Synthesis of New Chiral Building Blocks for Novel Peptide Nucleic Acids

WU, Jie(吴杰)   XU, Xiao-Yu(许笑宇)   LIU, Ke-Liang*(刘克良)
Beijing Institute of Pharmacology and Toxicology, Beijing 100850, China

N-Boc protected amino acids of analogues of peptide nucleic acid (PNA), which are a class of conformationally constrained building blocks based on 4-aminoproline backbone with chirality at 2-C and 4-C, have been synthesized. Those monomers can be used for the construction of novel peptide nucleic acid analogues.

Keywords   oligonucleotide, peptide nucleic acids, pyrrolidine, synthesis

Introduction

About ten years ago, PNA, a structural mimic of DNA in which the sugar-phosphate backbone is replaced by *N*-(2-aminoethyl)glycine (aeg) linkage emerged as a potential antisense therapeutic agent.[1] PNA has some advantages: (1) it is stable to cellular nucleases and proteases, (2) it hybridizes with complementary DNA or RNA (cDNA/RNA) sequences with high affinity, (3) it has low non-specific interaction with cellular contents and (4) it is easily synthesized by adoption of solid phase peptide synthesis chemistry. However, the major limitation of PNAs is their poor solubility in aqueous medium and achiral, and they bind to cDNA in both parallel (*N*-PNA/5′-DNA) and antiparallel (*N*-PNA/3′-DNA) modes. Based on the monomer synthesis from an amino acid, it seemed natural to substitute glycine with other amino acids in the preparation of the monomer.[2] Various groups have tried to put chirality to PNA molecules by linking other amino acids,[3] peptides,[4] or oligonucleotides,[5] at the terminus or by incorporation of chiral amino acids in place of glycine in the PNA backbone[6] to attempt discriminating between parallel and antiparallel modes of binding (Fig. 1).

Trans-4-hydroxy-*L*-proline is relatively easy to manipulate to access all four stereoisomers. The pyrrolidine ring in proline is a suitable unit for mimicking the ribose moiety in DNA. It has been proved to be a useful starting material to synthesize the conformationally constrained PNA monomers.[7] In our laboratory, 4-hydroxyproline has also been utilized to synthesize PNA monomers in which nucleobase substitution is at the position of 4-C of the pyrrolidine ring. Among the homo-oligomers aepPNA-LysT₁₀(C→N), aepPNA-LysA₁₂(C→

N), aapPNA-LysT₁₀(C→N) and aapPNA-LysA₁₀(C→N) synthesized using these monomers only aepPNA-LysA₁₂ and aapPNA-LysA₁₀ showed hybridization with complementary DNA.[8,9] We thought that the loss of affinity of these PNA analogues could be attributed to lack of flexibility of the backbone. In order to substantiate this assumption, we designed another PNA analogue: ampPNA. It has two differences from aepPNA and aapPNA we synthesized before: (1) nucleobase will not be directly but via a methylene attached to the pyrrolidine ring and (2) the carbon attached to nucleobase in ampPNA is achiral. Those will make the backbone in the ampPNA have more flexibility. Tertiary amino structure in each unit could be expected to obtain a good solubility and a high affinity for negatively charged natural nucleic acids (Fig. 2).
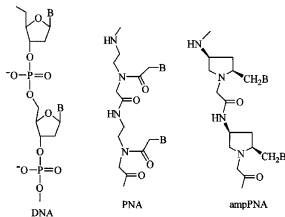

Fig. 1 Comparison of structure of DNA, PNA and ampPNA.

In this paper, the detailed synthesis of the new chiral PNA monomers containing all four natural bases (A, C, T, G) is described.

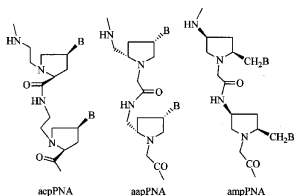## Results and discussion

As the key intermediates in the alkylation of four

Fig. 2 Comparison of the three chiral PNA.

nucleotide bases in our synthetic pathways, compound 7a or 7b was prepared from N-benzyl-*trans*-4-hydroxy-L-proline in 5 steps as shown in Scheme 1. The hydroxyl group was converted into an azido group via its sulfonylate. After reduction to the amine[10] and protection with di-*tert*-butyl dicarbonate, compound 4 was yielded in which the configuration of 4-C had been inverted. A positive nuclear overhauser effect (NOE) of 2-H on irradiation at 4-H was observed, indicating that inversion at 4-C had taken place as expected, giving the cis-product 4. Sodium borohydride[11] can turn compound 4 into corresponding alcohol 5 in mixed solvent of tetrahydrofuranmethanol in 70% yield. After removal of the benzyl-protecting group by catalytic hydrogenation, it reacted with ethyl

bromoacetate in the presence of Et₃N in CH₂Cl₂, followed by addition of CBr₄ and PPh₃ or methanesulfonyl chloride, affording the key intermediate 7a or 7b in 62% or yield 71% respectively, which was commonly used in the alkylation of nucleobase.
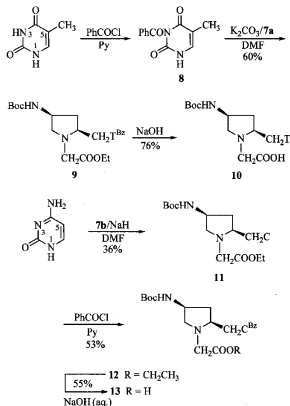
Derivatization of the pyrimidine base

Reaction of 7a with thymine gave a disubstituted product suggesting that protection of thymine at 3-N might be necessary. As a result, 3-N-benzoylthymine[12] (8) was synthesized. It reacted with 7a to form the desired compound 9. Finally, the ethyl ester was removed by hydrolysis, giving the thymine monomer[13] 10 in 76% yield (Scheme 2).

For the synthesis of the cytosine monomer, introduction of a protecting group for the 4-N amino group of cytosine was necessary in order to prevent chain extension from this position in the later peptide-coupling steps. A Bz group was selected for this purpose since it was also found to render the intermediate sufficiently soluble for chemical manipulation. Thus, the exocyclic amino group of cytosine was protected prior to alkylation by treatment with benzoyl chloride to give 4-N-Bz-cytosine. The subsequent alkylation of C^Bz was attempted by firstly generating the anion, using sodium hydride in anhydrous DMF, followed by addition of 7a.

Unfortunately, the product was not the desired, but a disubstituted one, which was confirmed by ¹H NMR spectra. The desired product 11 was achieved by heating a mixture of 7b, cytosine and NaH in anhydrous DMF at 60 ℃ for one

Scheme 1   Synthesis of the key intermediate 7a or 7b
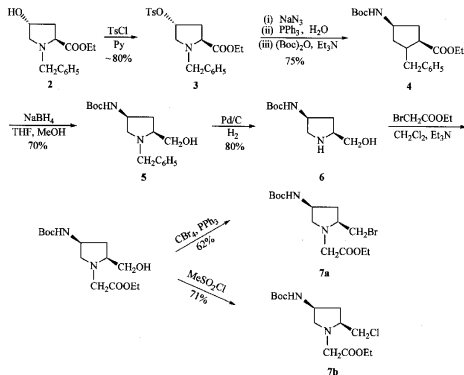
Scheme 2   Synthesis of pyrimidine monomers

week.[14] Those results may be attributed to the lower activity of **7b** compared with **7a**. After protection with benzoyl chloride of **11** and treatment with 2 mol/L NaOH, the cytosine monomer **13** was obtained (Scheme 2).

Derivatization of the purine base

For the synthesis of the adenine monomer, the 6-N amino group of adenine needs to be protected, and once again a Bz-protecting group was chosen. However here, unlike the case for the cytosine monomer, protection of the exocyclic amino group was performed after alkylation of adenine with **7b** in order to ensure that substitution occurred mainly at the 9-N position, rather than **7a** which has higher activity that would lower the rigio-selectivity. Although alkylation of adenine is notoriously nonregiospecific, it has been reported that alkylation of sodium adenine in DMF gives rise primarily to 9-N substituted products,[15] therefore alkylation of adenine with **7b** was carried out by firstly generating sodium adenine *in situ* in anhydrous DMF, followed by addition of **7b**. The structure of the product **14** was determined by 2D ^{1}H-^{13}C HMBC (heteronuclear multibond correlation). Since a long-range correlation between protons of the methylene at the pyroline 2-C position and 6-C, 8-C of the adenine has been observed in the HMBC spectrum, it can be concluded that the product is the desired 9-N isomer.

Rapoport[16,17] reported on the CBz-protection of trimethylsilyl-protected adenine and found that treatment with benzyl chloroformate under all usual conditions did not give clean or efficient acylation of the 6-N amino group. Rapport's reagent 1-(benzyloxycarbonyl)-3-ethylimidazolium tetrafluoroborate ($PhCH_2OCOImEt^+ \cdot BF_4^-$) can make the acylation markedly improved in literature, but it is not convenient for us to synthesize, so we decided to adopt the benzoyl group (Bz) used in conventional DNA chemistry as the protecting group. Thus, the exocyclic amino group in **14** was protected by treatment with a 6 molar excess of benzoyl chloride in pyridine. After purification, **15** was obtained in *ca*. 50% yield. Final removal of the ethoxyl group by hydrolysis gave the Bz-protected adenine monomer **16** in 53% yield.

As opposed to adenine, guanine can not be alkylated to give only one 9-N substituted product, because of its poor solubility in all kinds of solvents and competent reactions between 7-N and 9-N. However, since 2-amino-6-chloropurine[18] (**17**) gives almost exclusively 9-N isomer, it is generally employed instead as the starting material for the synthesis of guanine derivatives. Next step in the synthetic pathway involves replacement of the 6-chloro group with an oxygen functionality, which would yield the corresponding guanine carbonyl moiety at a later stage. The simplest and most obvious approach is that to exchange the chloro group for an alkoxy group and **18** was synthesized as shown in Scheme 3. Its alkylation with **7b** and LiH in DMF was achieved. After workup and purification by flash chromatography, the 9-N isomer **19a** and 7-N isomer **19b** were obtained in 3:1 ratio. The structure of the product was also determined by HMBC spectrum and further confirmed by comparison of the ^{1}H NMR spectra of 7-N (**19b**) and 9-N (**19a**) substituted guanine, and it was shown that $\Delta\delta$ value between 8-H and NH_2 signals in 7-N isomer was much greater than that for 9-N isomer[19]—the signal of 8-H for the 9-N isomer (**19a**) shifted upfield (δ 7.71) relative to the corresponding 8-H signal for 7-N isomer (**19b**, δ 7.73) and the corresponding NH_2 signal shifted downfield for the 9-N isomer (**19a**, δ 5.94) relative to the corresponding NH_2 signals for 7-N isomer (**19b**, δ 5.55). Attempt to let **19a** react with either isobutyric anhydride or acetyl chloride failed. Since this particular amino group could not be acylated with these reagents, it was implied that no protection is needed.[20]

Experimental

Reagents and solvents were obtained commercially and used without further purification unless otherwise indicated. DMF was re-distilled from P_2O_5 under reduced pressure. Tetrahydrofuran (THF), 1,4-dioxane and pyridine were distilled from sodium wire and stored over 4A molecular sienes. ^{1}H and ^{13}C NMR spectra were measured on a JNM-GX 400M NMR spectrometer. Chemical shifts were recorded in ppm relative to tetramethylsilane (TMS). FAB-MS spectra were recorded on a Zabspec mass spectrometer. Melting points were uncorrected.

Scheme 3 Synthesis of purine monomers
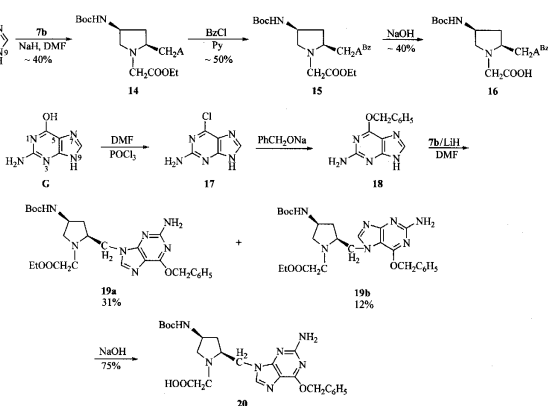
N-Benzyl-4R-(4-p-tolylsulfonyloxy)-L-proline-2S-ethyl ester (3)

N-Benzyl-trans-4-hydroxy-L-proline-ethyl ester (37.3 g, 0.15 mol) was dissolved in anhydrous pyridine (250 mL). This solution was stirred and cooled in an ice bath. Toluene-p-sulfonyl chloride (43 g, 0.23 mol) was added by portion, kept for 24 h after the temperature of the solution had allowed to warm to r. t.. The reaction mixture was poured into a solution of citric acid in ice-cold water and extracted with ethyl acetate. The combined organic phase was dried with anhydrous MgSO₄ and evaporated to give the crude product as an oil, which was purified by flash chromatography using petroleum ether (60—90 ℃):acetone = 5 :1 as the eluting solvent. Compound **3** was obtained as oil (46.5 g), in yield 77% , $[\alpha]_D^{20}$ - 15.2 (c 1.0, CH₂Cl₂); ¹H NMR (CDCl₃) δ; 1.22 (t, J =7.2 Hz, 3H, OCH₂**CH₃**), 2.23— 2.27 (m, 2H, tetrahydropyrrole-3-H), 2.43 (s, 3H, **CH₃**C₆H₄), 2.62—2.66 (m, 1H, tetrahydropyrrole-5-H), 3.21—3.22 (m, 1H, tetrahydropyrrole-5-H), 3.52—3.56 (m, 1H, tetrahydropyrrole-2-H), 3.61 (d, J = 13.2 Hz, 1H, C₆H₅**CH₂**), 3.89 (d, J =13.2 Hz, 1H, C₆H₅**CH₂**), 4.12 (q, J =7.2 Hz, 2H, OCH₂CH₃), 4.97—4.99 (m, 1H, tetrahydropyrrole-4-H), 7.24—7.76 (m, 9H, Ph); IR (KBr) ν; 1735 (C = O) cm⁻¹; MS (FAB) m/z (%): 404.0 (M+1, 20), 330.0 (M — COOEt, 25), 248.1 (M — C₇H₇ — SO₂, 5), 91.0 (CH₂C₆H₅, 100). Anal. calcd for C₂₁H₂₅NO₅S; C 62.51, H 6.24, N 3.47 S 7.95; found C 62.15, H 6.16, N 3.37, S 7.77.

N-Benzyl-4S-(tert-butoxycarbonyl)-L-proline-2S-ethyl ester (4)

NaN₃(10.4 g, 0.165 mol) was added to the solution of **3** (43.0 g, 0.11 mol) in dry DMF (250 mL) while stirring at 50 ℃. After 36 h, the solids in the mixture were filtered and washed with DMF, then the combined filtrate was concentrated under reduced pressure. The residue was dissolved in 200 mL of CH₂Cl₂ with stirring at 0 ℃, and after adding PPh₃(27 g, 0.1 mol), reacted for 24 h. 25 mL of H₂O was added to the reaction mixture, which was refluxed for 2 h then cooled to r. t. followed with addition of (Boc)₂O (20 g, 0.92 mol) and Et₃N (22.3 mL, 0.16 mol). The mixture was left overnight, then washed with brine (3 × 100 mL) and the organic layer was dried with anhydrous Na₂SO₄. After the solvent was evaporated, the residue was purified by column chromatography using petroleum ether (60—90 ℃):acetone = 3 :1 as the eluting solvent. Compound **4** was obtained as oil (27 g), in yield 70.5% , $[\alpha]_D^{20}$ -24.2 (c 1.0, CH₂Cl₂), yield 70.5% , 1.34 (s, 9H, C(CH₃)₃), 1.73—1.77 (m, 1H, tetrahydropyrrole-3-H), 2.34—2.39 (m, 1H, tetrahydropyrrole-3-H), 2.49—2.50 (m, 1H, tetrahydropyrrole-5-H), 2.58—2.59 (m, 1H, tetrahydropyrrole-5-H), 3.27—

3.30 (m, 1H, pyrrole-2-H), 3.49 (d, $J = 13.2$ Hz, 1H, $C_6H_5CH_2$), 3.84 (d, $J = 13.2$ Hz, 1H, $C_6H_5CH_2$), 3.93—3.94 (m, 1H, tetrahydrolline-4-H), 4.01—4.07 (m, 2H, OCH$_2$CH$_3$), 6.76 (s, 1H, NH), 7.21—7.31 (m, 5H, Ph); IR (KBr) ν: 3367 (N—H), 1704 (C=O) cm^{-1}; MS (FAB) m/z (%): 349 (M + 1, 45), 291.3 (M − t-Bu, 45), 275.3 (M − COOEt, 60), 219.2 (M − t-Bu − COOEt, 25), 91.1 (CH$_2$C$_6$H$_5$, 100). Anal. calcd for $C_{19}H_{28}N_2O_4$: C 65.49, H 8.10, N 8.04; found C 65.48, H 7.98, N 8.00.

N-Benzyl-4-S-tert-butoxycarbamido-tetrahydropyrrol-2S-ylmethanol (5)

MeOH (0.8 mL) was added dropwise over a period of 20 min to a mixture of ester 4 (0.35 g, 1 mmol) and NaBH$_4$ (94.5 mg, 2.5 mmol) in THF (4 mL) at 50—55 ℃. The mixture was stirred for 4 h, then water (2.0 mL) was added. Most of the organic solvents were evaporated under reduced pressure. Brine (3 mL) was added, and the mixture was extracted with ethyl acetate (10 × 5 mL). The extracted layer was washed with brine (3 × 2 mL), dried over anhydrous sodium sulfate, then evaporated, giving white solid which was recrystallized with ethyl acetate in 84% yield, m.p. 92—93 ℃, $[\alpha]_D^{20} − 53.5$ (c 1.0, CH$_3$OH); ^{1}H NMR (CDCl$_3$) δ: 1.41 (s, 9H, C(CH$_3$)$_3$), 2.38—4.09 (m, 10H), 7.24—7.34 (m, 5H, C$_6$H$_5$); IR (KBr) ν: 3348 (O—H), 1686 (C=O) cm^{-1}; MS (FAB) m/z (%): 307.2 (M + 1, 65), 275.2 (M − CH$_2$OH, 30), 251.2 (M − t-Bu, 45), 91.1 (CH$_2$C$_6$H$_5$, 100). Anal. calcd for C$_{17}$H$_{26}$N$_2$O$_3$: C 66.64, H 8.55, N 9.14; found C 65.99, H 8.57, N 8.99.

4-S-tert-Butoxycarbamido-tetrahydropyrrol-2-S-ylmethanol (6)

10% Pd-C (1.2 g) was added to a solution of 5 (6.1 g, 0.02 mmol) dissolved in methanol (50 mL). Then the mixture was stirred at 40—50 ℃ under hydrogen until 5 was not detectable by TLC. After Pd-C was filtrated, the filtration was evaporated under reduced pressure, giving white slice solid (3.1 g), in yield 76%, m.p. 169—171 ℃, $[\alpha]_D^{20} − 17.4$ (c 1.0, CH$_3$OH); ^{1}H NMR (CDCl$_3$ + CD$_3$OD) δ: 1.39 (s, 9H, C(CH$_3$)$_3$), 1.77—3.42 (m, 5H), 3.75 (d, $J = 8.8$ Hz, 1H, CH$_2$OH), 3.89 (d, $J = 8.8$ Hz, 1H, CH$_2$OH), 4.31—4.32 (m, 1H, pyrrole-4-H), 7.25 (s, 1H); IR (KBr) ν: 3349 (O—H), 3391 (N—H), 1679 (C=O) cm^{-1}; MS(FAB) m/z (%): 217.2 (M + 1, 100), 161.1 (M − t-Bu, 65), 117.1 (M − Boc, 10). Anal. calcd for C$_{10}$H$_{20}$N$_2$O$_3$: C 55.53, H 9.32, N 12.95; found C 55.83, H 9.38, 12.68.

2-S-Bromomethyl-4-S-(tert-butoxycarbamido)-tetrahydropyrrole-N-acetic acid ethyl ester (7a)

Et$_3$N (4.2 mL, 0.03 mmol) was added dropwise into a mixture of 6 (4.32 g, 0.02 mmol) and ethyl bromoacetate (2.67 mL, 0.024 mmol) in CH$_2$Cl$_2$(50 mL) in an ice-bath.

After titration, the reaction mixture was warmed to r.t. for 12 h followed by addition of CBr$_4$ (7.96 g, 0.024 mmol) and PPh$_3$ (6.3 g, 0.024 mmol). Then it was left overnight at r.t. The solvent was removed under reduced pressure and the residue was purified by column chromatography using petroleum ether (60—90 ℃) : aceton = 3 : 1 as eluting solvent to give 7a as a solid (4.1 g) in yield 62%, m.p. 99—101 ℃, $[\alpha]_D^{20} + 28.2$ (c 1.0, CH$_2$Cl$_2$); ^{1}H NMR (CDCl$_3$) δ: 1.28 (t, $J = 7.2$ Hz, 3H, OCH$_2$CH$_3$), 1.43 (s, 9H, C(CH$_3$)$_3$), 2.36—2.41 (m, 1H, tetrahydropyrrol-3-H), 2.54 (d, $J = 12$ Hz, 2H, CH$_2$Br), 2.64—2.74 (m, 1H, tetrahydropyrrol-3-H), 3.07—3.19 (m, 3H), 3.36—3.39 (m, 1H, tetrahydropyrrol-4-H), 3.72—3.78 (m, 1H, N-CH$_2$CO), 4.07—4.12 (m, 1H, N-CH$_2$CO), 4.18 (q, $J = 7.2$ Hz, 2H, OCH$_2$CH$_3$), 4.55 (s, 1H, NH); IR (KBr) ν: 3367 (N—H), 1741, 1678(C=O) cm^{-1}; MS (FAB) m/z (%): 365.2 (M + 1, 85), 309.2 (M − t-Bu, 85), 285.3 (M − Br, 80), 229.2 (M − Br − t-Bu, 80). Anal. calcd for C$_{14}$H$_{25}$N$_2$O$_4$Br: C 46.04, H 6.90, N 7.67, Br 21.88; found C 46.26, H 7.04, N 7.50, Br 21.42.

2-S-Chloromethyl-4-S-(tert-butoxycarbamido-tetrahydropyrrole-N-acetic acid ethyl ester (7b)

Et$_3$N (4.2 mL, 0.03 mmol) was added dropwise into a mixture of 6 (4.32 g, 0.02 mmol) and ethyl bromoacetate (2.67 mL, 0.024 mmol) in CH$_2$Cl$_2$(50 mL) in an ice-bath. After titration, the reaction mixture was warmed to r.t. for another 12 h followed by addition of MeSO$_2$Cl (2.35 mL, 0.03 mmol). Then it was standing overnight at r.t.. The solvent was removed under reduced pressure and the residue was purified by column chromatography using petroleum ether (60—90 ℃) : aceton = 3 : 1 as eluting solvent to give 7b as a solid (4.5 g), in 71% yield, m.p. 95—97 ℃, $[\alpha]_D^{20} + 12.5$ (c 1.0, CH$_2$Cl$_2$); ^{1}H NMR (CDCl$_3$) δ: 1.26 (t, $J = 7.2$ Hz, 3H, OCH$_2$CH$_3$), 1.43 (s, 9H, C(CH$_3$)$_3$), 2.26—2.31 (m, 1H, tetrahydropyrrol-3-H), 2.54 (d, $J = 11$ Hz, 2H, CH$_2$Cl), 2.54—2.66 (m, 1H, tetrahydropyrrol-3-H), 3.17—3.22 (m, 3H), 3.46—3.59 (m, 1H, tetrahydropyrrol-4-H), 3.74—3.80 (m, 1H, N-CH$_2$CO), 4.11—4.5 (m, 1H, N-CH$_2$CO), 4.21 (q, 2H, OCH$_2$CH$_3$), 4.63 (s, 1H, NH); IR (KBr) ν: 3328 (N—H), 1732, 1680 (C=O) cm^{-1}; MS m/z (%): 321.1 (M + 1, 65), 285.1.1 (M − Cl, 100), 204.3 (M − Cl − t-Bu, 35). Anal. calcd for C$_{14}$H$_{25}$N$_2$O$_4$Cl: C 52.41, H 7.85, N 8.73, Cl 11.05; found C 52.56, H 7.55, N 8.54, Cl 10.89.

3-N-Benzoyl thymine (8)[12]

Thymine (6.3 g, 50 mmol) was stirred with a larger excess of benzoyl chloride (110 mmol) in 150 mL of acetonitrile-pyridine (5:2, V/V) at r.t. for 16 h, it was converted into its 1-N, 3-N-dibenzoyl derivative. The 1-N, 3-N-dibenzoyl derivative was treated with 0.25 mol/L potassium car-

bonate in dioxane-water (1 : 1, V/V) for 1 h to give 3-N-benzoyl thymine (8 g), in 70% yield, m.p. 148—150 ℃.

4S-(tert-Butoxycarbamido)-2S-(3-N-benzoylthymin-1-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (9)

A stirred mixture of **7a** (1.4 g, 4 mmol), T^{Br} (1.01 g, 4.4 mmol) and anhydrous K_2CO_3 in anhydrous DMF was heated at 80 ℃ for 4 h. The solution was evaporated to dryness and the residue was purified on silica gel with petroleum ether (60—90 ℃) : acetone = 3 : 1 as eluting solvent. Compound **9** was obtained as oil, which was recrystallized with ethyl acetate, giving a white solid (1.06 g), in yield 52%, m.p. 102—104 ℃, $[\alpha]_d^{20}$ -14 (c 1.0, CH$_3$OH); ^{1}H NMR (CDCl$_3$) δ: 1.28 (t, J = 7.2 Hz, 3H, OCH$_2$CH$_3$), 1.41 (s, 9H, C(CH$_3$)$_3$), 1.60—1.62 (m, 1H), 1.97 (s, 3H, thymine-5-CH$_3$), 2.35—3.18 (m, 4H), 3.35 (d, J = 16.8 Hz, 1H), 3.55 (d, J = 16.8 Hz, 1H), 3.59 (d, J = 11.4 Hz, 1H), 4.11 (d, J = 12.6 Hz, 1H), 4.19 (q, 2H, OCH$_2$CH$_3$), 4.48 (s, 1H), 7.47—7.91 (m, 5H), 7.92 (s, 1H, thymine-6H); IR (KBr) ν: 3347 (N—H), 1756, 1694, 1642 (C = O) cm^{-1}; MS (FAB) m/z (%): 515.0 (M + 1, 70), 458.9 (M – t-Bu, 15), 411 (M – Bz, 45), 336.9 (M – Bz – COOEt, 80). Anal. calcd for C$_{26}$H$_{34}$N$_4$O$_7$: C 60.69, H 6.66, N 10.89; found C 60.60, H 6.53, N 10.93.

4S-(tert-Butoxycarbamido)-2S-(cytosin-1-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (11)

Sodium hydride (60% disp. 0.5 g, 12.5 mmol) was added to a vigorously stirred suspension of cytosine (1.52 g, 13.6 mmol) in anhydrous DMF (100 mL) at r.t.. After hydrogen production ceased, the mixture was heated to 50—60 ℃ for 3 h followed by addition of **7b** (2.0 g, 6.2 mmol). Then the reaction mixture was allowed to stir for a week at 50 ℃. After filtration followed by evaporation under reduced pressure, the residue was purified by flash chromatography using ethyl acetate : methanol = 20 : 1 as the eluting solvent. Compound **11** was afforded as a white foam (0.9 g), in yield 36%, $[\alpha]_d^{20}$ -9.3 (c 1.0, CH$_3$OH); ^{1}H NMR (CDCl$_3$) δ: 1.27 (t, J = 7.2 Hz, 3H), 1.41 (s, 9H, C(CH$_3$)$_3$), 1.56—4.03 (m, 10H), 4.15 (q, J = 6.6 Hz, 2H, OCH$_2$CH$_3$), 4.96 (s, 1H, NH), 5.74 (d, J = 7.2 Hz, 1H, cytosine-5-H), 7.48 (d, J = 7.2 Hz, 1H, cytosine-6-H); IR (KBr) ν: 3363 (N—H), 1725, 1694, 1653 (C = O) cm^{-1}; MS (FAB) m/z (%): 791.3 (2M + 1, 5), 396.2 (M + 1, 35), 296 (M – Boc, 10). Anal. calcd for C$_{18}$H$_{29}$N$_5$O$_5$: C 54.67, H 7.39, N 17.71; found C 54.60, H 7.53, N 17.93.

4S-(tert-Butoxycarbamido)-2S-(4-N-benzoylcytosin-1-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (12)

A solution of benzoyl chloride (1.5 mL, 10 mmol) in CH$_2$Cl$_2$(2 mL) was carefully added dropwise to a stirred solu-

tion of **11** (1.0 g, 2.5 mmol) in anhydrous pyridine (15 mL) at 0 ℃. The reaction was allowed to warm slowly to r.t. before being left to stir overnight. Subsequently, the reaction mixture was poured into a cooled saturated solution of NaH-CO$_3$(20 mL) and extracted with ethyl acetate (3 × 20 mL). The combined organic extracts were dried over anhydrous Mg-SO$_4$. Filtration followed by solvent evaporation afforded a crude product which was purified by flash chromatography using petroleum ether (60—90 ℃) : ethyl acetate = 1 : 1 as the eluting solvent, giving **12** as a light yellow oil (0.66 g), in yield 53%, $[\alpha]_d^{20}$ +12.5 (c 1.0, CH$_2$Cl$_2$); ^{1}H NMR (CDCl$_3$) δ: 1.25 (t, J = 7.2 Hz, 3H, OCH$_2$CH$_3$), 1.42 (s, 9H, C(CH$_3$)$_3$), 1.93—4.31 (m, 10H), 4.18 (q, 2H, OCH$_2$CH$_3$), 4.32 (s, 1H), 4.64 (s, 1H), 6.47 (d, J = 7.8 Hz, 1H, cytosine-5-C-H), 7.53 (d, J = 7.8 Hz, 1H, cytosine-6-C-H), 7.35—7.53 (m, 5H, Ph); IR (KBr) ν: 3367 (N—H), 1725, 1684, 1653 (C = O) cm^{-1}; MS (FAB) m/z (%): 500.1 (M + 1, 100), 444.0 (M – t-Bu, 15), 396 (M – Bz, 20). Anal. calcd for C$_{25}$H$_{33}$N$_5$O$_6$: C 60.11, H 6.66, N 14.02; found C 60.20, H 6.53, N 14.23.

4S-(tert-Butoxycarbamido)-2S-(adenin-9-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (14)

Sodium hydride (60% disp., 0.45 g, 11.2 mmol) was added to a vigorously stirred suspension of adenine (1.55 g, 11.2 mmol) in anhydrous DMF (100 mL). The reaction mixture was stirred at 50—60 ℃ for 3 h followed by addition of **7b** (3.25 g, 10 mmol). Then the reaction mixture was allowed to stir for a week at 50 ℃. After filtration followed by evaporation under reduced pressure, the residue was purified by flash chromatography using ethyl acetate as the eluting solvent. Compound **14** was afforded as white solid (1.55 g), in yield 36%, $[\alpha]_d^{20}$ -4.5 (c 1.0, CH$_3$OH); ^{1}H NMR (CDCl$_3$) δ: 1.17 (t, J = 7.2 Hz, 3H, OCH$_2$CH$_3$), 1.28 (s, 9H, C(CH$_3$)$_3$), 1.41—3.96 (m, 10H), 4.08 (q, 2H, OCH$_2$CH$_3$), 4.29—5.04 (m, 3H), 6.65 (s, 2H, adenine-6-NH$_2$), 7.92 (s, 1H, adenine-2-H), 8.22 (s, 1H, adenine-8-H); IR (KBr) ν: 3336, 3160 (N—H), 1735, 1694 (C = O) cm^{-1}; MS (FAB) m/z (%): 420.3 (M + 1, 100), 285.0 (M – A, 25), 229 (M – A – t-Bu, 30). Anal. calcd for C$_{19}$H$_{29}$N$_7$O$_4$: C 54.40, H 6.97, N 23.37; found C 54.60, H 7.03, N 23.13.

4S-(tert-Butoxycarbonylamino)-2S-(6-N-benzoyladenin-9-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (15)

A solution of benzoyl chloride (1.5 mL, 10 mmol) in CH$_2$Cl$_2$(2 mL) was carefully added dropwise to a stirred solution of **14** (1.0 g, 2.5 mmol) in anhydrous pyridine (15 mL) at 0 ℃. The reaction was allowed to warm slowly to r.t. before being left to stir overnight. Subsequently, the reaction mixture was poured into a cooled saturated solution of NaH-CO$_3$(20 mL) and extracted with ethyl acetate (3 × 20 mL).

The combined organic extracts were dried over MgSO$_4$. After filtration and evaporation of the solvent, a crude product was obtained and was purified by flash chromatography using petroleum ether (60~90 ℃)：ethyl acetate = 1：1 as the eluting solvent, giving **15** as a light yellow oil (0.66 g), in yield 53%，[α]$_d^{20}$ +5.6 (c 1.0, DMSO); ^{1}H NMR (CDCl$_3$) δ: 1.25 (t, J = 6 Hz, 3H, OCH$_2$CH$_3$), 1.38 (s, 9H, C(CH$_3$)$_3$), 1.69~4.01 (m, 10H), 4.11 (q, 2H, OCH$_2$CH$_3$), 7.26 (s, 1H, adenine-2-H), 7.46~7.54 (m, 3H, Ph), 8.28 (s, 1H, adenine-8H), 8.33~8.46 (m, 2H, Ph); IR (KBr) ν: 3357, 3160 (N—H), 1741, 1673, 1632 (C = O) cm^{-1}; FAB-MS m/z (%): 546.3 (M + Na$^+$, 40), 524.4 (M + 1, 100), 468.4 (M – t-Bu, 10), 446.4 (M + Na$^+$ – Boc, 100), 285.2 (M – A^{Bz}, 20). Anal. calcd for C$_{26}$H$_{33}$N$_7$O$_5$: C 59.64, H 6.35, N 18.73; found C 59.60, H 6.28, N 18.90.

4S-(tert-Butoxycarbonido)-2S-(2-amino-6-O-benzylpurin-9-yl)-methyl-tetrahydropyrrole-N-acetic acid ethyl ester (19a)

18 (0.96 g, 3.7 mmol) was dissolved in anhydrous DMF (50 mL) and lithium hydride 30 mg was added. The suspension was stirred for 2 h at r.t. and then **7b** (1.0 g, 3.1 mmol) was added. The reaction was heated at 80 ℃ overnight. Solvent was removed under reduced pressure and the residue was purified by flash chromatography using petroleum ether (60~90 ℃)：ethyl acetate = 4：7 as the eluting solvent, giving **19a** (0.5 g), in yield 31% and **19b** (0.2 g), in 12%, m.p. 62~64 ℃, [α]$_d^{20}$ –64.2 (c 1.0, CH$_2$Cl$_2$); ^{1}H NMR (CDCl$_3$) δ: 1.24 (t, J = 7 Hz, 3H, OCH$_2$CH$_3$), 1.43 (s, 9H, C(CH$_3$)$_3$), 2.21~4.13 (m, 10H), 4.18 (q, J = 7.1 Hz, 2H, OCH$_2$CH$_3$), 5.07 (s, 2H, purine-2-NH$_2$), 5.54 (s, 2H, OCH$_2$C$_6$H$_5$), 5.96 (s, 1H, NH), 7.29~7.56 (m, 5H, Ph), 7.71 (s, 1H, purine-8-H); IR (KBr) ν: 3337, 3160 (N—H), 1704, 1632 (C = O) cm^{-1}; MS (FAB) m/z (%): 526.2 (M + 1, 100), 436.2 (M – Bn, 55), 380.1 (M – Bn – t-Bu, 20). Anal. calcd for C$_{26}$H$_{35}$N$_7$O$_5$: C 59.41, H 6.71, N 18.65; found C 59.60, H 6.74, N 18.36.

General procedure for preparing the compounds 10, 13, 16, 20

A 0.67 mol/L (aq.) solution of sodium hydroxide (2.5 equiv.) was added to a stirred solution of **9, 12, 15** or **19a** (1 equiv.) in methanol (2 mmol/mL) at r.t. The reaction was standing for 0.5 h before being acidified to pH 4 with 1 mol/L HCl (aq.).

4S-(tert-Butoxycarbonylamino)-2S-(thymin-1-yl)-methyl-tetrahydropyrrole-N-acetic acid (10)

The above procedure was followed using **9** (1.03 g, 2 mmol). After acidification, the solvent was removed under reduced pressure and the residue was dissolved in methanol followed by filtrating the solid. The filtrate was concentrated

under reduced pressure again and the residue was purified by flash chromatography using CH$_2$Cl$_2$：methanol = 9：1 as eluting solvent, giving **10** as a white solid (0.58 g), in yield 76%, m.p. > 153 ℃ (dec.), [α]$_d^{20}$ +41.6 (c 1.0, CH$_3$OH); ^{1}H NMR (D$_2$O) δ: 1.43 (s, 9H, C(CH$_3$)$_3$), 1.86 (s, 3H, thymine-5-CH$_3$), 1.97~4.20 (m, 8H), 4.28 (d, J = 15.6 Hz, 1H, NCH$_2$CO), 4.59 (d, J = 15.6 Hz, 1H, NCH$_2$CO), 7.5 (s, 1H, thymine-6H); IR (KBr) ν: 3554 (COOH), 3357 (N—H), 1704, 1699 (C = O) cm^{-1}; MS (FAB) m/z (%): 405.3 (M + Na$^+$, 70), 383.3 (M + 1, 85), 304.1 (M + Na$^+$ – Boc, 100). Anal. calcd for C$_{17}$H$_{26}$N$_4$O$_6$: C 53.39, H 6.85, N 14.65; found C 51.83, H 7.29, N 13.64.

4S-(tert-Butoxycarbamido)-2S-(4N-benzoylcytosin-1-yl)-methyl-tetrahydropyrrole-N-acetic acid (13)

The above procedure was followed using **12** (0.6 g, 1.2 mmol). After acidification, the solvent was removed under reduced pressure and the residue was dissolved in methanol followed by filtration of the solid. The filtrate was concentrated under reduced pressure again and the residue was purified by flash chromatography using CH$_2$Cl$_2$：methanol = 9：1 as eluting solvent, giving **13** as a white solid (0.31 g), in yield 55%，m.p. > 93 ℃ (dec.), [α]$_d^{20}$ –27.8 (c 1.0, DM-SO); ^{1}H NMR (D$_2$O) δ: 1.32 (s, 9H, C(CH$_3$)$_3$), 1.39~3.88 (m, 10H), 7.55 (d, J = 7.8 Hz, 1H, cytosine-5H), 7.59~7.97 (m, 5H, Ph), 8.45 (d, J = 7.8 Hz, 1H, cytosine-6-H); IR (KBr) ν: 3316 (COOH), 1699, 1642 (C=O) cm^{-1}; MS (FAB) m/z (%): 472.2 (M + 1, 5), 416.1 (M – t-Bu, 5), 105 (Bz, 100). Anal. calcd for C$_{23}$H$_{29}$N$_5$O$_6$: C 58.59, H 6.20, N 14.85; found C 51.83, H 4.99, N 10.64.

4S-(tert-Butoxycarbamido)-2S-(6-N-benzoyladenin-9-yl)-methyl-tetrahydro-pyrrole-N-acetic acid (16)

The above procedure was followed using **15** (1.3 g, 2.5 mmol). After acidification, the solvent was removed under reduced pressure and the residue was dissolved in methanol followed by filtration of the solid. The Filtrate was concentrated under reduced pressure again and the residue was purified by flash chromatography using CH$_2$Cl$_2$：methanol = 9：1 as eluting solvent, giving **16** as a white solid (0.66 g), in yield 53%，m.p. > 273 ℃ (dec.), [α]$_d^{20}$ –43.1 (c 1.0, DMF); ^{1}H NMR (CD$_3$OD) δ: 1.32 (s, 9H, C(CH$_3$)$_3$), 2.02~4.40 (m, 10H), 6.67 (s, 1H), 7.55~8.05 (m, 5H, Ph), 8.47 (s, 1H, adenine-2-H), 8.71 (s, 1H, adenine-8-H), 11.12 (s, 1H, COOH); IR (KBr) ν: 3305 (COOH), 1689, 1611 (C = O) cm^{-1}; MS (FAB) m/z (%): 518.0 (M + Na$^+$, 100), 496.1 (M + 1, 10), 418.0 (M + Na$^+$ – Boc, 40), 105 (Bz, 100). Anal. calcd for C$_{24}$H$_{29}$N$_7$O$_5$: C 58.17, H 5.90, N 19.79; found C 52.83, H 5.46, N 17.57.

4S-(tert-Butoxycarbamido)-2S-(2-amino-6-O-benzylpurin-9-yl)-methyl-tetrahydro-pyrrole-N-acetic acid (**20**)

The above procedure was followed using **19a** (0.7 g, 1.3 mmol). After acidification, the solvent was removed under reduced pressure and the residue was dissolved in methanol followed by filtration of the solid. Filtration was concentrated under reduced pressure again and the residue was purified by flash chromatography using CH_2Cl_2 : methanol = 9 : 1 as eluting solvent, giving **20** as a white solid (0.4 g), in yield 75%, m.p. > 105 ℃ (dec.), $[\alpha]_d^{20} + 41.6$ (c 1.0, CH_3OH); ^1H NMR (CD_3OD) δ: 1.26 (s, 9H, $C(CH_3)_3$), 1.53—4.82 (m, 12H), 5.47—5.56 (m, 2H, purine-2-NH_2), 7.29—7.51 (m, 5H, Ph), 7.81 (s, 1H, purine-8-H); IR (KBr) ν: 3336 (COOH), 1704, 1616 (C =O) cm^{-1}; MS (FAB) m/z (%) 542.5 (M + 2Na$^+$, 20), 520.4 (M + Na$^+$, 100), 498.5 (M + 1, 50), 420.5 (M + Na$^+$ – Boc, 45). Anal. calcd for $C_{24}H_{31}N_7O_5$: C 57.94, H 6.28, N 19.71; found C 46.87, H 5.23, N 15.51.

References

1 (a) Nielsen, P. E.; Egholm, M.; Buchardt, O. Science **1991**, 254, 1497.
   (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. Bioconjugate Chem. **1994**, 5, 3.
   (c) Hyrup, B.; Egholm, M.; Nielsen, P. E. Biomed. Chem. Lett. **1996**, 6, 5.
   (d) Good, L.; Nielsen, P. E. Antisense Nucleic Acids Drug Dev. **1997**, 7, 431.

2 Dueholm, K.; Nielsen, P. E. New J. Chem. **1997**, 21, 19.

3 Kim, S. H.; Nielsen, P. E.; Egholm, O.; Buchardt, O.; Berg, R. H.; Norden, B. J. Am. Chem. Soc. **1993**, 115, 6477.

4 Koch, T.; Naesby, M.; Wittung, P.; Jorgensen, M.;

Buchardt, O. Tetrahedron Lett. **1995**, 36, 6933.

5 (a) Peterson, K. H.; Jensen, D. K.; Nielsen, P. E.; Egholm, M.; Larsson, C.; Buchardt, O.; Stanley, C. J.; Norden, B.; Nielsen, P. E. Biomed. Chem. Lett. **1995**, 5, 1119.
   (b) Bergmann, F.; Bannwarth, W.; Tam, S. Tetrahedron Lett. **1995**, 36, 6823.

6 Haaima, G.; Lohse, O.; Buchardt, O.; Nielsen, P. E. Angew. Chem., Int. Ed. Engl. **1996**, 35, 1939.

7 Gangamani, B. P.; Kumar, V.-A.; Ganesh, K. N. N. Tetrahedron **1996**, 52, 15017.

8 Li, Y.; Jin, T.; Liu, K. L. Nucleosides, Nucleotides Nucleic Acids **2001**, 20, 1705.

9 Bai, J. O.; Li, Y.; Liu, K. L. Chin. J. Chem. **2001**, 19, 276.

10 Girard, F.; Demassio, C.; Lee, M. G.; Agrofoglio, L. A. Tetrahedron **1998**, 54, 8745.

11 (a) Soai, K.; Oyamada, H.; Takase, M.; Ookawa, A. Bull. Chem. Soc. Jpn. **1984**, 57, 1948.
   (b) Soai, K.; Oyamada, H.; Takase, M.; Ookawa, A. Bull. Chem. Soc. Jpn. **1984**, 57, 2327.

12 Cruickshank, K. A.; Jiricny, J.; Reese, C. B. Tetrahedron Lett. **1984**, 25, 681.

13 Howarth, N. M.; Wakelin, L. P. G. J. Org. Chem. **1997**, 62, 5441.

14 Lewis, A. F.; Revankar, G. R.; Rando, R. F. J. Heterocycl. Chem. **1995**, 32, 1513.

15 Browne, D. T.; Eisenger, J.; Leonard, N. J. J. Am. Chem. Soc. **1995**, 90, 7302.

16 Watkins, B. E.; Kiely, T. S.; Rapoport, H. J. Am. Chem. Soc. **1982**, 104, 5702.

17 Watkins, B. E.; Rapoport, H. J. Org. Chem. **1982**, 47, 4471.

18 Iki, M.; Hayashi, T. EP 0 543 095 A2, **1993**.

19 Kjellberg, J.; Johansson, N. G. Tetrahedron **1986**, 42, 6541.

20 Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L. J. Org. Chem. **1994**, 59, 5767.