Pergamon

Tetrahedron, Vol. 52, No. 47, pp. 15017-15030, 1996

0040-4020/96 $15.00 + 0.00

PII: S0040-4020(96)00924-6

Synthesis of N^{α}-(Purinyl/Pyrimidinyl acetyl)-4-Aminoproline Diastereomers with Potential Use in PNA Synthesis#

Bargur P. Gangamani, Vaijayanti A. Kumar* and Krishna N. Ganesh*

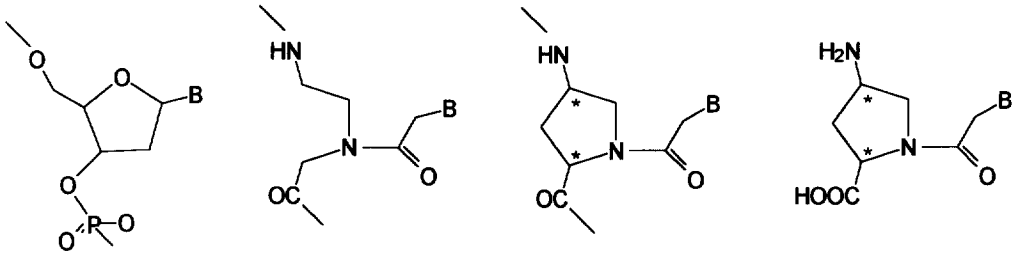
Division of Organic Chemistry (Synthesis), National Chemical Laboratory, Pune 411008, INDIA

Abstract: This paper describes an approach to introduce conformational constraint and chirality into the PNA backbone by bridging the ethylenediamine and glycine components of the same unit by a methylene group which leads to PNA based on 4-aminoprolyl backbone with chirality at C-4 and C-2. The synthesis and characterisation of all four diasteroisomers with thymine (T) as the sidechain nucleobase (**3a-d**) and the synthesis of one of the stereoisomer (2*S*, 4*R*) linked to each of the four nucleobases (**10-13**) are described. Using these monomeric units, two model dimers (**17, 18**) containing four chiral centres but differing in stereochemistry at only one site were prepared and CD data on these indicate considerable structural differences in base stacking induced by chiral backbone among these.

INTRODUCTION

Peptide Nucleic Acids (PNAs) are designed structural composites in which the nucleobases of natural DNA (**I**) are linked to a pseudopeptide backbone (**II**) of repeating units of N-(2-aminoethyl)glycine.[1] These compounds have recently emerged as novel nucleic acid mimics that efficiently bind DNA/RNA in a sequence specific manner[2] and hold potential for development as antisense therapeutic agents.[3] The extraordinary tight binding of PNA to complementary DNA and RNA has been attributed to a lack of negative charge on PNA, rigidity of the peptide backbone due to the presence of sp^2 hybridization and tertiary amide linkages. Attempts aimed at structural modifications to improve PNA properties such as binding affinity, specificity and cell permeability have centered around chemical modifications of PNA backbone,[4] and the introduction of asymmetry into achiral PNA[5-9] may influence its direction of binding with complementary DNA and RNA. The latter attribute was imparted by (i) covalent linking of a chiral amino acid,[5] an oligonucleotide[6] or a polypeptide[7] at the PNA terminus, (ii) constructing a PNA backbone with an amino acid carrying a nucleobase in the side chain[8] and (iii) incorporation of various amino acids in PNA backbone itself.[9]

An effective way to explore contribution of backbone rigidity to PNA binding with DNA and RNA, is to employ PNA monomers containing stereospecific chemical modifications. We envisaged the introduction of



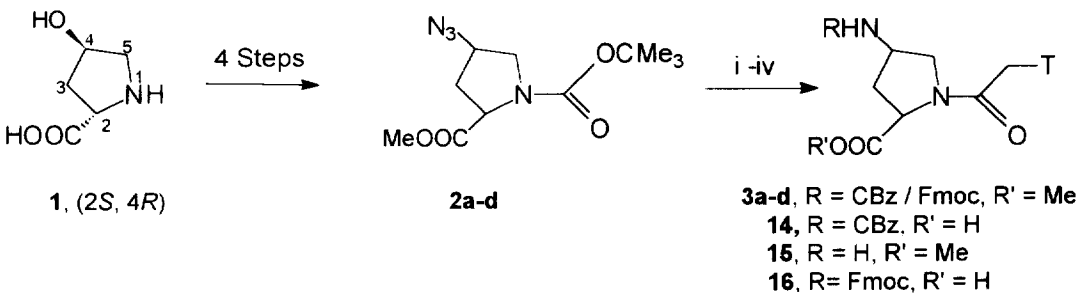
I II III IV

chemical modifications into PNA backbone in a way that not only makes them conformationally rigid but also chiral by a strategy as shown in structure **III** where the α-C of the glycine and β-C of ethylenediamine of thesame monomeric unit are bridged with a methylene group to generate five membered heterocyclic rings along the backbone. The monomer precursors to build such oligomers are 4-(*R/S*)-amino-N^{α}-(A/G/C/T-ylacetyl)-2-(*R/S*)-prolines **IV** containing two asymmetric centers at C-4 and C-2. A systematic oligomerisation of all the four possible diastereomers would lead to polychiral PNA having wide conformational diversity in the backbone.[9] In this paper, we report the preparation of appropriately protected T-monomers of all four diastereomers (**3a-d**) from a single starting material 4*R*-hydroxy-2*S* proline (**1**) and the synthesis of one stereoisomer (2*S*, 4*R*) linked to each of the four nucleobases (A, G, C and T) via a common precursor. Further, in order to demonstrate the utility of such monomers, synthesis of two of the derived prolyl-PNA dimers **17** and **18**, each containing four chiral centers but differing in stereochemistry at only one site, is described. Preliminary spectroscopic studies indicate that considerable differences in base stacking are induced by configurational change even at a single site in the backbone.

RESULTS AND DISCUSSION

The naturally occurring 4*R*-hydroxy-2*S*-proline (**1**) was transformed to the designed targets **3a-d** by four independent pathways (**A-D**), each involving a similar set of reactions, but differing in the order of inversion steps at C-4 and C-2 (Scheme-1). Path **A** consisted of esterification of carboxyl group in **1** and N^{α}-protection with BOC, followed by conversion into 4*R*-tosyl derivative.[10a] Treatment of this with NaN_3 in DMF resulted in S_N2 displacement (inversion) of 4*R*-tosyl group by N_3 to 4*S*-N_3 derivative[10b] (**2a**) that was reduced to 4*S*-NH_2 compound which was subsequently protected with either CBz or Fmoc group.[11] Deprotection of N1 followed by DCC-HOBT coupling with thymin-1-ylacetic acid[12] gave the fully protected 4*S*,2*S* T-monomer (**3/3a**). One of its epimer **2b** (4*R*,2*S* derivative) was obtained by path **B** which involved two successive inversion steps: first tosylation of 4*R*-OH function under Mitsunobu conditions[10a] to yield 4*S*- OTs derivative

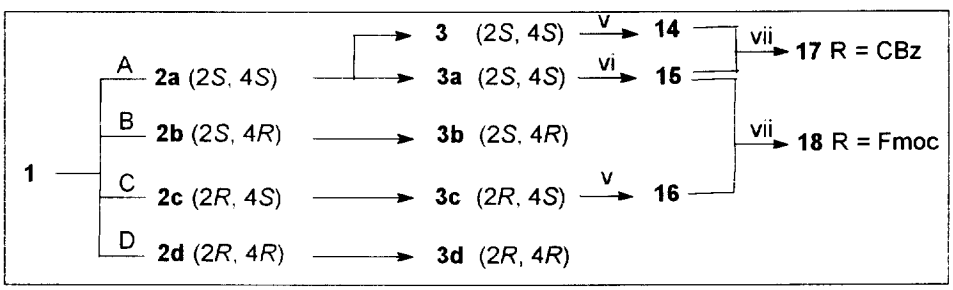
Scheme 1

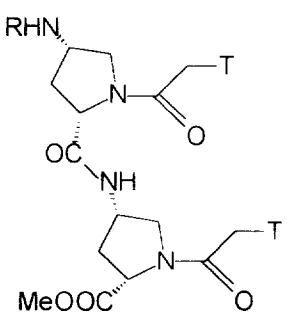


1, (2*S*, 4*R*)

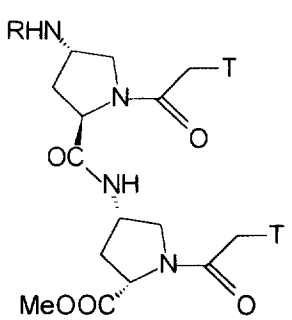
2a-d

3a-d, R = CBz / Fmoc, R' = Me
14, R = CBz, R' = H
15, R = H, R' = Me
16, R= Fmoc, R' = H





17 R = CBz
i
17a R = H



18 R = Fmoc
vi
18a R = H

i. H_2 / Pd C ii. FmocCl / CBzCl, Na_2CO_3 iii. TFA:DCM (1:1) iv. Thymin-1-ylacetic acid, DCC HOBT
v. 1M NaOH vi. Piperidine:DMF (2:8), vii. DCC HOBT

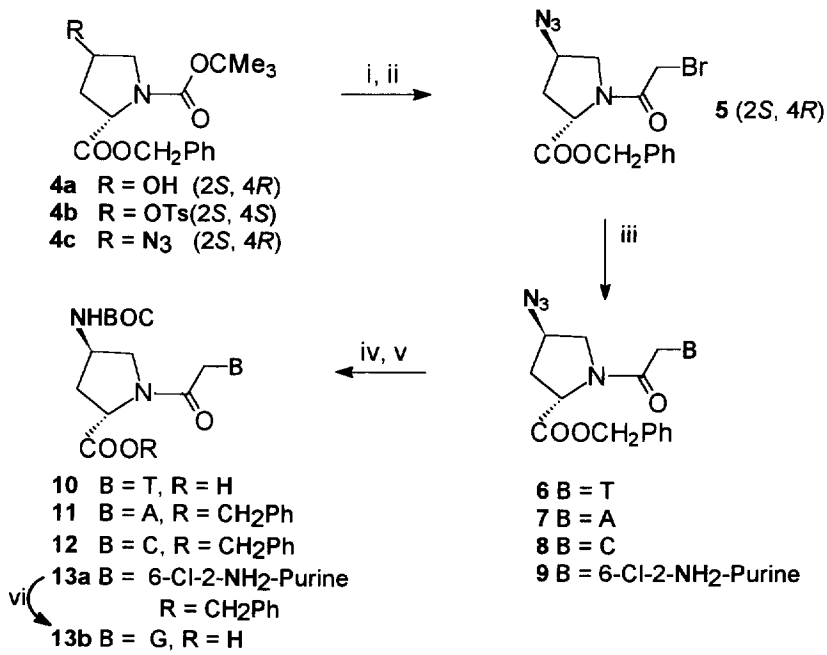
which in subsequent reaction with NaN_3 underwent an inversion to yield the azide 4*R*-N_3 (**2b**), with an overall retention of configuration at C4. This was transformed to the fully protected 4*R*,2*S* T-monomer (**3b**) by a similar set of reactions as described for 4*S*,2*S* isomer (**3a**). Paths **C** and **D** leading to 4(*S*/*R*),2*R*-derivatives (**3c,d**) begin with an inversion step at C-2 by treatment of **1** with AcOH-Ac_2O to obtain 4*R*-hydroxy-2*R*-proline[13], followed by an identical set of reactions as in paths **A** and **B** respectively. Thus all possible stereoisomers of 4-amino-N^{α}-thymin-1-ylacetyl prolines which are the T-monomeric units for the designed 4-aminoprolyl nucleic acids were obtained starting from **1** with overall yields in the range of 28-30%.

Although the above strategy involving condensation of prolyl ring NH with N-carboxy-methylated nucleobase was successful for synthesis of T-monomers, it is fraught with experimental difficulties for synthesis of monomers of other nucleobases, in particular that of purines.[12] We therefore sought to employ an alternative route towards a general synthesis of prolyl nucleic acid monomers which consisted of direct coupling of individual nucleobases (T, C, G and A) with suitably protected N^{α}-(bromoacetyl) proline (**5**). Such a strategy has been reported to be efficient in large scale synthesis of monomers of open chain peptide nucleic acids.[14] In the present work, from a synthetic point of view, the 4*R*-azido-N^{α}-(bromoacetyl) proline benzyl ester **5** is convenient, since in a subsequent catalytic hydrogenation step, simultaneous transformation of 4-azido to 4-amino and deprotection of 2-benzyl ester to carboxylic acid can be achieved. It is also possible to selectively reduce 4-azido to 4-amino by catalytic transfer hydrogenation using ammonium formate/Pd-C to obtain the C-protected monomeric units. With this working rationale, we attempted the synthesis of prolyl nucleic acid monomers corresponding to A,G,C and T nucleobases by this alternative method and illustrate the feasibility of the method by a synthesis of all 4 monomers (**10-13**) corresponding to 4*R*,2*S* series (Scheme II).

4*R*-Hydroxyproline was converted into its benzyl ester by reaction with benzyl alcohol and *p*-toluenesulphonic acid[11] followed by protection of N^{α} with BOC[11] to get **4a** which was transformed in two further steps to the 4*R*-azido derivative **4c**. Deblocking of N^{α}-Boc with TFA afforded the free amine, which upon treatment with bromoacetylchloride in presence of aq. sodium bicarbonate, afforded the desired common intermediate **5**. This reaction gave poor yields when triethylamine was employed as a base.[14] The individual alkylation of the four nucleobases thymine, adenine, cytosine and 2-amino-6-chloropurine with **5** was done by reaction of the two components in DMF in presence of a suitable base, followed by purification by chromatography to obtain **6**, **7**, **8** and **9** respectively, in 50 to 80% yield. While the alkylation of thymine and cytosine is regiospecific, providing only N1 substitution products, that of purines is known to be non regiospecific leading to significant N7 substituted product along with the desired major product from N9-substitution.[14] In the present work with adenine, no N7 regioisomer was observed as determined by NMR (see later). For synthesis of G-monomer, direct alkylation of guanine is not a facile reaction and hence 2-amino-6-chloropurine was alkylated with **5** to obtain exclusively the N9 regioisomer **9**. 4*R*-Azido group in

6-9 was then hydrogenated and protected with BOC[11] to get the protected monomers **10-13a** respectively. **13a** was then hydrolysed to the desired G-monomer **13b** by 1N NaOH.[14] This formally completed the synthesis of

Scheme 2



i, TFA:DCM (1:1); ii, $BrCH_2COCl$/ Na_2CO_3; iii, Thymine (T) and 6-Cl-$2NH_2$ purine/K_2CO_3, Cytosine (C) and Adenine (A)/NaH; iv, ammonium formate/Pd-C, v, BOC-N_3, Na_2CO_3, vi, 1M NaOH

all the nucleobase 4*R*,2*S* monomers required for building prolyl nucleic acids. This methodology coupled with that in Scheme-1 enables the synthesis of all possible 4-amino proline stereoisomers, corresponding to N1 (T and C) and N9 (A and G) substitution with the four nucleobases.

All compounds including those obtained as the intermediates in pathways **(A-D)** and in Scheme-2 were unambiguously characterized for structural purity by ^{1}H and ^{13}C NMR. The NMR spectra of N1-acyl substituted compounds (both BOC and carboxymethylated nucleobases) showed characteristic presence of *syn* and *anti* rotamer mixtures originating from the restricted rotation about the tertiary amide bond (>N-CO). The energy barriers for such rotation are high, leading to resolution of *syn* and *anti* rotamers at ambient

temperature on NMR time scale, as shown by a doubling of relevant resonances. An equilibrium composition of approximately 80:20 was noticed in most of the compounds in $CDCl_3$, although the assignment of specific resonances to individual rotamers was not possible in the present compounds due to complex and overlapping spectral patterns for prolyl ring protons in 200MHz spectra.

^{13}C NMR is valuable in identification of N7/N9 regioisomers in alkylation reactions of purines. In case of 7- and 9- methyladenines it has been established that the N9 isomer exhibits downfield shifts of 6 and 4 ppm for C-5 and C-6 respectively and upfield shifts of 10 and 4.5 ppm for C-4 and C-6 compared to corresponding chemical shifts of N7 regioisomer.[15] This pattern was also noticed with A-monomer of peptide-nucleic acids[14]. The ^{13}C NMR spectral data of compounds **7** and **11** prepared presently by alkyation of adenine was consistent with the data reported for N9 isomer. Further, the data for compound **7** prepared by two different routes (alkylation of adenine by **5** or by condensation of 4*R*-azido-2*S* proline benzyl ester with adenine-9-yl acetic acid[12]) was in complete agreement, thus supporting the assigned structures.The chiral purity of the synthesised compounds was ascertained by following optical rotation measurements at every step. The enantiomeric pairs **(2a/2d)/(2b/2c)** and **(3a/3d)/(3b/3c)** exhibited equal but opposite optical rotations, thus establishing the optical purity of compounds **2a-d** and **3a-d**.

Synthesis and characterization of model prolyl-nucleic acid chiral dimers **17** *and* **18**

To examine the efficacy of synthesised monomers in polymer building reactions, the syntheses of dipeptides **17** and **18** were accomplished from **3, 3a** and **3c** (Scheme 1). This was done first by hydrolysis of **3** and **3c** with alkali to yield the carboxylic components **14** and **16**. C-4 amino deprotection of **3a** with piperidine gave the common amino component **15** which was then condensed with **14** and **16** in presence of DCC-HOBT[11] to give the prolyl PNA dipeptides **17** (NH-4*S*,2*S*;4*S*,2*S*-OH) and **18** (NH-4*S*,2*R*;4*S*,2*S*-OH) respectively. The structural integrity of these two dipeptides was established by ^1H NMR and the diagnostic mass peaks (M+1) at 723 and 811 respectively in FAB-MS. The dipeptides **17** and **18** upon deprotection with appropriate reagents gave the free amines **17a** and **18a** which were characterised by ^1H NMR .

For a preliminary examination of the structural effects induced by different relationships of the chiral centres in dipeptides **17a** and **18a**, the CD spectra of the prolyl PNA are shown in Figure 1. These show characteristically different Cotton effects in the region 240-300 nm, with **17a** having only a negative CD band at 265 nm and **18a** both negative and positive CD bands at 278 and 255 nm respectively. These bands though weak, are completely absent in open chain PNA molecules[5] and must arise in **17a** and **18a** due to asymmetry in nucleobase stacking induced by the chiral peptide backbone. Substantial differences are also noticed in relative CD patterns in the region 210-230 nm, originating from stereochemical differences in the peptide linkage. The two dipeptides differ in prolyl ring stereochemistry at only one centre and it is remarkable that this single change affects the sign, amplitude and band position of the CD spectrum. This is a clear evidence

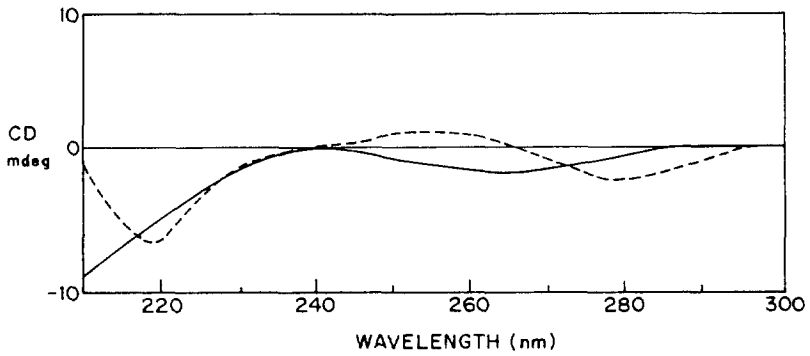


Figure 1. CD spectra of **17a** (___) and **18a** (----) in Tris buffer (10 mM, pH 7.2) at 25°C

for an ordered conformation of bases in dimers which senses the backbone chirality. In addition, stacking of bases could also be a function of syn/anti rotamer population due to restricted rotation around >N1-CO bond.

CONCLUSION

In this paper, we have demonstrated the design and preparation of stereochemically defined monomer building blocks, based on 4-aminoproline backbone, which are required for the synthesis of conformationally rigid, polychiral PNA. The possible effects of stereostructural constraints on backbone is preliminarily seen as induced chirality in base stacking in the model dipeptides which exhibit significant differences in their CD patterns upon change of stereochemistry even at a single site on backbone. Achiral PNA (**II**) has enormous structural freedom to assume a variety of conformations in solution (arising from the intervening methylene groups in the backbone) with a high statistical probability of those productive in DNA binding. The details of conformational preferences of PNA for efficient DNA binding are not yet deciphered in literature[16,17] and the availability of PNAs with predefined rigid conformations such as those derived from **IV** would be very valuable to delineate these factors. The conformationally defined building blocks reported here along with different nucleobases (A,T,C and G) when used in appropriate rational combinations, would lead to a library of chiral PNAs with diverse backbone geometry, that are useful for hybridisation screening with appropriate DNA/RNA targets. We are currently engaged in solid phase synthesis of homooligomers from these chiral monomeric units to generate polychiral PNAs for evaluating their biophysical properties.

EXPERIMENTAL

Solvents were obtained from standard commercial sources, distilled and dried before use. Nucleobases, 4-hydroxyproline, t-butoxycarbonylchloride, Fluorenemethoxycarbonylchloride, benzyloxy carbonyl chloride, dicyclohexylcarbodiimide and hydroxybenzotriazole were obtained from Aldrich chemical company. ^{1}H and ^{13}C NMR spectra were recorded using Bruker AC 200 MHz FTNMR spectrometer. Optical rotations were measured on JASCO DIP-181 polarimeter and CD spectra were obtained on a JASCO J600 instrument. Silica gel column chromatographic analyses were followed by pre-coated silica gel TLC plates (E. Merck 5554) and visualised by ninhydrin spray (amino compounds). All TLCs were run using DCM containing appropriate amounts of methanol.

4-Azido-N$^{\alpha}$-(tertbutoxycarbonyl) proline methyl ester 2a-d: N$^{\alpha}$-(tert-butoxycarbonyl-4-(*p*-toluene sulfonyloxy) proline methyl esters (4.0 g, 10 mmole) of appropriate stereochemistry (4*R*,2*S*; 4*S*,2*S*; 4*R*,2*R* and 4*S*,2*R* synthesised according to ref 10), were dissolved in dry DMF (20 mL) to which was added NaN_3 (1.2 g). The reaction mixture was stirred at 50^{0} C for 5 h after which the solvent was removed under reduced pressure and the residue worked up with ether. The products obtained **2a-d** respectively were purified by silica gel column chromatography using pet. ether containing increasing amounts of ethyl acetate.

Compound **2a:** ^{1}H NMR, $CDCl_3$, δ, 4.43 (dq, J=6.0, 8.8, 5.3 Hz, H-2), 4.33 (brm, 1H, H-4), 3.76 (s, 3H, OCH_3), 3.74 (m, 1H, H-5a), 3.48 (m, 1H, H-5b), 2.47 (m, 1H, H-3a), 2.17 (dt, J=4.3, 1.3 Hz, 1H, H-3b), ^{13}C NMR, $CDCl_3$ δ, 172.2 (ma) and 171.9 (mi) (COOMe), 154.0 (mi) and 153.4 (ma) (OCON<), 80.5 (CMe_3), 59.3 (mi) and 58.3 (ma) (C-4), 57.7 (ma) and 57.4 (mi) (C-2), 52.3 (OMe), 51.3 (mi) and 50.8 (ma) (C-5), 36.0 (ma) and 35.1 (mi) (C-3), 28.3 [$C(CH_3)_3$], $[\alpha]_D^{20}$ = -31.3 (c = 0.4, MeOH)

Compound **2b,** ^{1}H NMR, $CDCl_3$, δ, 4.36 (dt, J = 6.9, 7.6 Hz), 4.19 (m, 1H, H-4), 3.74 (s, 3H, OCH3), 3.64 (dd, J = 5.3, 12 Hz, H-5a), 3.47 (dt, J = 3.8, 12 Hz, H-5b), 2.32 (m, 1H, H-3a), 2.16 (m, 1H, H-3b), 1.46 (ma) and 1.41 (mi) [s, 9H, $C(CH)_3$]. ^{13}C NMR, $CDCl_3$, δ 172.5 (ma) and 172.2 (mi) (CO), 153.6 (mi) and 152.9 (ma) (NCO), 80 ($CCH_3)_3$, 59.04 (mi) and 58.5 (ma) (C-4), 51.7 (OCH_3), 51.1 (mi) and 50.9 (ma), C-5, 35.9 (ma) and 35.0 (mi) (C-3), 27.8 (C-CMe_3), $[\alpha]_D^{20}$ -60.9 (c = 0.4, MeOH)

Compound **2c**, ^{1}H NMR and ^{13}C NMR same as 2b, $[\alpha]_D^{20}$ +61.2 (c = 0.4, MeOH)

Compound **2d**,^{1}H NMR and ^{13}C NMR same as 2a, $[\alpha]_D^{20}$ +32.0 (c = 0.4, MeOH)

4N-(Fluorenylmethoxycarbonyl)-N$^{\alpha}$-(thymin-1-ylacetyl) proline methyl ester 3a-d: The compounds **2a-d** of stereochemistry 4*S*,2*S*; 4*R*,2*S*; 4*S*,2*R* and 4*R*,2*R* respectively were converted to **3a-d** by the following set of reactions: (i) hydrogenation of substrate (1.4g) over Pd-C (0.15 g) at 30 psi in MeOH for 4 h to yield 4-aminoproline derivatives, (ii) conversion of these to corresponding 4-N-(fluorenylmethoxycarbonyl)-N-(t-

benzyloxycarbonyl chloride (0.34 g, 2 mmole) in a mixture of dioxane (3 mL) and 10% aq. sodium carbonate (5 mL) and stirring at ambient temperature for 6 h. (iii) deprotection of N-BOC with 50% TFA:DCM[11] (1 h, 25° C) and (iv) condensation of resulting 4-N(fluorenylmethoxycarbonyl) proline methyl esters (2.0 g, 5 mmole) with thymine-1-ylacetic acid (1.1 g, 6 mmole) in dry DMF:DCM (1:1, 40 mL) in presence of DCC (1.6 g, 6.5 mmole) and HOBT (0.75 g, 6 mmole) at 4^0C for 24 h. The precipitated DCU was filtered off, washed with DCM and the residue after concentration was worked up in DCM (2x50 mL) by successively washing with aq. $NaHCO_3$ (2x20 mL), aq. $KHSO_4$ (20 mL), brine (20 mL) and water (20 mL). The dried organic extract upon concentration and column purification (elution DCM:MeOH) gave the products **3a-d** in 30% overall yield.

Compound **3**: ^{1}H NMR, $CDCl_3$ δ 9.02 (mi) & 8.95 (ma) (s, 1H,*T* NH), 7.4 (s, 5H, Ar-H), 7.09 (mi) & 7.0 (ma) (s, 1H, *T* H-6), 6.0 (d, 1H, C-4 NH), 5.1 (s, 2H, $ArCH_2$), 4.65 (d, 1H, J = 16 Hz, *T* CHa) 4.5-4.44 (m, 2H, H-2, H-4), 4.25 (d, 1H, J = 16 Hz, *T* CHb), 4.0-3.5 (m, 2H, H-5), 3.75 (ma), 3.7 (mi) (s, 3H, OCH_3), 2.5 (ma) & 2.35 (mi) (m, 1H, H-3b), 2.05 (mi) & 2.0 (ma) (br, 1H, H-3a), 1.9 (s, 3H, *T* CH_3) ^{13}C NMR, $CDCl_3$ δ 155.8 (*T* C-2), 151.4 (*T* C-4), 141.2 (*T* C-6), 110.5 (*T* C-5), 58.0 (OCH_3), 52.67 (C-2), 48.7 (*T* CH_2), 50.9 (C-4), 43.6 (C-5), 35.0 (C-3), 12.2 (*T* CH_3).

Compound **3a**: ^{1}H NMR, $CDCl_3$ δ, 8.62 (s, 1H, *T* NH), 7.76-7.26 (m, 8H, Fmoc), 7.07 (mi) & 7.05 (ma) (s, 1H, *T* H-6) 5.98 (ma) (d, J = 11 Hz, C-4 NH) & 5.3 (mi, br, C-4 NH), 4.73-4.16 (m, 7H, H-2, H-4, *T* CH_2, Fmoc CH and OCH_2), 4.0-3.57 (br-m, 2H, H-5), 3.78 (ma) & 3.76 (mi) (s, 3H, OCH_3), 2.5 (br-m, 1H, H-3b), 2.04 (m, 1H, H-3a) & 1.93 (s, 3H, *T* CH_3). ^{13}C NMR, $CDCl_3$ δ, 155.8 (*T* C-2), 151.3 (*T* C-4), 141.2 (*T* C-6), 110.4 (*T* C-5), 57.9 (OCH_3), 52.6 (C-2), 51.9 (*T* CH_2), 50.8 (C-4), 48.6 (C-5), 34.5 (C-3), 12.2 (*T* CH_3). $[\alpha]_D^{20}$ (c = 0.4, MeOH) = -38.8°, FAB-MS:[M+1] = 533.

Compound **3b**: ^{1}H NMR $CDCl_3$ δ,: 9.75 (s, 1H, T NH), 7.76-7.25 (m, 8H, Fmoc), 7.02 (mi) & 6.96 (ma) (s, 1H, *T* H-6), 6.34 (ma) & 6.12 (mi) (d, 1H, J = 7.4 Hz, C-4 NH), 4.82-3.89 (m, 7H, H-2, H-4, *T* CH_2, Fmoc CH, OCH_2), 3.72 (s, 3H, OCH_3), 3.8-3.66 (m, 2H, H-5), 2.44 (mi) & 2.17 (ma) (br-m, 2H, H-3a,b), 2.1 (s, 3H, *T* CH3). ^{13}C NMR $CDCl_3$ δ, 156.1 (*T* C-2), 151.6 (*T* C-4), 141.3 (*T* C-6), 110.8 (*T* C-5), 57.9 (OCH_3), 52.6 (C-2), 51.4 (*T* CH_2), 50.7 (C-4), 48.9 (C-5), 35.0 (C-3), 12.2 (*T* CH3), $[\alpha]_D^{20}$ (c = 0.4, MeOH) = +22.5°, FAB-MS: [M+1] = 533.

Compound **3c**: ^{1}H NMR ($CDCl_3$) δ: 9.82 (s, 1H, *T* NH), 7.76-7.25 (m, 8H, Fmoc), 7.02 (mi) & 6.94 (ma) (s, 1H, *T* H-6), 6.4 (ma) & 6.18 (mi) (d, J = 7.4 Hz, C-4 NH), 4.83-3.89 (m, 7H, H-2, H-4, *T* CH_2, Fmoc CH, CH_2), 3.71 (ma) & 3.73 (mi) (s, 3H, OCH_3), 3.78-3.72 (m, 2H, H-5), 2.44 (mi) & 2.22 (ma) (brm, 2H, H-3a,b), 2.11 (s, 3H, *T* CH_3). ^{13}C NMR ($CDCl_3$) δ: 156.4 *T* C-2), 151.9 (*T* C-4), 141.5 (*T* C-6), 111.3 (*T* C-5),

58.2 (OCH_3), 52.9 (C-2), 51.9 (*T* CH_2), 51.0 (C-4), 49.2 (C-5), 35.4 (C-3), 12.6 (*T* CH_3). $[\alpha]_D^{20}$= -22.8 (c = 0.4, MeOH), FAB-MS: [M+1] = 533.

Compound **3d:** ^{1}H NMR , $CDCl_3$ δ: 7.76-7.26 (m, 8H, Fmoc), 7.08 (mi) & 7.0 (ma) (s, 1H, *T* H-6), 6.05 (ma) & 5.4 (mi) (d, 2H, J = 7.4 Hz, C-4 NH), 4.75-4.15 (m, 7H, H-2, H-4, *T* CH_2, Fmoc CH, CH_2), 3.98-3.56 (br-m, 2H, H-5) 3.77 (ma) & 3.75 (mi) (s, 3H, OCH_3), 2.5 (br-m, 2H, H-3a,b), 1.93 (s, 3H, *T* CH_3). ^{13}C NMR $CDCl_3$ δ: 156.0 (*T* C-2), 151.6 (*T* C-4), 141.4 (*T* C-6), 110.6 (*T*-C5), 58.1 (OCH_3), 52.8 (C-2), 52.0 (*T* CH_2), 51.0(C-4), 48.9 (C-5), 34.8 C-3), 12.3 (*T* CH_3). $[\alpha]_D^{20}$ = +38.63, (c = 0.4, MeOH), FAB-MS: [M+1]=533.

4*S*-(*p*-Toluenesulfonyl)-N$^\alpha$-(tertbutoxycarbonyl) proline benzyl ester 4b: N-(t-Butoxycarbonyl)-4*R*-hydroxy-2*S* proline benzylester[11] **4a** was converted to corresponding 4*S*-tosyl derivative **4b** under Mitsunobu conditions using triphenyl phosphine (9.2 g), DEAD (6.2 g) in THF (300 mL) and methyl-*p*-toluenesulphonate (6.6 g) in THF (60 mL) The product was purified by column chromatography, eluting with pet ether-EtOAc to obtain **4b** in 60% yield. ^{1}H NMR, $CDCl_3$, δ, 7.25-7.85 (q, 9H, overlapping signals ArH), 5.3 (m, 3H, overlapping signals for $ArCH_2$ and H-4), 4.45 (m, 1H, H-2), 3.65 (m, 2H, H-5a,b), 2.45 (s, 3H, $ArCH_3$), 2.3-2.55 (m, 2H, H-3), 1.38 and 1.28 [2 x s, 9H, $C(CH_3)_3$]

4*R*-Azido-N$^\alpha$-(tertbutoxycarbonyl)-2*S*-proline benzyl ester 4c: Compound **4b** (4.6 g, 10 mmole), was dissolved in DMF (25 mL) containing NaN_3 (1.2 g, 20 mmole) and stirred at 50^0 C for 3 h. Work up and column chromatography afforded **4c** (2.8 g, 80%) ^{1}H NMR, ($CDCl_3$), δ, 7.35 (s, 5H, ArH), 5.15 (m, 2H, $ArCH_2$), 4.45 (mi) and 4.34 (ma) (2 x t, 1H, J = 7.5 Hz, H-2), 4.13 (brm, 1H, H-4), 3.66 (m, 1H, H-5a), 3.56 (ma, dd, J=2.8, 11.4 Hz, H-5b), 3.42 (mi, dd, J=3.5, 11.4 Hz, H-5a), 3.56 (ma, dd, J=2.8, 11.4 Hz, H-5b), 3.42 (mi, dd, J=3.5, 11.4 Hz, H-5b), 2.3 (m, 1H, H-3a), 2.13 (m, 1H, H-3b).

4*R*-Azido-N$^\alpha$-(bromoacetyl) proline benzyl ester 5: Compound **4c** (3.5 g) was treated with 50% TFA in DCM for 1 h and the residue obtained after evaporation of solvent was dissolved in dioxane (15 mL) containing 10% aq. Na_2CO_3 (25 mL, pH 8.0). The mixture was cooled in ice bath for addition of bromoacetyl chloride (7 equiv). The pH was then adjusted to 8.0 by addition of aq. Na_2CO_3. The mixture was concentrated to half its volume and extracted with DCM. The dried organic layer upon concentration and column purification gave **5** (3.1 g, 80% yield) ^{1}H NMR, ($CDCl_3$) (2 rotamers), δ, 7.37 (s, 5H, ArH), 5.16 (m, 2H, $ArCH_2$), 4.72 (mi) and 4.63 (ma) (2xt, 1H, J = 7.0 Hz, H-2), 4.3 (m, 1H, H-4), 3.92 (m, 1H, H-5a), 3.65 (m, 1H, H-5b), 3.81 (s, 2H, CH_2Br), 2.36 (m, 1H, H-3a), 2.18 (m, 1H, H-3b)

4*R*-Azido-N$^\alpha$-(thymin-1-ylacetyl) proline benzyl ester 6: A mixture of **5** (1.25 g, 0.3 mmole), thymine (0.6 g, 0.3 mmole) and solid K_2CO_3 (0.7 g, 3 mmole) in dry DMF (20 mL) was stirred at 25°C for 3 h. The residue obtained after aqueous work up was purified by column chromatography by elution with DCM-MeOH to obtain **6** (yield 90%), ^{1}H NMR, ($CDCl_3$) δ, 9.45 (brs, 1H, NH), 7.38 (mi) and 7.32 (ma) (s, 5H, Ar-H), 7.05 (ma) and 6.85 (mi) (s, 1H, T-CH), 5.2 (m, 2H, $ArCH_2$), 4.84 (mi, q, J = 5.8, 7.8 Hz, H-2), 4.65 (ma, q, J = 13.7, 2.0 Hz, H-2), 4.6 (d, 1H, J=16.2, *T* C-Ha), 4.37 (d, 1H, J=16.2 Hz, *T* C-Hb), 4.35 (ma) and 4.2 (mi) (m,

1H, H-4), 3.92 (dd, 1H, J = 11.6, 5.8 Hz, H-5a), 3.7 (dt, J = 11, 4.0 Hz, 1H, H-5b), 2.35 (m, 2H, H-3a.b), 1.95 (s, 3H, *T* CH_3).

4*R*-Azido-N^{α}-(adenin-9-ylacetyl) proline benzyl ester 7, 4*R*-azido-N^{α}-(cytosin-1-ylacetyl) proline benzyl ester 8 and 4*R*-Azido-N^{α}-(6chloro-2-aminopurin-9-ylacetyl) proline benzyl ester 9: These were synthesised from **5,** appropriate nucleobase and either NaH (for *C* and *A*) or K_2CO_3 (for 2-amino-6-chloropurine) by following a similar procedure as above. The products **7, 8** and **9** were purified by column chromatography by elution with DCM-MeOH (yields 50-70%).

Compound **7**, ^{1}H NMR,($CDCl_3$) δ 8.31 (ma) and 8.3 (mi) (s, 1H, A H-8), 7.93 (ma) and 7.81 (mi) (s, 1H, A H-2), 7.38 (mi) and 7.33 (ma) (s, 5H, ArH), 6.36 (ma) and 6.26 (mi) (brs, Ad NH_2), 5.25 (mi) and 5.15 (ma) (2 x dd, J =12.5 Hz, 2H, $ArCH_2$), 4.95 (ma, s, A CH_2) and 4.92 (mi, t, J = 7.5 Hz, *A* CH_2), 4.65 (t, 1H, J = 9 Hz, H-2), 4.35 (ma) and 4.18 (mi) (m, 1H, H-4), 3.92 (q, 1H, J = 5.2, 10 Hz, H-5a), 3.71 (q, 1H, J=2.5, 10 Hz, H-5b), 2.35 (m, 2H, H-3a,b) ^{13}C NMR, ($CDCl_3$) δ, 170.6 (COOBn), 165.2 (mi) and 164.7 (ma) (>NCO), 155.5 (*A* C-6), 152.7 (*A* C-2), 149.7 (*A* C-4), 141.1 (*A* C-8), 134.9-127.9 (Ar), 118.5 (*A* C-5), 68.02 (mi) and 67.1 (ma) ($ArCH_2$), 59.3 (C-2), 57.9 (ma) and 57.2 (mi) (C-4), 51.3 (*A* CH_2), 44.3 (ma) and 43.8 (mi) (C-5), 36.4 (mi) and 34.08 (ma) (C-3), $[\alpha]_D^{20}$ -44.0 (c = 0.3, MeOH)

Compound **8**, ^{1}H NMR, ($CDCl_3$+D_2O), δ 7.45-7.15 (m, 6H, ArH and *C* H-6), 5.9 (d, J = 8.0 Hz, *C* H-5), 4.95-5.53 (m, 2H, $ArCH_2$), 4.82-4.45 (overlapping multiplets, 3H, N-CH_2, H-2), 4.35 (brm, 1H, H-4), 3.95 (brm, 1H, H-5b), 3.7 (brm, 1H, H-5a), 2.05-2.55 (overlapping multiplet, 2H, H-3a,b), $[\alpha]_D^{20}$ -51.3 (c = 0.3, MeOH)

Compound **9**, ^{1}H NMR, ($CDCl_3$), δ 8.0 (mi) and 7.8 (ma) (s, H-8), 5.31 (brs, 2H, NH_2), 5.23 (mi, m, $ArCH_2$), 5.11 (dd, J=12.2, $ArCH_2$), 4.6 (mi, m, NCH_2) and 4.61 (ma, t, J = 7.6 Hz, NCH_2), 4.35 (ma, m, H-4) and 4.2 (mi, m, H-4), 3.89 (ma, dd, J = 5.4, 10.2 Hz, H-5a) and 3.81 (mi, dd, J = 3.1, 12.9 Hz, H-5a), 3.64 (dd, J = 3.3, 10.1 Hz, H-5b), 2.4 (mi, m, H-3a), 2.37 (ma, dq, J = 14.0, 7.5 Hz, H-3a), 2.17 (dq, J = 14, 7.5 Hz, H-3b), $[\alpha]_D^{20}$ -42.5 (c = 0.4, MeOH).

4*R*-N-(ter*t*butoxycarbonyl)-N^{α}-(thymin-1ylacetyl) proline, 10: Compound **6** (2 g) in MeOH was hydrogenated under pressure (30 psi) using Pd-C (10-12%, 200 mg) as catalyst for 15 h. The catalyst was filtered and the free amino acid product **10,** was isolated by evaporation of solvent. This compound (1.4 g, 5 mmole) was dissolved in dioxane:water (1:1, 20 mL) and treated with tertbutoxycarbonyl azide (0.8 mL) and kept stirred at 25° C for 24 h. with pH maintained at 9.0 by addition of 4N NaOH. The reaction mixture was concentrated to 10 mL, neutralised with Dowex 50 H^+ and filtered. Removal of solvent from the filtrate afforded the required product which was purified by crystalisation from MeOH. ^{1}H NMR, (D_2O), δ 7.39 (d, J = 1.2 Hz, *T* H-6), 4.7 (t, J = 12.1 Hz, N-CH_2), 4.52 (t, J = 9.2 Hz, H-2), 4.36 (m, 1H, H-4), 4.2 (mi) and 3.95 (ma) (dd, H-5b), 2.25-2.5 (bt, 2H, H-3a,b), 1.9 (s, 3H, *T* CH3), 1.46 [s, 9H, $C(CH_3)_3$], $[\alpha]_D^{20}$ -14.7 (c = 0.3, MeOH).

4*R*-N-(tertbutoxycarbonyl)-N^{α}-(A/G/C-ylacetyl) proline benzyl ester 11-13a: The 4*R*-azido compounds **7-9** (1.0 g) were individually dissolved in methanol (50 m L) and reacted with ammonium formate (4 eq) and Pd-

C (10-12%) for 24 h. after which the catalyst was removed by filtration over celite. The filtrate was evaporated under vacuo and the residue dissolved in water:dioxane (1:1, 15 mL), was treated with 1-butoxycarbonylazide (1.5 eq) and TEA (1.5 eq) at 50° C for 15 h, after which the products **11-13a**, were isolated by chromatographic purification (yields 60-70%).

Compound **11**, ^{1}H NMR,($CDCl_3$), δ 8.29 (s, 1H, *A* H-8), 7.82 (ma) and 7.72 (mi) (s, 1H, H-2), 7.32-7.36 (m, 5H, ArH, 6.48 (brs, 2H, NH_2), 5.75 (brd, ma) and 5.52 (br, mi, 1H, NHBOC), 5.25 (mi) and 5.13 (ma) (dd, 2H, $ArCH_2$), 4.9 (dd, J = 15.6 Hz, 2H, Ad CH_2), 4.64 (t, 1H, J = 5.8 Hz, H-2), 4.36 (br, 1H, C-4), 3.96 (dd, 1H, H-5a), 3.78 (brm, 1H, H-5b), 2.5-2.05 (brm, 2H, H-3a,b), 1.49 [s, 9H, $C(CH_3)_3$], ^{13}C NMR, ($CDCl_3$), δ 171.4 (ma) and 171.1 (mi) (COOR), 165.5 (CONH), 155.4 (C6 and OCON<), 152.4 (Ad C-2), 149.3 (*A* C-4), 141.6 (*A* C-8), 135.1-127 (Ar), 118.2 (*A* C-5), 79.8 (CMe_3), 67.9 (mi) and 67.1(ma) ($ArCH_2$), 57.8 (C-2), 51.4 (C-4), 50.0 (*A* NCH_2), 44.7 (C-5), 34.8 (C-3), 28.3 [$C(CH_3)_3$], $[\alpha]_D^{20}$ -20.7 (c = 0.3, MeOH).

Compound **12**, ^{1}H NMR, ($CDCl_3$+D_2O), δ 7.25-7.4 (overlapping signals, ArH, *C* H-6), 5.9 (d, J = 6.36 Hz, 1H, *C* H-5), 5.1 (overlapping q, 2H, $ArCH_2$), 4.55-4.95 (brm, 2H, N-CH_2), 4.15-4.5 (brm, 2H, H-3a,b), 1.45 [s, 9H, $C(CH_3)_3$], $[\alpha]_D^{20}$ -49.6 (c = 0.3, MeOH).

Compound **13a**, ^{1}H NMR, ($CDCl_3$), δ 7.85 (ma) and 7.7 (mi) (s, 1H, *G* H-8), 7.3 (m, 5H, ArH), 4.97-5.45 (overlapping signals, 4H, NH_2 and $ArCH_2$), 4.82 (dd, 2H, N-CH_2), 4.65 t, J = 4.2 Hz, H-2), 4.4 (brm, 1H, H-4), 4.05 (brm, 1H, H-5a), 3.65 (brm, 1H, H-5b), 2.45 (mi, m) and 2.22 (ma, t) (2H, H-3a,b), 1.47 [s, 9H, $C(CH_3)_3$], $[\alpha]_D^{20}$ -20.5 (c = 0.2, MeOH).

4*R*-N-(tertbutoxycarbonyl)-N$^{\alpha}$- (guanin-9-ylacetyl) proline, 13b: A mixture of 13a (0.5g, 1mmole) and 1N aq. NaOH (2mL) was stirred at r.t for 2 h after which it was neutralised with ion-exchange resin, Dowex H^+. MeOH (5 mL) was added, the resin was filtered off and the product that slowly precipitated was collected (yield 60%).^{1}H NMR, (d_6-DMSO:D_2O), δ 7.7 (s, 1H, *G* H-8), 4.92 (q, 2H, J = 19.5 Hz, N-CH_2), 4.28 (q, 1H, J = 8.1, 4.8 Hz, H-2), 4.18 (t, 1H, J = 7.2 Hz, H-4), 3.84 (q, 1H, J = 9.6, 5.8 Hz, H-5a), 3.4 (q, 1H, J = 9.6, 7.2, H-5b), 2.1 (dq, 1H, H-3a,3b), 1.4 [s, 9H, $C(CH_3)_3$]

Dipeptide Nucleic acids 17a and 18a: *General method*: The thymine monomers **14** and **16** (carboxyl component, 0.12 mmole) were condensed separately with the amino components **15** (30 mg, 0.1 mmole) in DMF (500 μL) in presence of DCC (25 mg, 0.12 mmole) and HOBT (15 mg, 0.1 mmole) at 25° C for 4 h. The usual work up afforded the protected dipeptide products **17** and **18** respectively, which were quantitatively deprotected in one step using either H_2/Pd-C or piperidine:DMF to yield the free dipeptides **17a** and **18a** respectively.

Compound **17a**: ^{1}H NMR (500 MHz, D_2O) δ: 7.38 (2H, *T* CH), 4.7-4.28 (8H, *T* CH_2 x 2, H-4, H-2 x 2), 4.16-3.67 (m, 4H, H-5a,b x 2), 3.75 (s, 3H, OCH_3), 2.8-2.12 (m, 4H, H-3a,b x 2), 1.88 (s, 6H, *T* CH_3 x 2). UV λ_{max} = 269 nm.

Compound **18a**: ^{1}H NMR (500 MHz, D_2O) δ: 7.42,7.38 (2H, *T* CH), 4.8-4.4 (8H, *T* CH_2 x 2, H-4, H-2 x2), 4.2-3.7 (m, 4H, H-5a,b x 2), 3.77, 3.73 (s, 3H, OCH_3), 2.8-2.12 (m, 4H, H-3a,b x 2), 1.89,1.87 (s, 6H, *T* CH_3 x 2). **UV** λ_{max}=270 nm.

Compound **17 :** ^{1}H NMR d_6-DMSO δ: 11.3 (s, 2H, *T* NH x 2, 5.03 (s, 2H, CBz CH_2), 4.57-3.81 (m, 8H, H-2, H-4 x 2, *T* CH_2 X 2), 3.56m (s, 3H, OCH_3), 3.4-3.2 (br, 4H, H-5a,b x 2 HO), 2.3-2.55 (br, 4H, H-3a,b, x 2) 1.75 (s, 6H, *T* CH_3 x 2) FAB-MS: [M+1] = 723.

Compound **18:** ^{1}H NMR d_6-acetone δ: 7.95-7.25 (m, 10H, Fmoc-Ar, *T* CH), 4.67-4.16 (m, 10H, H-2, H-4 x 2, Fmoc CH, CH_2, *T* CH_2 x 2), 3.91-3.54 (m, 4H, H-5a,b x 2), 3.7(s, 3H, OCH_3), 2.65-2.2 (br, 4H, H-3a,b x 2), 1.82,1.8 (s,6H, *T* CH_3 x2) FAB-MS: [M+1] = 811

Acknowledgment: BPG thanks UGC for a fellowship, KNG acknowledges Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore of which he is a Honorary Senior Fellow. We thank Indian Institute for Chemical Technology, Hyderabad for FABMS and Tata Institute of Fundamental Research, Bombay for 500 Mz NMR spectra and CD data.

REFERENCES

1. Nielsen P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497.
2. Egholm, M.; Buchardt, O.; Chiristensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature*, **1993**, *365*, 566.
3. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anticancer. Drug Design.* **1993**, *8*, 53. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O.; *Bioconj.Chem.* **1994**, *5*, 3. (c) Egholm, M.; Nielsen, P. E.; Buchardt, O., *Innovation Perspect Solid. Phase. Synth. Collect Papers Int. Symp.* **1994**, 145; CA: *123*, 170115. (d) Christensen, L.; Fitzpatrick, R.; Gildoa, B.; Coull, J.; Barren, B. *ibid*, 149; CA: *123*, 170116. (e) Zekanowski, C. *Postepy Biochem.*, **1995**, *41*, 32; CA: *123*, 191284. (f) Nielsen, P. E., *Ann. Rev. Biophys. Biomol. Struct.*, **1995**, *24*, 167. (f) Hyrup, B.; Egholm, M.; Nielsen, P. E. *BioMed. Chem. Lett.* **1996,** *4*, 5.
4. Hyrup, B.; Egholm, M.; Nielsen, P. E. *J. Am. Chem. Soc.* **1994**, *116*, 7964.
5. (a) Kim, S. H.; Nielsen, P. E.;, Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477. (b) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature*. **1994**, *368*, 561.
6. Peterson, K. H.; Jensen, D. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1119. (b) Bergmann, F.; Bannwarth, W.; Tam, S.; *Tetrahedron Lett.* **1995**, *36*, 6823.
7. Koch, T.; Naesby, M.; Wittung, P.; Jorgensen, M.; Larsson, C.; Buchardt, O.; C.J. Stanley, B. Norden, Nielsen, P. E.; *Tetrahedron Lett.* **1995**, *36*, 6933.

8. (a) Garner, P.;Yoo, J. U.; *Tetrahedron Lett.* **1993**, *34*, 1275. (b) Lenzi, A.; Reginato, G.; Taddei, M. *ibid*, **1995**, *36*, 1713, 1717. (b) Diederichsen, U.; Schmitt, H. W. *ibid*, **1996**, *37*, 475.

9. (a) Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173. (b) Dueholm, K. L.; Peterson, K. H.; Jensen, D. K.; Nielsen, P. E.; Egholm, M.;Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1077. (c) Diederischen, U.; Schmidt, H. W. *Tetrahedron Lett.* **1996**, *37*, 475 (d) Lobberding, A.; Schwemler, C.; Schwenner, E.; Stropp, W.; Kretschner, A EP 0646595 A1 (1995).

10. (a) Peterson, M. L.; Vince, R. *J. Med. Chem.* 1991, *34*, 2787 (b) Eva Ng, K.; Orgel, L. E. *J. Med. Chem.* **1989**, *32*, 1754.

11. Bodanszky, M. and Bodanszky, A. *The Practice of Peptide Synthesis*, Springer - Verlag, 1984

12. Dueholm, K. L.; Egholm, M.; Beherns, 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.

13. Stille, J. R.; Frietschel, S. J.; Baker, G. L. *J. Org. Chem.* **1981**, *46*, 2954.

14. Meltzer, P. C.; Liang, A. Y.; Matsudaira, P. *J. Org. Chem.* **1995**, *60*, 4305.

15. Chenon, M. -T.; Pugmire,R. J.; Grant, D. M.; Tanzica, R. P.; Townsend, L. B. *J. Am. Chem. Soc.*, **1975**, *97*, 4627.

16. (a) Leijon, M.; Graslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry*, **1994**, *33*, 9820. (b) Chen, S. -M; Mohan, V.; Kiefly, J. S.;Griffith, M. C.; Griffey,R. H. *Tetrahedron Lett.* **1994**, *35*, 5105.

17. Wittung, P.; Eriksson, M.; Lyng, R.; Nielsen, P. E.; and Norden, B. *J. Am. Chem. Soc.* **1995**, *117*, 10167

NCL Communication No 6360

(*Received in UK* 8 *July* 1996; *revised* 8 *October* 1996; *accepted* 10 *October* 1996)