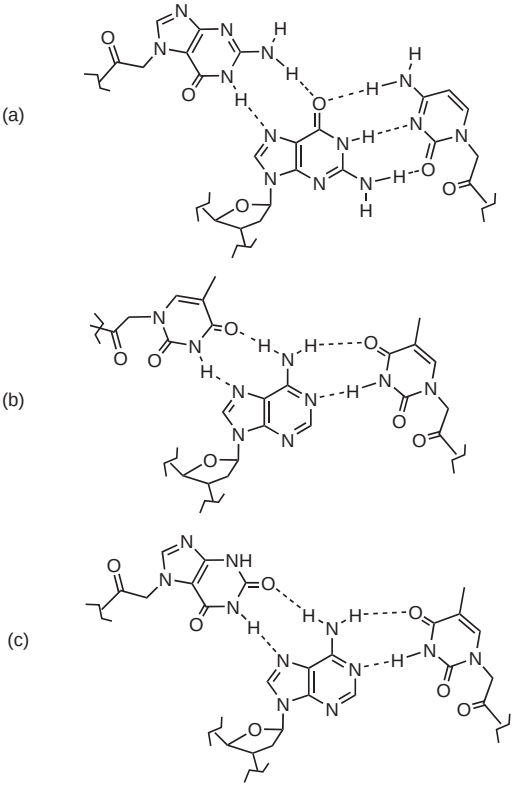


Figure 1 Proposed nucleobase interactions for a $(PNA)_2$:DNA triple helix. (a) Base triplet with N^7-guanine; (b) nonisomorphous base triplet with N^1-thymine; (c) base triplet with N^7-xanthine.

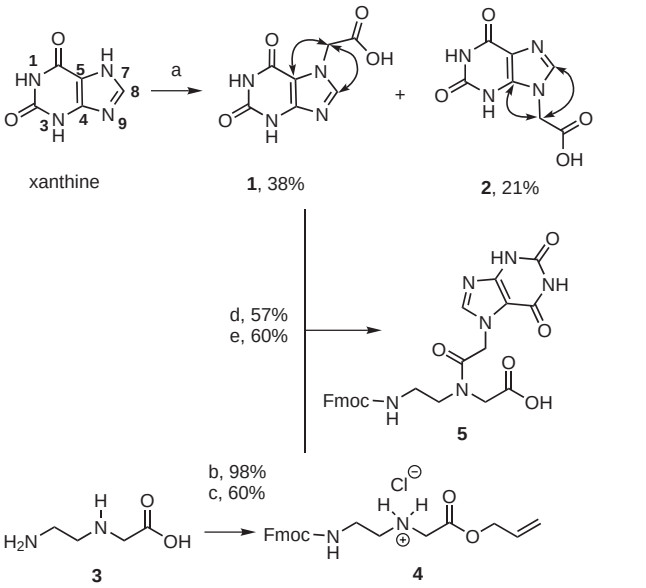
SYNLETT 2005, No. 9, pp 1442–1446
Advanced online publication: 29.04.2005
DOI: 10.1055/s-2005-868503; Art ID: S07604ST

For this work, the xanthine-7-acetic acid derivative **1** was required (Scheme 1). Our attempt to adapt the procedure of Müller et al.[10] to produce N^7-xanthine led exclusively to the N^3,N^7-bis(carboxymethylene) product. Although simple alkylated derivatives of xanthine are well known, no direct monoalkylation method to form xanthine-7-acetic acid has been described. In fact, this relatively simple compound has not been previously reported.

We have developed a technically simple alkylation of unmodified xanthine by using chloroacetic acid in aqueous base. The reaction afforded a mixture of the monoalkylated N^7- and N^9-regioisomers **1** and **2** where required N^7-derivative **1** prevailed in the ratio 38:21, respectively. Fractional recrystallization from water was used to separate these isomers on a 5 g scale resulting in ca. 98% purity (N^7-isomer). The overall chemical yield of approximately 60% is likely due to hydrolysis of chloroacetic acid, which is employed at a 1:1 stoichiometry with xanthine. This moderate yield is acceptable in view that xanthine may be recovered from the reaction mixture and the other starting materials are inexpensive. Since the procedure is very simple and inexpensive, no further attempts at optimization of the reaction yield were considered necessary.

Assignment of N^7- and N^9-isomers **1** and **2** was made by use of 2D NMR (gHMBC) spectroscopy. This experiment clearly showed a difference between the N^7- and N^9-isomers with respect to the coupling of methylene protons of the -CH_2-COOH fragment with C4 and C5 positions. Methylene protons of N^7-derivative **1** (singlet at δ = 4.99 ppm) are coupled to both C5 and C8 carbons and not to C4. The analogous methylene protons of N^9-isomer **2** (singlet at δ = 4.44 ppm) showed coupling to C8 and C4 but not C5. Although xanthine may be alkylated at the N^3 position, this possibility was discounted by comparison to the available data for N^3-xanthine acetic acid.[16]

Initially, we chose to condense compound **1** with methyl *N*-(2-Fmoc-aminoethyl)glycinate (**10**, Scheme 2) based upon ongoing work in our laboratory. However, the poor solubility of xanthine derivatives led to poor selectivity in the saponification of the methyl ester versus elimination of the Fmoc-protecting group. Thus, we changed to the allyl ester (**4**, Scheme 1), which affords orthogonal removal condition to Fmoc-deprotection and also increases the solubility of the backbone submonomer in organic solvents. While this compound has been previously synthesized by transesterification from *tert*-butyl ester,[17] we have employed a convenient 2-step method starting from 2-*N*-aminoethylglycine[18] (**3**) followed by carbamylation with Fmoc-*N*-hydroxysuccimide (Fmoc-NHS). We have found compound **3** to be a very useful precursor that may be produced and esterified easily on large scale (ca. 20 g) for synthesis of Boc-, Mmt- and Fmoc- backbone submonomers. When possible, we prefer to use the methyl ester, as it is less susceptible than the allyl ester to 2-oxopiperazine formation during the carbamylation reaction as observed by Seitz and coworkers.[17]



Scheme 1 *Reagents and conditions*: (a) $ClCH_2CO_2H$, NaOH, H_2O, reflux 5 h, key HMBC correlations shown by double-headed arrows; (b) allyl alcohol, HCl (g), reflux 2.5 h; (c) i. Fmoc-NHS, DIEA, CH_2Cl_2, 0–20 °C, 4 h; ii. HCl–Et_2O, 0 °C; (d) DCC, HOBt, DMSO–DMF, 18 h; (e) i. $Pd[(PPh_3)]_4$, MeOH–THF, *p*-$CH_3C_6H_4SO_2Na$, 20 °C, 17 h; ii. HCl, H_2O.

The monomer synthesis was next advanced by DCC/HOBt-mediated condensation between acid **1** and backbone submonomer **4**. However, due to poor solubility of xanthine-7-acetic acid a 1:1 DMSO–DMF solvent system was employed. Once the nucleobase derivative was condensed with the backbone submonomer, the solubility improved substantially. The Fmoc-protected monomer allyl ester was transformed into target acid **5** by cleavage of allyl ester with palladium triphenylphosphine complex in the presence of sodium *p*-toluenesulfinate, as previously reported.[19] The sodium salt of **5** was surprisingly well soluble in water.[20]

Next we turned our attention to the preparation of the guanine PNA monomers. The alkylation of guanine is known to give a mixture of N^9/N^7-regioisomers where N^9-isomer is the main product. With respect to formation of *N*-glycosides, the N^9-isomer may be favored by use of a bulky O^6-protecting group such as the diphenylcarbamoyl derivative,[21] but this effect can be less reliable for alkylation reactions.[22] Alternatively, 6-chloro-2-aminopurine is often used to access the N^9-guanine derivatives selectively.[23]

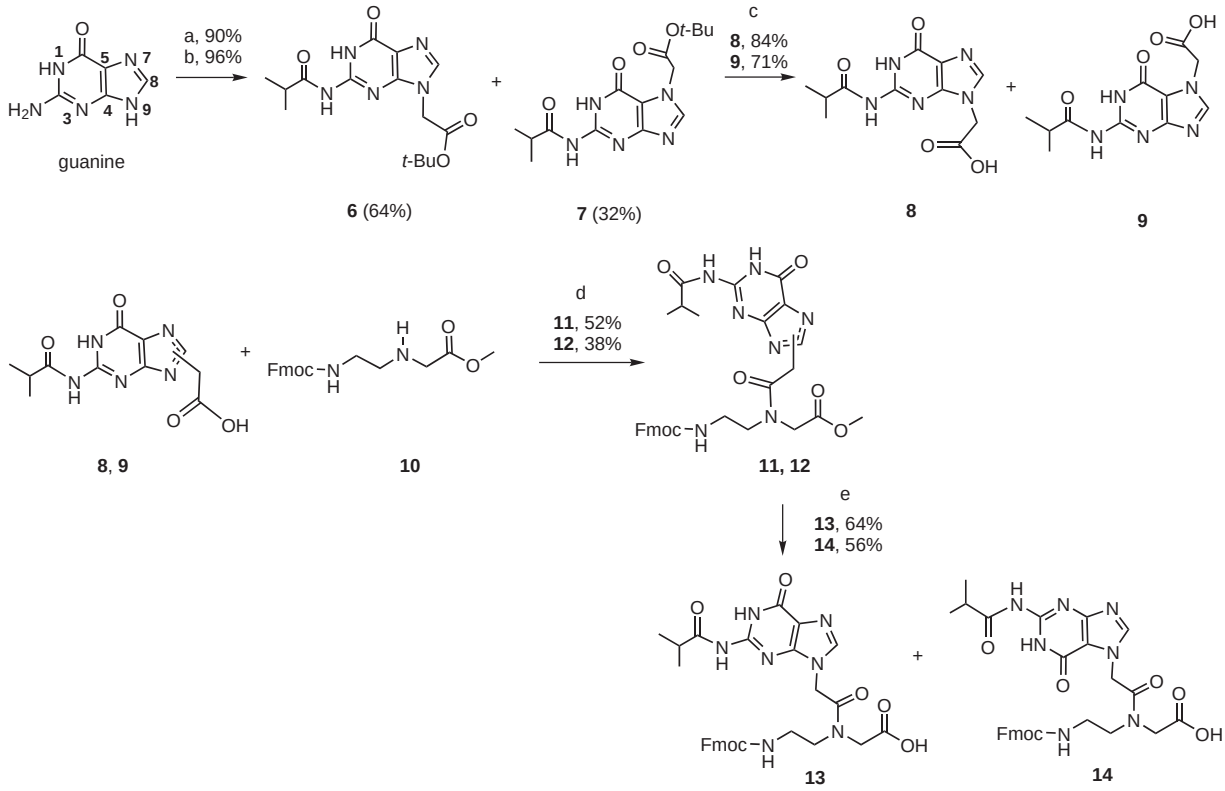
Since we would find both N^7- and N^9-carboxymethylguanine useful, we pursued the synthesis starting with unmodified guanine. An isobutyryl group was installed on the 2-amino position in 90% yield by reaction with isobutyryl anhydride in DMF at 150 °C (Scheme 2).[22] Reaction of 2-*N*-isobutyrylguanine with *tert*-butyl bromoacetate at ambient temperature in the presence of sodium hydride afforded a mixture of N^9/N^7-isomers **6** and **7** with overall 94% yield and 62:32 ratio, respectively, whereas reaction at elevated temperatures favored the formation of the N^9-regioisomer. Compounds **6** and **7** were separated by ex-

ploiting their differential solubility and column chromatography and were subsequently identified by 2D-NMR experiments.[24]

Regiochemical assignment of ^{13}C signals in substituted purines is known from INEPT experiments.[25] Thus, the signal for C5 was identified as the lone upfield peak at δ = 120.3 ppm (N^9-isomer **6**) and δ = 112.4 ppm (for N^7-isomer **7**). Signals at δ = 141.0 ppm (**6**) and δ = 145.5 ppm (**7**) in carbon spectra were assigned to C8 on the basis of the residual signal from attached protons. The same effect was used for assignment of signals in carbon spectra for methylene, *tert*-butoxy and isobutyryl groups in the high field area of spectra. Interaction of attached protons of isobutyryl and *tert*-butoxy groups gave the ground for assignment of signals at δ = 180.9 ppm (**6**) and δ = 180.7 ppm (**7**) as the carbonyl carbon of *i*-PrCONH-fragment and δ = 167.4 ppm (**6**) and δ = 167.5 (**7**) as for carbonyl carbon of COO*t*-Bu groups. Given that signals of C2 and C6 in the purine ring[8] are expected in the range of δ = 150–160 ppm, we assigned to these atoms the peaks at δ = 149.6 ppm and δ = 155.5 ppm (**6**) and the peaks at δ = 153.3 ppm and δ = 157.5 ppm (**7**) accordingly. Thus, the signals at δ = 148.8 ppm (**6**) and δ = 147.8 ppm were assigned for C4 atoms in accordance with literature data for purines.[23] As done for the substituted xanthines, gHMBC NMR spectroscopy was used to assign the regioisomers based on coupling of the CH_2COOt-Bu fragment methylene protons with the C4 and C5 atoms. Thus, methylene protons of N^9-derivative **6** with singlet at δ = 4.86 ppm have mutual coupling signals with carbonyl of COO*t*-Bu fragment, C4 and C8 carbons. However, while analogous methylene protons of N^7-isomer **7** (singlet at δ = 5.06 ppm) showed similar coupling with COO*t*-Bu carbonyl and C8 carbon, they did not show interaction with C4 carbon but instead showed coupling to C5. In summary, these compounds were found to be consistent with those recently reported.[22]

Once purified, N^9- and N^7-isomers **6** and **7** were used separately in further transformations. Cleavage of *tert*-butoxy group was made by the treatment in TFA–CH_2Cl_2 followed by DCC/HOBt-mediated coupling of carboxymethyl guanines **8** and **9** with methyl *N*-(2-Fmoc-aminoethyl)glycinate (**10**). This reaction afforded esters **11** and **12**, which were purified chromatographically. Saponificaton of Fmoc-protected monomer methyl esters revealed the monomer acids **13** and **14**.[26] The hydrolysis was completed by using eight-fold excess of NaOH in H_2O–THF medium at 0 °C for five minutes. Prolonged reaction time or poorly soluble monomer esters led to an unacceptable degree of Fmoc removal, in which case the allyl ester was employed.

Since the use of N^7-guanine derivatives in triplex formation is well established, we have first tested the ability of N^7-xanthine to form triple helices. We synthesized Ac-T_6-lys-NH_2 (**15**) and Ac-X_6-lys-NH_2 (**16**) by standard methods[27] and investigated their hybridization with



Scheme 2 N^9/N^7-Guanine Fmoc/acyl monomers. *Reagents and conditions*: (a) $(i\text{-PrCO})_2O$, DMF, 150 °C, 7 h, 90%; (b) NaH, DMF, $BrCH_2COOt$-Bu, 0–20 °C, 17 h; (c) CF_3COOH, CH_2Cl_2, $(C_2H_5)_3SiH$, 18 h; (d) DCC, HOBt, DMF, 18 h; (e) THF, H_2O, NaOH, 0 °C, 5 min.

poly(rA).[28] Thermal melting analysis (T_m = 62.0 °C) and determination of the stoichiometry of binding confirmed the expected triplex formation for (**15**) with poly(rA). Xanthine oligomer **16** also showed a single cooperative transition (Tm = 59.5 °C, pH 7.0) that was slightly stabilized at acidic pH (Tm = 66.0 °C, pH 5.5). A Job plot revealed that oligomer **16** binds in a 2:1 fashion with poly(rA), consistent with the supposition that N^7-xanthine is a suitable mimic of thymine (Figure 2).

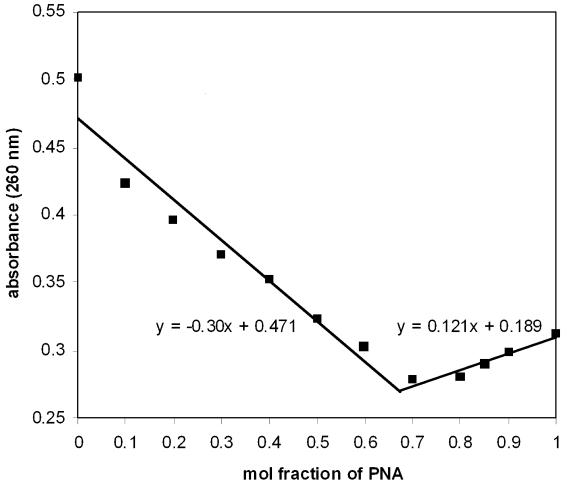


Figure 2 Job plot for Ac-X_6-lys-NH_2 indicating $(PNA)_2$:DNA triple helix formation.

Currently, we are investigating the use of monomers **14** and **5** in the Hoogsteen strand of clamp-PNAs for the recognition of mixed pyrimidine sequence DNA.

Acknowledgment

We thank the Natural Sciences and Engineering Research Council (of Canada) for funding this work. APW is the recipient of an NSERC Undergraduate Summer Research Award and FW is partially supported by the Faculty of Graduate Studies, University of Western Ontario.

References

(1) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
(2) Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.
(3) Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E. A.; Sprankle, K. G.; Griffey, R. H.; Kiely, J. S.; Freier, S. M. *J. Am. Chem. Soc.* **1995**, *117*, 831.
(4) Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 737.
(5) Hoogsteen, K. *Acta Crystallogr*. **1963**, *16*, 907.
(6) Hunziker, J.; Priestly, E. S.; Brunar, H.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 2661.
(7) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2003**, *68*, 4439.
(8) Egholm, M.; Christensen, L.; Duholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217.
(9) Molecular models were constructed using HyperChem 5.1 from existing crystallographic data. The xanthine triplet was initially geometry optimized using molecular mechanics (Amber force field) and subsequently refined by use semi-empirical methods (PM3), as previously described: Fenyö, R.; Tímár, Z.; Pálinkó, I.; Penke, B. *J. Mol. Struc.–Theochem.* **2000**, *496*, 101.
(10) Müller, C. E.; Deters, D.; Dominik, A.; Pawlowski, M. *Synthesis* **1998**, 1428.
(11) Bridson, P. K.; Richmond, G.; Yeh, F. *Synth. Commun.* **1990**, *20*, 2459.
(12) Müller, C. E.; Shi, D.; Manning, M. Jr.; Daly, J. W. *J. Med. Chem.* **1993**, *36*, 3341.
(13) Marzilli, L. G.; Epps, L. A.; Sorrell, T.; Kistenmacher, T. J. *J. Am. Chem. Soc.* **1975**, *97*, 3351.
(14) For example: Hoffmann, M. F. H.; Brückner, A. M.; Hupp, T.; Engels, B.; Diederichsen, U. *Helv. Chim. Acta* **2000**, *83*, 2580.
(15) (a) Sanjayan, G. J.; Pedireddi, V. R.; Ganesh, K. N. *Org. Lett.* **2000**, *2*, 2825. (b) Timár, Z.; Bottka, S.; Kovács, L.; Penke, B. *Nucleosides Nucleotides* **1999**, *18*, 1131.
(16) N^7- and N^9-xanthine acetic acids were separated by differential solubility in water, the N^7-derivative being less soluble. N^7-isomer: white solid, mp 320 °C (dec). ^{1}H NMR (400 MHz, DMSO): δ = 13.30 (br s, 1 H), 11.61 (s, 1 H), 10.91 (s, 1 H), 7.91 (s, 1 H), 4.99 (s, 2 H). ^{13}C NMR (100 MHz, DMSO): δ = 169.8, 156.3, 151.9, 149.6, 143.9, 107.3, 47.7. HRMS (EI): *m/z* calcd for $C_7H_6N_4O_4$: 210.0389; found: 210.0397. N^9-isomer: white solid, mp >290 °C (dec). ^{1}H NMR (400 MHz, DMSO): δ = 10.82 (s, 1 H), 7.60 (s, 1 H), 4.84 (s, 2 H). ^{13}C NMR (100 MHz, DMSO): δ = 170.1, 158.9, 152.2, 142.8, 138.5, 115.7, 48.4. HRMS (EI): *m/z* calcd for $C_7H_6N_4O_4$: 210.0389; found: 210.0451. N^3-isomer: see ref.[11]
(17) Seitz, O.; Köhler, O. *Chem.–Eur. J.* **2001**, *7*, 3914.
(18) Heimer, E. P.; Gallo-Torres, H. E.; Felix, A. M.; Ahmad, M.; Lambros, T. J.; Scheidl, F.; Meienhofer, J. *Int. J. Pept. Res.* **1984**, *23*, 203.
(19) Honda, M.; Morita, H.; Nagakura, I. *J. Org. Chem.* **1997**, *62*, 8932.
(20) Selected data.
Allyl ester precursor to **5**: white solid, mp 168–170 °C (dec). ^{1}H NMR (400 MHz, DMSO): δ = 11.53 (br s, 1 H), 10.87 (br s, 1 H), 7.87–7.20 (m, 10 H), 5.90 (m, 1 H), 5.38–5.09 (m, 4 H), 4.66–4.19 (m) and 4.09 major (s, 7 H), 3.46 major (m, minor rotamer overlapping with H_2O), 3.09 minor (m, major rotamer overlapping with H_2O). HRMS (ESI-TOF): *m/z* calcd for sodium adduct $C_{29}H_{28}N_6O_7Na$: 595.1917; found: 595.1912.
Compound **5**: off-white solid, mp 162–164 °C (change in appearance), 205–208 °C (dec). ^{1}H NMR (400 MHz, DMSO): δ = 12.81 (br s, 1 H), 11.58 major and 11.56 minor (1 H), 10.86 major and 10.84 minor (1 H), 7.88–7.28 (m, 10 H), 5.27 major and 5.07 minor (s, 2 H), 4.34–4.20 (m)and 3.98 major (s, 5 H), 3.43 major (m, minor rotamer overlapping with H_2O), 3.10 minor (m, major rotamer overlapping with H_2O). HRMS (ESI-TOF): *m/z* calcd for sodium adduct $C_{26}H_{24}N_6O_7Na$: 555.1604; found: 555.1602.
(21) Robins, M. J.; Zou, R.; Guo, Z.; Wnuk, S. *J. Org. Chem.* **1996**, *61*, 9207.
(22) Timár, Z.; Kovács, L.; Kovács, G.; Schmél, Z. *J. Chem. Soc., Perkin Trans. 1* **2000**, 19.
(23) For PNA: Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.
(24) Data for **6** and **7**.
Compound **6** (N^9-isomer): white solid, mp 310–331 °C. ^{1}H NMR (400 MHz, DMSO): δ = 12.10 (s, 1 H), 11.66 (s, 1 H), 7.94 (s, 1 H), 4.87 (s, 2 H), 2.76 (sept, ^3J = 6.8 Hz, 1 H) 1.40 (s, 9 H), 1.09 (d, ^3J = 6.9 Hz, 6 H). ^{13}C NMR (100 MHz,

DMSO): δ = 180.9, 167.4, 155.5, 149.6, 148.8, 141.0, 120.3, 83.0, 45.5, 35.3, 28.4, 19.6. HRMS (EI): *m/z* calcd for $C_{15}H_{21}N_5O_4$: 335.15937; found: 335.15937.
Compound **7** (N^7-isomer): white solid, mp 202–204 °C. ^{1}H NMR (400 MHz, DMSO): δ = 12.13 (s, 1 H), 11.56 (s, 1 H), 8.10 (s, 1 H), 5.06 (s, 2 H), 2.72 (sept, ^3J = 6.8 Hz, 1 H), 1.40 (s, 9 H), 1.10 (d, ^3J = 6.8 Hz, 6 H). ^{13}C NMR (100 MHz, DMSO): δ = 180.7, 167.6, 157.5, 153.3, 147, 9, 145.5, 112.4, 82.7, 48.6, 35.4, 28.3, 19.6. HRMS (EI): *m/z* calcd for $C_{15}H_{21}N_5O_4$: 335.15937; found: 335.15937.

(25) Osterman, R. M.; McKittrick, B. A.; Chan, T. M. *Tetrahedron Lett.* **1992**, *33*, 4867.

(26) Data for **13** and **14**.
Compound **13** (N^9-isomer): white solid, mp 234–235 °C. ^{1}H NMR as reported in ref.[22]; HRMS (ESI-TOF): *m/z* calcd for sodium adduct $C_{30}H_{31}N_7O_7Na$: 624.2183; found: 624.2213.
Compound **14** (N^7-isomer): white solid, mp 188–190 °C (dec). ^{1}H NMR (400 MHz, DMSO): δ = 12.13 (br s, 1 H), 11.59 major and 11.56 minor (s, 1 H), 8.18 minor and 8.15 major (s, 1 H), 7.87–717 (m, 10 H), 5.38 major and 5.20 minor (s, 2 H), 3.46–3.12 (m, 4 H), 2.71 (m, 1 H), 1.10 major and 1.05 minor (d, ^3J = 6.8 Hz, 6 H). HRMS (ESI-TOF): *m/z* calcd for sodium adduct $C_{30}H_{31}N_7O_7Na$: 624.2183; found: 624.2165.

(27) Oligomers were synthesized on Rink amide resin by an ABI 433a peptide synthesizer at the 5 μmol scale according to the manufacturer-supplied cycles and purified by RP-HPLC.
Data for -T_6-lys-NH_2 (**15**) HRMS (MALDI-TOF): *m/z* calcd for $C_{74}H_{101}N_{27}O_{26}$: 1783.7411; found: 1784.6012.
Data for Ac-X_6-lys-NH_2 (**16**) HRMS (ESI-TOF): *m/z* calcd for $C_{74}H_{89}N_{39}O_{26}$: 1940.76; found: 1940.55 [MH^+]

(28) Thermal denaturation was measured at strand concentration of 1.3 μM in base pairs with 150 mM NaCl, 10 mM Na_2HPO_4, 1 mM ETDA, pH 7.0 at a 2:1 PNA:RNA ratio. All transitions were well-behaved and monophasic. The first derivative method was used to estimate the Tm. The stoichiometry of binding was determined by the method of continuous variations: Job, P. *Ann. Chim. (Paris)* **1928**, *9*, 113.