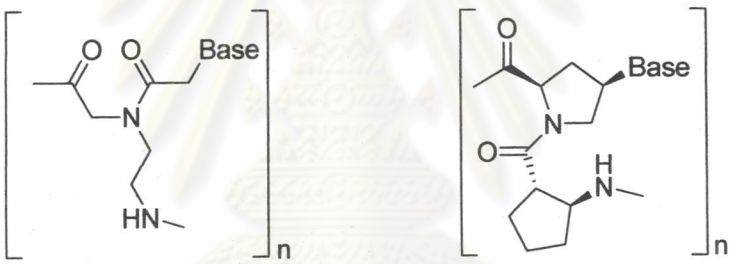
CHAPTER III

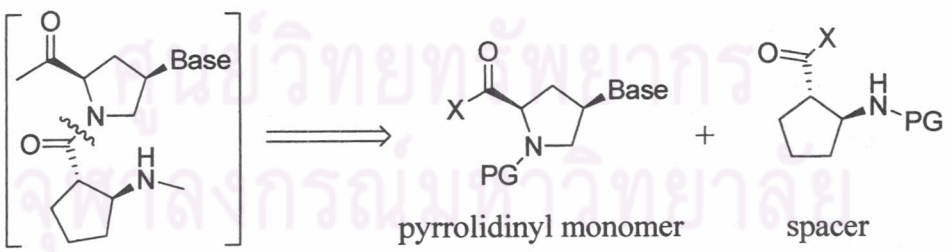
RESULTS AND DICUSSION

The structure of the *ss*ACPC-PNA (**12**) was designed based on a modification of the original *aeg*PNA (**1**) by rigidifying the backbone with proline and (1*S*,2*S*)-2-aminocyclopentanecarboxylic acid (**Figure 3.1**). Preliminary experiments suggested that this PNA system containing all thymine binds specifically to complementary DNA. As a result, the PNA (**12**) containing all four nucleobases was the selected primary target for the synthesis and biophysical properties study, mainly by UV-visible and circular dichroism spectroscopy.



(1) **(12)**

Figure 3.1 Structure of *aeg*PNA (**1**) and *ss*ACPC-PNA (**12**)



(12)

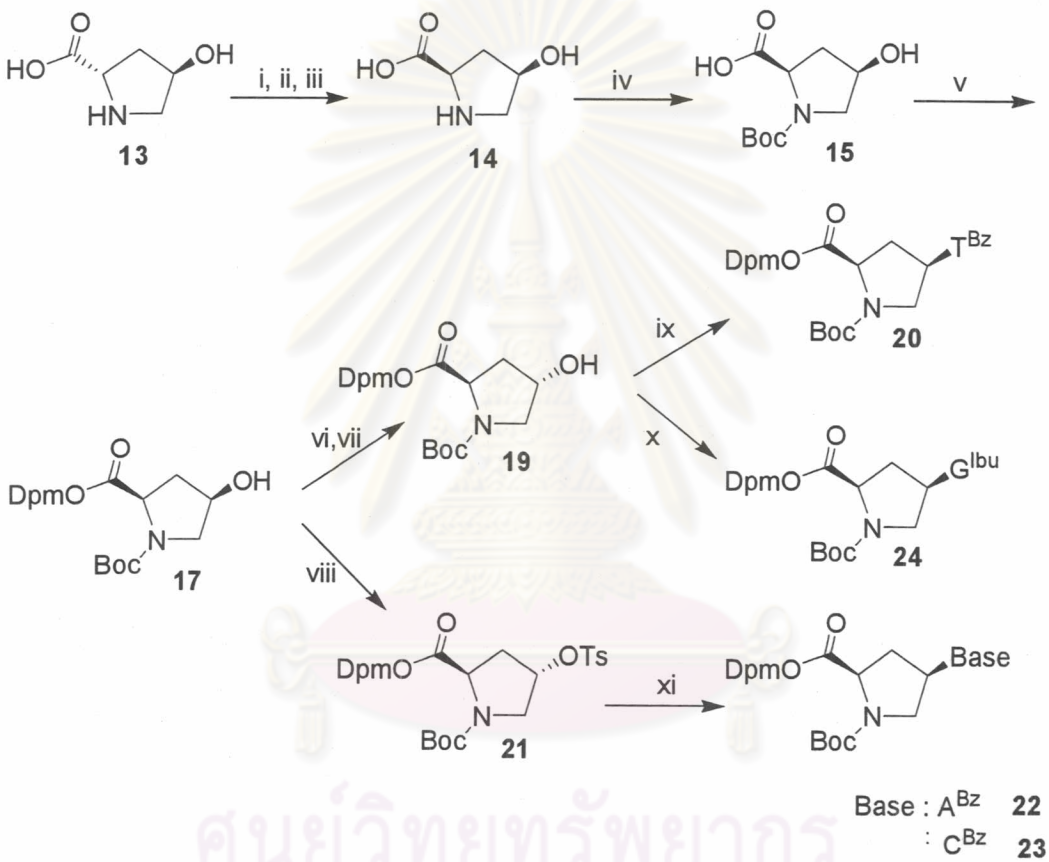
Figure 3.2 Retrosynthetic analysis for synthesis *ss*ACPC-PNA (**12**)

From retrosynthetic analysis (**Figure 3.2**), the structure of PNA (**12**) may be separate into two parts, namely the pyrrolidinyl monomer and the ACPC β-amino acid spacer. Synthesis of both types of monomers and their assembling via Solid Phase Peptide Synthesis (SPPS) will be discussed in the next sections.

3.1 Synthesis of pyrolidine monomers

3.1.1 Synthesis of intermediates

The Boc-protected monomers **20**, **22**, **23** and **24** which had been reported by Vilaivan et al.[41] (**Figure 3.3**) were the starting points for the synthesis of the pyrrolidine monomers.



Condition: i. Ac_2O, heat 90 °C 16 h; ii. 2 M HCl, reflux 5 h; iii. propylene oxide , MeOH; iv. Boc_2O, ^tBuOH, NaOH (aq), 8 h; v. Ph_2CN_2, EtOAc, 8 h; vi. HCO_2H, Ph_3P, DIAD, THF, 8 h; vii. NH_3, MeOH, 2 h; viii. MeOTs, Ph_3P, DIAD, THF, 8 h; ix. N^3-BzT, Ph_3P, DIAD, THF, 8 h; x. a) N^2-IbuG(ONpe), Ph_3P, DIAD, dioxane, 8 h; b) DBU, pyridine 2 h; xi. N^6-BzA or N^4-BzC, K_2CO_3, DMF, heat 90 °C 5 h.

Figure 3.3 A synthetic scheme for the intermediate proline derivative of thymine (**20**), adenine (**22**), cytosine (**23**) and guanine (**24**) according to Vilaivan et al.

The literature method was adopted essentially without modification. The identity of each compound was ascertained by comparison of NMR spectra with reference materials previously prepared in this laboratory. The commercially available *trans*-4-hydroxy-L-proline was used as the starting material for all pyrrolidine monomers. This was epimerized at the α-position using acetic anhydride followed by hydrolysis to get *cis*-4-hydroxy-D-proline. However, the conversion was not complete and the product was obtained as a 2:1 ratio mixture of epimers (*cis*-D:*trans*-L). The *cis*-D-product was obtained pure by recrystallization. The ^{1}H NMR spectra of the *trans*-L and *cis*-D forms are quite different therefore both can be distinguished by this mean as shown in **Figure 3.4**

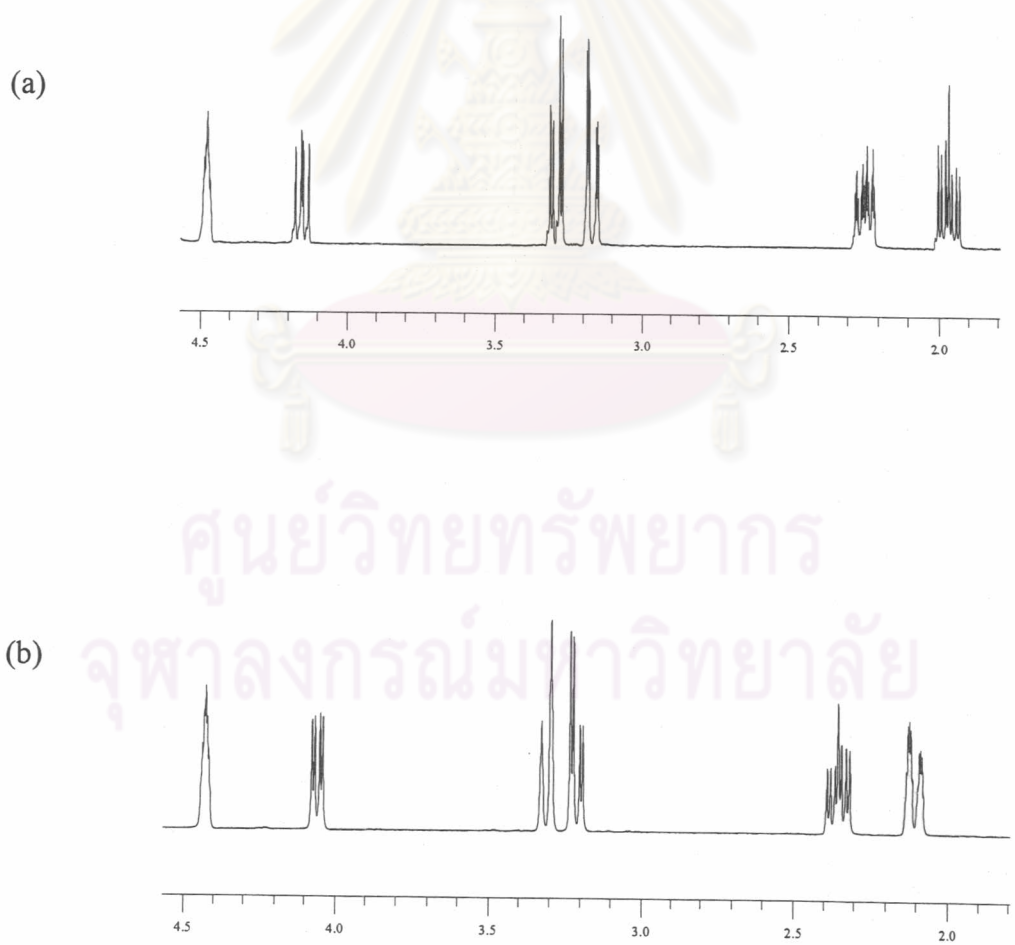


Figure 3.4 Comparison of ^{1}H NMR spectra (D_2O, 400 MHz) between (a) *trans*-4-hydroxy-L-proline and (b) *cis*-4-hydroxy-D-proline

The mechanism of epimerization from *trans*-L to *cis*-D *via* a mixed anhydride was shown in **Figure 3.5**.

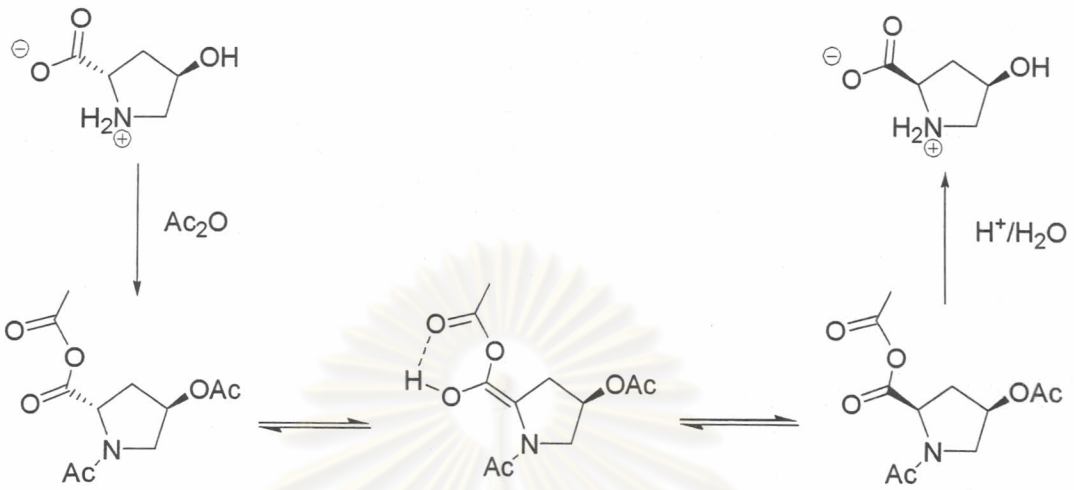
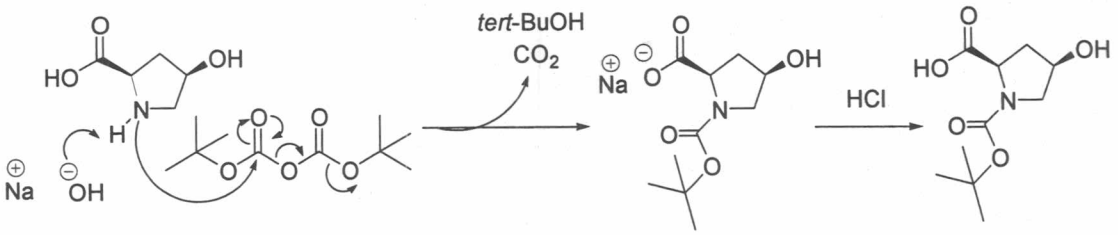


Figure 3.5 Epimerization from *trans*-4-hydroxy-L-proline to *cis*-4-hydroxy-D-proline

After the *cis*-4-hydroxy-D-proline was obtained, the highly nucleophilic sites of the compound (NH and COOH) were protected with temporary protecting groups in order to avoid the undesired reaction in subsequent steps. The *tert*-butoxycarbonyl (Boc) group and diphenylmethyl (Dpm) groups were chosen as the protecting groups for the amino and carboxylic groups respectively in order to provide a stable intermediate for the next Mitsunobu reaction step. The reaction of *cis*-4-hydroxy-D-proline (**14**) and di-*tert*-butyl dicarbonate (Boc_2O) in aqueous sodium hydroxide/*tert*-butanol gave the sodium salt of the desired compound, carbon dioxide and *tert*-butanol as by-product (**Figure 3.6**). The structure of *N*-*tert*-butoxycarbonyl-*cis*-4-hydroxy-D-proline **(15)** was confirmed by ^1H NMR, which shows two sharp singlet peaks at 1.35 and 1.41 ppm due to rotamer of the Boc group. The two rotamer coexist in equilibrium and give rise to two different sets of signals for protons in each of them. This phenomenon is common for tertiary amides, which exhibit two rotameric forms due to resonance **(Figure 3.7a).** The partial double bond character of the C-N bonds gives rise to a high rotational barrier (in the order of 75-90 kJ/mol) for interconversion between the *cis*- and *trans*- amide bond (**Figure 3.7b**).[64] This restricted rotation resulted in a slow rotation on the NMR time scale at room temperature so that two sets of peaks due to each rotamer were observed.



(15)

Figure 3.6 Mechanism of *N-tert*-butoxycarbonylation of *cis*-4-hydroxy-D-proline to give (**15**).

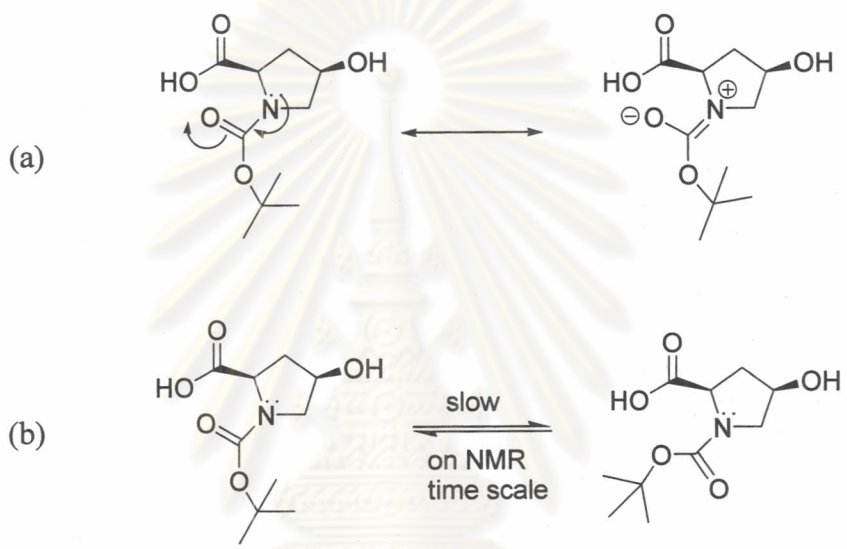


Figure 3.7 (a) The resonance structure of amides (b) Restricted rotation of around C-N bond in *N*-Boc hydroxyproline.

For protection of carboxylic acid group, diphenylmethyl (Dpm) protecting group was used since it is compatible with the Boc group and resistant to acid and basic conditions. Furthermore it can be easily removed simultaneously with the Boc group by TFA/anisole condition in the final step. The crystalline Boc/Dpm intermediate (**17**) was synthesized from *N-tert*-butoxycarbonyl-*cis*-4-hydroxy-D-proline (**15**) by treatment with freshly prepared diphenyldiazomethane (**16**) as shown **Figure 3.8**

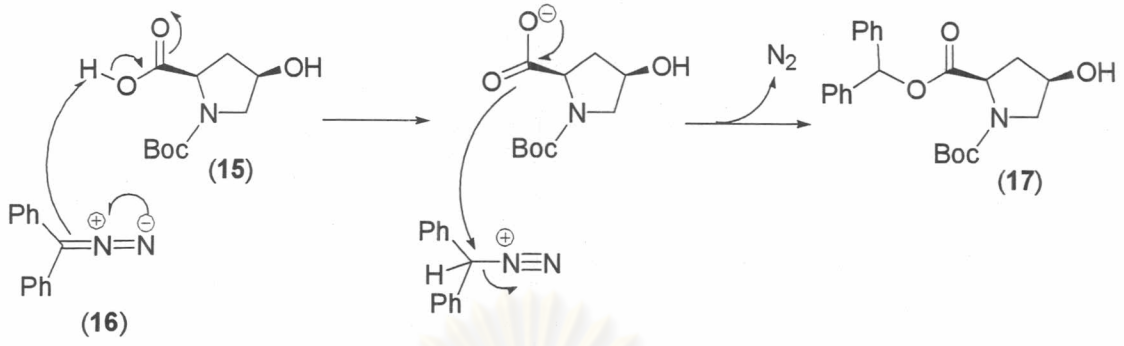


Figure 3.8 The mechanism of protecting the carbonyl group by diphenyldiazomethane.

In order to introduce nucleobase into the proline ring at the C-4 position with the configuration *cis*-D formed, inversion of the configuration at the 4-OH group was necessary. In case of thymine and guanine, the protected *cis*-D-proline was first converted into the corresponding *trans*-D-isomer by Mitsunobu reaction using formic acid as nucleophile followed by hydrolysis to give *N*-*tert*-butoxycarbonyl-*trans*-4-hydroxy-D-proline diphenylmethyl ester (**19**) (**Figure 3.9**). The structure of compound (**19**) was confirmed by ^{1}H NMR.[41] Next, alkylation with the nucleobase was carried out by Mitsunobu reaction using N^3-benzoylthymine as nucleophile to give *N*-*tert*-butoxycarbonyl-*cis*-4-(N^3-benzoylthymin-1-yl)-D-proline diphenylmethyl ester (**20**) with inversion of the C-4 center as desired. The same reaction using N^2-isobutyryl-O^6-(4-nitrophenylethyl)-guanine should in principle give *N*-*tert*-butyloxycarbonyl)-*cis*-4-(N^2-isobutylguanin-9-yl)-D-proline diphenylmethyl ester (**24**) after removal of the nitrophenylethyl group. However, this compound was not synthesized in this study but was a generous gift from Miss Patcharee Ngamwiriyawong.

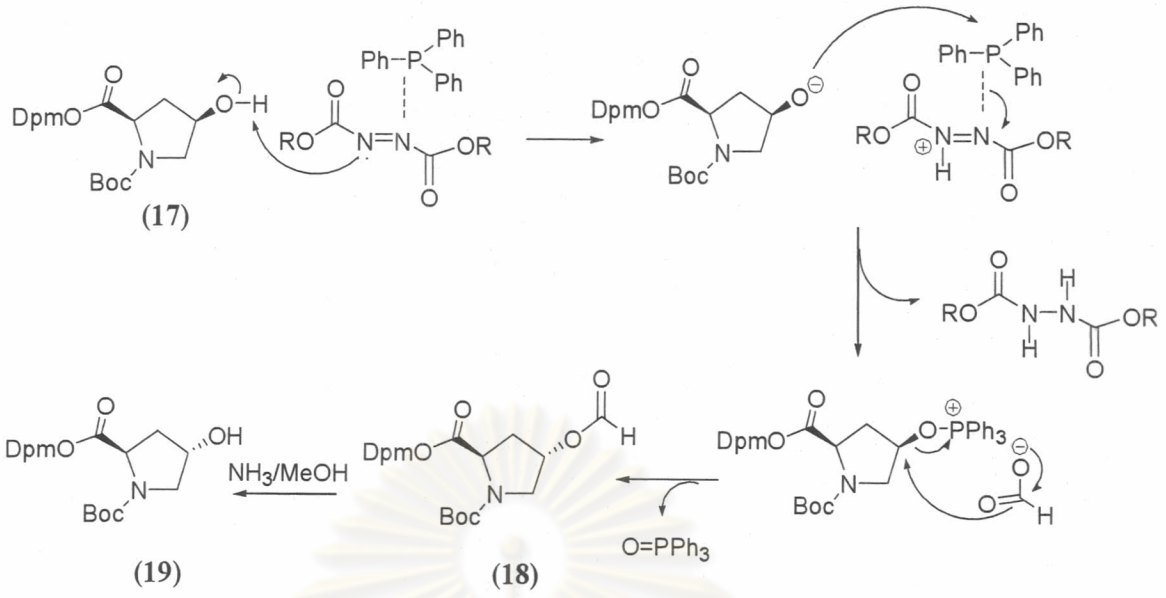


Figure 3.9 Mitsunobu reaction of *N-tert*-butoxycarbonyl-*cis*-4-hydroxy-D-proline diphenylmethyl ester **(17)** to form *N-tert*-butoxycarbonyl-*trans*-4-formyl-D-proline diphenylmethyl ester **(18)** and its hydrolysis to *N-tert*-butoxycarbonyl-*trans*-4-hydroxy-D-proline diphenylmethyl ester **(19)**

For the A and C monomers, the Mitsunobu reactions using N^6-benzoyladenine (BzA) and N^4-benzoylcytosine (BzC) were reported to be unsuccessful [42]. These two compounds must be prepared through an inverted tosylate which was prepared by Mitsunobu reaction of (**17**) using methyl tosylate (MeOTs) as the nucleophile. The tosylate was then reacted with the nucleobases to cause a displacement *via* S_N2 reaction. The reactions were successful for both protected adenine and cytosine (**Figure 3.10**). The structures of the protected adenine (**22**) and protected cytosine (**23**) monomers were confirmed by comparison of the NMR spectra with reference materials previously prepared in this laboratory.

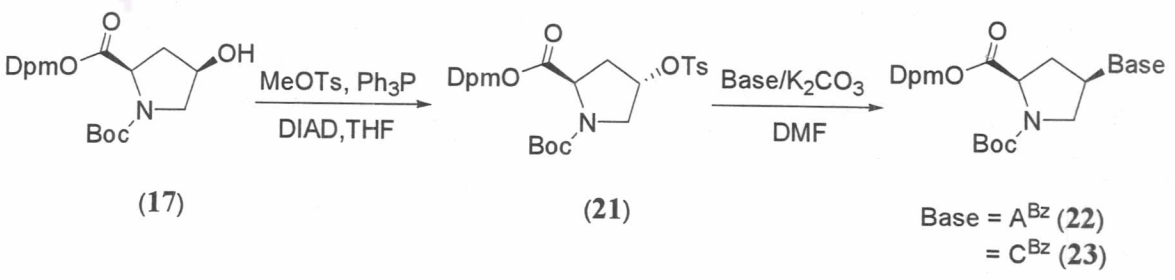


Figure 3.10 Introduction of nucleobases into proline ring *via* tosylation and S_N2 displacement when the nucleobases were adenine and cytosine.

3.1.2 Synthesis of activated PNA monomers

In the solid phase synthesis of oligo-PNA containing all four nucleobases, the activated PNA monomers have to be prepared. Previous experiences in this group suggested that Fmoc chemistry had a number of advantages over Boc chemistry [42]. This is mainly due to the mild conditions for the deprotection *N*-Fmoc group (20 % piperidine) which results in a simpler design of equipment that is capable of handing small-scale synthesis (50 μmol or less).

To obtain the Fmoc-protected monomers, the Boc-protecting group must be replaced by a fluorenylmethoxycarbonyl (Fmoc) group. To deprotect the Boc group and diphenylmethyl (Dpm) ester group of the intermediate proline compounds (**20, 22, 23** and **24**), they were treated with trifluoroacetic acid (TFA) in the presence of anisole at 30 °C for 8 h to give the free acids in the form of TFA salts (**Figure 3.11**). The role of anisole was to act as a scavenger for the highly electrophilic diphenylmethyl cation released upon deprotection of the diphenylmethyl group. The crude TFA salts were used for the next step without further purification. Under this condition, the benzoyl protecting group of N^3-benzoylthymine compound **(20)** was also cleaved by TFA as shown in **Figure 3.12**. However, the presence of unprotected thymine was shown to have no effect on the subsequent coupling steps [64] because the nucleophilicity of thymine N^3 was not high due to the resonance shown in **Figure 3.12**.

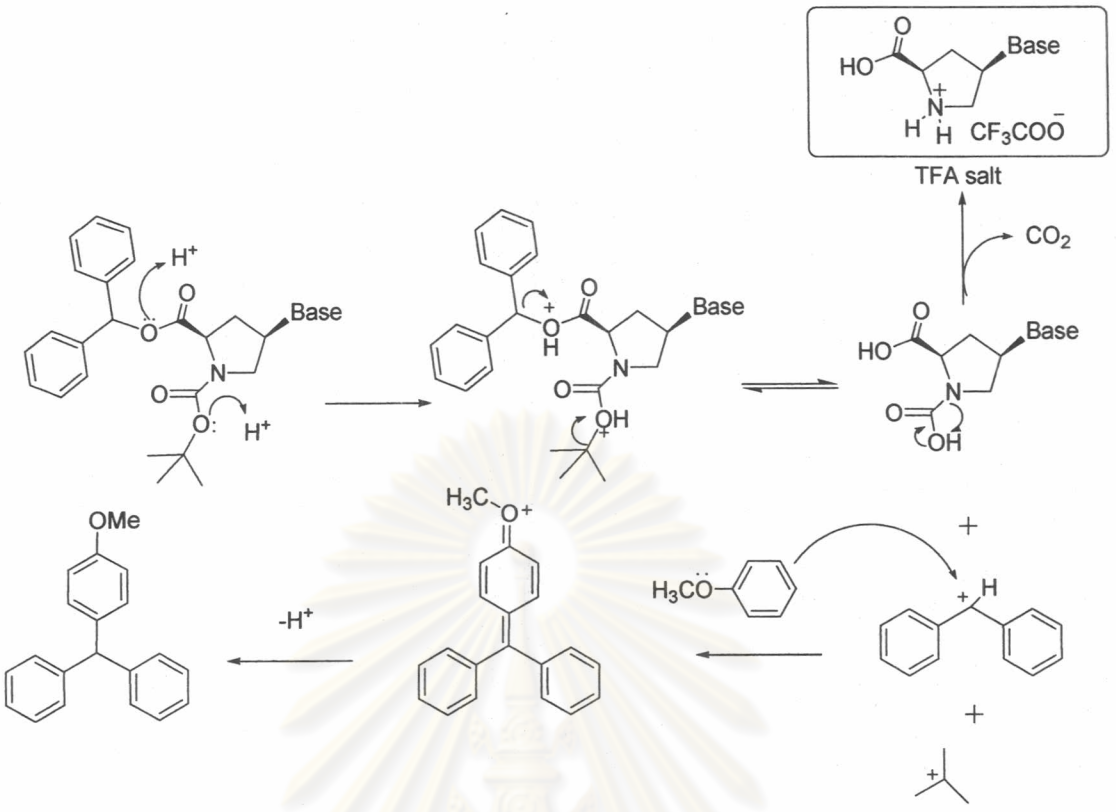


Figure 3.11 Mechanism for deprotection of Boc/Dpm protecting group in the presence of TFA and anisole.

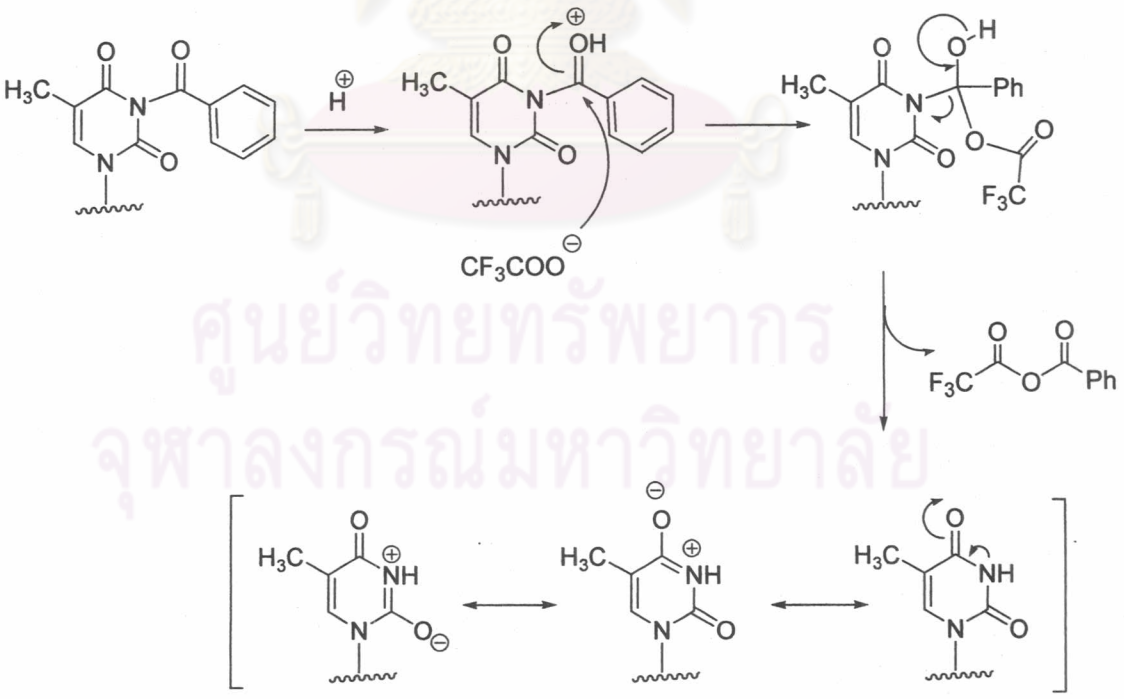


Figure 3.12 Deprotection benzoyl protecting group of *N-tert*-butoxycarbonyl-*cis*-4-(N^3-benzoylthymin-1-yl)-D-proline diphenylmethyl ester (**20**)

In the next step, the amino group was protected with Fmoc by reacting with 9-fluorenyl succinimidyl carbonate (FmocOSu) under basic conditions **(Figure 3.13)**. The FmocOSu was used instead of FmocCl because FmocCl was too reactive and may cause undesirable reactions such as dipeptide formation [65]. Furthermore, FmocOSu was also more soluble and more stable in aqueous-organic solvents than FmocCl. In all cases, the Fmoc protection took place smoothly at room temperature. The Fmoc amino acids were easily isolated by first extraction of the reaction mixture with diethyl ether to remove the excess FmocOSu followed by acidification with aqueous HCl. The Fmoc-acids (**25, 27, 29 and 31**) were precipitated as white solid which were collected by filtration and dried under vacuum. The yield of each compound is reported in **Table 3.1.**

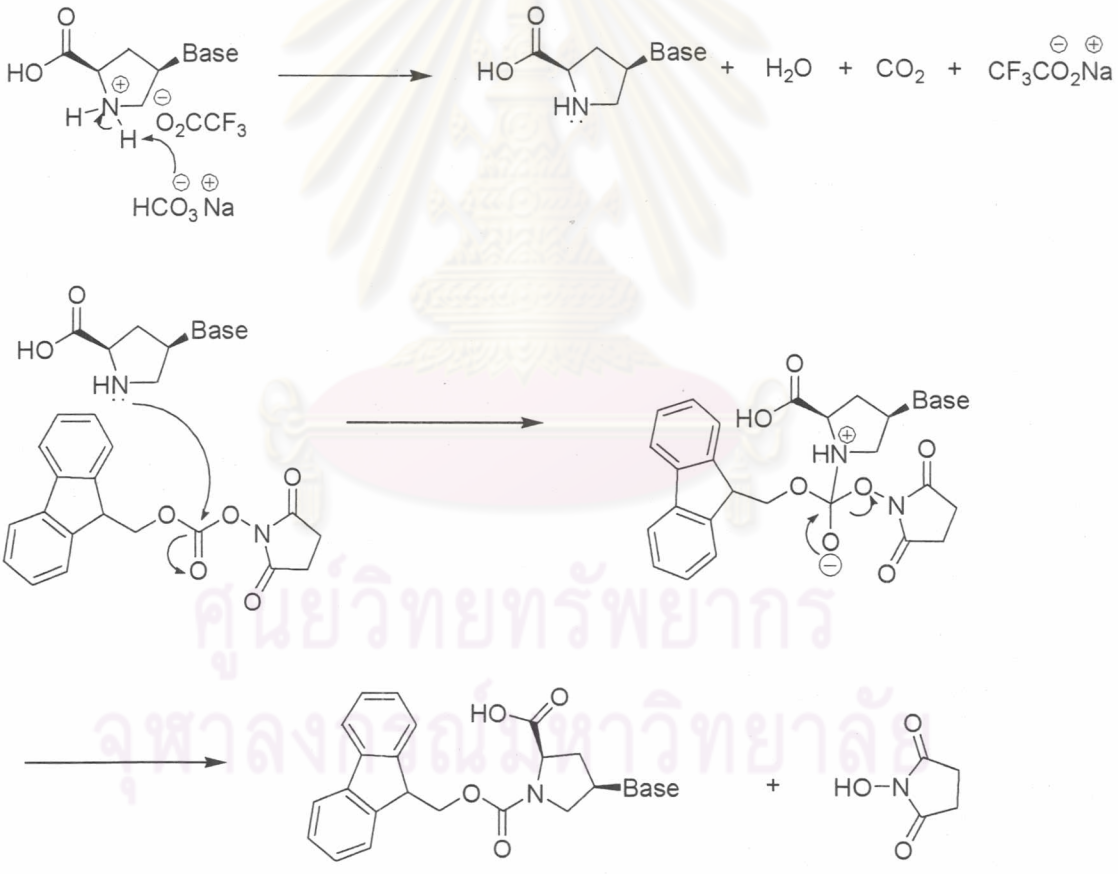
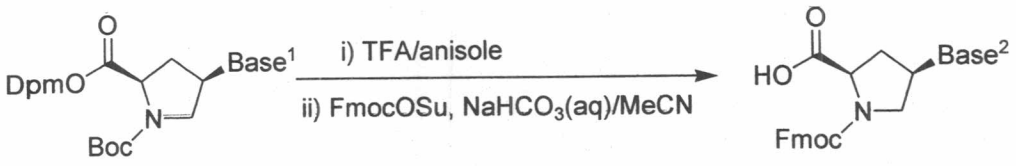


Figure 3.13 Mechanism for the protection of N atom with Fmoc group.

Table 3.1 Fmoc protection of PNA monomers



Starting material	Base[1]	Product	Base[2]	%yield
20	T^{Bz}	**25**	T	76
22	A^{Bz}	**27**	A^{Bz}	82
23	C^{Bz}	**29**	C^{Bz}	76
24	G^{Ibu}	**31**	G^{Ibu}	49

^{1}H NMR spectra of (**25, 27, 29** and **31**) show two sets of signals due to protons and carbons on the proline ring in approximately 1:1 ratio as shown in **Figure 3.15**. For example, the ^{1}H NMR spectrum of (*N*-fluoren-9-ylmethoxycarbonyl)-*cis*-4-(thymin-1-yl)-D-proline (**25**) showed the following important signals: 2.33 and 2.39 (1×CH_2 3'), 2.97 (1×CH_2 3'), 3.60 and 3.75 (1×CH_2 5'), 3.95 and 4.04 (1×CH_2 5'), 4.68 and 4.76 (CH 2'), 5.22 and 5.31 (CH 4') **(Figure 3.15)**. This can be explained by the formation of two rotamers due to the slow rotation of the Fmoc group as discussed previously for the corresponding Boc derivative (section 3.1.1) **(Figure 3.14)**.

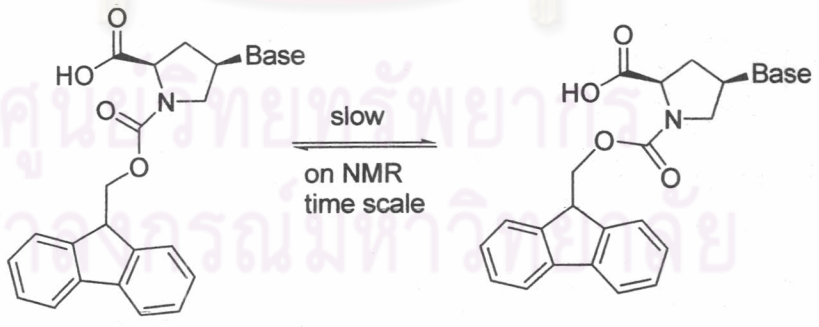
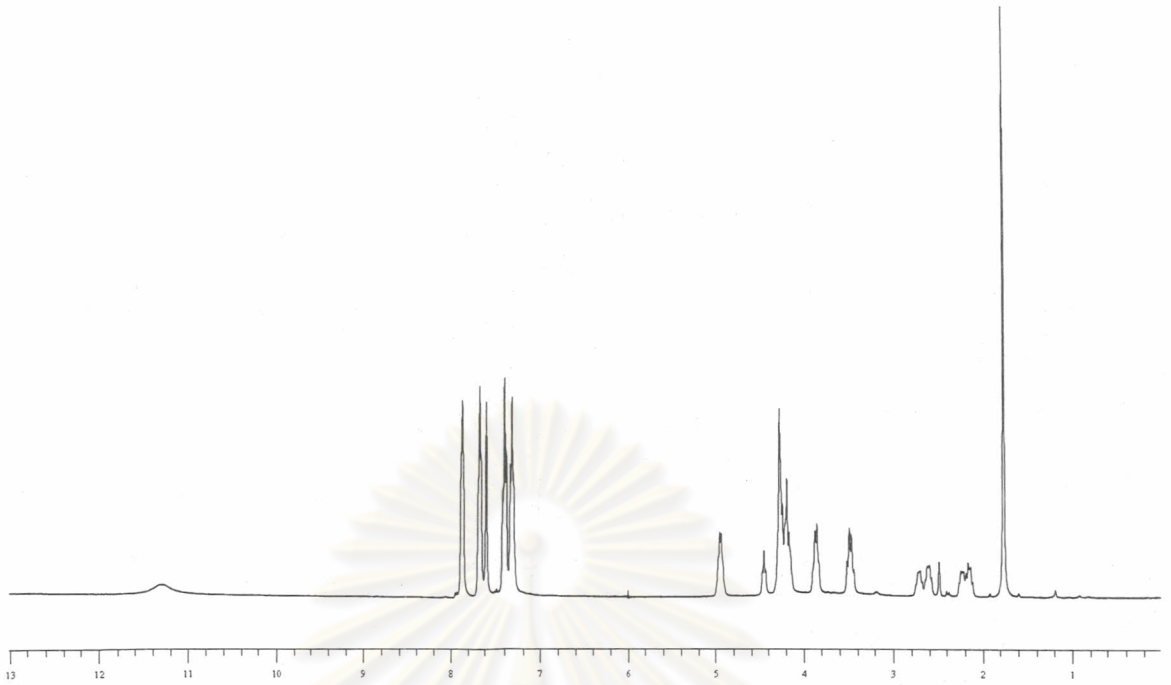
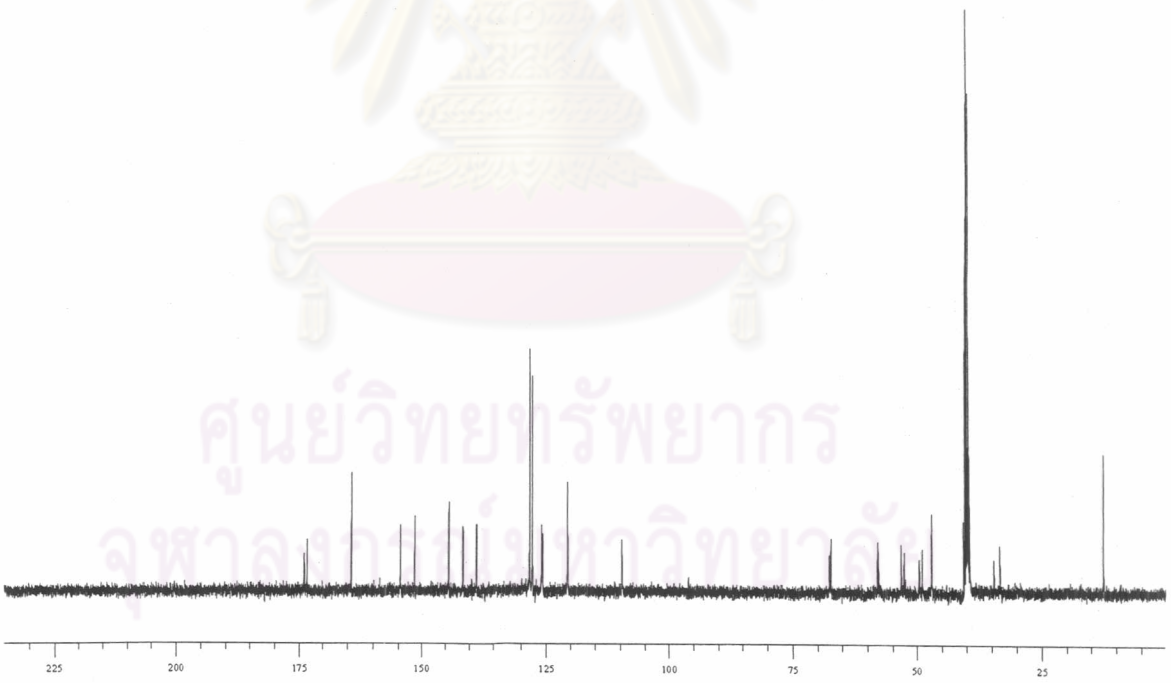


Figure 3.14 Rotamer of the Fmoc free acids (**25, 27, 29** and **31**)



(a)

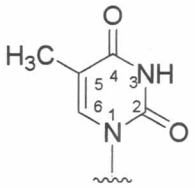


(b)

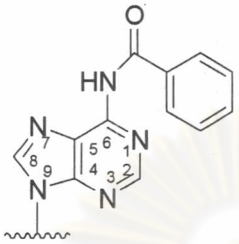
Figure 3.15 (a) ^{1}H NMR spectrum (400 MHz, DMSO-d_6) and (b) ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of (*N*-fluoren-9-ylmethoxycarbonyl)-*cis*-4-(thymin-1-yl)-D-proline (**25**).

^{13}C NMR chemical shifts of the Fmoc-protected monomers (**25, 27, 29** and **31**) were assigned as shown in **Table 3.2**. The values are fully consistent with the expected structures, indicating that the reactions were successful.

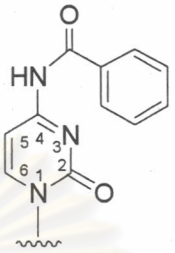
Table 3.2 ^{13}C NMR of Fmoc-protected monomers (**25, 27, 29** and **31**)



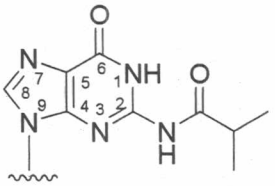
25



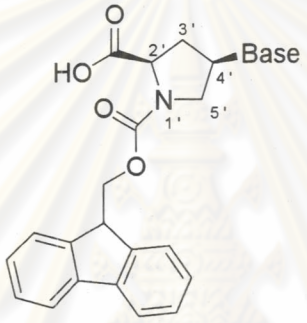
27



29



31



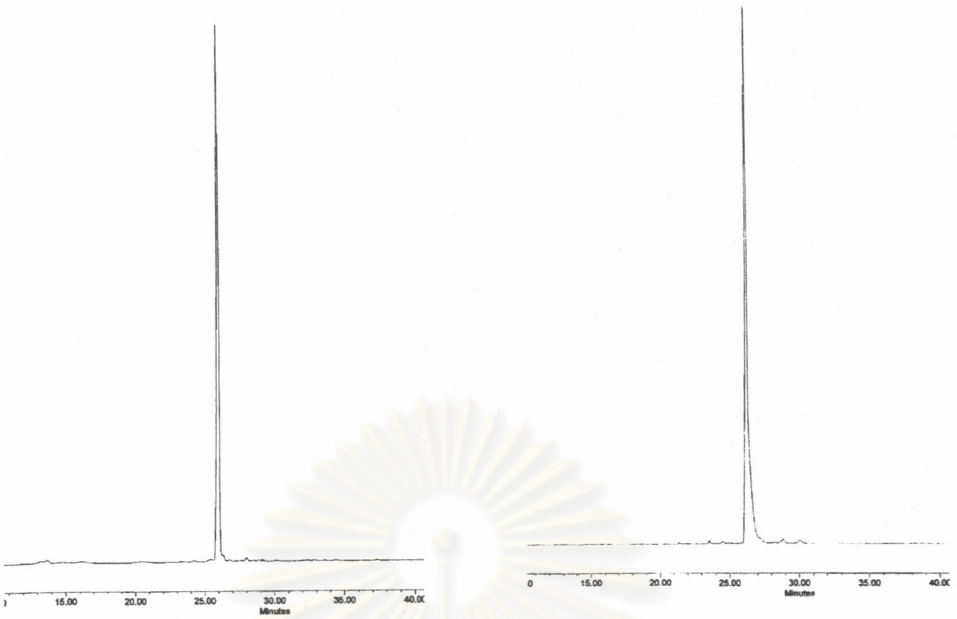
Position	Chemical shifts (ppm)			
	25	**27**	**29**	**31**
2	151.4	152.9	155.4	149.1
4	164.2	150.7	163.1	155.3
5	109.5, 109.6	126.2	96.6	120.5
6	138.3	151.8	147.9	148.2
8	-	141.1	-	138.3
2'	57.4, 57.8	57.7, 58.0	57.7, 58.0	57.8, 58.2
3'	33.5, 34.9	34.0, 35.2	33.5, 33.6	34.8, 35.6
4'	52.7, 53.3	52.4, 53.0	55.5, 56.1	52.1, 52.9
5'	48.8, 49.4	50.0, 50.5	49.2, 49.8	50.5, 51.1
$\underline{C}H_3$	12.6	-	-	19.3
$\underline{C}H(CH_3)_2$	-	-	-	35.1
Benzoyl <u>C</u>	-	132.9	133.1	-
Benzoyl <u>CH</u>	-	126.2, 127.6, 133.8	128.2, 128.8, 133.6	-

Position	Chemical shifts (ppm)			
	25	**27**	**29**	**31**
Benzoyl CO	-	166.0	167.9	-
Ibu CO	-	-	-	180.6
Fmoc CH	47.1	47.1	47.1	47.0
Fmoc CH_2	67.4, 67.7	67.5, 67.8	67.4, 67.7	67.4, 67.7
Fmoc Ar CH	120.6, 125.6, 127.7, 128.2	120.6, 125.8, 128.2, 128.9	120.6, 125.7, 127.6, 128.9	120.2, 125.8, 127.6, 128.1
Fmoc Ar C	141.2, 144.1	143.7, 144.1	141.1, 144.2	141.0, 144.1
Fmoc CO	154.4	154.3	154.4	154.4
Proline CO	173.1, 173.7	173.0, 173.5	173.2, 173.8	173.0, 173.6

The purity of compounds (**25, 27, 29** and **31**) were confirmed by reverse phase HPLC using the acetonitrile-water gradient system as described in section 3.3.2 (ix). The HPLC chromatograms of the compound (**25, 27, 29** and **31**) were shown in **Figure 3.16**. In all cases only one major peak was obtained indicating the purity of >99 % (t_R **25**:24.93; **27**:26.27, **29**:25.97 and **31**:26.15 min).

25

27

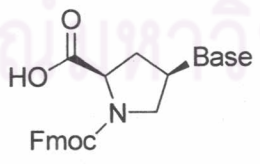


29 **31**

Figure 3.16 HPLC chromatograms of the compounds (**25, 27, 29** and **31**) Condition: C-18 column 3μ particle size 4.6 x 50 mm; gradient system of 0.01% TFA in acetonitrile/water 10:90 – 90-10 in 25 min; hold time 5 min; flow rate 0.5 mL/min.

Since these compounds are new, the identity of the compounds were further confirmed by High Resolution Mass Spectrometry (HRMS) analysis (**Table 3.3**). In all case the observed masses were in good agreement with the expected values (errors 5 ppm)

Table 3.3 HRMS analysis of Fmoc amino acids (**25, 27, 29** and **31**).



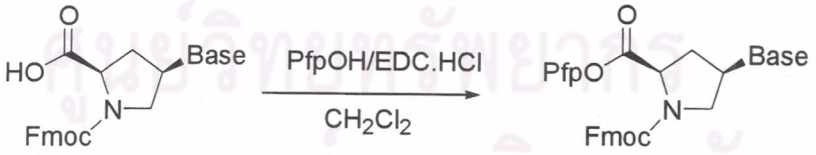
Compound	Base	Formula	M·H$^+$(found)	M·H$^+$(calcd.)
25	T	$C_{25}H_{23}O_6N_3$	462.1656	462.1666
27	A^{Bz}	$C_{32}H_{26}O_5N_6$	575.2041	575.2043
29	C^{Bz}	$C_{31}H_{26}O_6N_4$	551.1938	551.1910
31	G^{Ibu}	$C_{29}H_{28}O_6N_6$	557.2149	557.2149

The characterization by ^{1}H, ^{13}C NMR together with reverse phase HPLC and HRMS analysis suggested that the compounds (**25, 27, 29** and **31**) were successfully obtained in good purity suitable for the next steps.

Condensation reaction between a carboxylic acid and an amine to form amide bond requires activation of the carbonyl group. Although several activation methods are available, the pentafluorophenyl (Pfp) ester was selected as the method of choice due to the excellent ability of the Pfp group to act as a leaving group [66]. Furthermore, Pfp esters are usually stable crystalline solid which allow convenient storage inspite of their high reactivity. The reactivity of the Pfp esters may also be enhanced by an additive such as hydroxybenzotriazole (HOBt) or its derivatives [67].

The free acids were converted into the corresponding Pfp esters by reaction with pentafluorophenol (PfpOH) and *N*-(3-dimethylaminopropyl)-*N'*-ethyl-carbodiimide hydrochloride (EDC.HCl) in dichloromethane. This method is very convenient to perform and provide a good yield therefore it is preferred over the literature report [64] which used either PfpOH/DCC or PfpOTfa/DIEA. The reactions were usually completed within 1 h at room temperature according to TLC analysis. The products were purified by flash column chromatography. The chromatography must be performed quickly to avoid decomposition of the product on the column. In all cases the T, A, C and G monomers (**26**, **28**, **30** and **32**) were obtained as white solid which may be stored in a freezer for at least a few months **(Table 3.4)**.

Table 3.4 Activation of the Fmoc amino acids by formation of Pfp esters



Free acid derivative	Base	Pfp ester derivative	%yield
25	T	**26**	85
27	A^{Bz}	**28**	81
29	C^{Bz}	**30**	76
31	G^{Ibu}	**32**	66

^{1}H NMR spectra of the free acids and Pfp esters were quite similar. However, the major difference between polarity of the free acid and Pfp ester derivatives were easily identified by TLC. For example, The Pfp ester derivative (**26**) has R_f = 0.43 (60% ethyl acetate:hexane) while the free acid (**25**) has R_f = 0 under the same condition. In case of ^{13}C NMR spectra of Pfp esters, most of the peaks of Pfp ester were similar to the free acids. However, the Pfp esters showed additional peaks characteristic to Pfp CF. For example Fmoc-amino-*cis*-4-T-D-proline pentafluorophenyl ester (**26**) showed a characteristic pack of Pfp CF at 136.6-142.2 ppm. These peaks were not well resolved due to extensive ^{13}C-^{19}F couplings. In addition, the ^{13}C CO peak of the proline moiety shifted upfield by 5-6 ppm upon the formation of the Pfp ester. Comparison of the spectra of Pfp ester (**26**) and the free acid (**25**) was shown in **Figure 3.17**.

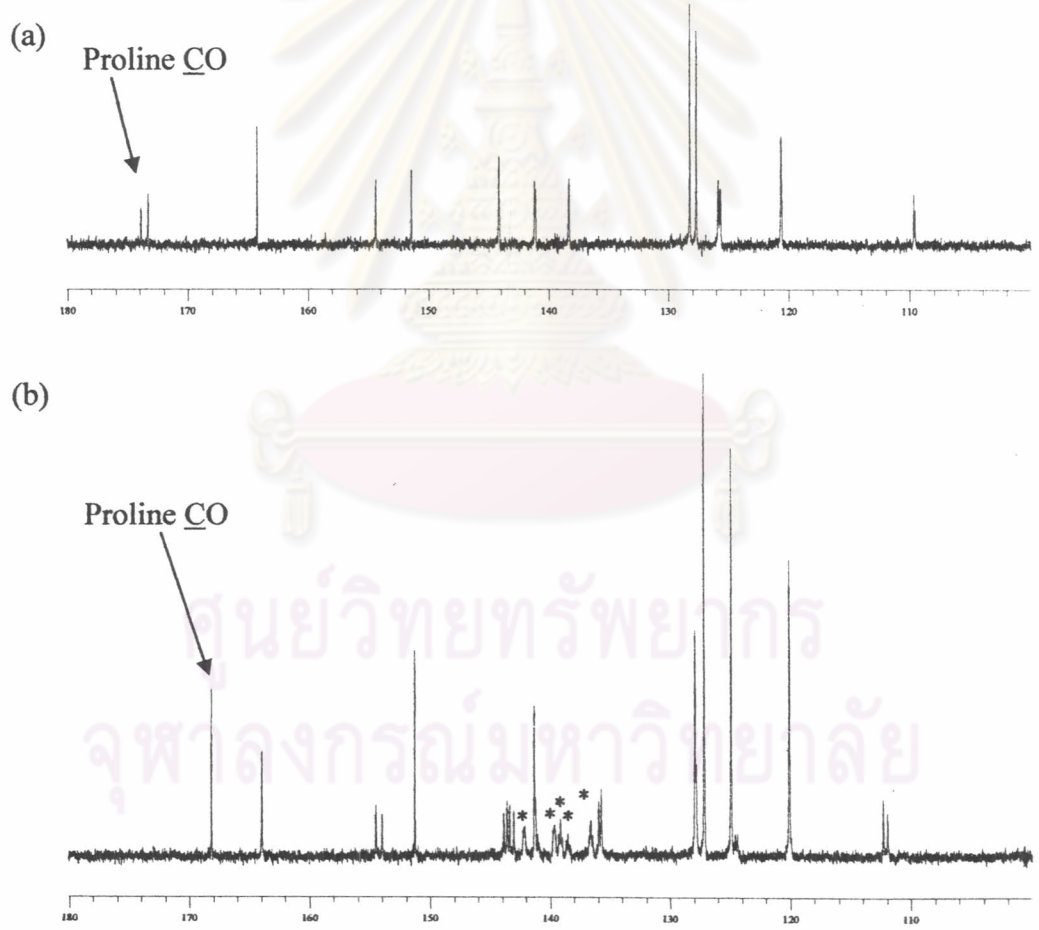


Figure 3.17 Comparison of ^{13}C NMR spectra between (a) (*N*-fluoren-9-ylmethoxycarbonyl)-*cis*-4-(thymin-1-yl)-D-proline (**25**) and (b) (*N*-fluoren-9-ylmethoxycarbonyl)-*cis*-4-(thymin-1-yl)-D-proline pentafluorophenyl ester (**26**)

The mechanism for the Pfp-ester formation is demonstrated in **Figure 3.18**.

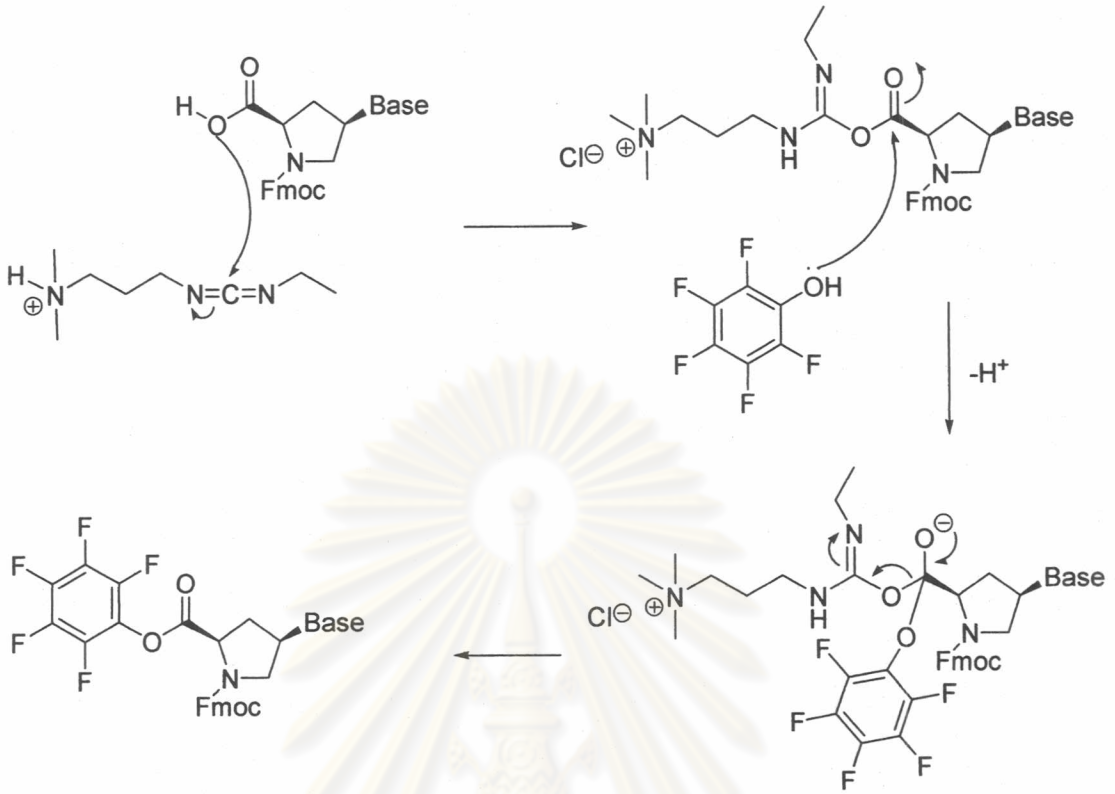
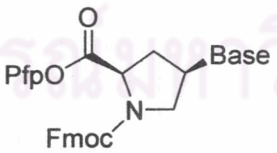


Figure 3.18 Reaction mechanism for the activation of the Fmoc free acid with Pfp.

Due to the sensitivity to water of the Pfp esters, the HPLC analysis under reverse-phase condition was not performed. However, according to the NMR and TLC analysis the activated monomers were sufficiently pure for the solid phase reactions. The success of the synthesis was also confirmed by HRMS analysis (**Table 3.5**)

Table 3.5 HRMS analysis of Pfp ester monomers



Compound	Base	Formula	M·H$^+$(found)	M·H$^+$(calcd.)
26	T	$C_{31}H_{23}O_6N_3F_5$	628.1505	628.1508
28	A^{Bz}	$C_{38}H_{26}O_5N_6F_5$	748.1898	748.1885
30	C^{Bz}	$C_{37}H_{26}O_6N_4F_5$	717.1758	717.1773
32	G^{Ibu}	$C_{35}H_{28}O_6N_6F_5$	723.2001	723.1991

The Pfp ester derivatives (**26, 28, 30** and **32**) were used as monomers in oligomerization which will be described in the next section.

3.1.3 Synthesis of ACPC spacer

Synthesis of the β-amino acid spacer (*ss*ACPC) is previously reported by Gellman [63]. Ethyl cyclopentanone-2-carboxylate (**33**) was reacted with (*S*)-(-)-α-methybenzylamine in the presence of glacial acetic acid to give an enamine intermediate. The enamine was stereoselectively reduced with sodium cyanoborohydride to give the (*S,S*)-*trans* β-aminoester **(34)** along with other stereoisomers, which could be separated by column chromatography. The correct isomer was identified by comparison of NMR spectrum with that reported by Gellman. At the end, 21 % yield of ethyl (1*S*,2*S*)-2-[(1'*S*)-phenylethyl]-aminocyclopentane carboxylate (**34**) was obtained (**Figure 3.19**).

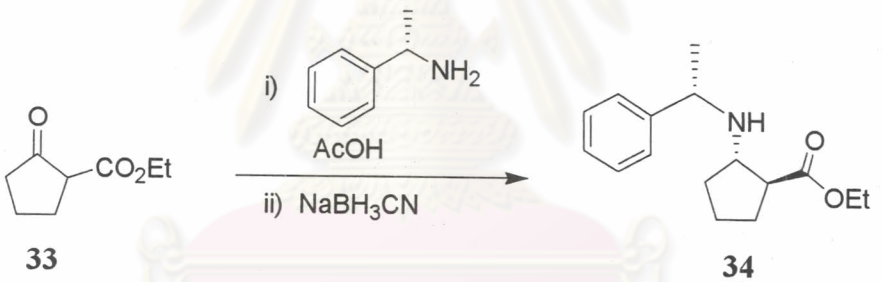


Figure 3.19 Synthesis of ethyl (1*S*,2*S*)-2-[(1'*S*)-phenylethyl]-aminocyclopentane carboxylate (**34**)

The *S*-methylbenzylamine chiral auxiliary was removed by hydrogenation over a palladium-carbon catalyst to afford the free amine (**35**). Acid hydrolysis give the amino acid as hydrochloride salt (**36**) in 64 % yield from (**34**). Protection of the amino group with FmocOSu was accomplished in the same way as described for protection of the activated PNA monomers. This resulted in the free acid (**37**) in 63 % yield (**Figure 3.20**).

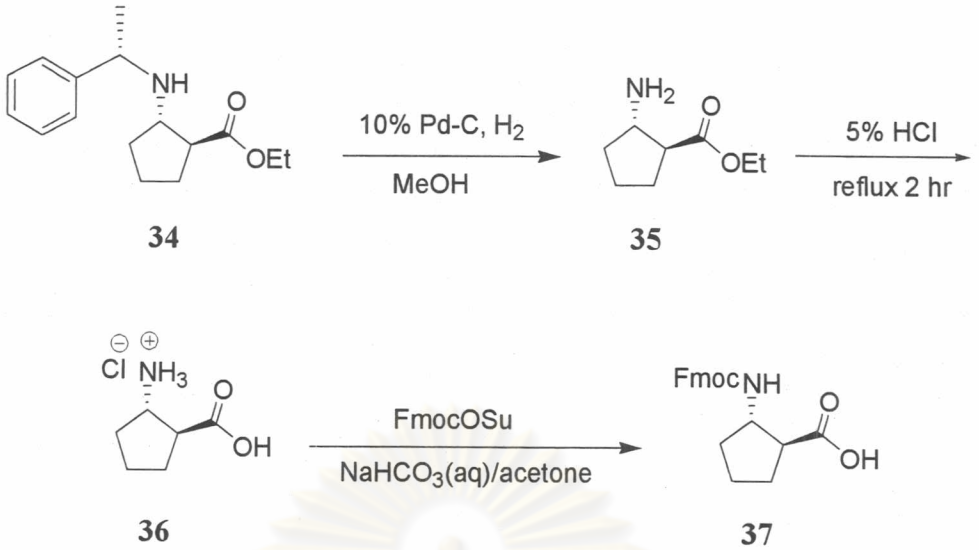


Figure 3.20 Synthesis of (1*S*,2*S*)-2-aminocyclopentane carboxylic acid (**37**)

The ^{1}H NMR spectra of compound (**37**) showed the following important signals: 4.26 ppm (Fmoc C<u>H</u>), 4.48 ppm (Fmoc C<u>H</u>$_2$), 7.34 ppm (Fmoc Ar C<u>H</u>), 7.42 ppm (Fmoc Ar C<u>H</u>), 7.62 ppm (Fmoc Ar C<u>H</u>) and 7.80 ppm (Fmoc Ar C<u>H</u>) which is in good agreement with the literature values [63]. The structure was further confirmed by comparison of ^{1}H NMR spectrum and the specific rotation value ($[\alpha]_D$ +36.4, c = 1.0 in MeOH) with the value reported in the literature ($[\alpha]_D$ +36.3, c = 1.21 in MeOH).

The Fmoc free acid (**37**) was reacted with PfpOH and EDC.HCl as decribed previously for the PNA monomers. This reaction was complete within 1 h according to TLC analysis. The product was again quickly purified by flash column chromatography to avoid decomposition of the product on the column. The product (**38**) was obtained in 78 % yield as a white solid **(Figure 3.21)**. The identity of the product was confirmed by ^{1}H, ^{13}C and elemental analysis which are fully consistent with the expected structure.

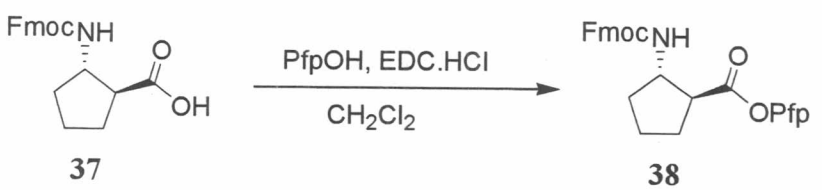


Figure 3.21 Synthesis of (1*S*,2*S*)-2-(*N*-fluoren-9-ylmethoxycarbonyl)-amino cyclopentane pentafluorophenyl ester (**38**)

3.2 Solid phase synthesis of *ss*ACPC-PNA (12)

3.2.1 Solid phase synthesis of *ss*ACPC-PNA (12) under standard conditions

Solid Phase Peptide Synthesis (SPPS) was introduced by Merrifield in 1963 [68]. The technique involves growing of a peptide chain from amino acid building blocks on a heterogeneous solid support such as polystyrene resin. The entire process occurs in the same reaction vessel. The reaction involves simple washing and filtration steps therefore minimizing product losses, producing higher-yield (milligram levels of material) and higher-purity peptides. Synthetic peptides are usually produced in a stepwise fashion starting from C (carboxyl) terminus to N (amino) terminus through a series of coupling cycles.

In practice, the manual microscale synthesis of PNA was carried out in pipette as described in section 2.3.1. For PNA syntheses, a Rink amide (RAM) linker (**Figure 3.22**), which is more acid labile than MBHA linker and Wang linker, was used. The peptide-linker bonds are easily cleaved with 60-95 % TFA to provide peptide amides. The RAM linker is stable to piperidine, it is therefore compatible with the Fmoc chemisty for solid phase synthesis. Moreover, the amino linker of RAM resin allows easy coupling with active esters to form amide bonds, therefore giving higher loading efficiency compared to resins containing hydroxyl linker such as Wang resin.

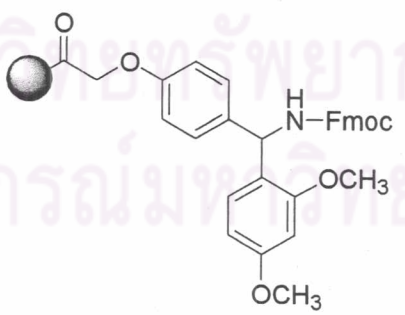


Figure 3.22 TentaGel S RAM Fmoc resin

The TentaGel resin containing a polystyrene matrix with a PEG modifier was chosen as the support for PNA synthesis because it is swellable in many solvents, especially in dimethylformamide (DMF) which is the only solvent used for Fmoc-SPPS in this study. The swelling property of the solid support will determine how the reactant can access reactive sites on the solid support which has a direct consequence on the synthesis efficiency.

Synthesis of PNA (**12**) consists of three important steps. These include i) depotection step, the Fmoc group protected amino was removing by using 20% piperidine in DMF to give free amine. ii) coupling step, the monomer or spacer was alternately coupled until to give the required sequence. iii) capping step, the resin-bound amino group that had not been successfully coupled was capped by an acetyl group for stopping incomplete reaction in the next cycle. The synthesis cycle of SPPS originally used in this study is exactly as described in the literature [60]. The details of which is shown in **Figure 3.23**.

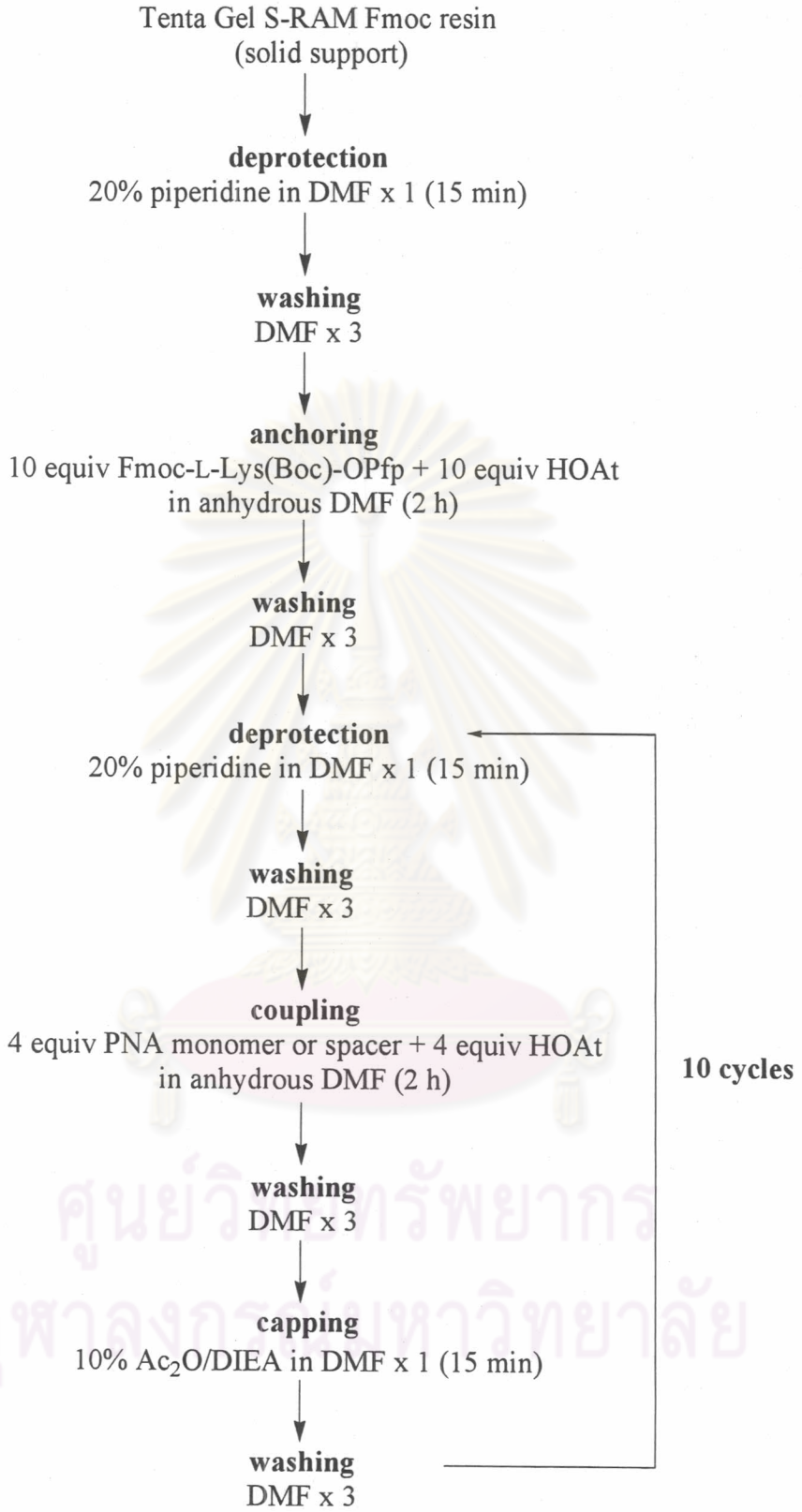


Figure 3.23 The original standard protocol for solid phase synthesis of PNA (**12c**)

First of all, the resin was swollen in anhydrous DMF. In the next step, the deprotection was carried out by treatment of the resin with 20% piperidine in DMF for 15 min in order to remove the terminal Fmoc protecting group. The deprotected Fmoc group was released from the resin-bond peptide as the piperidine-dibenzofulvene adduct by an E_1CB-type mechanism via the stabilized dibenzocyclopentadienyl anion as shown in **Figure 3.24**

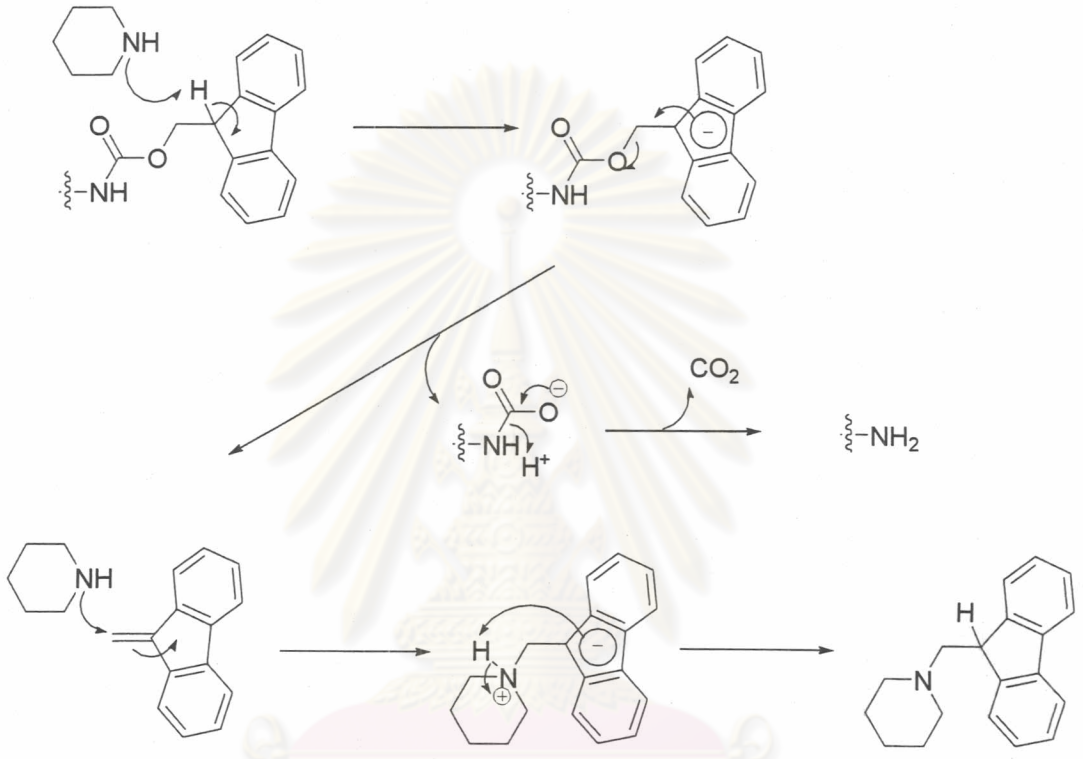


Figure 3.24 Mechanism for deprotection of Fmoc protecting group from resin bound peptide

It has been suggested that introducing a lysine at C-terminus of peptide chain could serve to prevent the self aggregation of the peptide chain due to the repulsion of the positively charged side chain [7,69]. It also increased the solubility of the peptide in aqueous medium therefore we chose to follow the same practice. After the first deprotection, the resin was then anchored with Fmoc-Lys(Boc)-OPfp as the first amino acid residue in the presence of 1-hydroxy-7-azabenzotriazole (HOAt) as an auxiliary nucleophile or "activator". The role of the activator was to increase the reactivity of the Pfp ester by a transient formation of the highly reactive HOAt ester (**Figure 3.25**). It has been reported that HOAt is more reactive and provide less

racemization than HOBt and is now being the most popular additive for peptide coupling reactions [66].

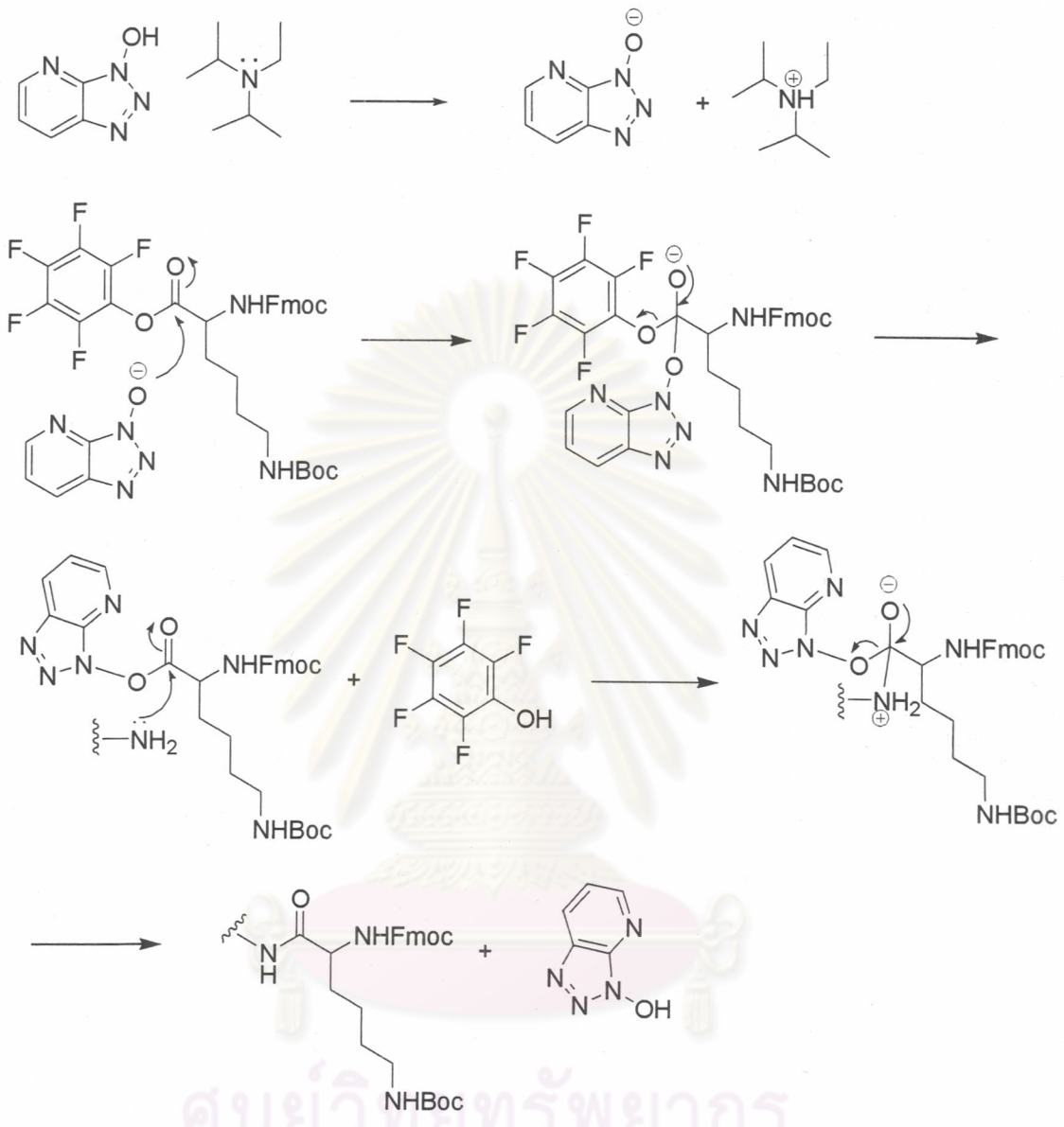


Figure 3.25 Mechanism for coupling of anchoring *via* HOAt

Because of the heterogeneous nature of SPPS, the reaction would not be completed simply with 1:1 stoichiometric proportion of reactant and would require a very long reaction time. As a result, in order to enhance the coupling efficiency within a reasonable time scale, excess and high concentration (2-10 equivalents) of the coupling reagents were generally used. Next, depotection Fmoc group was alternately coupling between activated PNA monomer (**26, 28, 30 and 32)** and the activated spacer (**38**). Extensive washing with DMF was performed after each step. The

coupling efficiency was monitored by measurement of the amounts of dibenzofulvene-piperidine adduct released upon deprotection at 264 nm.

The synthesis cycle was repeated united the growing peptide chain was extended up to 5-mer (10 cycles). The coupling efficiency in each step was calculated from UV-absorption and a typical example is shown in **Table 3.6.**

Table 3.6 UV-absorption data and percent coupling efficiency in the synthesis of T_5 PNA (**12c**)

Cycle No.	Monomer[a]	OD_{264}[b]	%efficiency
1	lysine	0.875	100.0
2	T_1	0.850	97.1
3	X_1	0.830	97.6
4	T_2	0.772	93.0
5	X_2	0.771	99.9
6	T_3	0.753	97.7
7	X_3	0.754	100.1
8	T_4	0.723	95.9
9	X_4	0.762	105.4
10	T_5	0.731	95.9
11	X_5	0.696	95.2
Overall coupling efficiency			79.5
Average coupling efficiency/step			97.9

[a] when X is the spacer (**38**)

[b] all OD_{246} were received from 100-fold dilution with absolute methanol

After addition of the final residue was completed, the N-terminal Fmoc group was removed by 20 % piperidine in DMF. Then, end-capping with Ac_2O/DIEA at the N-terminus was performed to prevent intramolecular cyclization follow in degradation initiated by the free amino group at N-terminus during subsequent treatments and long-term storage. For A, C and G-containing sequences, the nucleobase protecting groups (benzoyl and isobutyryl) were removed by treatment

with 1:1 ammonia/dioxane at 55 °C for 6 h prior to the cleavage from the resin [60]. The PNA was released from the resin by treatment with 95 % trifluoroacetic acid to give the peptide amide. The mechanism of which is shown in **Figure 3.26.** The acid-labile Boc group at the side chain of lysine was also simultaneously cleaved by the action of TFA.

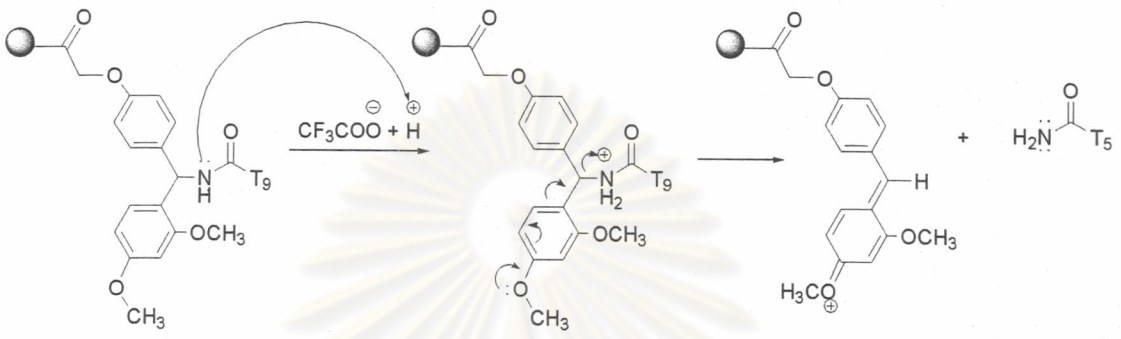


Figure 3.26 Mechanism for cleavage PNA from the resin in TFA

In case of bases sequences containing (A, C and T), the average coupling yields were approximately over 95 % per step under the standard protocol described above which was generally adopted for PNA synthesis in this research group. However, with sequences containing base G, the percent efficiency for next coupling steps of nucleobase containing monomer always drop to 80 %. In such cases, purification was difficult to achieve by HPLC. Therefore, for G-containing sequences, there need to be an optimization study to improve the coupling efficiency for this step, or a development of a better purification method. Optimizing the coupling conditions was first attempted as described in the next section.

3.2.2 Optimization of coupling step for G-containing sequences

A simple dimeric PNA Ac-TG-LysNH_2 was used as a model for study of optimizing the efficiency of the coupling step. The coupling condition was optimized by varying the coupling reagents and additives as shown in **Table 3.7**

Table 3.7 Optimization of the coupling step

monomer	Coupling method	%Efficiency			
		X→T	T→X	X→G	G→Lysine
Pfp ester	HOAt/DIEA	99	85	100	92
Pfp ester	HOAt/DIEA/LiCl	98	87	98	96
Free acid	HOAt/DIC	100	85	95	100
Free acid	HOAt/DIC/DMAP	100	84	98	95
Free acid	HATU/DIEA	97	72	100	94

Condition: Coupling step was used 4 eq Pfp ester monomer, 4 eq coupling reagent in 30 μL DMF; capping step was used 10 % acetic anhydride/DIEA in DMF.

Additives such as organic base or inorganic salts were shown to have beneficial effects to the coupling efficiency in many cases. In one experiment, DIEA was selected as organic base for deprotonation of hydroxyl group of HOAt or quaternary nitrogen in peptide bond. In another experiment, LiCl was added as a hydrogen bond destroying agent. It was reported that in some cases [70], the reason for low efficiency of peptide coupling is due to the presence of intramolecular hydrogen bonding of the growing peptide chain to form stable secondary structures. LiCl was reported to destroy these bonds by coordinating with the potential hydrogen bond donor. This will free up the peptide chain and allow it to react with the reagents better than when it is folded. However no improvement from the previous condition was observed for coupling of free acid derivatives, an activator in addition to the HOAt auxiliary nucleophile was also required to generate the active coupling species. Diisopropyl carbodiimide (DIC) in the presence of HOAt or *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate (HATU) in the presence of DIEA were selected as the potential activators. HATU was been widely used as a coupling reagent for coupling reaction of free carboxylic acid monomer and particularly for PNA synthesis [71-72]. Double coupling and increasing coupling time was preformed to ensure that the coupling efficiency was as high as possible were attempted. However, for coupling T monomer was to the spacer on G, the efficiency still dropped to ~80% regardless of the coupling conditions. However, comparably good efficiency (>95 %) were obtained for other couplings for all coupling conditions tested. In the light of these results, we chose to use the previous condition for

coupling (Pfp/HOAt) and moved to develop a purification step instead. This will be described in the next section.

3.2.3 Development of purification step

In SPPS, even after coupling with excess reagent, the coupling efficiency was not always approaching 100%. Some of the residual free amino groups on the resin have not been coupled and was still reactive enough to react with the monomer in the next coupling cycle. This situation was therefore resulted in skipping of one coupling cycle which brought about missing of a repeating unit in peptide chain. In fact, this might happen more than once and might occur anywhere in peptide chain. The resulting incomplete peptides are generally called "truncated sequence" or "deletion sequences" and would contaminate the peptide product. Furthermore, these truncated peptides generally possess similar polarity to the desired peptides product which led to difficulty in purification. Capping with a reactive reagent such as acetic anhydride/DIEA can reduce the formation of deletion sequences by stopping the incomplete peptide chains from growing (**Figure 3.27**). Nevertheless, separation of the acetylated incomplete peptide from the complete peptide are still not trivial. The alternative approach for synthesis of mixed base PNA was to develop a method for purification of the successful sequence from the incomplete sequences produced by solid phase peptide synthesis based on this idea. In principle, this should be possible by separation using a hydroprobic capping reagent to attach a hydrophobic tag to the incomplete sequences during the capping steps.

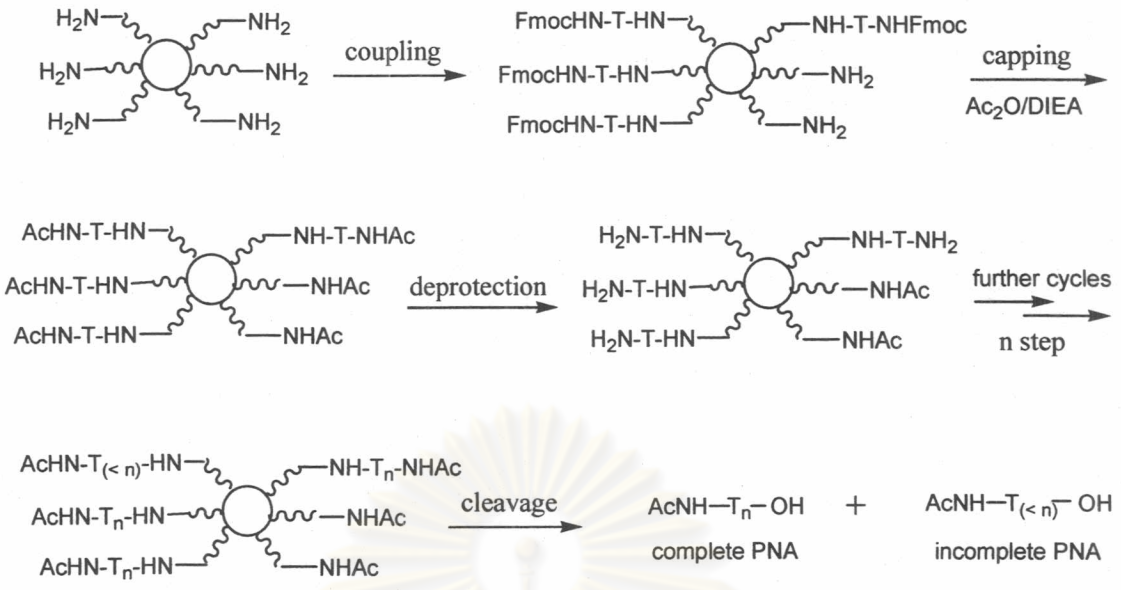


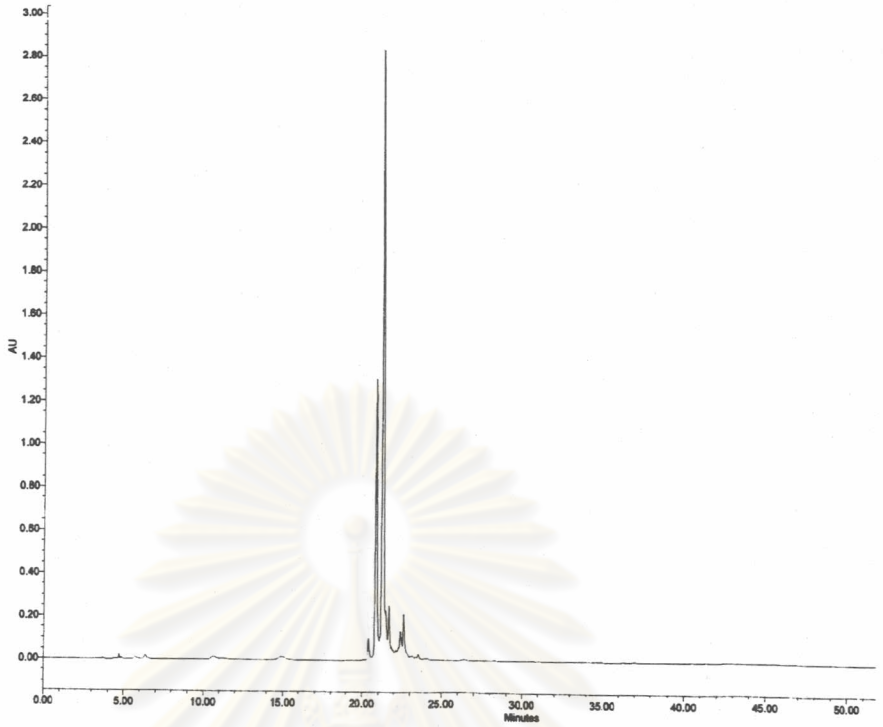
Figure 3.27 Capping of truncated peptide

In 2004, Kumar developed capping reagent as long chain fluorous for capping deletion sequence [73]. The products were simply purified by filration through fluorous silica gel to remove the fluorous-tagged incomplete peptide. However, the capping reagent described is expensive and not commonly available and require highly expensive fluorous silica. To make the concept more practical, inexpensive hydrophobic capping reagents such as long-chain carboxylic acid chlorides or anhydrides were attempted. It was hoped that the incomplete peptides after tagging with the hydrophobic group would be more hydrophobic than the complete peptide and could be easily separated by chromatography.

After many attempts, it was found that 10% lauroyl chloride/DIEA in DMF was the ideal hydrophobic capping reagent to change the polarity of incomplete PNA chain and also to diminish the nucleophilicity of the nitrogen atom at N-terminus as an lauroylamide which would not be able to grow in the next coupling cycles.

The PNA sequence Ac-TGAC-LysNH_2 (**12a**) was used as a model for studying of purification by this approach. A comparison of the crude HPLC chromatogram with that obtained from standard Ac_2O capping was shown in the **Figure 3.28**

(a)



(b)

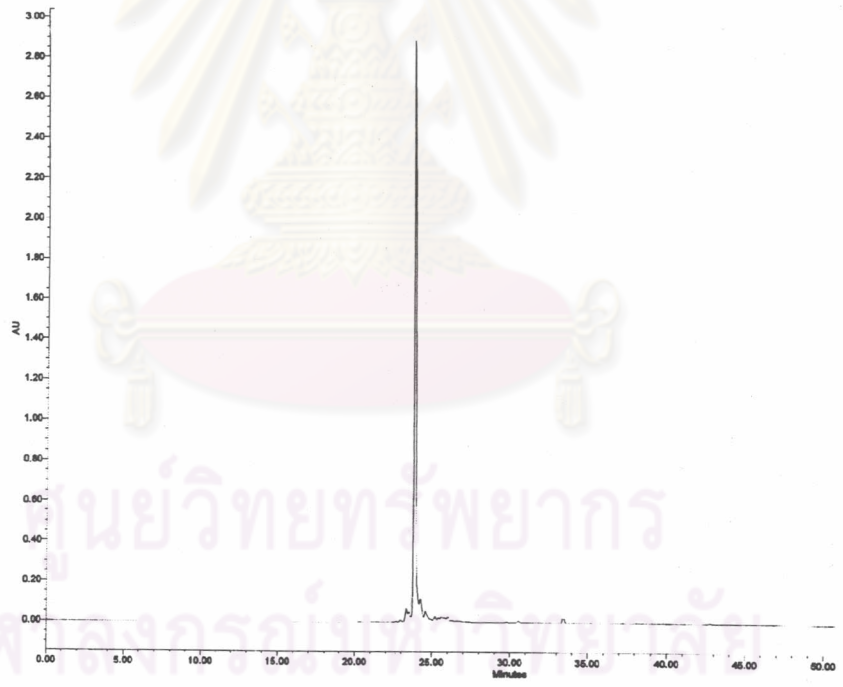


Figure 3.28 HPLC chromatogram of crude PNA Ac-TGAC-LysNH$_2$ (**12a**) (a) capped with Ac_2O (b) capped with lauroyl chloride

Surprisingly, reverse phase HPLC analysis revealed a single peak in the chromatogram of lauroyl chloride capped PNA while chromatogram of the reaction capped with Ac_2O showed a number of inseparable impurities due to incomplete

PNA. The crude PNA (**12a**) obtained from capping with Ac_2O and with lauroyl chloride were further characterized by MALDI-TOF mass spectrometry as shown **Figure 3.29**. The crude PNA (**12a**) from capping with Ac_2O showed a mass peak of the complete sequence at *m/z* 1537.7 and the T deletion sequence at *m/z* 1425.6 whereas the product from capping with lauroyl chloride showed only one peak at *m/z* 1536.7.

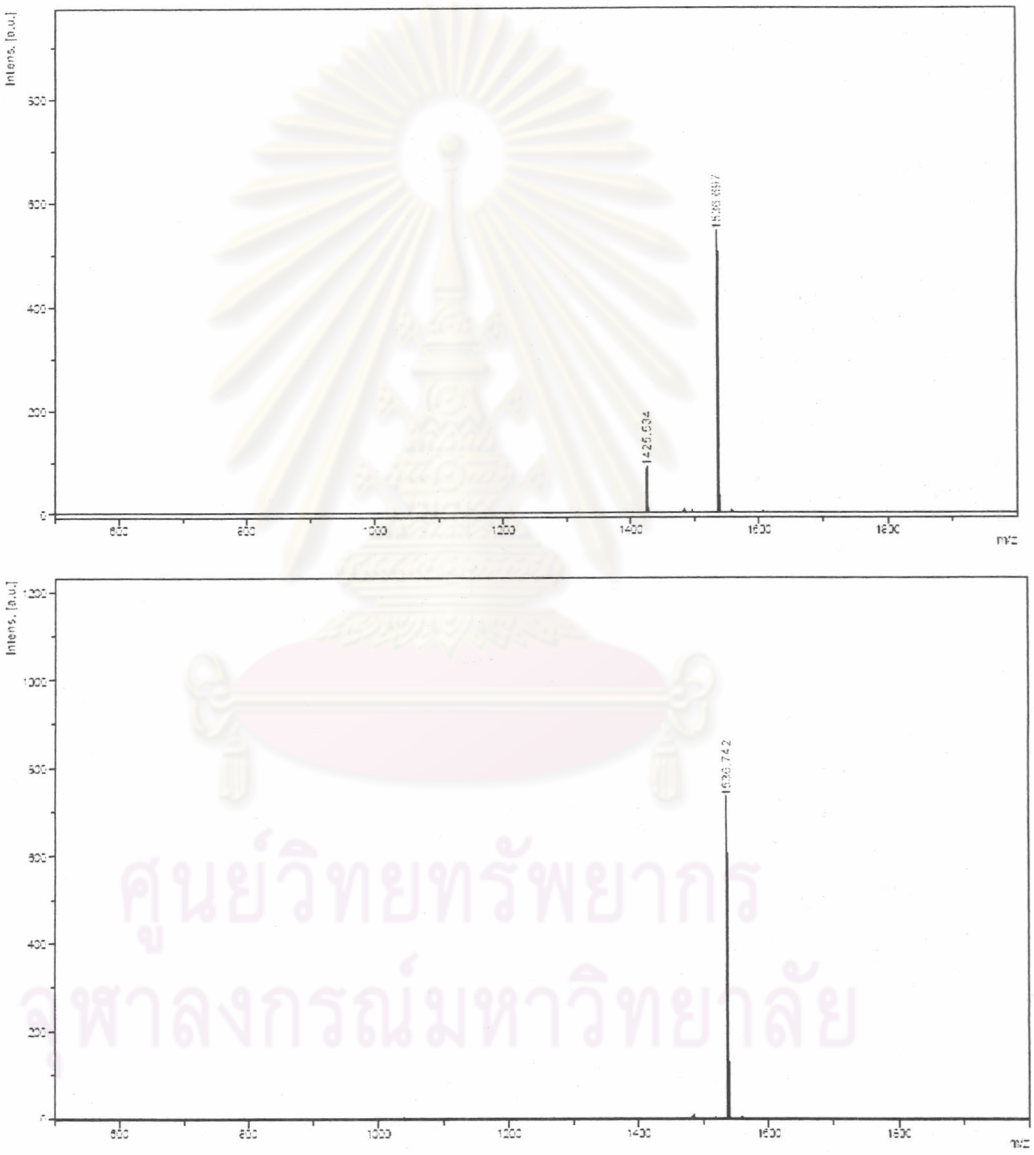


Figure 3.29 MALDI-TOF mass spectrum of crude PNA Ac-TGAC-LysNH_2 (**12a**) (a) capped with Ac_2O (b) capped with lauroyl chloride

We proposed that the lauroyl capped incomplete PNA were only slightly soluble in water. Furthermore, the soluble parts were hydrophobic enough to be

trapped on nylon the filter before HPLC analysis and purification. Therefore only complete PNA pass through nylon filter. At first it was hoped that this will provide a very efficient way for PNA purification without requiring HPLC. Unfortunately for longer sequences the impurities could pass through the nylon filter for a significant extent. Attmpts to remove these impurities by passing through a reverse phase silica was not successful and therefore HPLC purification were usually necessary. However, since the incomplete sequences now contain a hydrophobic tag, the retention times are distinctly longer than the complete PNA sequence (which contains only acetyl cap). It was thus concluded that the lauroyl chloride capped method that have been developed still offers an efficient solution for PNA purification which was not possible in case of traditional Ac_2O capping.

3.24 Synthesis of longer mixed-base PNA sequence

The mixed-base PNA was synthesized using the optimized condition obtained above. By employing lauroyl chloride as the capping reagent instead of Ac_2O, the complete PNA (**12**) could be better separated from the incomplete PNA by reverse phase HPLC. During the coupling step, we chose Pfp/HOAt method as originally used, but DIEA was also added as an additive to improve the reaction rate. In the presence of DIEA, over 90 % coupling efficiency was obtained in only 30 minutes therefore the original coupling time of 2 have been dramatically decreased. The new protocol of solid phase of PNA was shown in **Figure 3.30**.

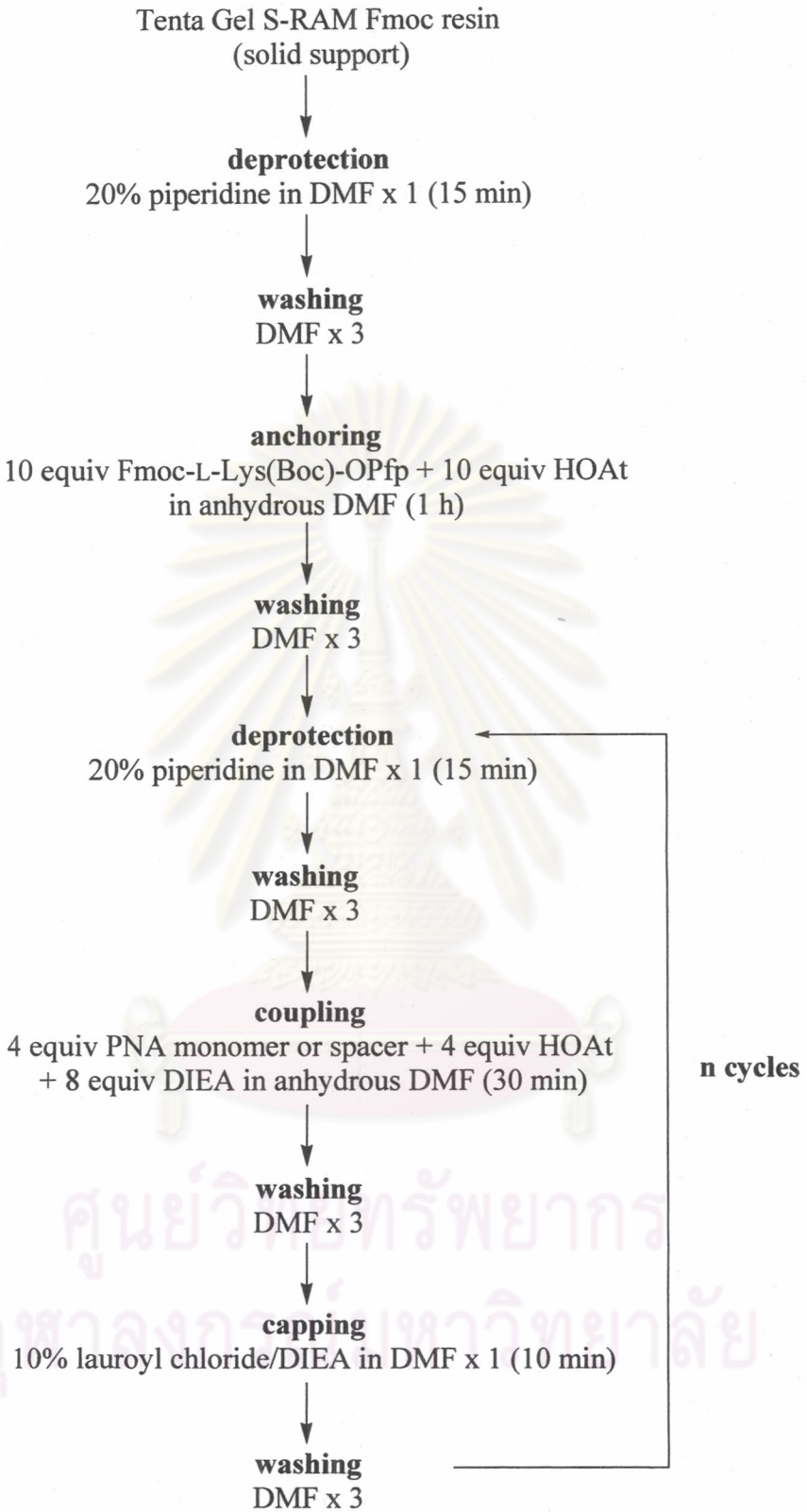


Figure 3.30 The protocol for solid phase synthesis for mixed-base PNA sequence

The new PNA synthesis and purification conditions allow straightforward access to almost any sequences of PNA with the base length between 5-15. A complete synthesis cycles for a 10 mer PNA can be readily performed in only a few days (excluding purification). The coupling efficiency data of all PNA synthesized in this work was shown in **Table 3.8.**

Table 3.8 Quantitative analyses showing the efficiency of a 1 μmol scale peptide synthesis

PNA	Sequence	Scale (μmol)	Coupling efficiency (%)	
			overall	average
12c	Ac-TTT TT-LysNH_2	1.0	79.5	97.9
12e	Ac-TTT TTT T-LysNH_2	1.2	96.1	99.6
12f	Ac-TTT TTT TTT-LysNH_2	5.0 (split in to 4 reactions)	82.8	98.1
12g	Ac-TTT TAT TTT-LysNH_2		85.4	98.4
12h	Ac-TTT TGT TTT-LysNH_2		77.8	97.5
12i	Ac-TTT TCT TTT-LysNH_2		96.0	99.6
12d	Ac-AAA AA-LysNH_2	1.0 (split in to 2 reactions)	87.3	98.6
12k	Ac-AAA AAA AAA-LysNH_2		84.6	97.9
12j	Ac-AAA ATA AAA-LysNH_2	2.0 (split in to 3 reactions)	91.0	99.1
12l	Ac-AAA AGA AAA-LysNH_2		87.7	98.7
12m	Ac-AAAA ACA AAA-LysNH_2		91.0	99.1
12n	Ac-TCA CTA CTA-LysNH_2	2.0 (split in to 4 reactions)	77.0	97.4
12o	Ac-TCA CAA CTA-LysNH_2		81.7	98.0
12p	Ac-TCA CGA CTA-LysNH_2		81.6	98.0
12q	Ac-TCA CCA CTA-LysNH_2		84.4	98.3
12r	Ac-GTA GAT CAC T-LysNH_2	1	79.5	98.9
12s	Ac-AGT GAT CTA C-LysNH_2	1	69.5	98.3
12t	Ac-CAT CTA GTG A-LysNH_2	1	57.2	97.4
12u	Ac-GAC ATG ACA T-LysNH_2	1	85.8	98.5
12v	Ac-TAT GTA CTA T-LysNH_2	1	63.5	97.9
12w	AcLys-GCT ACG TCG C-NH_2	1	74.7	98.6
12x	AcLys-TGT ACG TCA CAA CTA-NH_2	1	70.5	98.9

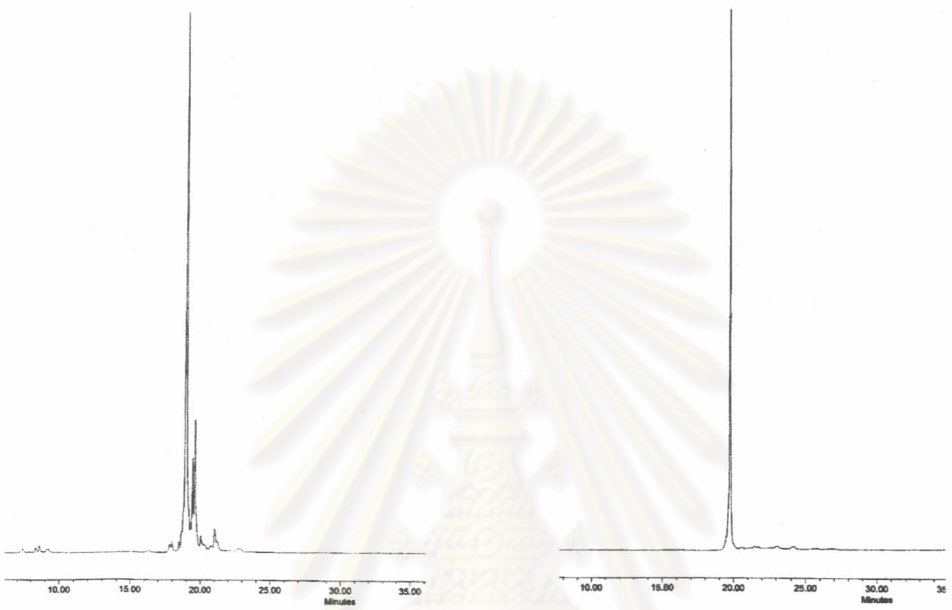
The successfully synthesized PNA were purified by reverse-phase HPLC and characterized by MALDI-TOF mass spectrometer as shown in **Table 3.9**.

Table 3.9 Characterization data of ACPC PNA sequences

PNA	t_R	$M{\cdot}H^+$ (found)	$M{\cdot}H^+$(calcd)	% error
12c	18.81[a]	1848.306	1848.480	0.010
12e	19.45[a]	2513.472	2513.176	0.012
12f	27.24[b]	3177.249	3177.472	0.007
12g	27.31[b]	3186.709	3186.483	0.007
12h	27.48[b]	3205.015	3202.478	0.014
12i	27.33[b]	3161.131	3162.472	0.042
12d	21.63[c]	1893.414	1893.938	0.028
12k	22.28[c]	3258.678	3258.577	0.003
12j	22.35[c]	3249.666	3249.565	0.003
12l	22.22[c]	3274.822	3274.572	0.008
12m	22.17[c]	3234.840	3234.566	0.008
12n	19.27[a]	3159.468	3159.508	0.001
12o	18.03[a]	3167.392	3168.520	0.036
12p	17.91[a]	3184.388	3184.515	0.004
12q	17.14[a]	3143.355	3144.508	0.036
12r	19.71[a]	3555.940	3556.669	0.020
12s	22.34[c]	3556.003	3556.669	0.019
12t	19.04[a]	3556.463	3565.669	0.020
12u	17.81[a]	3565.172	3565.680	0.014
12v	27.37[b]	3546.949	3546.662	0.008
12w	17.81[a]	3534.654	3533.653	0.028
12x	15.71[a]	5200.304	5205.433	0.099

Condition for reverse-phase HPLC: [a]C-18 column 3μ particle size 4.6 x 50 mm; gradient system of 0.01% TFA in acetonitrile/water 10:90 – 90-10 in 25 min; hold time 5 min; flow rate 0.5 mL/min. [b]C-18 column 5μ particle size 4.6 x 250 mm; gradient system of 0.01% TFA in acetonitrile/water 10:90 – 90-10 in 25 min. [c]C-18 column 3μ particle size 4.6 x 50 mm; gradient system of 0.01% TFA in acetonitrile/water 0:100 – 90-10 in 25 min.

The purified PNA exhibited only one major peak in the HPLC chromatogram (monitored at 260 nm) indicating that the products were obtained in good purity. Typical HPLC chromatogram and MALDI-TOF mass spectrum of synthesized PNA are shown in **Figure 3.31** and **Figure 3.32**. In all cases the quasi-molecular ions ($M{\cdot}H^+$) were clearly observed with a molecular weight in good agreement with the calculated values, thus confirming the identities of all PNA synthesized.



(a) (b)

Figure 3.31 HPLC chromatogram of a) crude and b) purified PNA GTAGATCACT (**12r**)

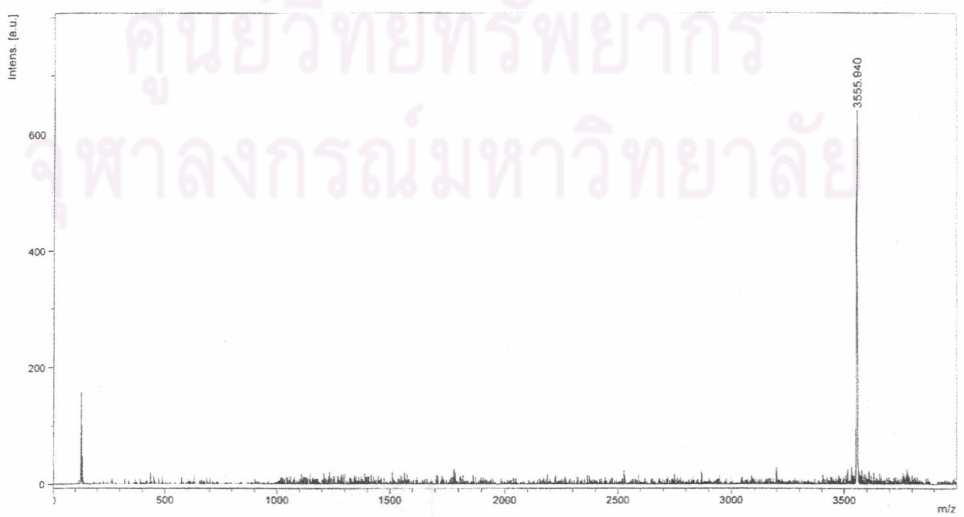
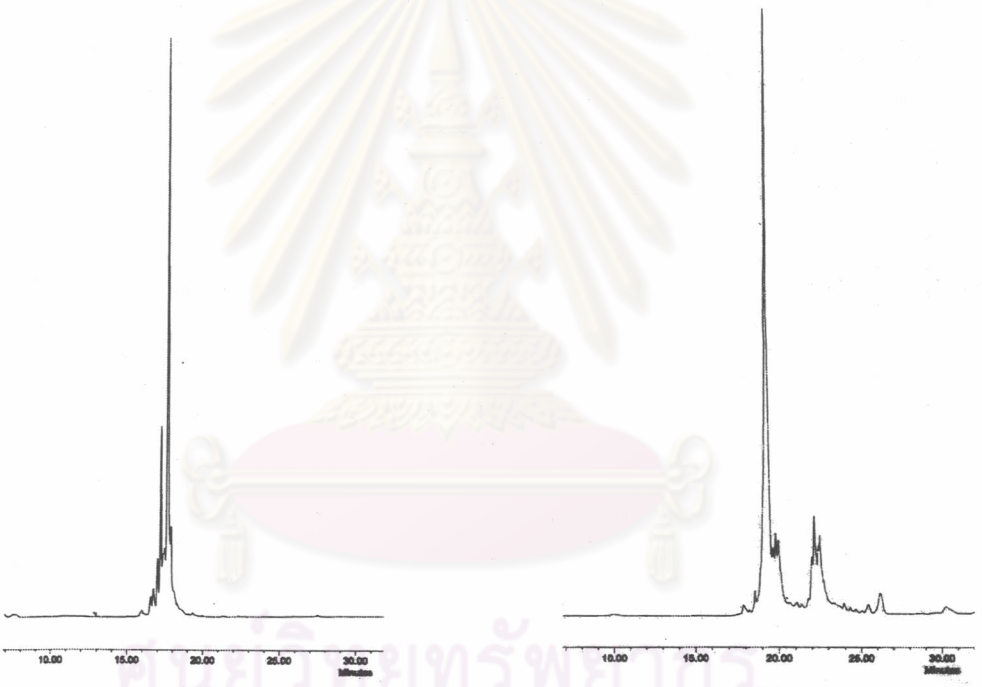


Figure 3.32 MALDI-TOF mass spectrum of purified PNA GTAGATCACT (**12r**)

During the synthesis of many test sequences, some are still problematic especially those containg multiple or adjacent purine bases. Synthesis and purification of the PNA (**12w**) and (**12x**) with CG rich and long (15 bases) sequences respectively proved to be problematic even when lauroyl chloride was used as the capping reagent. In such cases, the lysine was omitted from the C-terminus and was included at the N-terminus of the PNA after the synthesis was completed so that only the complete PNA will contain lysine. The complete PNA can now separate from lauroyl capped deletion products in the crude PNA by reverse phase HPLC due to the difference in polarity between incomplete and complete PNA as shown in **Figure 3.33**



(a) (b)

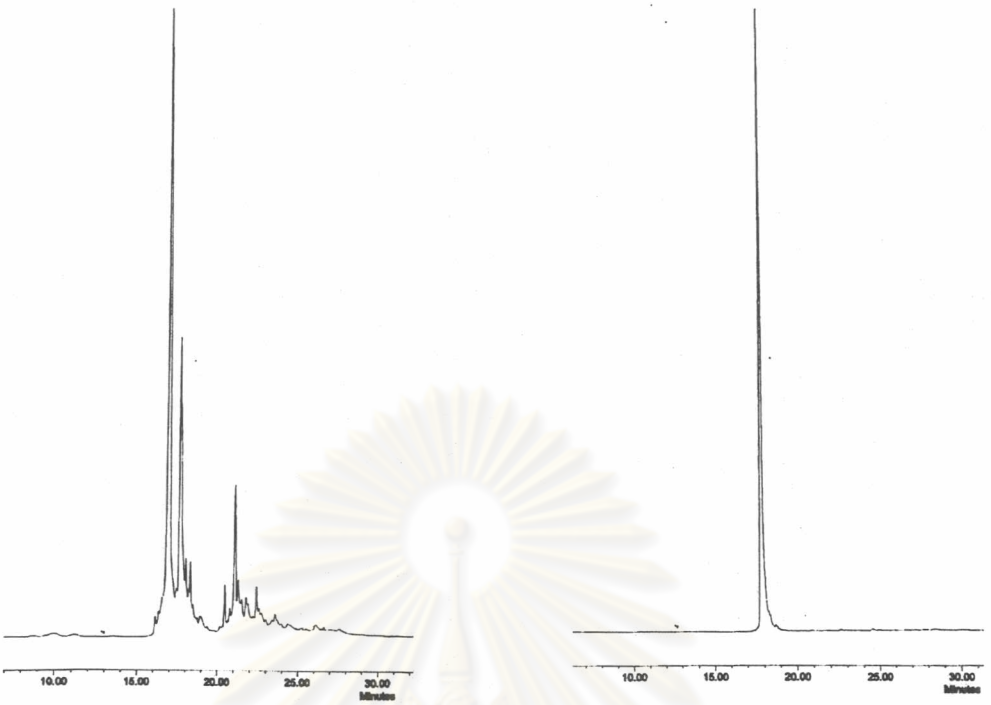


Figure 3.33 HPLC chromatogram of PNA GCTACGTCGC (**12w**), (a) lysine attached at the C-terminus and Ac_2O as capping reagent, (b) lysine attached at the C-terminus and lauroyl chloride as capping reagent, (c) lysine attached at the N-terminus and lauroyl chloride as capping reagent. (d) purified PNA obtained from (c).

In summary, several achievements have been made for the mixed-sequence PNA synthesis in this study. First of all, a new capping strategy using hydrophobic lauroyl chloride allows an efficient purification of the desired PNA from failure sequences. In difficult cases, moving the lysine to the N-terminal position can further facilitate the purification. Secondly, the improved coupling condition can reduce the time per coupling step from 2 h to 30 min without compromising yield. Thirdly, the synthesis and purification strategy are widely applicable to a variety of mixed base PNA.

3.3 Binding properties of PNA

3.3.1 Hybridization of PNA containing thymine only with DNA

The principal technique used for determining the formation and stability of PNA·DNA hybrids in this study was a thermal denaturation or melting temperature (T_m) measurement. Two strands of nucleic acid in a duplex are held together by hydrogen bonding and are also stabilized by additional hydrophobic, π-π and dipole-dipole interaction between the stacked base pairs. At high temperature, the DNA duplex separates into two single stranded DNA (DNA denaturation). Here, the DNA bases are unstracked, which results in a slight increase (10-20%) in the UV absorption of the nucleobase. This increased absorption is called hyperchromism. It can be conveniently monitored at 260 nm, the wavelength at which all four nucleobases strongly absorb. Measurement of the absorbance of a DNA complex at 260 nm while slowly increasing the temperature provides a convenient mean to observe denaturation of the complex. In a thermal denaturation experiment, the polynucleotide absorbance typically changes very slowly at first, and then rapidly rises to a maximum value (**Figure 3.34**). The temperature at which 50% of all duplexes are molten is called "melting temperature" (T_m). This value provides information on the stability of the duplex structure. Stable duplexes melt at a higher melting temperature than less stable duplexes. It should be noted that the melting process occurs abruptly within only 10-20 °C range. This is a result of cooperativity between base-base interactions. Once the input energy is sufficient to breach the first hydrogen-bonded base pair, the dissociation continue until the two strands completely separate.

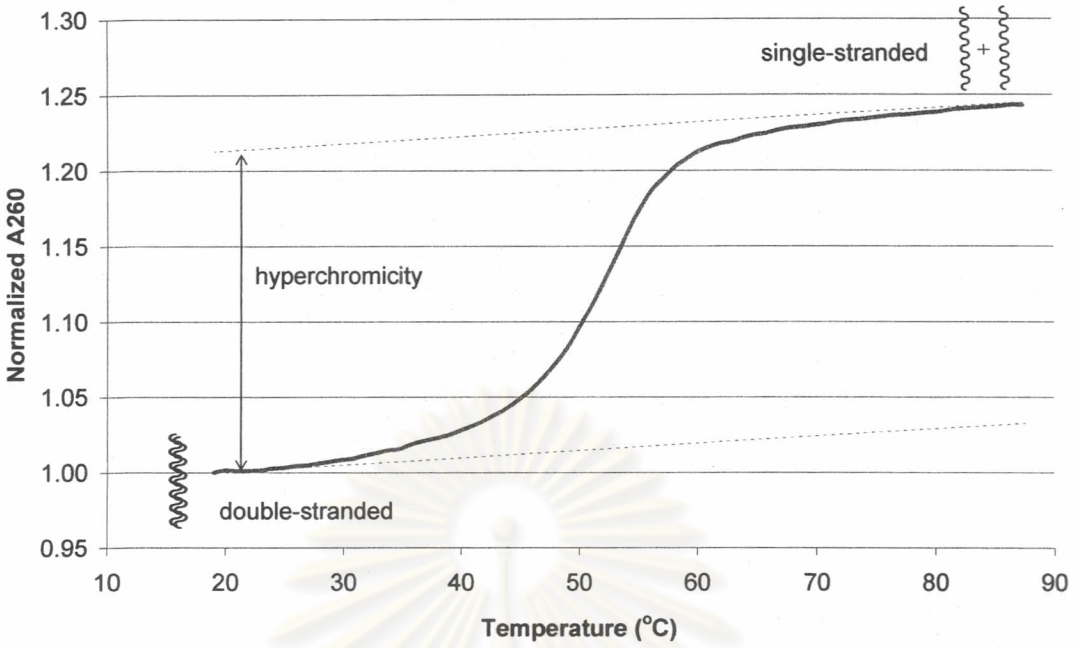


Figure 3.34 The change of UV absorbance in a typical thermal denaturation experiment of a DNA duplex.

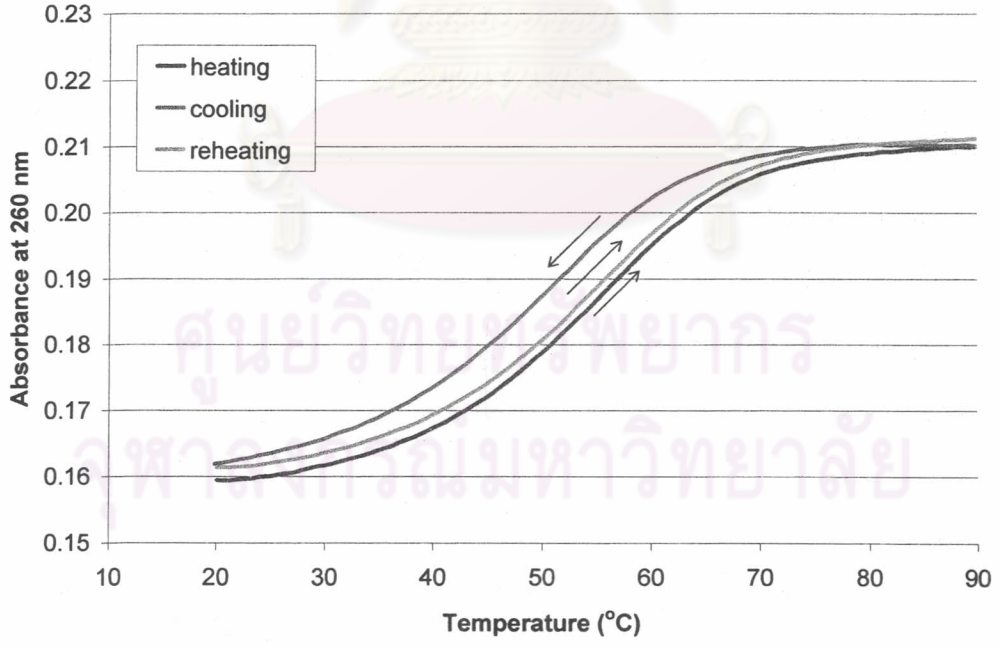


Figure 3.35 A typical T_m curve obtained from heating, cooling and reheating of a PNA·DNA hybrid (T_7·dA_7) at a heating and cooling rate of 1.0 °C/min and a hold time at the end of each cycle of 10 min.

Upon heating a 1:1 mixture of PNA (**12e**) with a T_7 sequence and its complementary DNA (dA_7), a typical sigmoidal melting curve was observed, as was evident from the sharp increase in the absorbance at 260 nm. The melting process was fully reversible and only little hyteresis was observed at the heating rate use (1.0 °C/min) (**Figure 3.35**). This indicates a fast kinetic of binding. The single-stranded PNA or DNA showed no such melting behavior in the absence of the DNA/PNA complementary strand. This suggests that the PNA can form a hybrid with DNA. In a preliminary experiment, the T_m value of a T_{10} PNA (**12b**) with $d(A_{10})$ was estimated to be >85 °C.[58] This indicates that the T_{10} (**12b**)·$d(A_{10})$ hybrid is much more stable than the corresponding DNA·DNA hybrid (T_m = 21 °C).[74] The presence of a single mismatch base in the DNA strand resulted in lowering of the T_m values of the complex by 18-25 °C. Two mismatch bases lowered the T_m further to 34 °C as shown in **Table 3.10**. The corresponding T_m data for *aeg*PNA (**1b**) [75] with the same T_{10} sequence is also included for comparison.

Table 3.10 T_m of T_{10} (**12b**) with single and double mismatched DNA. Condition: ratio of PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM, sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

PNA	DNA	T_m (°C)	Δ T_m (°C)	% hyperchromicity
12b	$d(A_{10})$	>85	-	-
12b	$d(A_4TA_5)$	67	18	38
12b	$d(A_4CA_5)$	68	17	34
12b	$d(A_4GA_5)$	61	24	27
12b	$d(A_3TATA_4)$	34	51	23
1b	$d(A_{10})$	72	13	-
1b	$d(A_4GA_5)$	59	26	-

The percentage hyperchromicity was also decreased when mismatches were introduced into the DNA strand indicating the reduction of base-base stacking in the resulting hybrid. Circular dichroism spectroscopy (CD) which is a powerful technique used to establish the interaction of nucleic acids [76] was also used for studying the interaction of PNA and DNA. If hybridization with DNA occurs, an induced CD signal should be observed especially where the nucleobase chromosphores absorb, i.e. 250-280 nm, due to the bases becoming re-oriented as they from the stack.[77] The 1:1 mixture of T_{10} (**12b**) with $d(A_{10})$ showed a significantly different CD spectrum from the calculated sum of the CD spectra of the individual components. This clearly indicates the conformational change induced by PNA·DNA hybridization.

All the synthesized PNAs were included lysine at the C-terminus to prevent self-aggregation of the oligomer as suggested by Egholm and co-worker.[75] The presence of an additional positively charged hydrophilic side chain should also improve water solubility of the oligomers. In some cases, PNAs containing lysine at the N-terminus were synthesized to facilitate purification (**Section 3.33**). It is therefore necessary to investigate the effect of position of lysine at C-terminus and N-terminus on T_m. Ac-T_9-LysNH_2 (**12f**) and Ac-Lys-T_9-NH_2 was used as a model. In a melting experiment, the Ac-T_9-LysNH_2 (**12f**) and AcLys-T_9-NH_2 bound $d(A_9)$ with similar stabilities (T_m = 76 and 77 °C respectively) (**Figure 3.36**). Therefore, the position of N-terminal lysine on the PNA has no effect on T_m value within the experimental limit. In three experiments of Tm measurement of (**12f**)·$d(A_9)$ were carried out to investigate precision of T_m value. The result showed that the T_m were 76.7, 76.8 and 76.8 °C in the same condition (**Figure 3.37**). Therefore, this method is reliable.

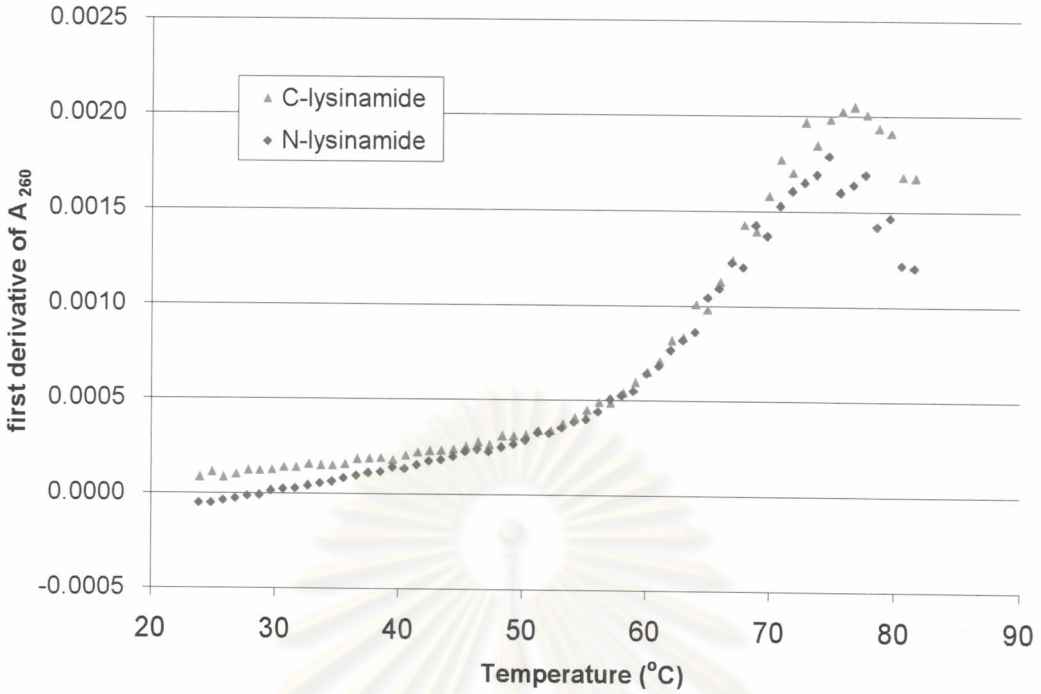


Figure 3.36 First –derivative normalized UV-T_m of C-lysinamide and N-lysinamide PNA T_9 with $d(A_9)$. The T_m were measured at a ratio of PNA:DNA = 1: 1, [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

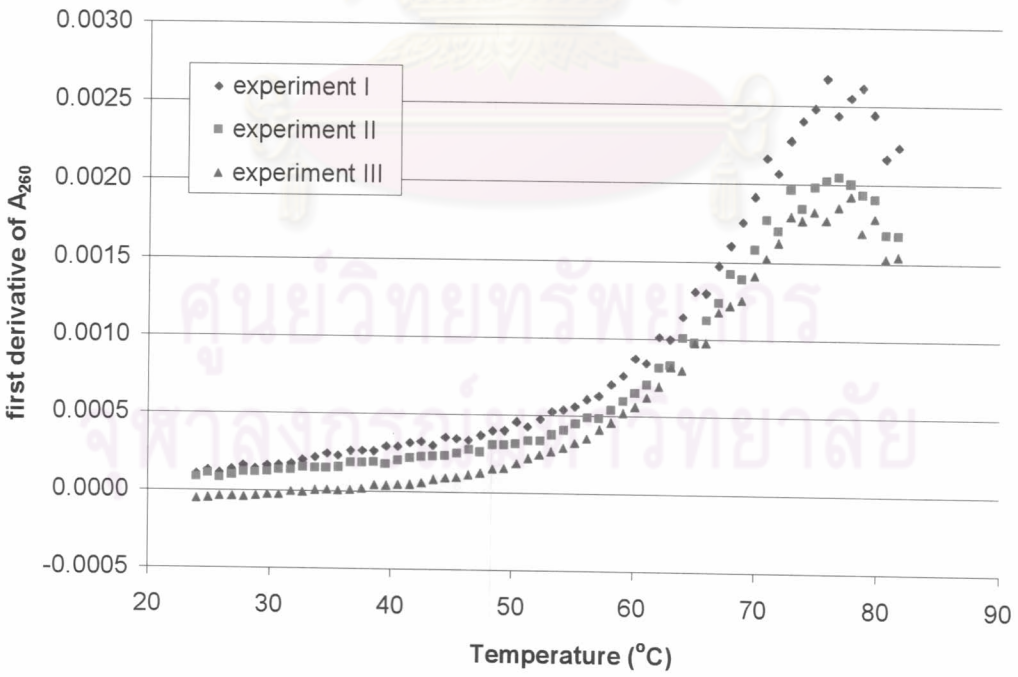


Figure 3.37 First –derivative normalized UV-T_m plots between (**12f**) and $d(A_9)$**.** The T_m were measured at a ratio of PNA:DNA = 1: 1, [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

To investigate the relationship between the base length and T_m of PNA·DNA hybrids, the melting curves of equimolar mixtures of PNA containing thymine and complementary DNA with different numbers of base pairs, n = 5, 7, 9, 10, were obtained as shown in **Figure 3.38**. Apart from n = 5, where no meeting was observed, All are much higher than the corresponding DNA·DNA duplexes with the same length. An increase in T_m of roughly 10 °C per number of additional base pair was observed. Interestingly, when the DNA component were poly(dA), an even higher T_m values were obtained. So that even the T_5 PNA hybrid showed a T_m of 37 °C (**Table 3.11**). In these cases, multiple PNA should bind simultaneously on the DNA strand and may cooperatively stabilize by additional interstrand base-base stacking.

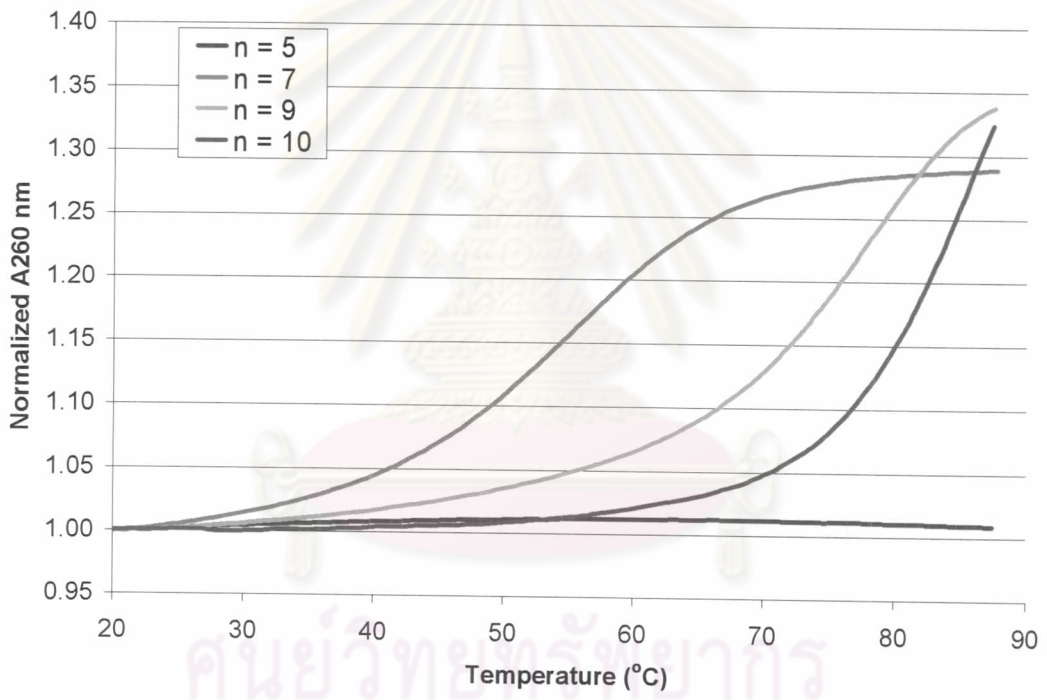
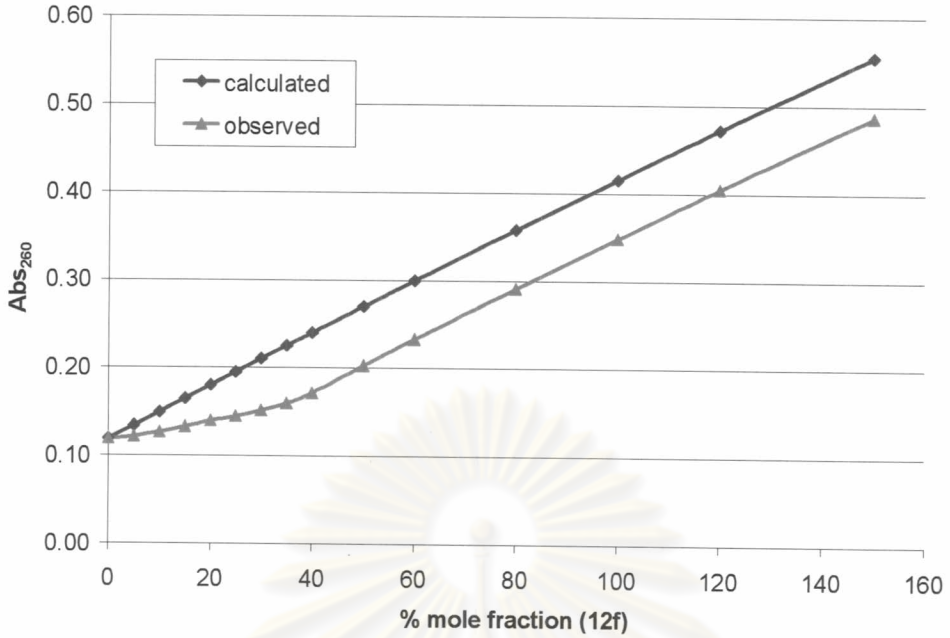


Figure 3.38 T_m curves of T_n·dA_n hybrid, when n = 5, 7, 9 and 10: Condition PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

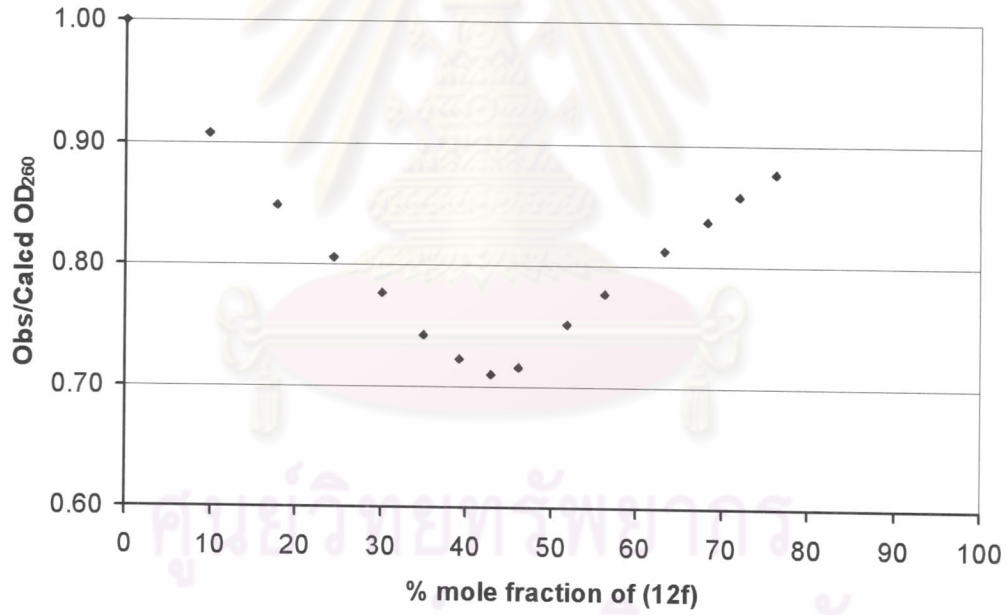
Table 3.11 T_m of PNA T_n·d(A_n) and Tn. Poly(dA) hybrid with different number of nucleobase. Condition: 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min, PNA:DNA = 1:1, [PNA] = 1 μM.

PNA	n	T_m with dAn (°C) (% hyper-chromicity)	T_m with poly dA (°C) (% hyper-chromicity)
12c	5	<20 (N/A)	37 (25)
12e	7	54 (24)	74 (>44)
12f	9	77 (>34)	>85 (N/A)
12b	10	>85 (>33)	-

The stoichiometry of the PNA·DNA complex was determined by UV titration.[78] In these experiments, the T_9 PNA (**12f**) was added to the solution of DNA solution phosphate buffered (10 mM, pH 7.0) in several aliquots and the UV absorbance at 260 nm was measured after each addition. A significant deviation of the observed UV absorbance at 260 nm from the calculated value due to hypochroism was observed when poly(dA) as analyst was titrated with T_9 (**12f**) (**Figure 3.39a**). The ratio of observed and calculated value at 260 nm was plotted against percent mole fraction of T_9 (**12f**). The result showed an inflection point around 50 %, which indicated the formation of 1:1 PNA·DNA complex. (**Figure 3.15b)**



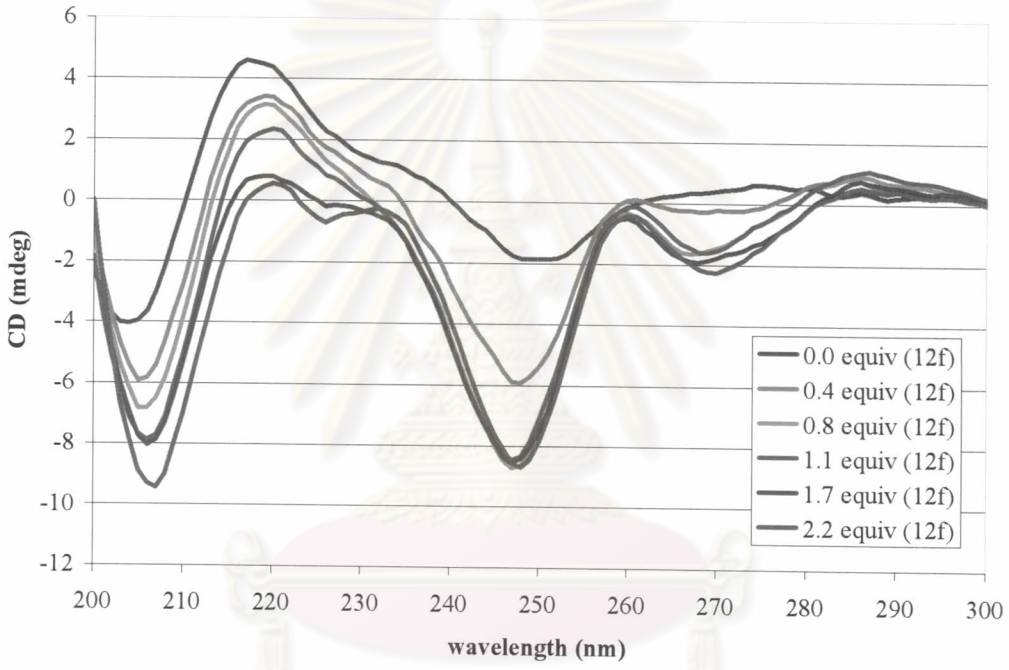
(a)



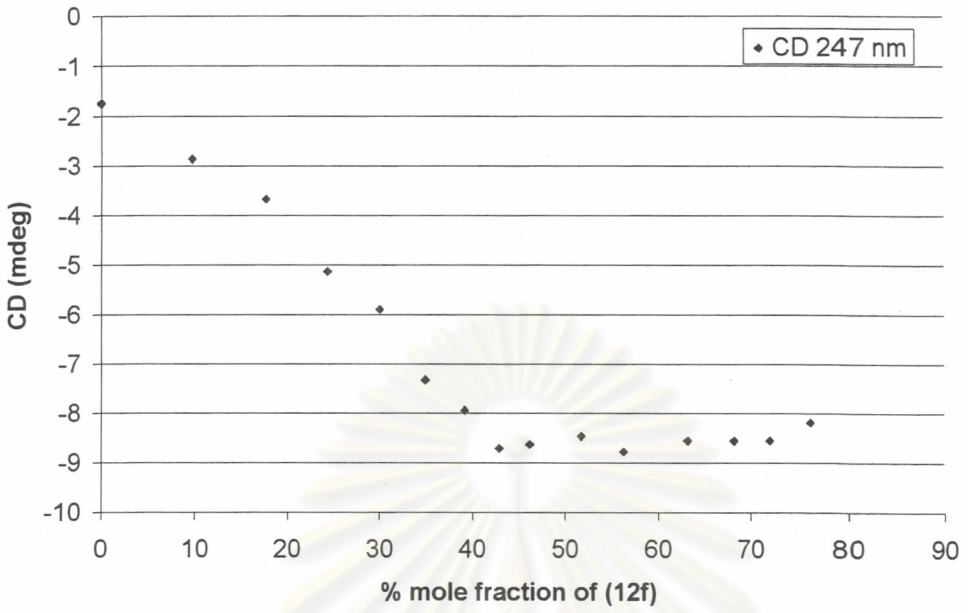
(b)

Figure 3.39 UV titration plot of poly(dA) and T_9 (**12f**).(a) a plot between A260 and percent mole of calculated and observed (b) a plot between the ratio of observed A260/calculated A260 and % mole of T_9 Condition: concentration of ply(dA) was constant at 11.1 μM, 10 mM sodium phosphate buffer, pH 7.0, 20 °C. The titrant was PNA T_9 (**12f**) (958.4 μM).

The stoichiometry was also studied by circular dichroism (CD) experiment. The CD spectra of the same T_9 PNA and poly(dA) mixture at different ratio were measured (**Figure 3.40**). The CD signal of the 1:1 hybrid formed between (**12f**) and poly(dA) showed negative bands at 207, 247, and 268 nm and positive bands at 221, 260, and 286 nm (**Figure 3.40a**) which are similar to those of the duplex formed between poly(dT) and poly(dA).[79] By plotting CD intensity at certain wavelength against molar ratio of PNA:DNA, the same 1:1 stoichiometry was established. This is in good agreement with the UV titration experiment above.



(a)



(b)

Figure 3.40 (a) A CD titration plot of poly(dA) and PNA T_9 (**12f**). (b) A plot between % mole fraction of PNA T_9 (**12f**) and the CD intensity at 247 nm. Condition: concentration of poly(dA) was constant at 11.1 μM, 10 mM sodium phosphate buffer, pH 7.0, 20 °C. The titrant was PNA T_9 (**12f**) (958.4 μM).

The effect of ionic strength was also studied. Different concentration of NaCl had relatively little effect on T_m of PNA·DNA hybrids (**Table 3.12**). This is not surprising considering that the PNA is an uncharged analogue of DNA. Nevertheless, a slight decrease in T_m was observed at higher salt concentration. There appears to be a linear relationship between T_m and logarithm of the salt concentration (**Figure 3.41**). This phenomenon is in opposite direction to that of DNA·DNA hybrid but similar to Nielsen's PNA·DNA hybrid. It has been suggested that the counterion release upon helix formation rather than counterion uptake as in the case DNA·DNA.[79]

Table 3.12 Effect of salt concentration on T_m of 1:1 T_9 (**12f**)·d(A_9). Condition: ratio of PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM, sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

[NaCl] (mM)	T_m (°C)	(%Hyperchromicity)
1	79	33
10	77	33
100	74	34
1000	72	34

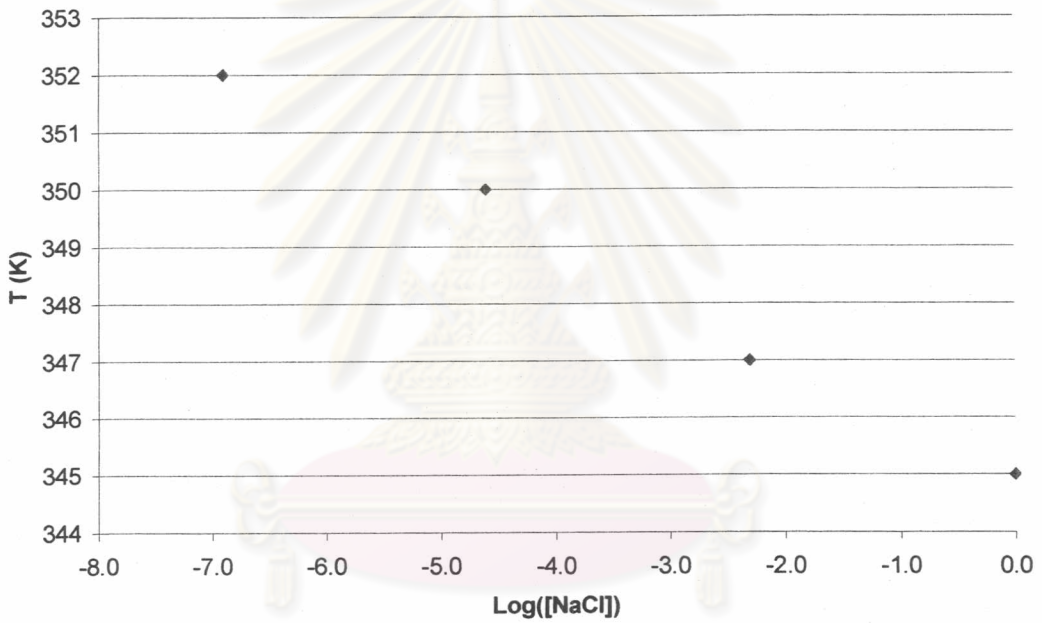


Figure 3.41 A plot between T_m of (**12f**)·d(A_9) hybrid at different salt concentration and logarithm of salt concentration.

The effect of pH was also investigated. The T_m value at low pH is similar to the values at pH = 7.0 whereas T_m value is decreased under higher pH condition (**Table 3.13**). This might be explained in terms of protonation of the lysine at low pH to form a positively charged side-chain which can interact electrostatically with the phosphate backbone of the DNA. At high pH the lysine side-chain became neutral therefore this secondary stabilization (in addition to the usual base pairing) is absent.

Table 3.13 Effect of pH on T_m of **(12f)**·d(A_9) hybrid. Condition: ratio of PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM, sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

pH	T_m (°C)	(% Hyperchromicity)
6.0	77	34
7.0	77	34
8.0	73	32

3.3.2 Stabilities of complementary and non complementary base pair in mixed base PNA

To determine the specificity of recognition of each of the four bases in PNA and DNA, the PNA oligomers (**12f**, **12g**, **12h**, **and 12i**) with the sequences $T_4\underline{X}T_4$ (X = T, A, G, C) were synthesized. The thermal stability of hybrids between these PNA and nonameric oligodeoxynucleotides with the sequences $A_4\underline{Y}A_4$ (Y = T, A, G, C) was measured (**Table 3.14**). These symmetrical sequences should form hybrids regardless of the preferred orientation of PNA·DNA binding (**Section 3.3.3**).

Table 3.14 T_m and % hyperchromicity of hybrids between PNA (T_4XT_4) and d(A_4YA_4). Condition: ratio of PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM, sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

PNA	X	Y in DNA			
		A	T	C	G
12f	T	**77** (37)	48 (25)	48 (26)	39 (19)
12g	A	40 (22)	**67** (35)	34 (21)	<20
12h	G	29 (13)	39 (21)	**64** (25)	<20
12i	C	35 (20)	26 (9)	27 (10)	**54** (25)

There are only 4 perfectly matched base pairs (A_{PNA}·T_{DNA}, T_{PNA}·A_{DNA}, C_{PNA}·G_{DNA}, and G_{PNA}·C_{DNA}) in the table (highlighted in underline). As expected, only

the hybrids formed from complementary PNA and DNA, i.e. (**12f**)·d(A_9), (**12g**)·d(A_4TA_4), (**12h**)·d(A_4CA_4) and (**12i**)·d(A_4GA_4) possessed the highest T_m values. These values ranges from 54 °C to 77 °C which are also much higher than the corresponding DNA·DNA hybrids (18-20 °C). The duplexes containing a single mismatch had a significantly decreased T_m values (ΔT_m = 19-47 °C). The maximum ΔT_m value is that of (**12g**) T_4AT_4·d(A_4GA_4) hybrid (47 °C) whereas the minimum ΔT_m value is that of (**12f**) T_9·d(A_4TA_4)(19 °C). The T_m of the mismatched PNADNA pairs was markedly lower when the both the mismatched bases are purines (i.e. G_{PNA}·G_{DNA}, G_{PNA}·A_{DNA}, A_{PNA}·G_{DNA}, A_{PNA}·A_{DNA}). The effect is most obvious when G is on the DNA strand. Such a dramatic decrease in T_m is probably explained by the sterically hindered nature of purine bases which cause destablization when they are forced together in the mismatched pairing.

In addition, percent hyperchromicity for the mismatch hybrids was also significantly decreased (about 25-37% for perfect match, and <10-27% for single mismatch) indicating a less favourable base stacking. Both the decreases in T_m and hyperchromicity for mismatched hybrids clearly indicate that the base pairing in PNA and DNA follows the normal Watson-Crick base-pairing specificity.

The above study focused on the base pairing specificity in polypyrimidine sequence. In the next study, the base pairing in polypurine PNA sequences was also investigated. Similar thermal denaturation experiments of hybrids between PNA **12j**, **12k**, **12l**, and **12m** with sequences A_4XA_4 (X = T, A, G, C) and 9-mers oligodeoxynucleotides with the sequences T_4YT_4 (Y = T, A, G, C) was carried out. As expected the hybrids formed from complementary PNA and DNA, namely (**12j**)·d(T_4AT_4), (**12k**)·d(T_9), (**12l**)·d(T_4CT_4), (**12m**)·d(T_4GT_4), possessed the highest T_m values (**Table 3.15**). These values range from 80 °C to 71 °C. In general these values are higher than the PNA T_4XT_4·d(A_4XA_4) hybrids and are also much higher than the corresponding DNA·DNA hybrids. The difference in T_m values between perfect match and single mismatch was also very large (ΔT_m = 22-37 °C).

Table 3.15 T_m of hybrids (°C, % hyperchromicity) between PNA (A_4XA_4) and d(T_4YT_4). Condition: ratio of PNA:DNA = 1:1, [PNA] = 1 μM, 10 mM, sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

PNA	X	Y in DNA			
		A	T	C	G
12j	T	**79** (>31)	53 (28)	45 (28)	41 (28)
12k	A	55 (33)	**80** (>43)	58 (30)	55 (33)
12l	G	45 (29)	48 (38)	**71** (46)	45 (34)
12m	C	48 (21)	44 (29)	45 (30)	**71** (28)

It should be noted that although the base sequence and composition are identical to the previous experiment, the T_m's were not equal. For example while the (**12i**) T_4CT_4·d(A_4GA_4) hybrid has a T_m of 54 °C, the analogous (**12l**) A_4GA_4·d(T_4CT_4) hybrid has much a higher T_m value of 71 °C. This is not surprising considering the inherent asymmetry of PNADNA hybrids. The results confirmed that the Watson-Crick base-pairing specificity in PNA·DNA hybrids are followed in both pyrimidine- and purine-rich environments.

3.3.3 Directional preference in mixed-base PNA·DNA binding

In the preceding section, only symmetrical sequences have been used in order to focus on the relative stabilities of different base pairs without having to concern about the direction of binding. On the other hand, unsymmetrical PNA sequences have two possible orientations in hybridization to DNA as shown in **Figure 3.42**. The antiparallel binding defined as when the 3' end of the DNA sequence is oriented toward the N-terminal of the PNA. The opposite binding orientation when the 5' end of the DNA sequence is oriented toward the N-terminal of the PNA is therefore defined as parallel.

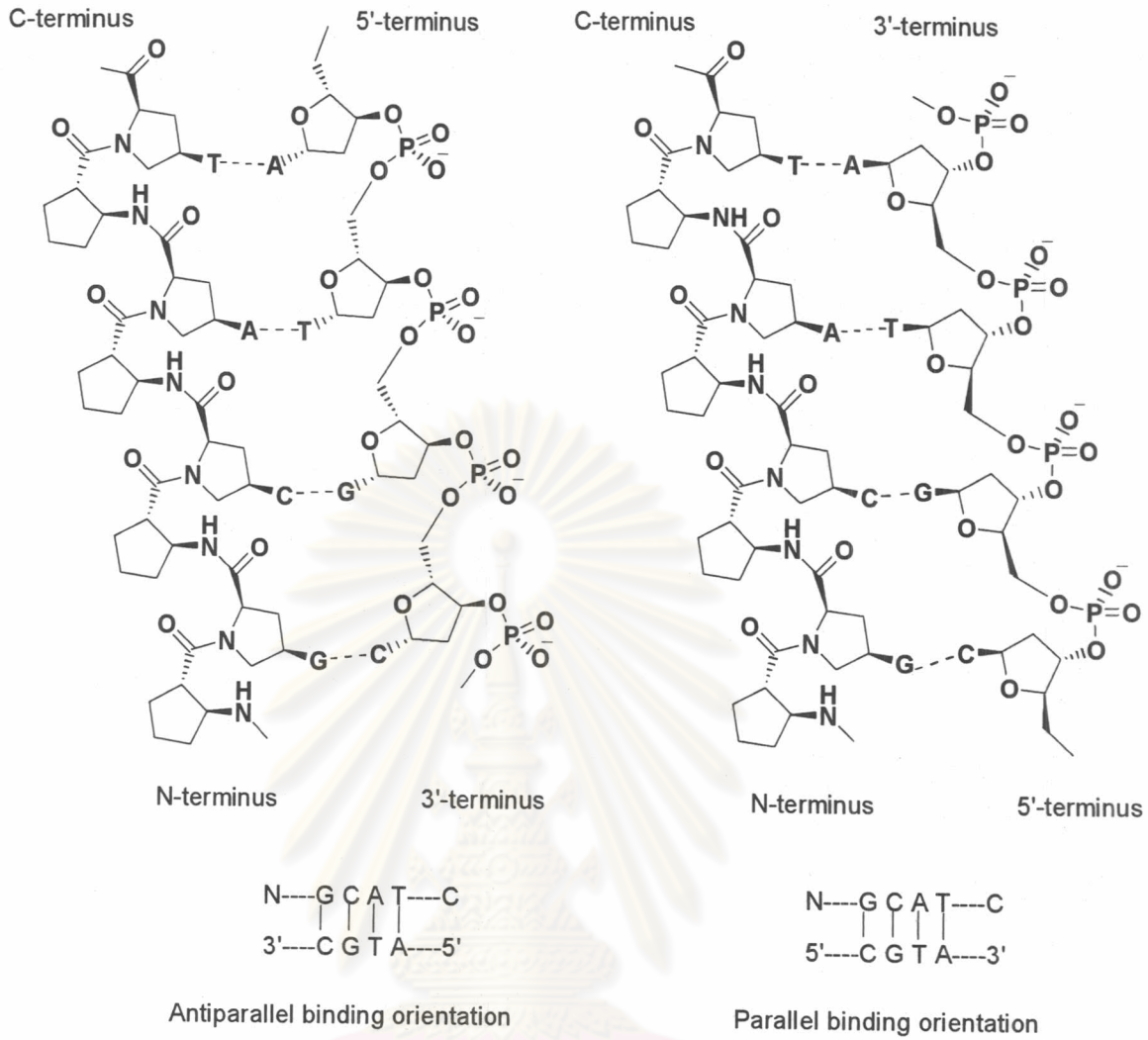


Figure 3.42 Definition of parallel and antiparallel orientations in PNA·DNA hybrids.

In order to study the orientation of PNA·DNA hybridization, a 10 mer mixed base PNA GTAGATCACT (**12r**) was synthesized. T_m of the two hybrids of (**12r**) with complementary antiparallel DNA (5'-AGTGATCTAC-3') and (5'-CATCTAGTGA-3') were measured. The antiparallel hybrid PNA (**12r**)·d(AGTGATCTAC) showed a T_m value of 58 °C, while the parallel hybrid PNA (**12r**)·d(CATCTAGTGA) showed no binding at all (T_m < 20 °C) under identical conditions (**Figure 3.43**).

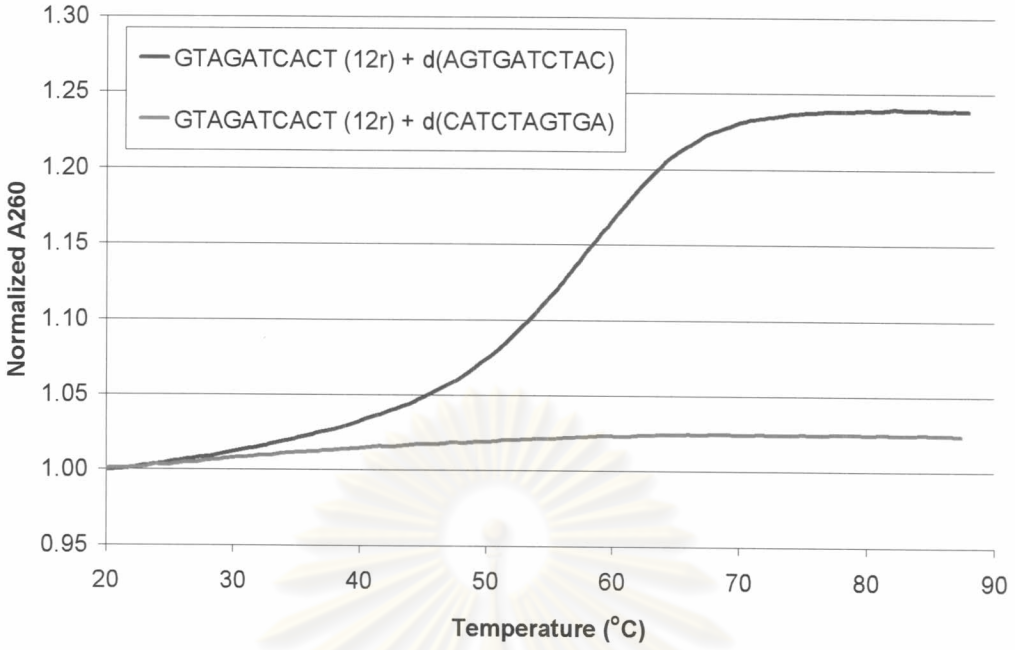


Figure 3.43 Melting curves of PNA Ac-GTAGATCACT-LysNH_2 (**12r**)·DNA hybrids (parallel and antiparallel). Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

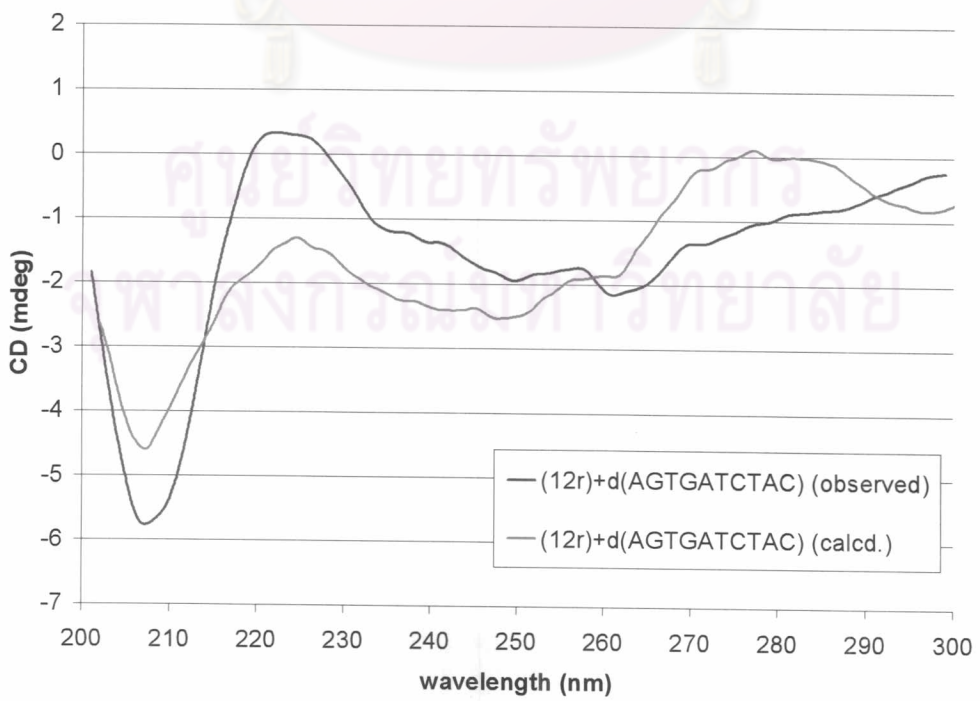
To further confirm the preference for antiparallel hybrid formation, another two 10 mer PNAs AGTGATCTAC (**12s**) and CATCTAGTGA (**12t**) were synthesized. The sequences were chosen in such a way that they are running in an opposite direction but otherwise identical. With this design, the antiparallel DNA probe for (**12s**) will be the parallel DNA probe of (**12t**) and *vice versa*. Indeed, meltings were only observed when antiparallel hybrids can form (**12s**)·d(GTAGATCACT) and (**12t**)·d(TCACTAGAGT). On the other hand, the parallel hybrids (**12s**)·d(TCACTAGAGT) and (**12t**)·d(GTAGATCACT) were not formed according to T_m measurement. This indicates that the antiparallel orientation is strongly preferred in hybridization of *ss*ACPC-PNA (**12**) and DNA.

Table 3.16 T_m of parallel and antiparallel PNA·DNA hybrids.

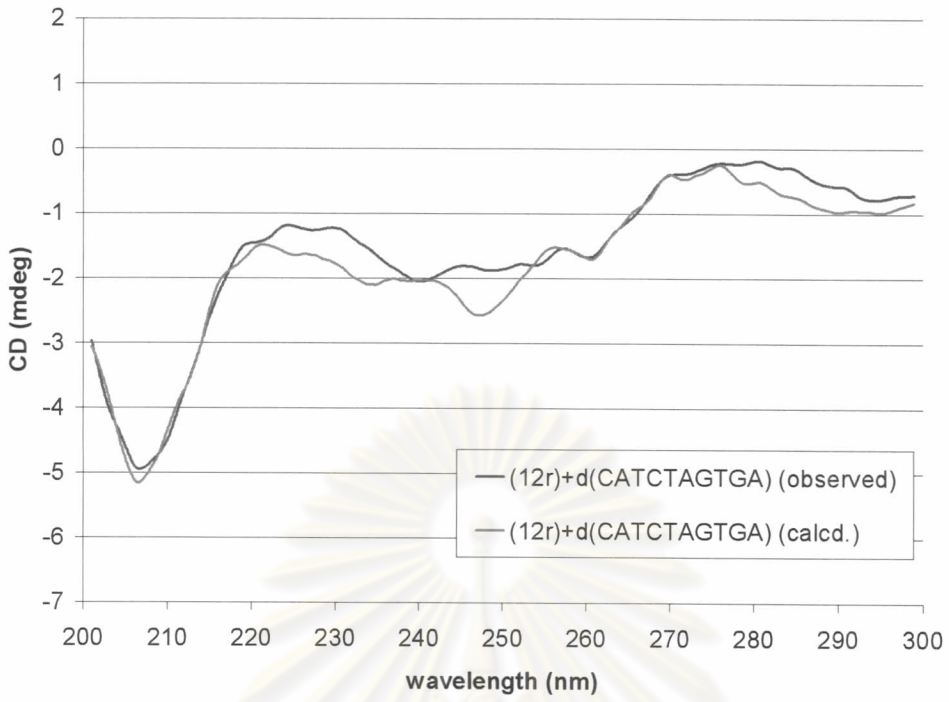
PNA	T_m (°C)[a]	
	5'-GTAGATCACT-3'	5'-TCACTAGATG-3'
AGTGATCTAC (**12s**)	61 (antiparallel)	<20 (parallel)
CATCTAGTGA (**12t**)	<20 (parallel)	61 (antiparallel)

[a] The T_m were measured at a ratio of PNA:DNA = 1: 1, [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

In addition, the direction preference is further studied by circular dichroism spectroscopy. The antiparallel PNA (**12r**)·DNA hybrid showed a distinct induced CD signal from the sum of CD spectrum of each component as shown in **Figure 3.44 (a)**. The notable feature of the CD spectrum of antiparallel hybrids includes a negative band at 209 nm and a positive band at 225 nm. The positive band near 280 nm originally present becomes flatten. On the other hand, the parallel hybrid exhibited essentially no spectral change compared to the sum of CD spectrum of each component as shown in **Figure 3.44 (b)**, confirming the absence of parallel hybrid formation The CD spectral change is more noticeable when difference spectra were calculated (**Figure 3.45**). The difference spectrum of the parallel hybrid is almost featureless but that of the antiparallel hybrid shows a definite change around 210, 225 and 275 nm as discussed earlier.



(a)



(b)

Figure 3.44 CD spectra of 1:1 mixture of PNA (**12r**) with (a) antiparallel complementary DNA and (b) parallel complementary DNA

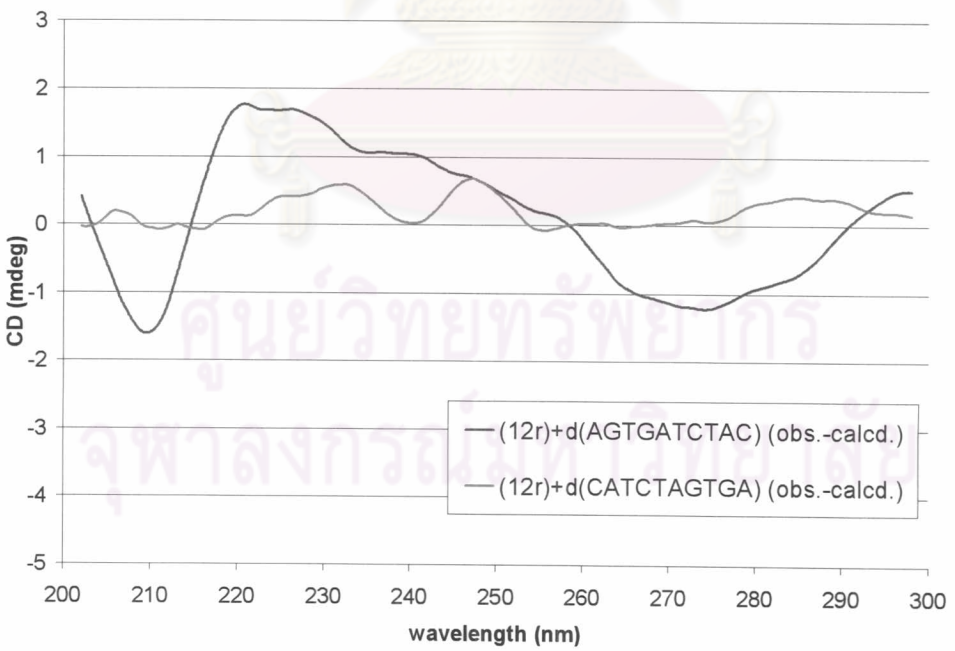


Figure 3.45 The difference CD spectra of 1:1 mixture of PNA (**12r**) (observed – sum of CD spectra of all components) and its parallel and anti parallel DNA. The concentration of the PNA strand was 1μM. The CD experiment was carried out in 10 mM sodium phosphate buffer pH 7.0 at 25 °C.

Other PNA (**12**) with unsymmetrical sequences also formed duplexes with DNA only in antiparallel mode as shown in **Table 3.17**. In all but one cases, the T_m's of the antiparallel hybrids were much higher than the parallel hybrids. No melting was observed at all of the parallel hybrids. The apparent equal stability of parallel and antiparallel hybrids of PNA (**12q**) and DNA appeared to violate this general observation at first sight. However, upon a closer look, it becomes apparent that the "parallel" DNA complementary to (**12q**) which have a sequence 5'-AGTGGTGAT-3' can in fact also bind to the PNA in antiparallel fashion to form a staggered hybrid (**Figure 3.46**). These hybrids contain no mismatches base although with one less base paring between terminals A-T. The T_m are not therefore much different from the antiparallel hybrids.

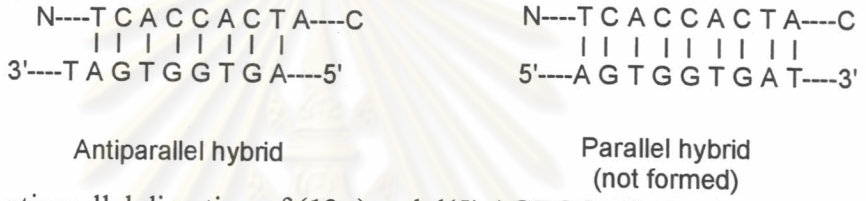


Figure 3.46 Antiparallel direction of (**12q**) and d(5'-AGTGGTGAT-3')

Table 3.17 T_m of mixed-base PNA with parallel and antiparallel complementary DNA. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

PNA	**PNA Sequence (N→C)**	**DNA (5'→3')**	**oreintation** (p = parallel) (ap = antiparallel)	**T_m (°C) (% hyper-chromicity)**
12n	TCACTACTA	TAGTAGTGA	ap	44 (18)
		AGTGATGAT	p	<20
12o	TCACAACTA	TAGTTGTGA	ap	49 (21)
		AGTGTTGAT	p	<20
12p	TCACGACTA	TAGTCGTGA	ap	45 (22)
		AGTGCTGAT	p	<20
12q	TCACCACTA	TAGTGGTGA	ap	37 (19)
		AGTGGTGAT	p	37 (12)
12r	GTAGATCACT	AGTGATCTAC	ap	58 (24)
		CATCTAGTGA	p	<20

PNA	PNA Sequence (N→C)	DNA (5'→3')	oreintation (p = parallel) (ap = antiparallel)	T_m (°C) (% hyper-chromicity)
12s	AGTGATCTAC	GTAGATCACT	ap	57 (25)
		TCACTAGATG	p	<20
12t	CATCTAGTGA	TCACTAGATG	ap	61 (16)
		GTAGATCACT	p	<20
12u	GACATGACAT	ATGTCATGTC	ap	61 (24)
		CTGTACTGTA	p	<20
12v	TATGTACTAT	ATAGTACATA	ap	61 (30)
		ATACATGATA	p	<20
12w	GCTACGTCGC	GCGACGTAGC	ap	61 (15)
		CGATGCAGCG	p	<20
12x	TGTACGTCACAACTA	TAGTTGTGACGTACA	ap	79 (32)
		ACATGCAGTGTTGAT	p	<20

The observation that PNA (**12**) strongly favors forming antiparallel hybrid with DNA is significant. The original *aeg*PNA (**1**) although slightly favours parallel hybrids, the difference instability are usually not very high (ΔT_m < 20 °C). A pronounced parallel-antiparallel selectivity is obtained only when at least a few chiral amino acids were added into the core structure of PNA (**1**). For example, a "chiralbox" containing three adjacent aminoethyllysine residues inserted into the middle of the PNA (**1**) strand resulted in an exclusive formation of antiparallel hybrid with DNA [80].

3.3.4 Relationships between T_m / base length and base composition

The relationship of T_m and base-length of PNA·DNA hybrids containing all four nucleobases is shown in **Figure 3.47**. For nonamers, four sequences of PNA were studied (**12n-12q**). Each sequence contian the same basic skeleton with one base different at the central position (T, A, G and C). The T_m's of these PNA were between 37-49 °C. The PNA (**12q**) containing two adjacent C residues in the central position

gave a markedly lower T_m (37 °C) whereas the others showed T_m in the range of 44-49 °C. For decamers, six different sequences were studied and all showed T_m in the same range (57-61 °C) regardless of the sequence (**Figure 3.47**). For the only one pentadecamer studied, a very high T_m (79 °C) was observed. It can be concluded that the longer sequence length, the higher the stabilities of the hybrids which is in good agreement with the homothymine sequences (Section 3.3.1).

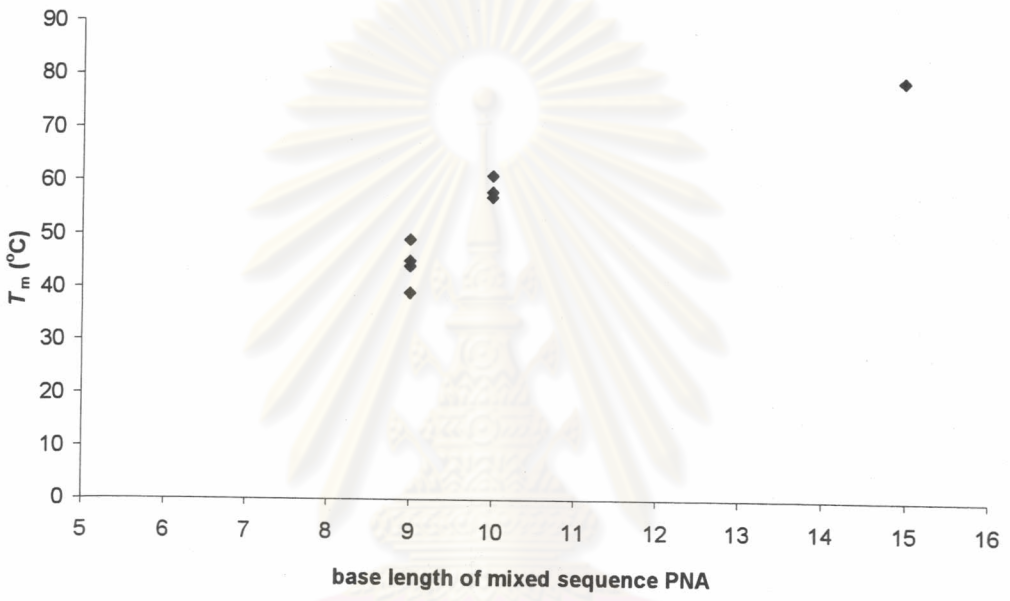


Figure 3.47 The relationship between T_m and base length of mixed sequence PNA·DNA hybrids.

In **Figure 3.48**, T_m values of ten mixed-base PNA·DNA hybrids were plotted as a function of CG content to to investigate the effect of percent CG content on T_m. Only PNA with the same base length are compared and mixed-sequences containing very high AT or CG content (>80 %) which may form unusual structures were excluded. Six of these sequences are decamer. Four of them (**12r**, **12s**, **12t**, **12u**) had CG content 40 %, have T_m in the range of 57-61 °C. Surprisingly, the sequences (**12w**) with high CG content (70 %) and (**12v**) with low CG content (20 %) also showed T_m's in the same range as PNA with lower % GC content (both = 61 °C). A similar trend was observed for the 9-mer sequences. The results indicated that T_m values depend on base length but not on percent CG content. These results are different from DNA whereby C·G pairing are stronger than AT pairing due to the

presence of three hydrogen bondings in C·G versus two in AT pairs. In addition, the stacking energy resulting from C·G pairs is higher than AT pairs [81]. The explanation for this unusual phenomenon is not yet clear without further experiments. At present we propose that the backbone of *ss*ACPC can accommodate A·T pairing better than C·G pairing. This might counterbalance the stronger hydrogen bonding and higher stacking energy of C·G pairing.

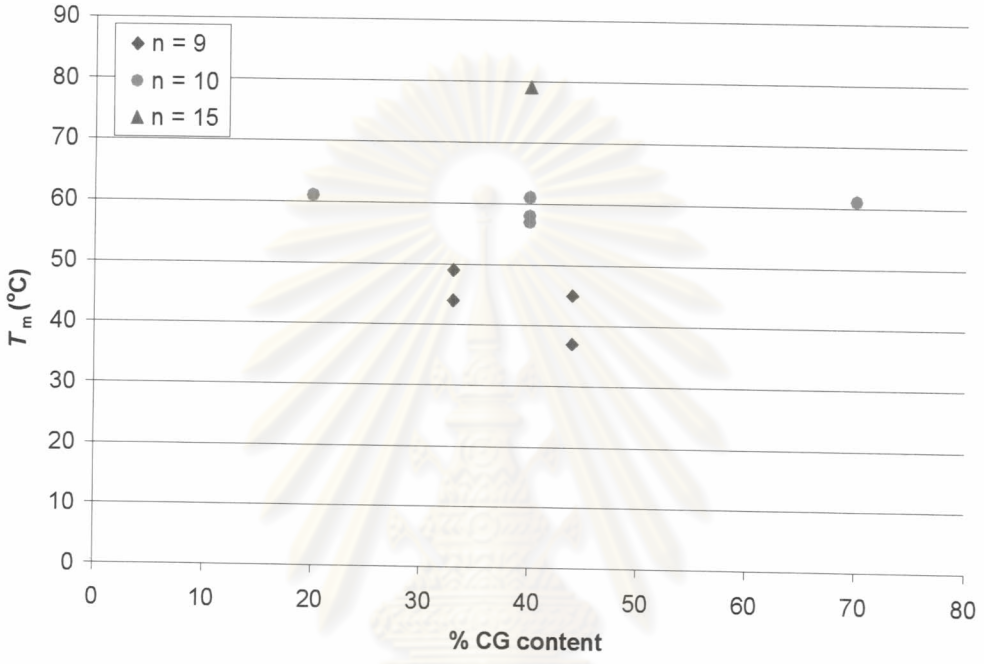


Figure 3.48 The relationship between T_m and % CG content.

The stoichiometry of the mixed base PNA·DNA hybridization was also studied by UV and CD titration. The DNA titrant d(TAGTTGTGACGTACA) was added to the solution of the 15mer PNA (**12x**) under the usual hybridization condition. The plot of the ratio of observed and calculated A_{260} as a function of % mole fraction of d(TAGTTGTGACGTACA) showed an inflection point around 50 % mole fraction of DNA, indicating the formation of 1:1 PNA·DNA hybrid (**Figure 3.49**). CD titration of PNA (**12x**) and d(TAGTTGTGACGTACA) was also studied. By plotting the CD intensity of mixtures at different ratio of reactants, the 1:1 stoichiometry of the PNA·DNA hybrid was also established (**Figure 3.50**).

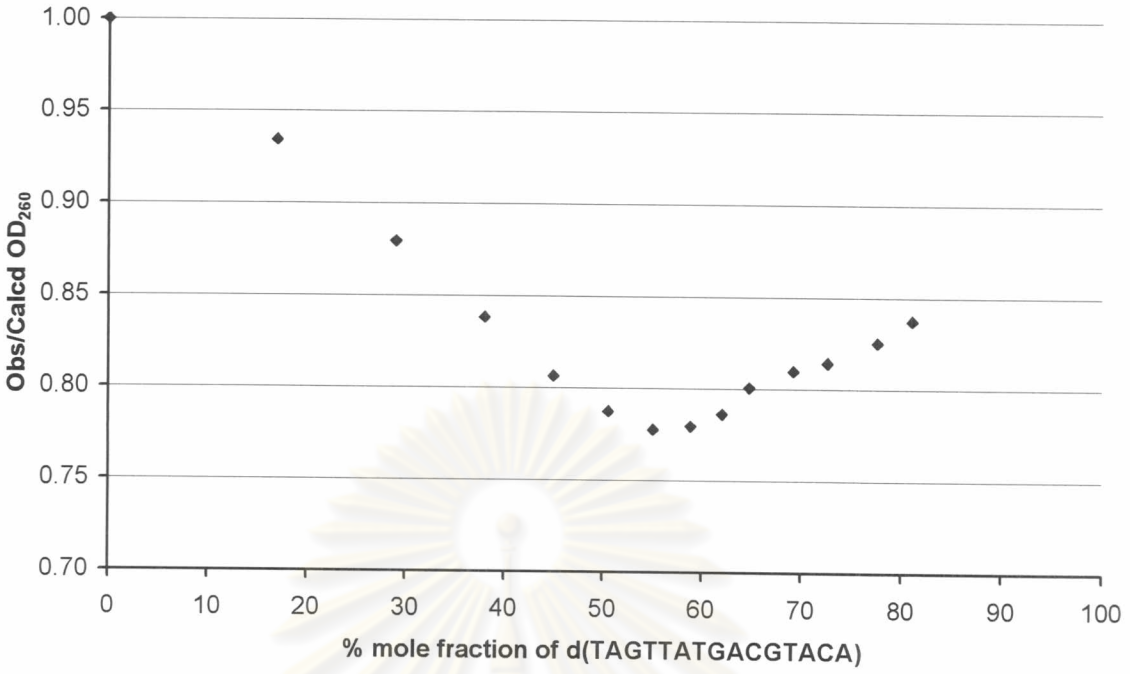
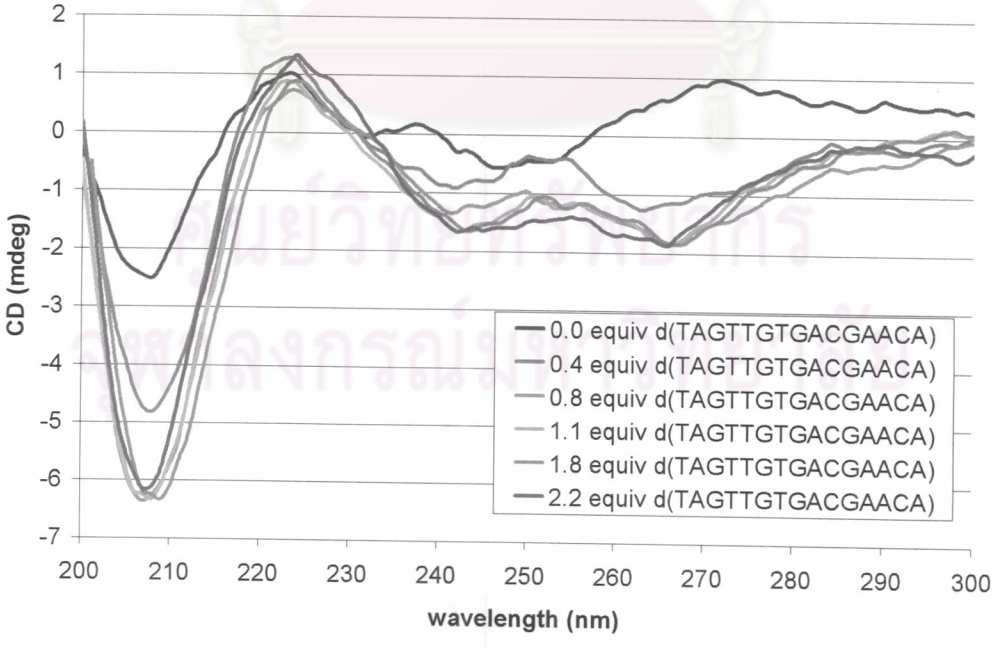
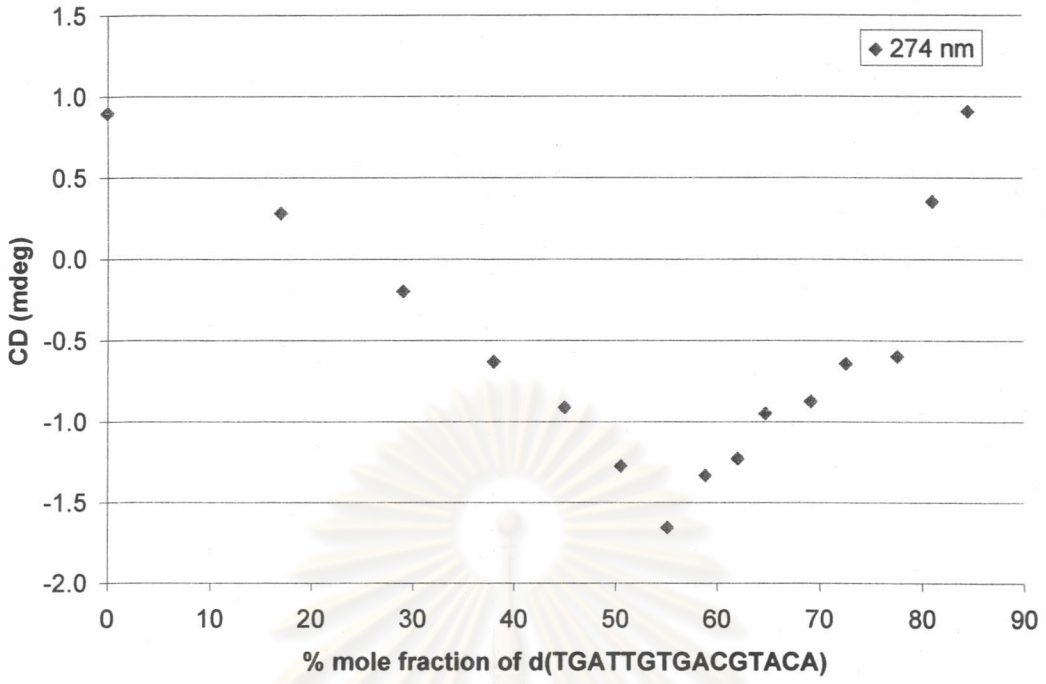


Figure 3.49 UV titration plot of PNA (**12x**) and d(TAGTTGTGACGTACA) Condition: concentration of PNA (**12x**) was constant at 8.5 μM, 10 mM sodium phosphate buffer, pH 7.0, 20 °C. The titrant was d(TAGTTGTGACGTACA) (830.9 μM).



(a)



(b)

Figure 3.50 (a) CD titration plot of PNA (**12x**) and d(TAGTTGTGACGTACA). (b) Plot between % mole fraction of d(TAGTTGTGACGTACA) and the elliplicity at 247 nm. Condition: concentration of PNA (**12x**) was constant at 8.5 μM, 10 mM sodium phosphate buffer, pH 7.0, 20 °C. The titrant was d(TAGTTGTGACGTACA) (830.9μM).

3.3.5 Effect of base mismatching in mixed sequence PNA

The presence of a mismatched base in the DNA sequence caused a reduction of T_m of the PNA·DNA hybrid. The effect of the position of mismatch bases was first studied. A series of DNA complementary to the decamer PNA (**12r**) containing a single mismatch base at different positions along the DNA strand were synthesized and the stability of the hybrids formed with (**12r**) were investigated by T_m experiment (**Table 3.18**).

Table 3.18 T_m of mixed-base PNA GTAGATCACT (**12r**) with single mismatch DNAs. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

DNA (5'→3')[a]	T_m (°C) (% Hyperchromicity)	ΔT_m (°C)
AGTGATCTAC	58 (24)	-
AGTG**T**TCTAC	34 (14)	-24
AGTGATCT**T**C	43 (15)	-15
TGTGATCTAC	59 (22)	+1

[a] Position of mismatch indicated by underlining

When the mismatch position is at the middle of the sequence, the highest change of T_m value (-24 °C) was observed. Interestingly the mismatch base located at the end of the complex does not seem to have a significant effect on T_m. The same effect was previously observed for the "parallel" hybrid between (**12r**) and DNA. This can be explained by the unzipping of the terminal base pair to compensate for strains introduced by a non-optimal internal base-pairing.

The ability of mixed base PNA to discriminate mismatching in DNA is very important in the use of PNA in diagnostic and therapeutic applications. The stability of single-mismatched hybrids should be decreased relative to the completely matched hybrid in such a way that at a specific temperature range, only the completely matched hybrids are stable and a single mismatch are not. Ideally this temperature window should be as wide a possible. The specificity of recognition between the decamer PNA (**12r**) and single-mismatched DNA probes with variation of a base at the four middle positions was determined. The melting temperatures of all 13 possible single-mismatched hybrids were shown in **Table 3.19**. In all cases the T_m values of single mismatched hybrids were lower than the perfect match around 19-29 °C. The percent hyperchromicity was also decreased when mismatches were introduced into the DNA strand indicating a less favorable base stacking.

Table 3.19 T_m of hybrids of mixed-base PNA GTAGATCACT (**12r**)with fully matched and single mismatched DNA. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

DNA sequence (5'→3')	Type of mismatch	T_m (°C)	(% Hyper-chromicity)	ΔT_m (°C)
AGTGATCTAC	no mismatch	58	24	0
AGTGAT**A**TAC	C_{PNA}-A_{DNA}	29	15	-29
AGTGAT**T**TAC	C_{PNA}-T_{DNA}	34	19	-24
AGTGAT**G**TAC	C_{PNA}-G_{DNA}	32	16	-26
AGTGA**A**CTAC	T_{PNA}-A_{DNA}	37	18	-21
AGTGA**C**CTAC	T_{PNA}-C_{DNA}	39	20	-18
AGTGA**G**CTAC	T_{PNA}-G_{DNA}	34	12	-24
AGTG**T**TCTAC	A_{PNA}-T_{DNA}	34	14	-24
AGTG**C**CTAC	A_{PNA}-C_{DNA}	31	12	-27
AGTG**G**TCTAC	A_{PNA}-G_{DNA}	29	8	-29
AGT**A**ATCTAC	G_{PNA}-A_{DNA}	34	20	-24
AGT**T**ATCTAC	G_{PNA}-T_{DNA}	33	17	-25
AGT**C**ATCTAC	G_{PNA}-C_{DNA}	39	21	-18

The same experiment was also carried out with a longer 15mer sequence, AcLys-TGTACGTCACAACTA-NH_2 (**12x**). The sensitivity towards mismatches was also investigated by T_m measurement (**Table 3.20**). The 15mer PNA (**12x**) bind to complementary oligonucleotides with a T_m of 79 °C. A single mismatch at the central positions in the DNA strands caused a drop in T_m of 10-18 °C. Furthermore the presence of double mismatches caused a further reduction in T_m to 52 °C. From the results it was concluded that *ss*ACPC-PNA (**12**) system can form hybrid with complementary DNA in a highly sequence-specific manner.

Table 3.20 T_m of mixed-base PNA TGTACGTCACAACTA (**12x**) with single mismatch DNA. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

DNA sequence (5'→3')	Type of mismatch	T_m (°C)	(% Hyper-chromicity)	ΔT_m (°C)
TAGTTGTGACGTACA	no mismatch	79	32	-
TAGTTGTGA**A**GTACA	G_{PNA}-A_{DNA}	61	28	-18
TAGTTGTGA**T**GTACA	G_{PNA}-T_{DNA}	66	30	-13
TAGTTGTGA**G**GTACA	G_{PNA}-G_{DNA}	65	27	-14
TAGTTGTG**T**CGTACA	T_{PNA}-T_{DNA}	65	30	-14
TAGTTGTG**C**CGTACA	T_{PNA}-C_{DNA}	65	27	-14
TAGTTGTG**G**CGTACA	T_{PNA}-G_{DNA}	65	27	-14
TAGTTGT**A**ACGTACA	C_{PNA}-A_{DNA}	66	32	-13
TAGTTGT**T**ACGTACA	C_{PNA}-T_{DNA}	65	31	-14
TAGTTGT**C**ACGTACA	C_{PNA}-C_{DNA}	65	30	-14
TAGTTG**A**GACGTACA	A_{PNA}-A_{DNA}	65	28	-14
TAGTTG**C**GACGTACA	A_{PNA}-C_{DNA}	69	29	-10
TAGTTG**G**GACGTACA	A_{PNA}-G_{DNA}	65	29	-14
TAGTTG**A**G**T**CGTACA	T_{PNA}-T_{DNA}, A_{PNA}-A_{DNA}	52	25	-27

In addition to the effect of mismatching, the effect of insertion and deletion of bases in the DNA complementary to PNA TGTACGTCACAACTA (**12x**) was studied (**Table 3.21**).

Table 3.21 T_m of TGTACGTCACAACTA (**12x**) with deletion and insertion of complementary DNA. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

DNA (5'→3')	T_m (°C)	(% hyperchromicity)	Remark
TAGTTGTGACGTACA	79	>31	perfect match
TAGTTGTG**T**ACGTACA	65	15	T insertion
TAGTTGTG**G**ACGTACA	67	15	G insertion
TAGTTG**TA**CGTACA	61	14	G deletion

The results showed that in all cases the T_m were decreased from that of the perfect match hybrid (79 °C). Hybrids containing base insertion or deletion in the DNA component have a significantly lower percent hyperchromicity (14-15 %) compared to the perfect match hybrids due to the lower degree of base stacking. However, both T-insertion and G-insertion hybrids exhibited an unexpectedly high T_m of 65 °C. When the T-insertion DNA sequence was examined, the hexadecamer may be separated into two complementary parts (TAGTTGTG and ACGTACA) by a base **T**. This DNA sequence should bind with PNA by using only 8-base moiety of only one sequence leaving the remaining part free. However from **Figure 3.47** and **Figure 3.48** mixed base PNA·DNA hybrid with only 8-base long should have a T_m value much below 65 °C. It is proposed that the strand containing the extra base insertion makes a loop around the insertion region to force the whole sequences to bind together (**Figure 3.51**). The same explanation should be applicable in cases of G-insertion which gave T_m around 67 °C and 15 % hyperchromicity and G-deletion which gave of a 61 °C and hyperchromicity of 14 °C. In the latter case, the loop should take place on the PNA strand instead.

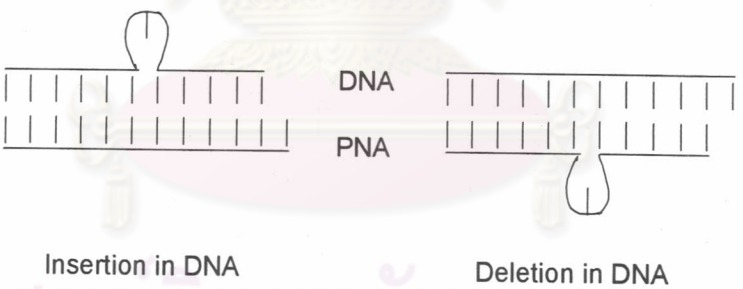


Figure 3.51 Binding model of PNA with insertion and deletion of complementary DNA

Table 3.22 T_m of *ss*ACPC-PNA (**12**)·DNA, Nielsen's PNA (**1**)·DNA and DNA·DNA. Condition: 10 mM sodium phosphate buffer, pH 7.0; 100 mM NaCl; [PNA] = [DNA] = 1.0 μM.

Sequence of PNA (N→C) or DNA (5'→3')	**Sequence of DNA probe (5'→3')**	**T_m (°C) (ΔT_m)[a]**		
		***ss*ACPC PNA (12)**	**Nielsen[b] PNA (1)**	**DNA[b]**
GTAGATCACT	AGTGATCTAC (antiparallel perfect match)	53	51	34
	CATCTAGTGA (parallel perfect match)	<20	38	<20
TGTACGTCACAACTA	TAGTTGTGACGTACA (antiparallel perfect match)	74	70	53
	ACATGCAGTGTTGAT (parallel perfect match)	<20	56	<20
	TAGTTGT**A**ACGTACA (C_{PNA}-A_{DNA} mismatch)	61 (-13)	50 (-20)	38 (-15)
	TAGTTGT**T**ACGTACA (C_{PNA}-T_{DNA} mismatch)	59 (-15)	49 (-21)	37 (-16)
	TAGTTGT**C**ACGTACA (C_{PNA}-C_{DNA} mismatch)	61(-13)	51 (-19)	38 (-15)

[a] ΔT_m = T_m of perfect match hybrid - T_m of single mismatched hybrid. The greater the numbers indicate higher sensitivity to mismatching. [b] Data taken from reference [8]

Attempts have been made to compare the stability and specificity in DNA binding between Nielsen's aegPNA (**1**), ssACPC-PNA (**12**) and DNA. Due to the low stability of DNA·DNA hybrids and the lack of comparable T_m data of PNA (**1**) under low salt conditions, the T_m of the decamer (**12r**) and 15mer (**12x**) with their complementary and single mismatched DNA were measured again at 100 mM NaCl. As observed earlier in the study of homothymine PNA (**12f**), the T_m of the mixed base PNA·DNA hybrids were slightly lower than without the add salt. For decamer, PNA (**12r**) and PNA (**1r**) gave almost the same T_m values (53 and 51 °C) with antiparallel DNA whereas the corresponding DNA hybrids melts at only 34 °C. For 15 mer, four

sequences with variation of the base at the central position were studied. The T_m of perfect match complementary PNA (**12x**)-DNA hybrid was higher than PNA (**1x**) and DNA by 4 °C and 21 °C respectively. Hybrids of PNA (**1x**) with complementary parallel DNA are slightly less stable than the antiparallel hyrids, but no parallel hybrid formation was observed for both (**12**) and DNA. With single mismatched base pairing was present (C-A, C-T and C-C), all three mismatched **12x**·DNA hybrids showed higher T_m values than the corresponding DNA hybrids of (**1**) and of DNA. However, the ΔT_m values compared to the perfect match (**12x**)·DNA hybrid are somewhat lower (5-6 °C) than those of (**1**) and are in the same order as DNA. The result from this experiment suggested that this PNA sequence bind to DNA with higher affinity but less selectivity than PNA (**1**). Nevertheless, the selectivity is at least in the same order as DNA hence this PNA (**12**) should be useful as a probe for DNA sequence determination. Furthermore, the PNA (**12**) also show a stronger preference for binding to DNA in antiparallel orientation than (**1**).

Finally, the specificity of Watson-Crick base pairing was investigated in another randomly mixed base PNA system. Four PNA (**12n-12q**) with sequences Ac-TCAC<u>X</u>ACTA-LysNH_2 (X= A, T, C, G) were synthesized and hybridized with the antiparallel DNA probes d(TAGT<u>Y</u>GTGA), when Y= A, T, C and G (**Table 3.23**). Remarkably, only the fully complementary duplexes exhibit an observable melting with a T_m between 37-49 °C. No melting was observed for any of the single mismatch hybrids. The result confirmed the high affinity and specificity in binding between PNA (**12**) and DNA.

Table 3.23 T_m and % hyperchromicity of nonamer mixed-base PNA (**12**) TCAC<u>X</u>ACTA and d(TAGT<u>Y</u>GTGA) hybrids. Condition: 10 mM sodium phosphate buffer, pH 7.0; [PNA] = [DNA] = 1.0 μM.

PNA	X	Y in DNA			
		A	T	C	G
12n	T	**<u>44</u>** (18)	<20	<20	<20
12o	A	<20	**<u>49</u>** (21)	<20	<20
12p	G	<20	<20	**<u>47</u>** (22)	<20
12q	C	<20	<20	<20	**<u>37</u>** (19)

3.3.6 Selectivity among different types of nucleic acids

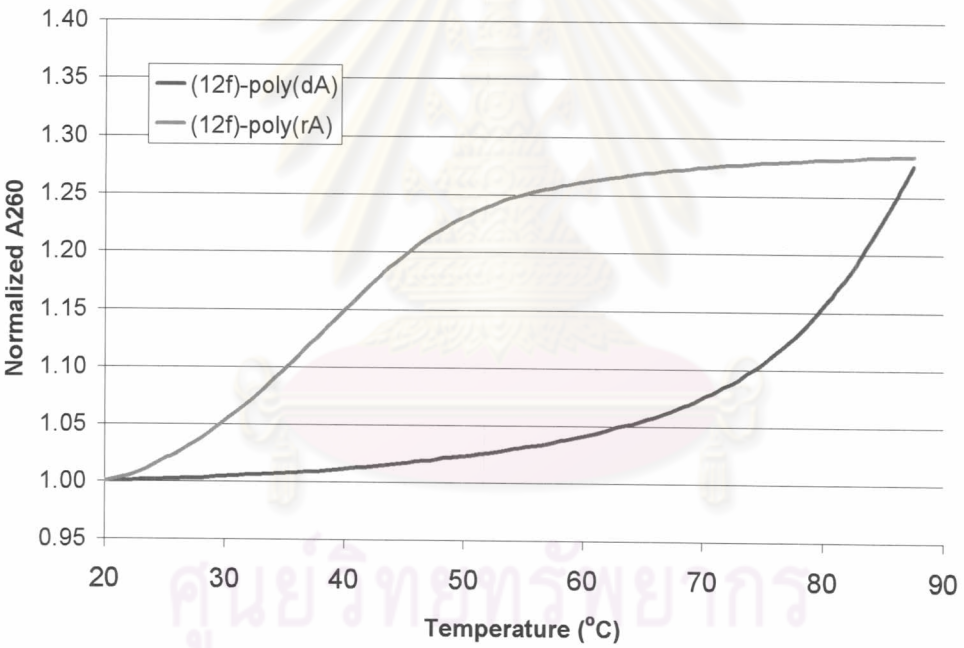
The selectivity, stability and affinity of PNA (**12**) in binding with different types of nucleic acids were investigated by melting temperature. The T_m experiment PNA (**12**) with T_5, T_7 and T_9 sequences were selected to check the selectivity among DNA and RNA. The T_m of T_5, T_7 and T_9 with poly(dA) were 37, 74 and >85 °C respectively. When the same experiment was performed using poly(rA) instead of poly(dA) the T_m in all cases were substantially less (maximum 38 °C for T_9) (**Table 3.24** and **Figure 3.52**). The results showed an interesting higher stability of the hybrids formed from (**12**)-DNA than (**12**)-RNA indicating that the *ss*ACPC PNA (**12**) prefers binding to DNA over RNA. For another set of experiment, PNA A_5 (**12d**) and A_9 (**12k**) was used as a model to examine the binding selectivity with poly(dT) and poly(rU) (**Table 3.25** and **Figure 3.52**). The apparent absence of melting of **12k**-poly(dT) is explained by the fact that the T_m value of this hybrid must be extremely high (>90 °C). In fact, even **12d** with only T_5 sequence formed a hybrid of T_m = 72 °C already with poly(dT).The higher affinity towards DNA over RNA. This phenomenon has been previously observed in a related PNA with D-aminopyrrolidinecarboxylic acid (D-Apc) in place of *ss*ACPC spacer [60]. It should be noted that D-Apc is in fact the aza-analogue of *ss*ACPC therefore the similar behavior is not unexpected.

Table 3.24 The melting temperature of homothymine PNA (**12f**) with poly(dA) and poly(rA). Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

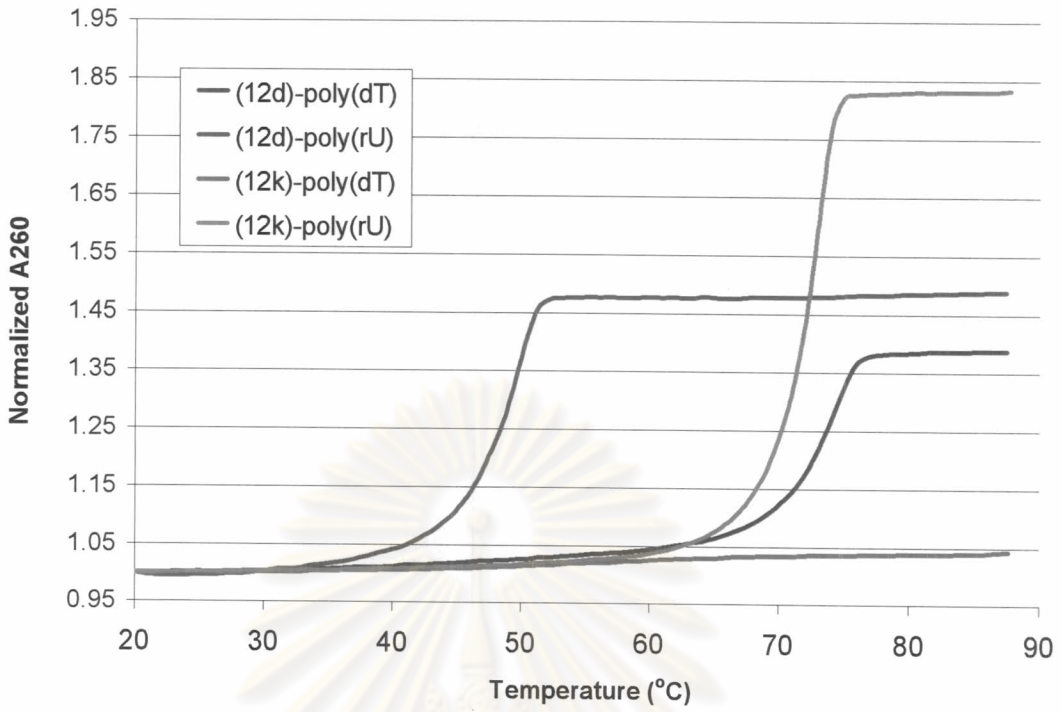
PNA	T_m (°C) (% Hyperchromicity)		ΔT_m (°C) (PNA·DNA-PNA·RNA)
	Poly(dA)	Poly(rA)	
12d	37 (25)	<38 (8)	N/A
12e	74 (>44)	<38 (14)	N/A
12f	>85 (N/A)	38 (26)	>47

Table 3.25 The melting temperature of homoadenine PNA (**12d**) and (**12k**) with poly(dT) and poly(rU). Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

PNA	T_m (°C) (% Hyperchromicity)		ΔT_m (°C) (PNA·DNA-PNA·RNA)
	Poly(dT)	Poly(rU)	
12d	72 (37)	46 (47)	36
12k	>90 (N/A)	70 (83)	>20



(a)



(b)

Figure 3.52 (a) T_m curves of (**12f**)·poly(dA) and (**12f**)·poly(rA) hybrids. (b) T_m curves of (**12d**)·poly(dT), (**12k**)·poly(dT), (**12d**)·poly(rU) and (**12k**)·poly(rU) hybrids. Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

Table 3.26 The melting temperature of PNA(**12x**) with poly(dT) and(rU). Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

PNA	**T_m (°C) (% Hyperchromicity)**		**ΔT_m**
	d(TAGTTGTGACGTACA)	r(UAGUUGUGACGUACA)	**(°C)**
12x (0 mM NaCl)	74 (>31)	61 (28)	13
12x (100 mM NaCl)	79 (31)	63 (28)	16

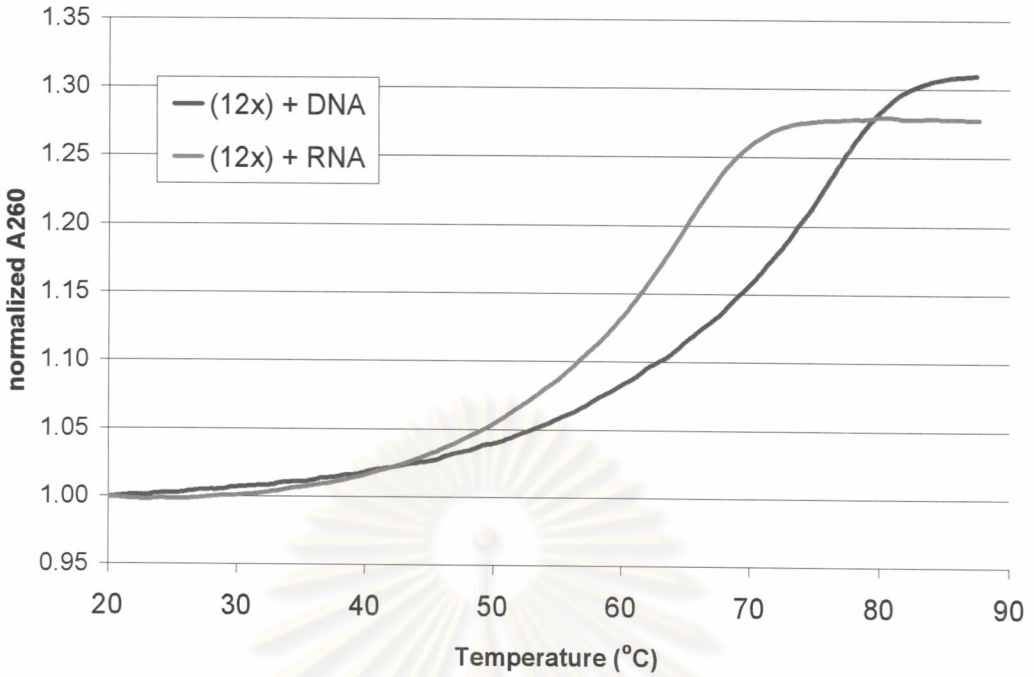
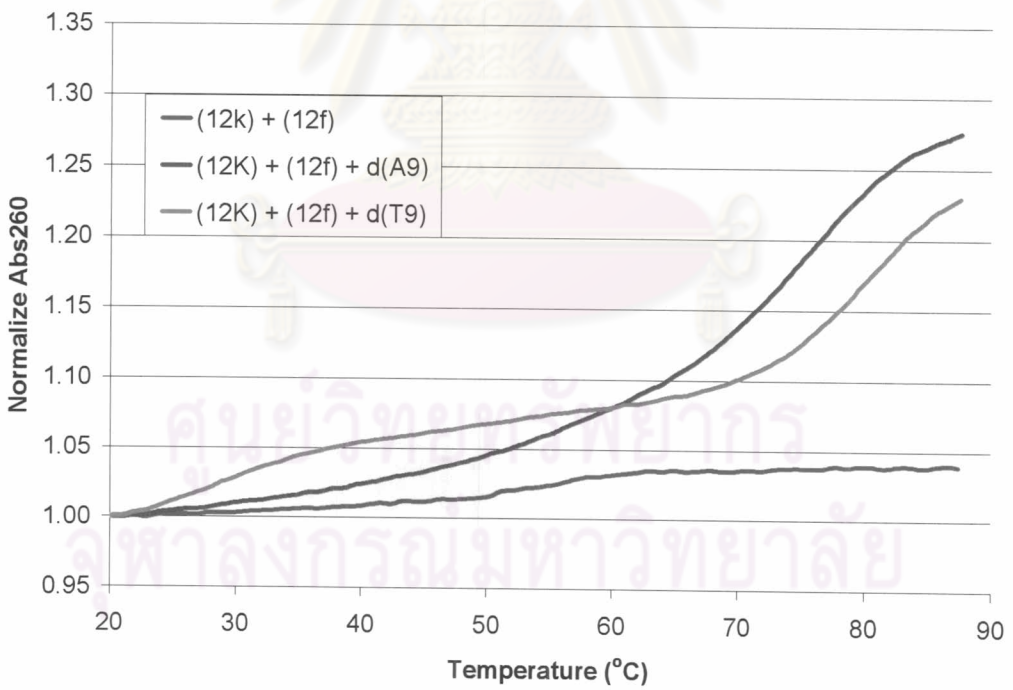


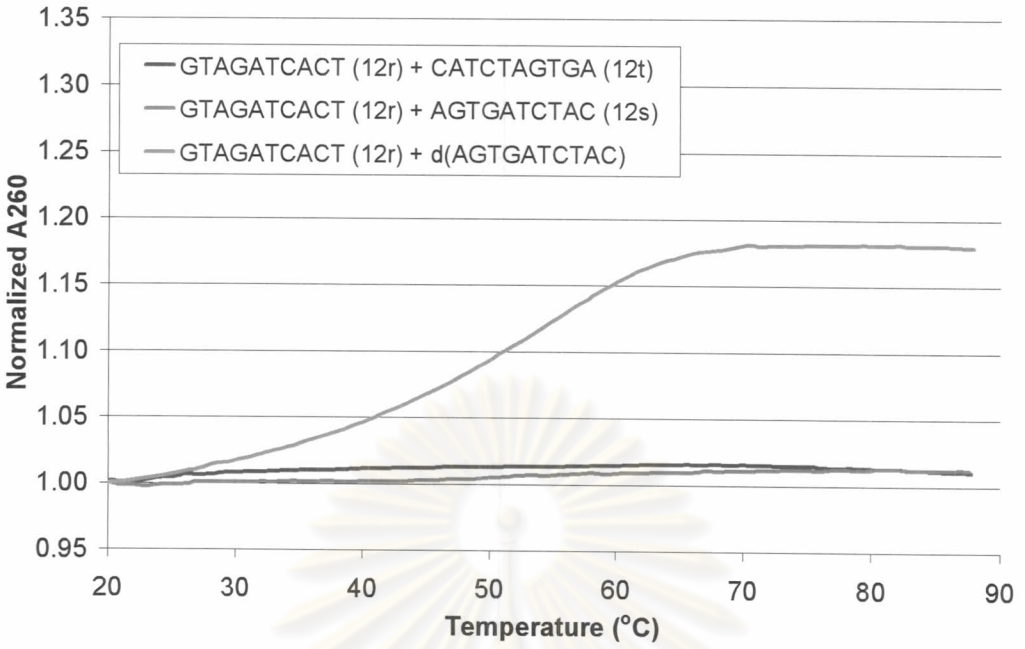
Figure 3.53 T_m curves of hybrids between (**12x**) and complementary DNA and RNA. Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, 100mM NaCl, heating rate 1.0 °C/min.

To investigate further the generality of this DNA/RNA binding selectivity, the T_m experiment was also carried out between the 15mer mixed base PNA (**12x**) and a complemenmtary antiparallel RNA r(UAGUUGUGACGUACA) under both high salt (100 mM NaCl) and low salt (0 mM NaCl) conditions. The results were compared with the corresponding PNA·DNA hybrid as shown in **Table 3.26** and **Figure 3.53**. In both cases, the T_m of PNA·DNA hybrids were higher than the PNA·RNA hybrid by 13-16 °C therefore the DNA preference over RNA is also applicable in case of mixed base PNA (**12**). The reason for the large difference between T_m of (**12x**)·DNA and T_m of (**12x**)·RNA in the homo-oligomer compared to the mixed sequences is not fully understood at present, but we have confirmed by UV-titration that in all cases only duplexes were formed. The preference for DNA binding of this particular PNA system is in sharp contrast with (**1**) and many other PNA developed to date which usually prefer RNA binding.[80] The structural basis of this remarkable selectivity will be the subject of future investigation.

Finally we also investigated the possibility of formation of self-pairing hybrids between two complementary strands of PNA (**12**). In the first experiment, the homothymine PNA (**12f**) and the homoadenine PNA (**12k**) were mixed together and a T_m was determined. Surprisingly, no melting was observed at all indicating the absence of hybrid formation between the two strands of PNA (**12**). However, addition of either dA_9 or dT_9 to the mixture of (**12f**) and (**12k**) resulted in formation of the PNA·DNA hybrid as evidenced by the melting observed. This confirms that both strands of PNA are free to bind with the complementary DNA even in the presence of the complementary PNA strands. Next melting experiments were performed on mixtures of the decamer PNA (**12r**) with its antiparallel (**12s**) and parallel (**12t**) PNA. Again no melting was observed in both cases confirming that the inability of PNA (**12**) to form self-pairing hybrids. The lack of self-pairing between two complementary strands of **2** is rather unexpected and, to the best of our knowledge, has not been reported in other PNA systems. The basis of this unique selectivity again will be the subject of our future investigation.



(a)



(b)

Figure 3.54 (a) T_m curves of **(12f)**+**(12k)**, **(12f)**+**(12k)**+dA_9 and **(12f)**+**(12k)**+dT_9. (b) T_m curves of **(12r)**+**(12t)**, **(12r)**+**(12s)** and **(12r)**·antiparallel DNA hybrids. Conditions: PNA:DNA or RNA = 1:1 (by nucleotide), [PNA] = 1 μM, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.