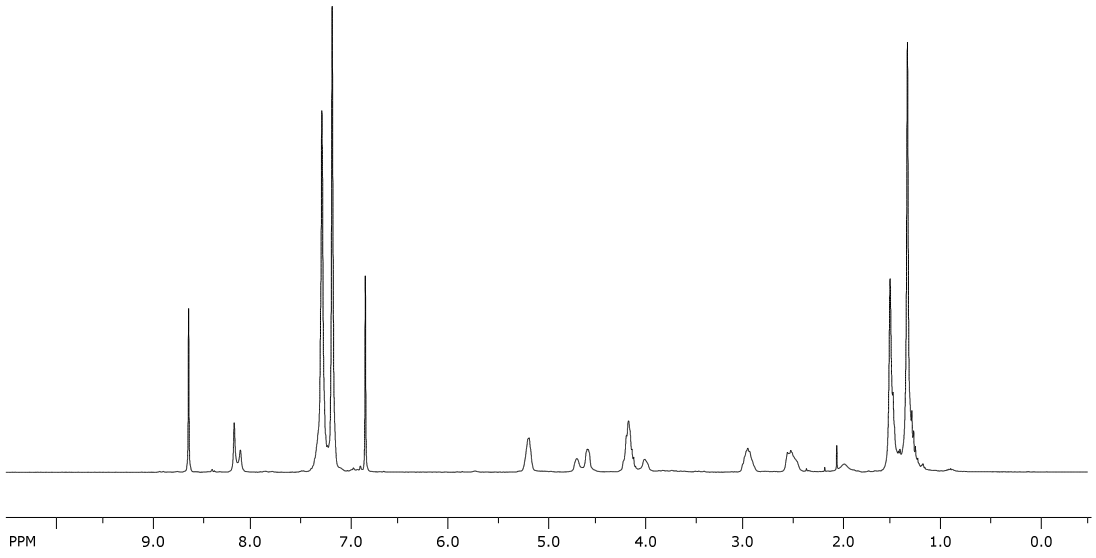
Electronic Supplementary Information (ESI)

Specific recognition of cytosine by hypoxanthine in pyrrolidinyl peptide nucleic acid

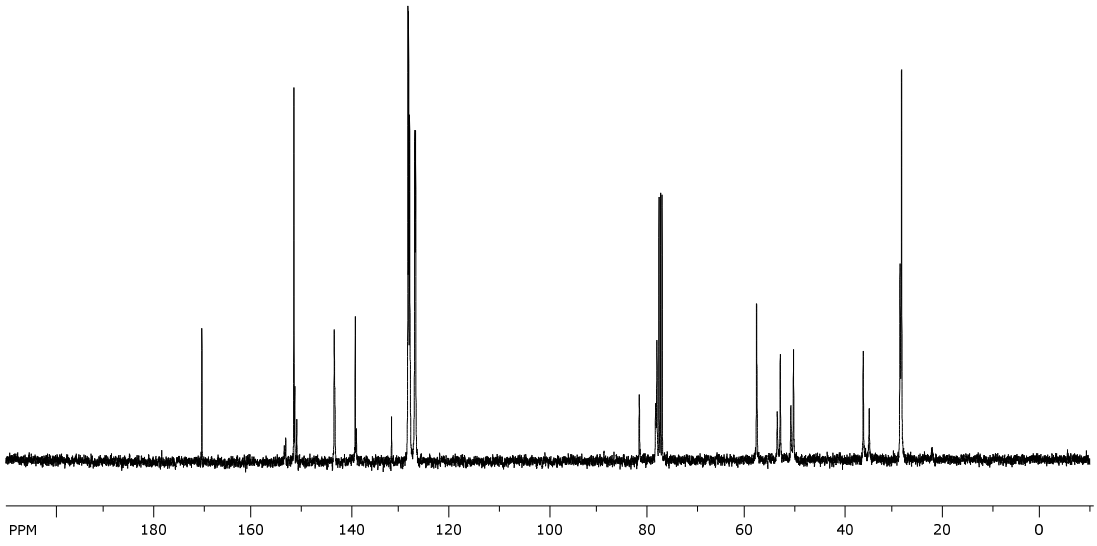
Chotima Vilaivan,[a] Wimonmas Srinarang,[a] Nattawut Yotapan,[a] Woraluk Mansawat,[a] Chalotorn Boonlua,[a] Junji Kawakami,[b] Yoshie Yamaguchi,[b] Yuko Tanaka[b,c] and Tirayut Vilaivan[a,*]

[a]Organic Synthesis Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai Road, Patumwan, Bangkok 10330, Thailand
[b]Department of Nanobiochemistry, FIRST, Konan University and [c] Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

Content **Page**

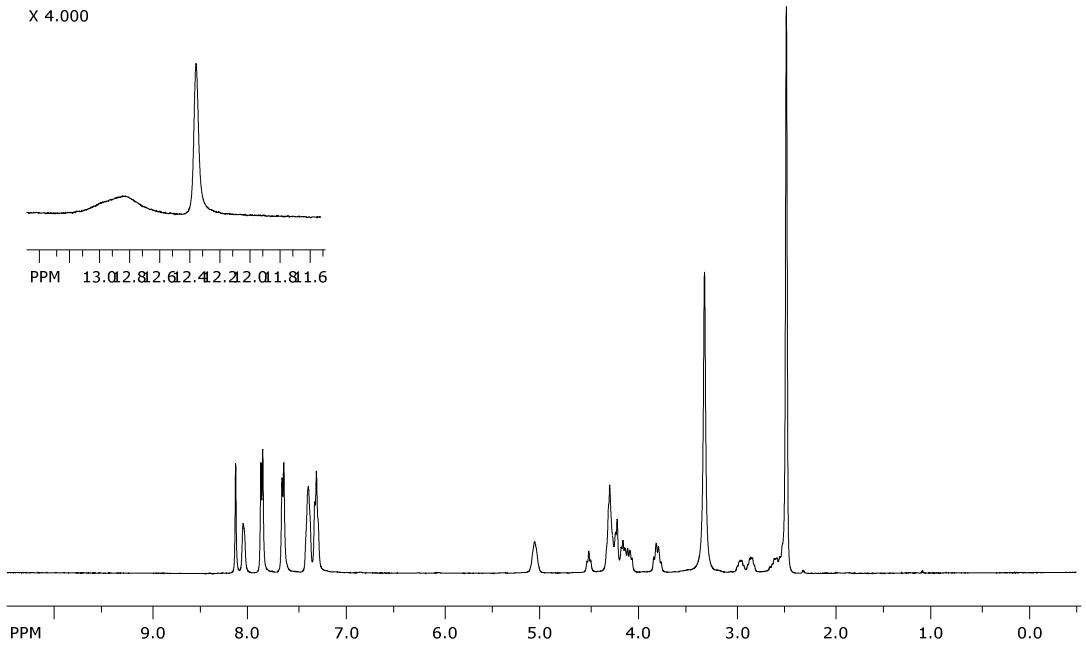


(a)

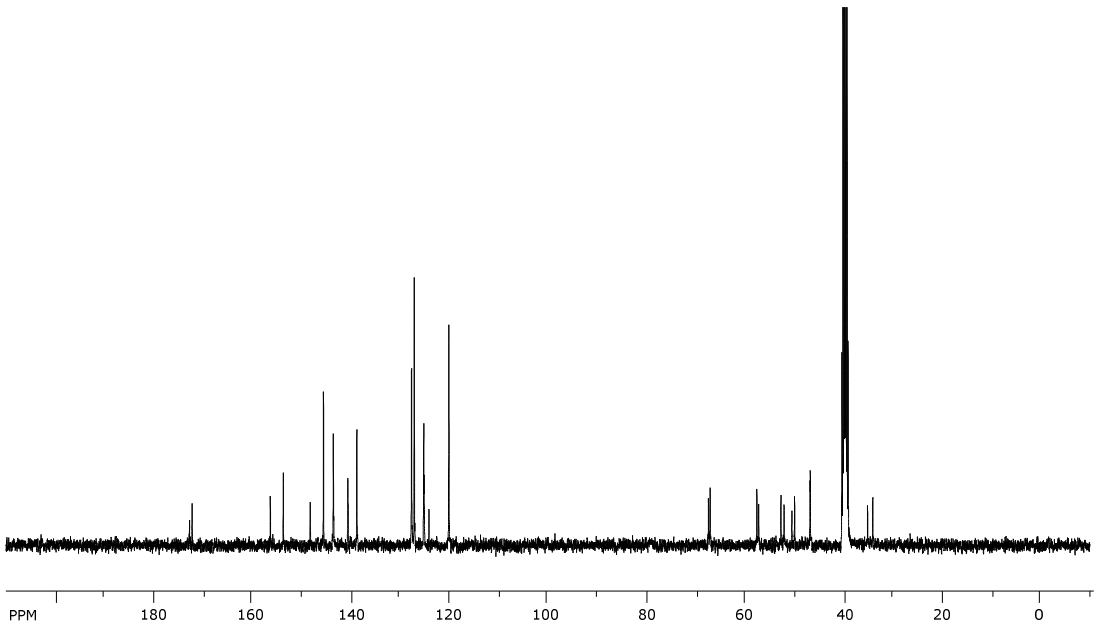


(b)

Fig. S1 (a) ^{1}H NMR spectrum (400 MHz, $CDCl_3$) and (b) ^{13}C NMR spectrum (100 MHz, $CDCl_3$) of *N*-*tert*-Butoxycarbonyl-(4′*R*)-(6-chloro-9H-purin-9-yl)-(2′*R*)-proline diphenylmethyl ester (**2**)



(a)



(b)

Fig. S2 (a) ^{1}H NMR spectrum (400 MHz, DMSO-d_6) and (b) ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of *N*-9H-fluoren-9-ylmethoxycarbonyl-(4′*R*)-[6-oxo-1H-purin-9(6H)-yl]-(2′*R*)-proline (**3**)

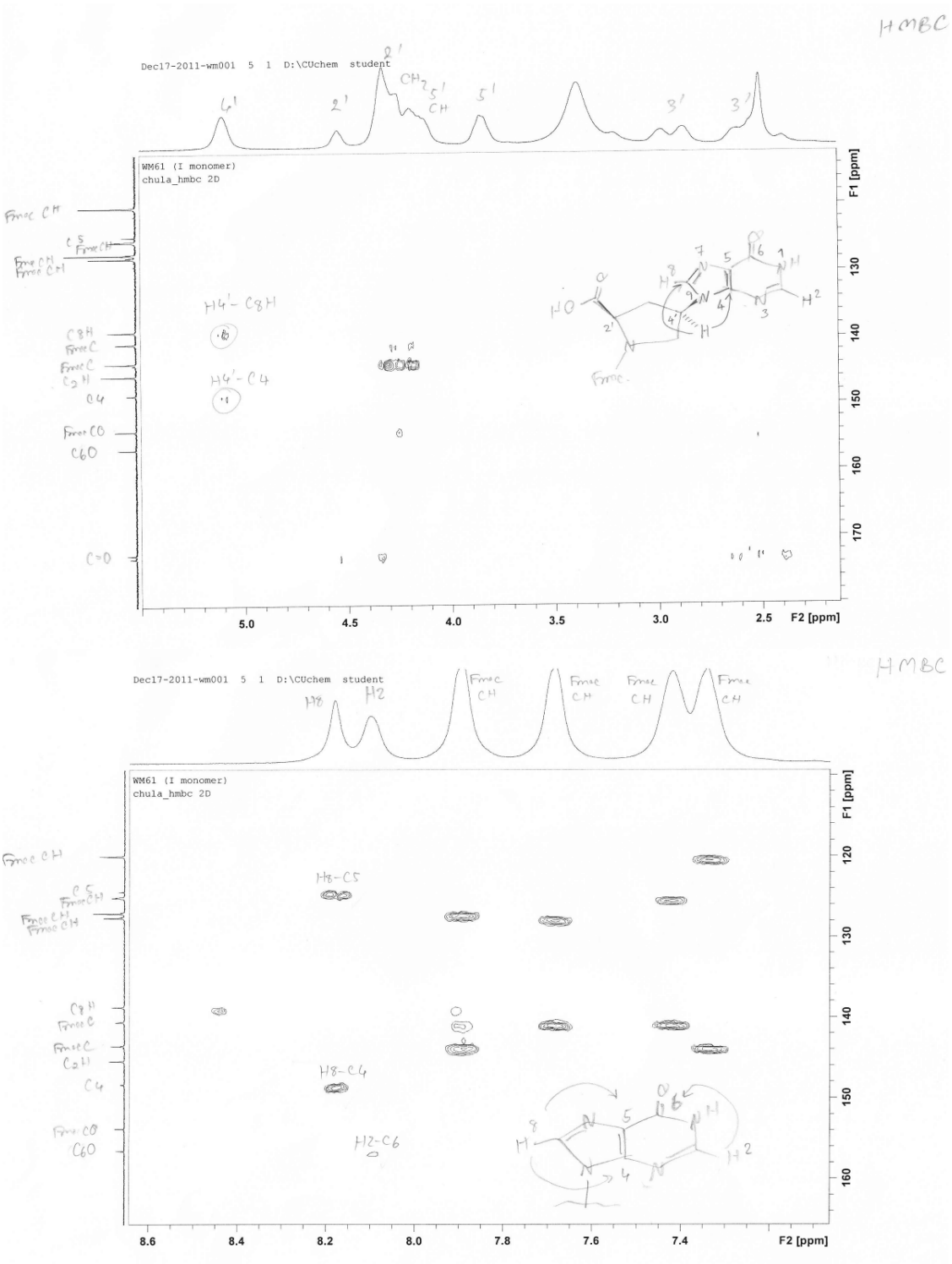


Fig. S3 Selected expanded regions of ^{1}H-^{13}C HMBC spectra (400 MHz, DMSO-d_6) of *N*-9H-fluoren-9-ylmethoxycarbonyl-(4′*R*)-[6-oxo-1H-purin-9(6H)-yl]-(2′*R*)-proline (**3**)

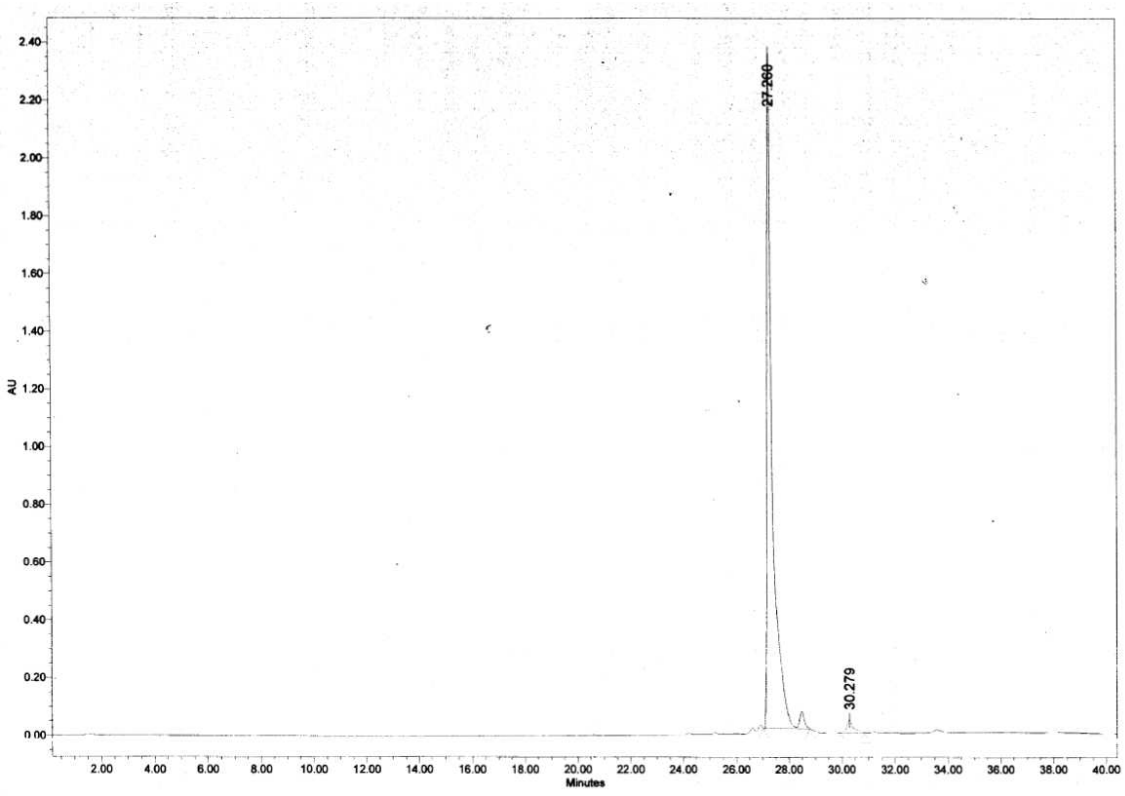
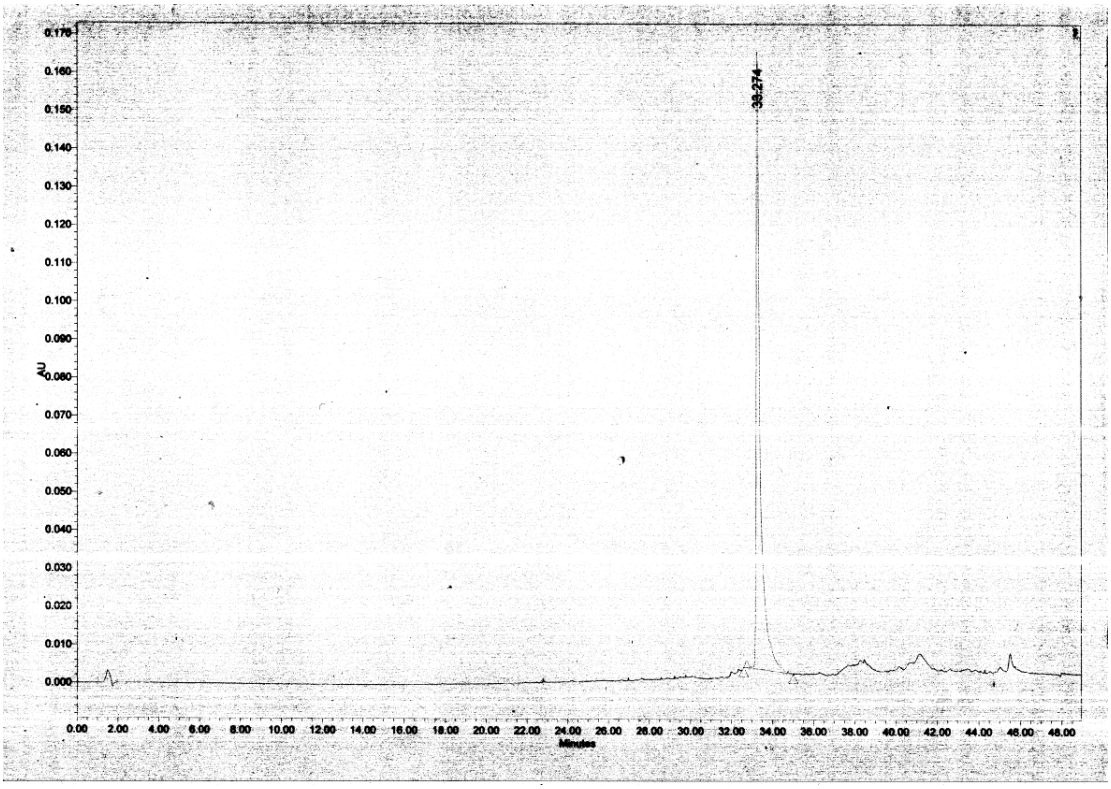
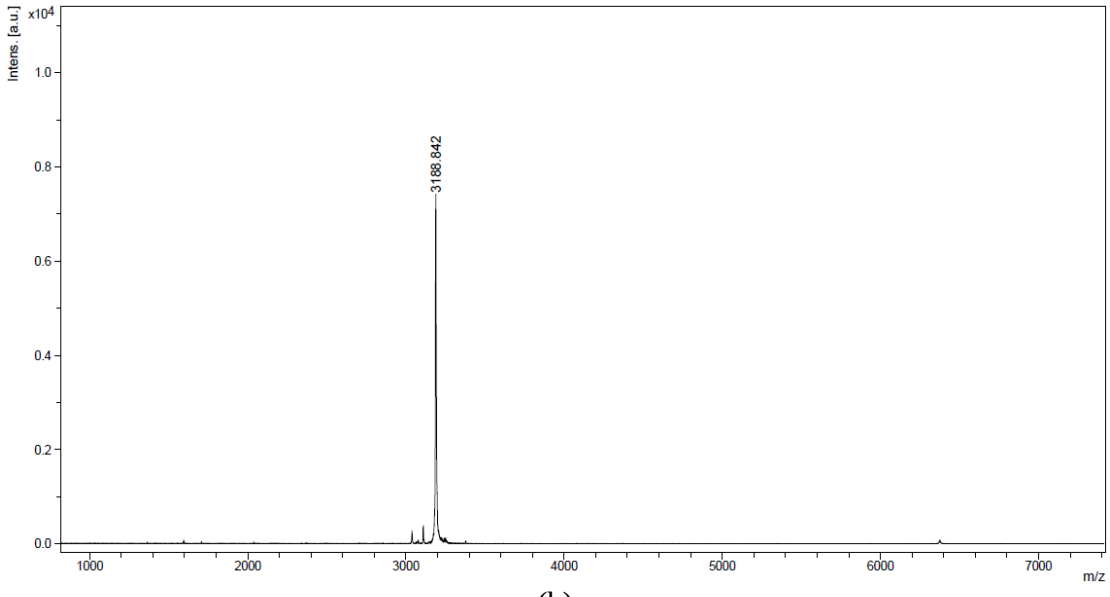


Fig. S4 Analytical HPLC chromatogram of *N*-9H-fluoren-9-ylmethoxycarbonyl-(4′*R*)-[6-oxo-1H-purin-9(6H)-yl]-(2′*R*)-proline (**3**).

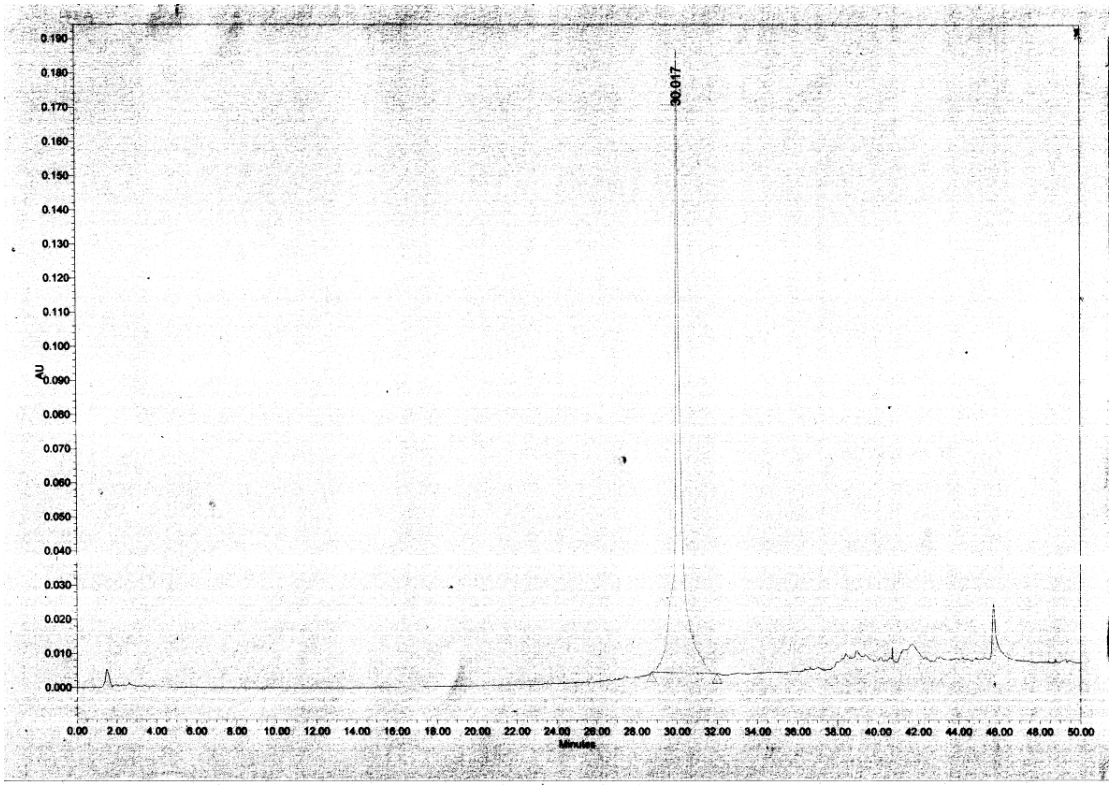


(a)

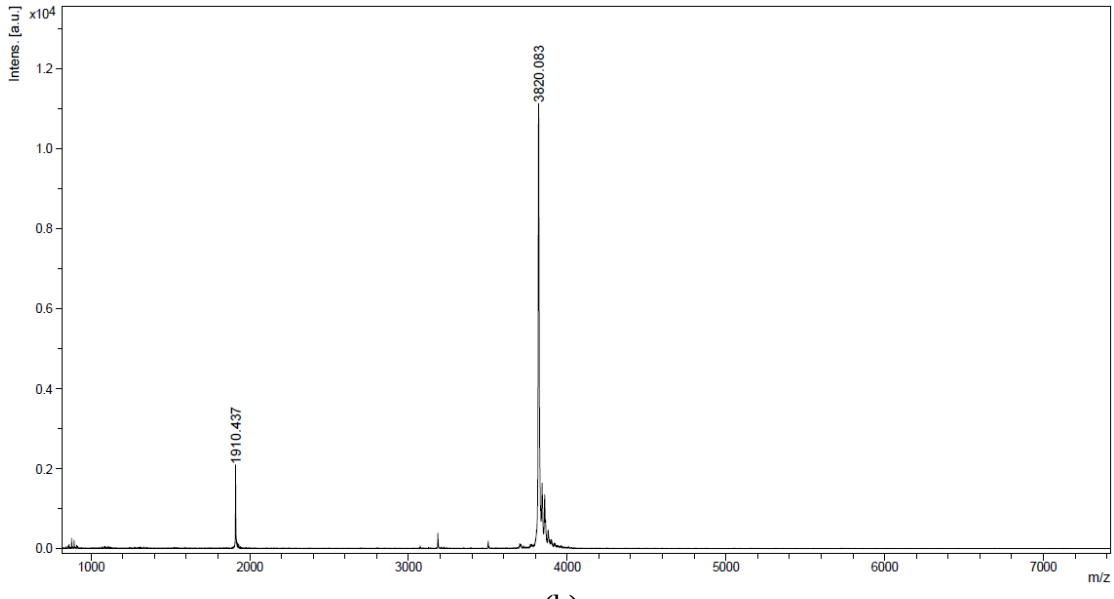


(b)

Fig. S5 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA1** (Ac-TTTTITTTT-LysNH_2) (calcd. *m/z* 3189.4)

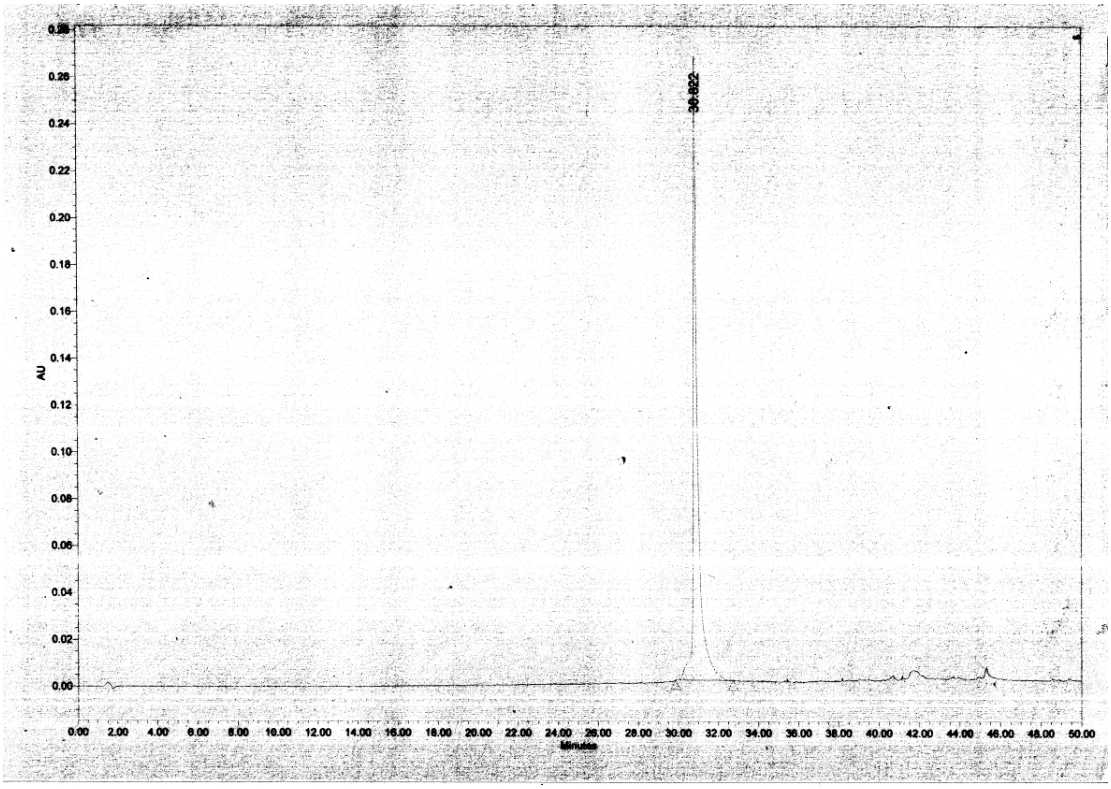


(a)

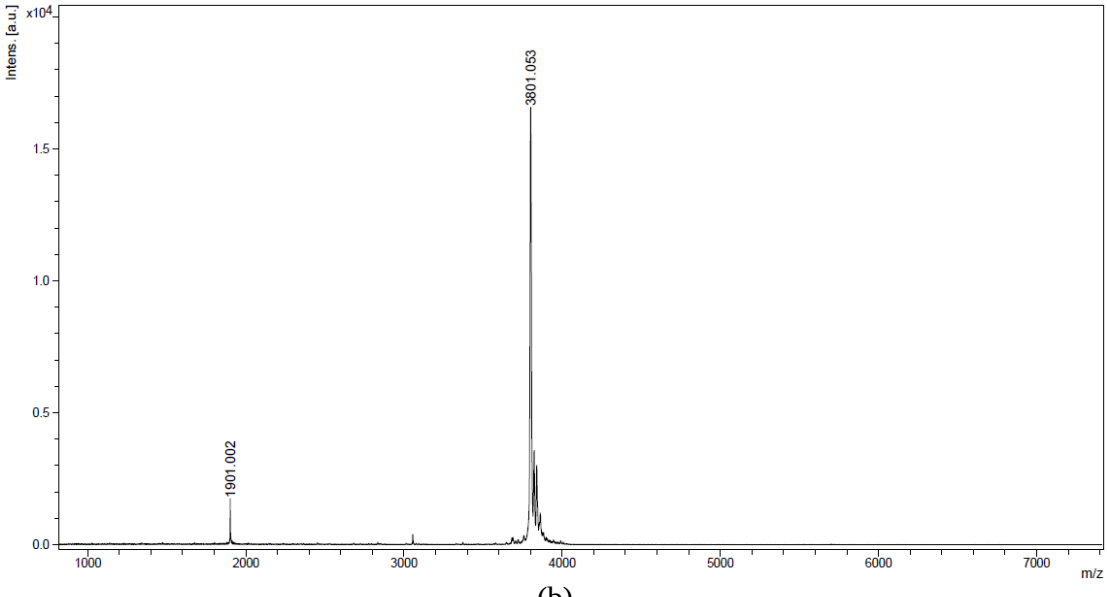


(b)

Fig. S6 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA2** (Ac-CCTTAIACATC-LysNH_2) (calcd. *m/z* 3821.1)

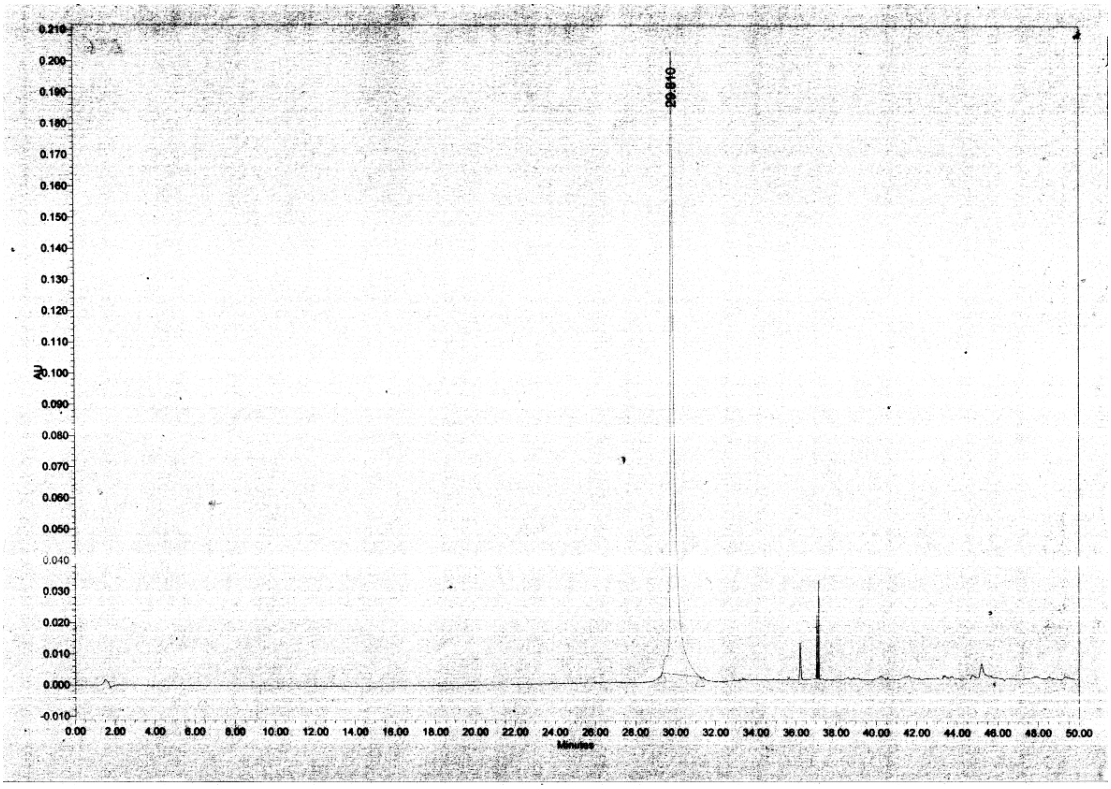


(a)

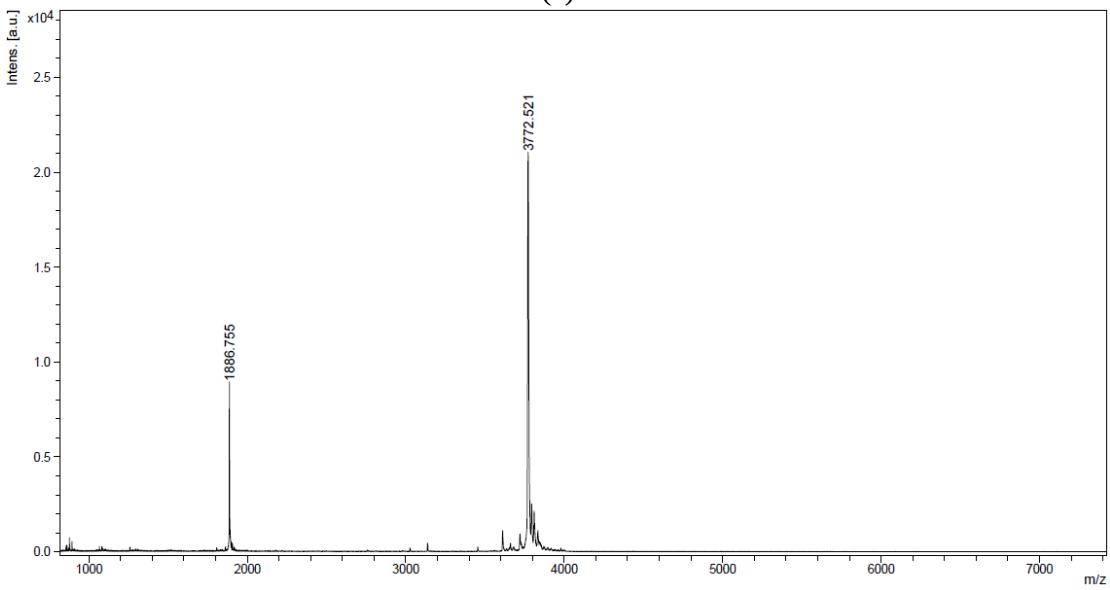


(b)

Fig. S7 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA3** (Ac-CCTTTTITCATC-LysNH$_2$) (calcd. *m/z* 3803.1)

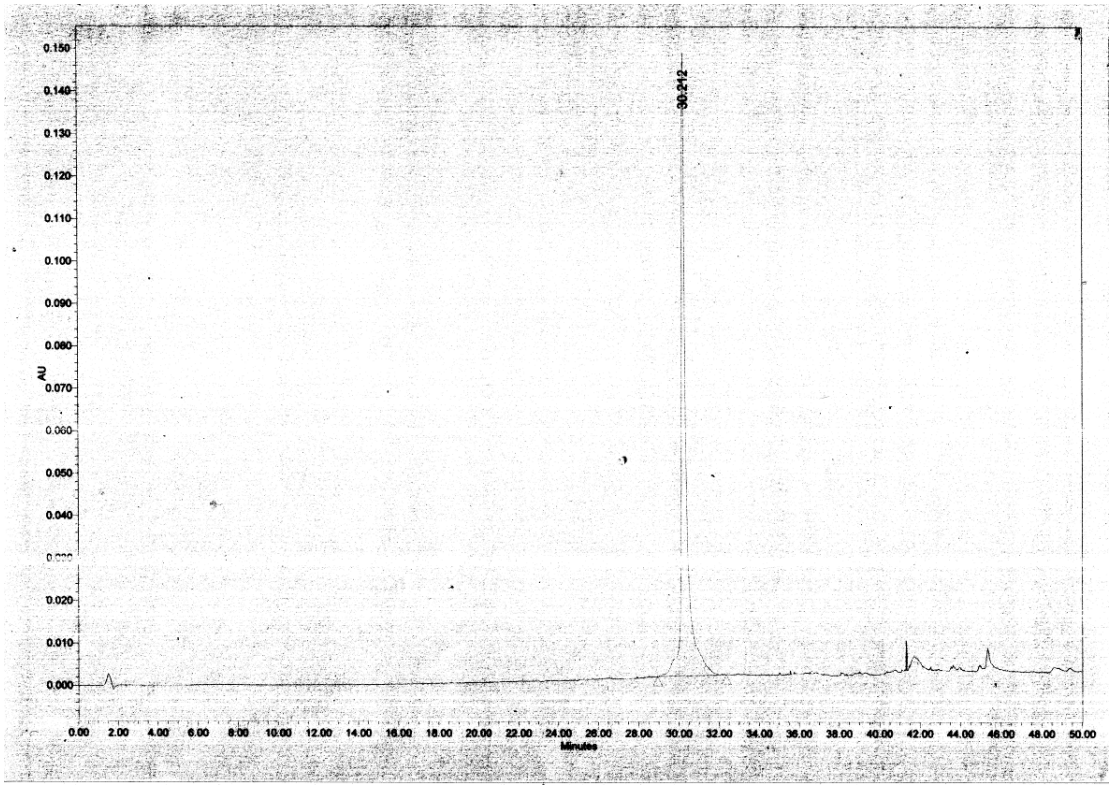


(a)

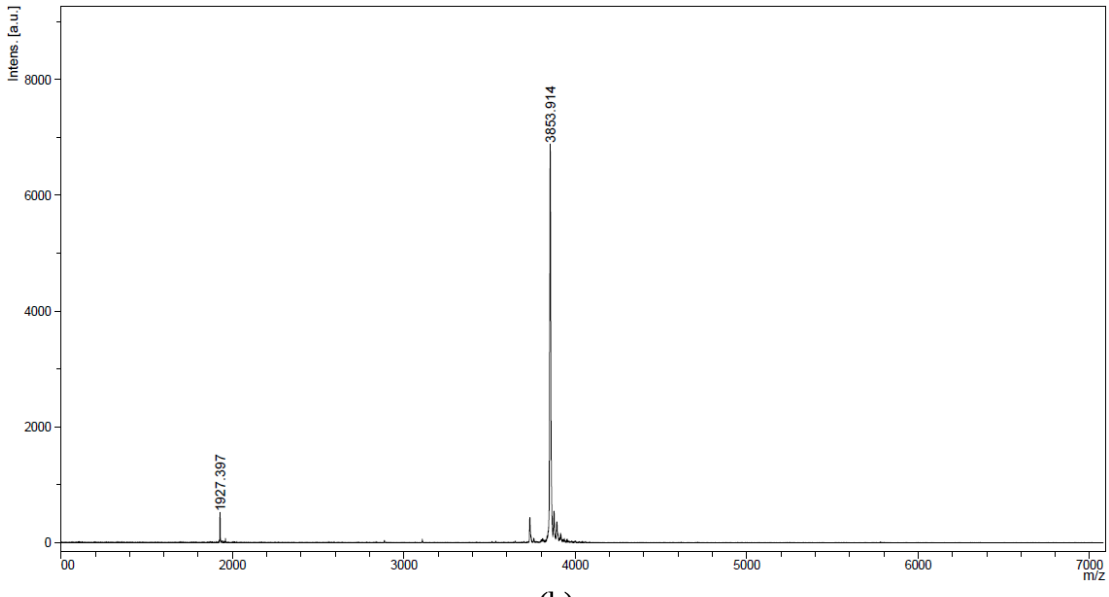


(b)

Fig. S8 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA4** (Ac-CCTTCICCATC-LysNH_2) (calcd. *m/z* 3773.1)

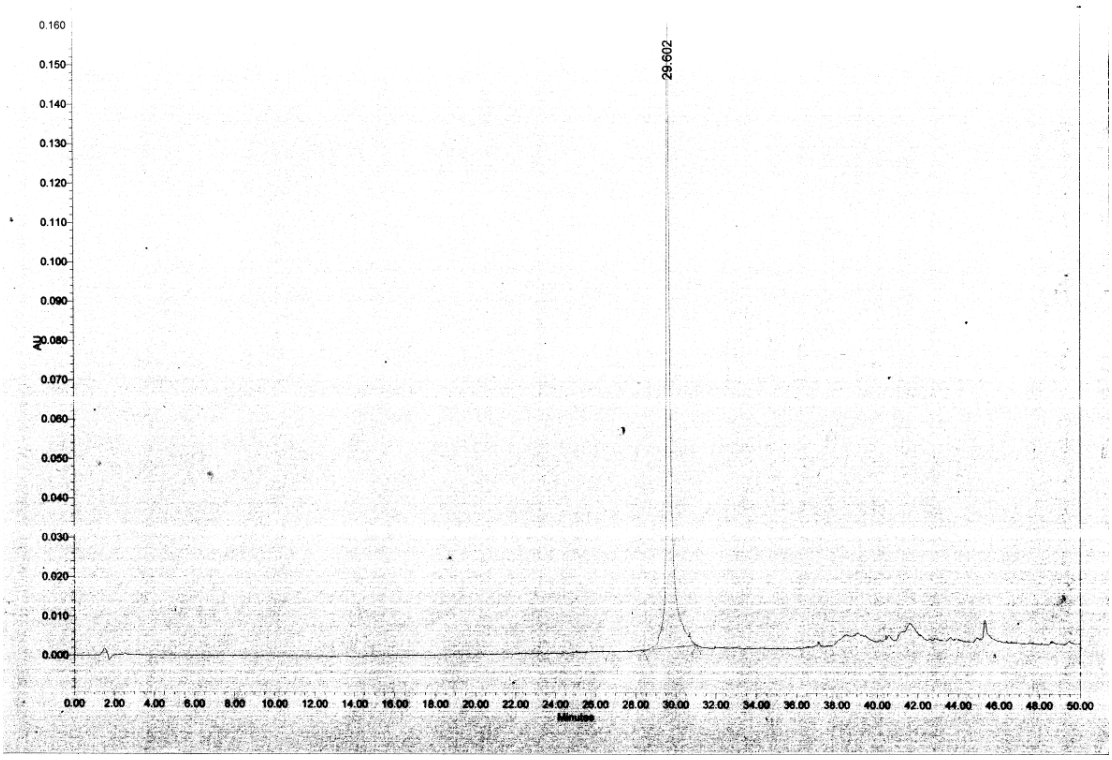


(a)

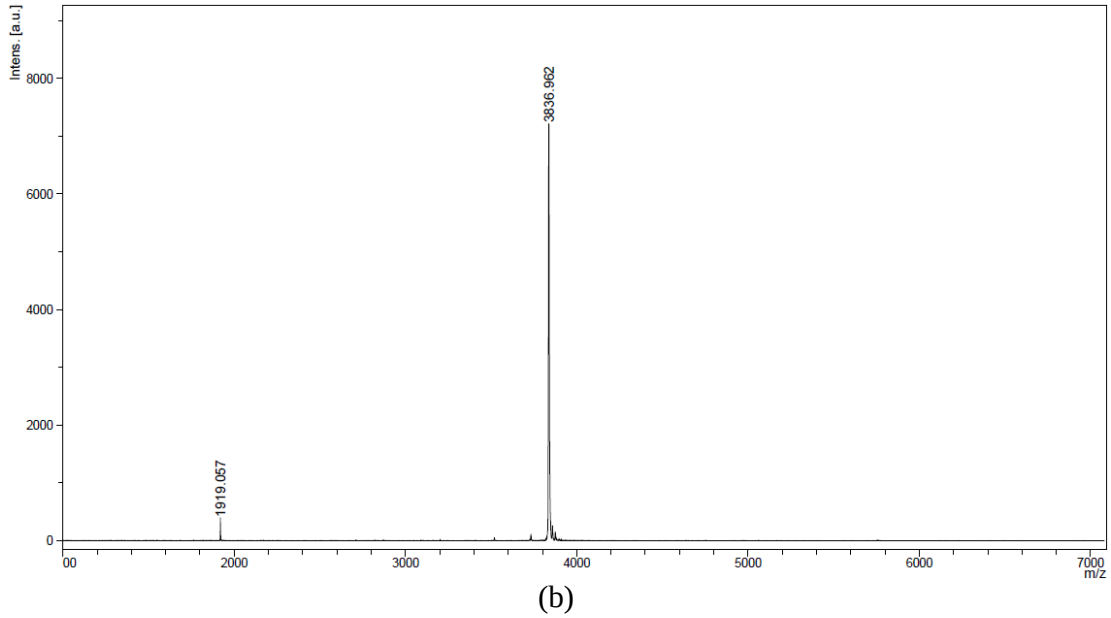


(b)

Fig. S9 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA5** (Ac-CCTTGIGCATC-LysNH_2) (calcd. *m/z* 3853.1)

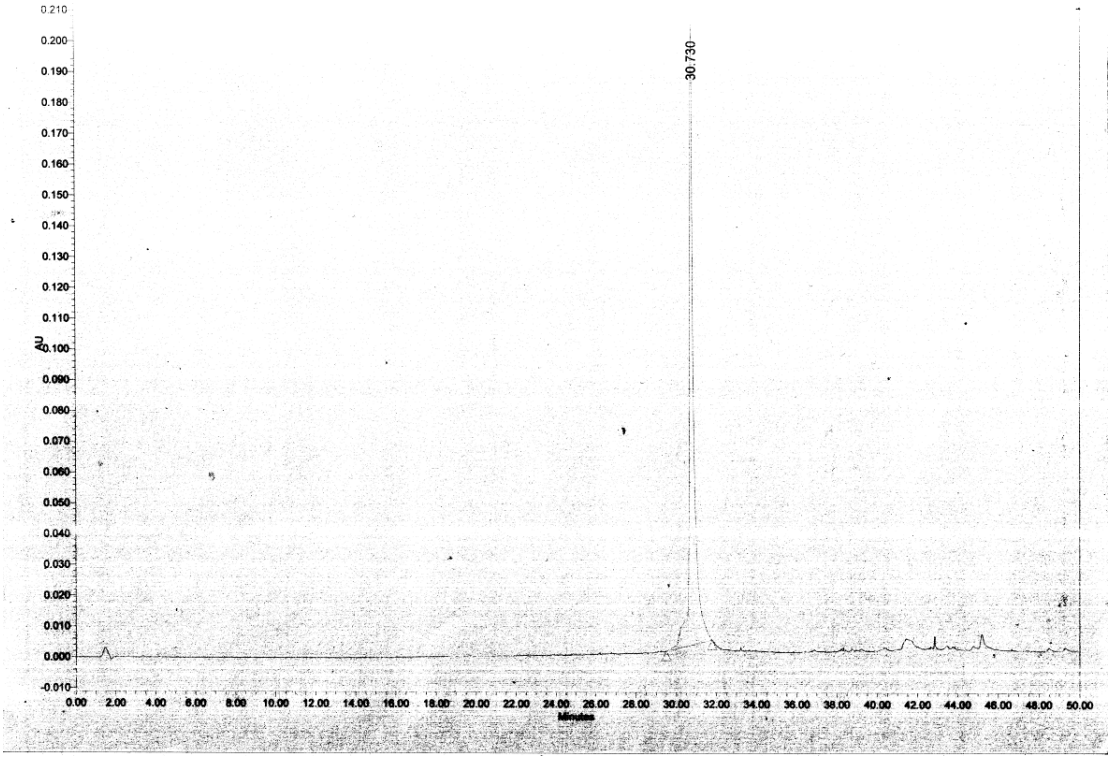


(a)

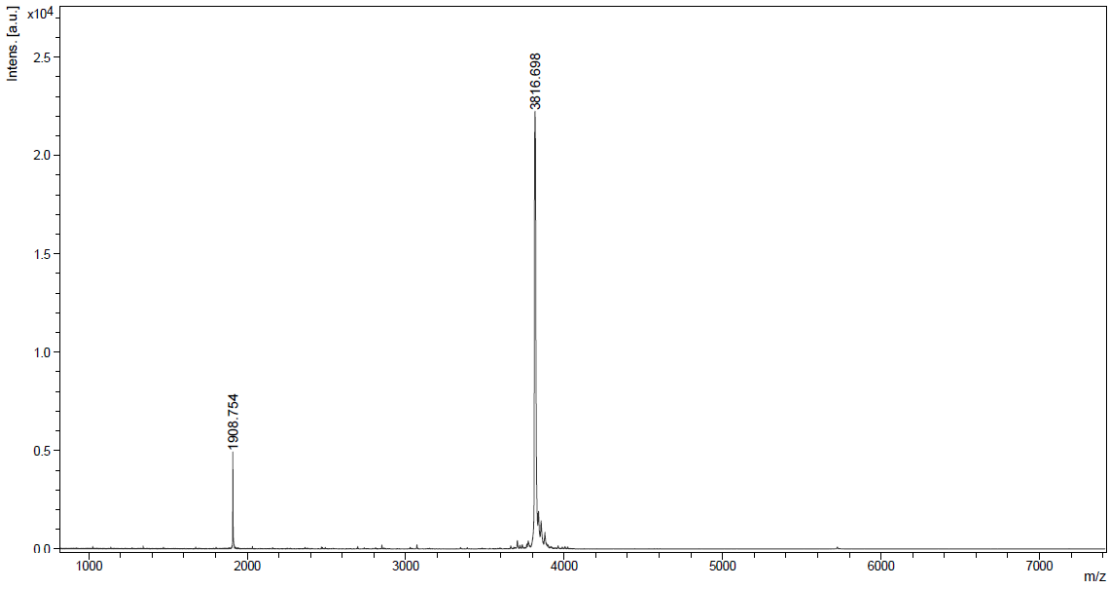


(b)

Fig. S10 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA6** (Ac-CCTTAGACATC-LysNH_2) (calcd. *m/z* 3836.2)

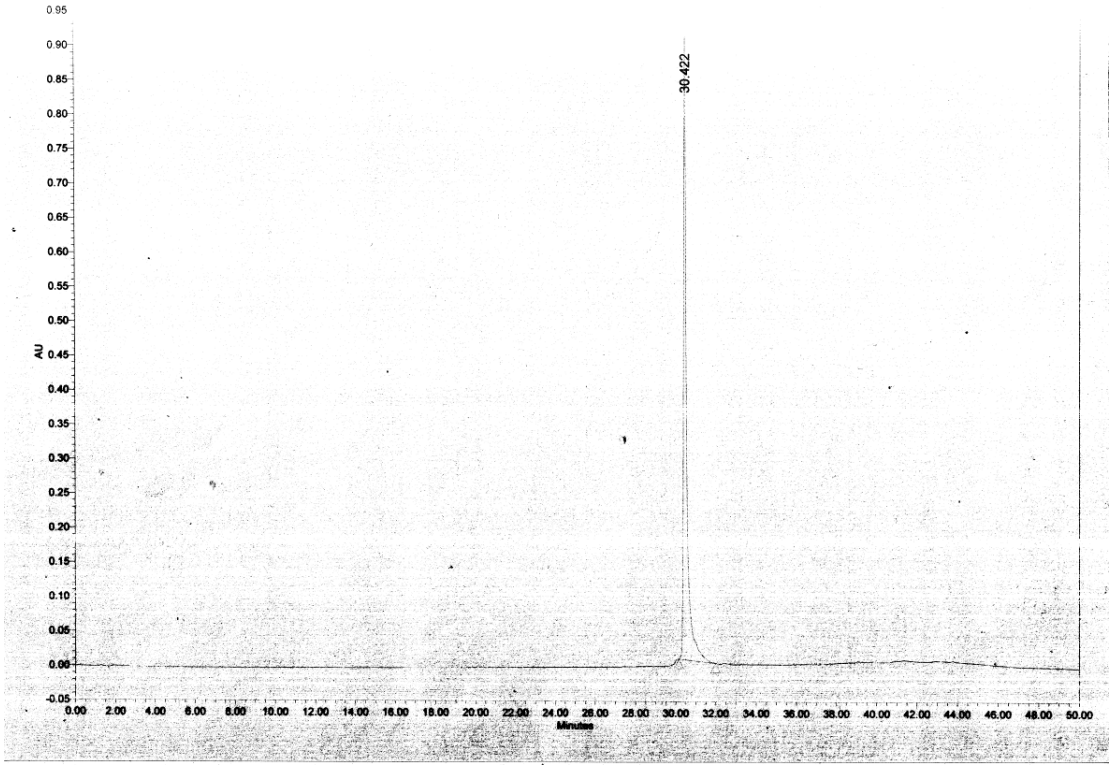


(a)

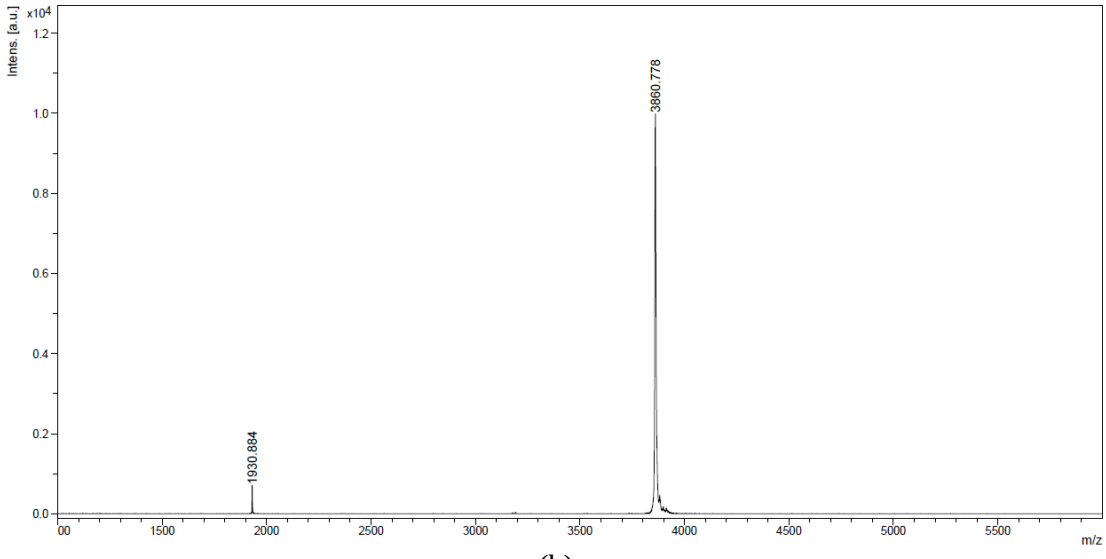


(b)

Fig. S11 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA7** (Ac-CCTTTGTCATC-LysNH_2) (calcd. *m/z* 3818.1)

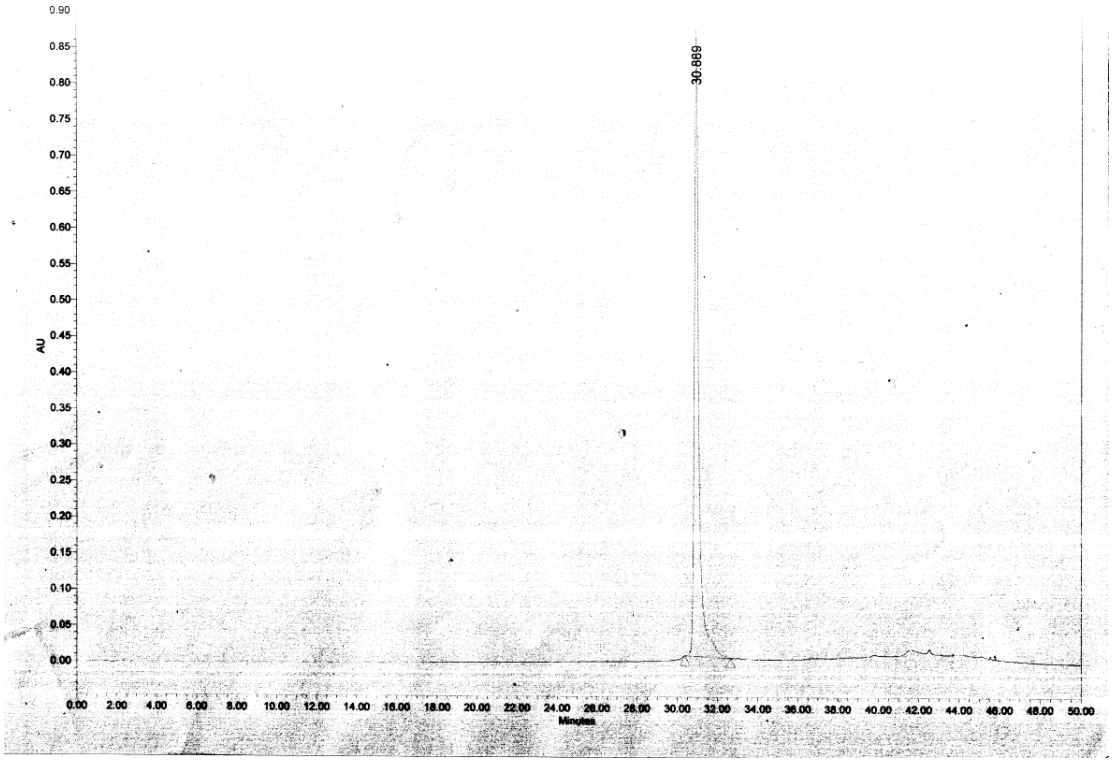


(a)

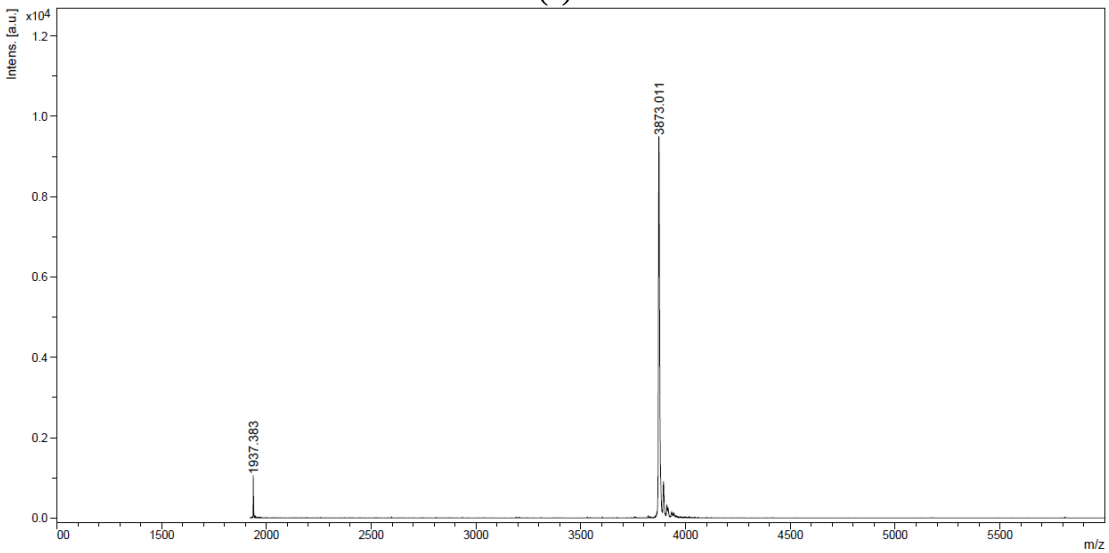


(b)

Fig. S12 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA8** (Ac-AATTTICATCA-LysNH_2) (calcd. *m/z* 3860.2)

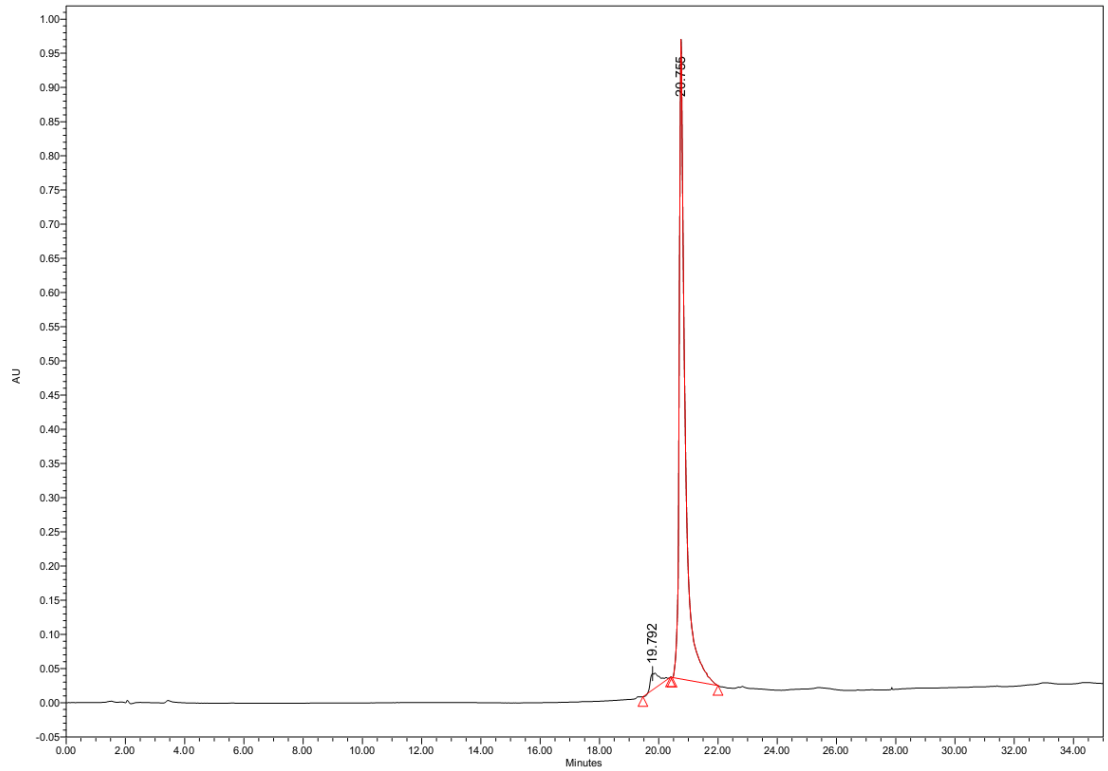


(a)

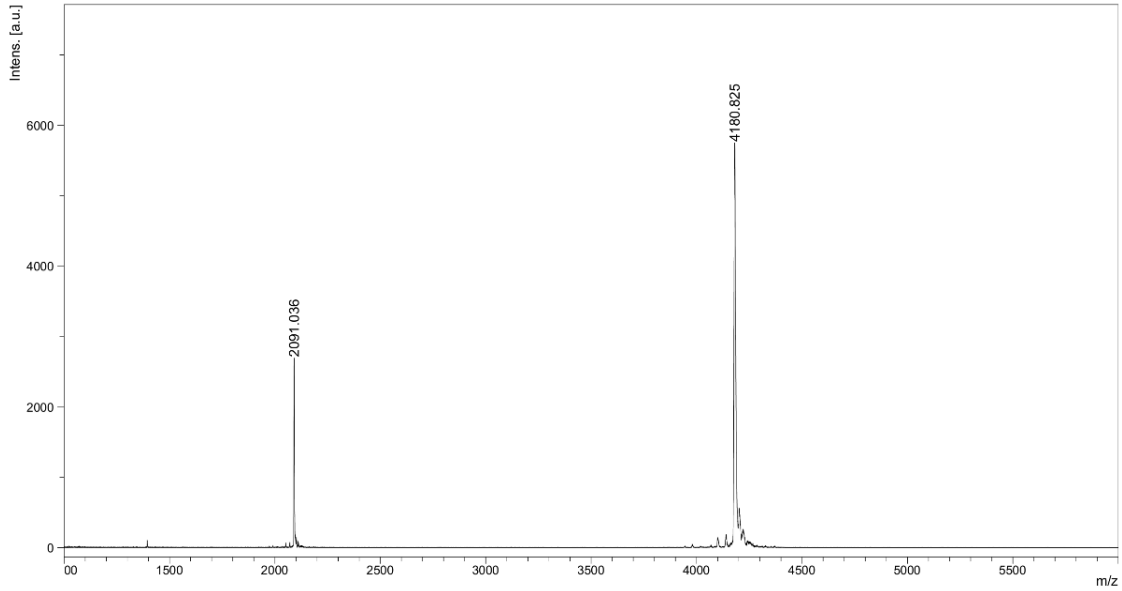


(b)

Fig. S13 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA9** (Ac-AATTTGCATCA-LysNH$_2$) (calcd. *m/z* 3875.2)

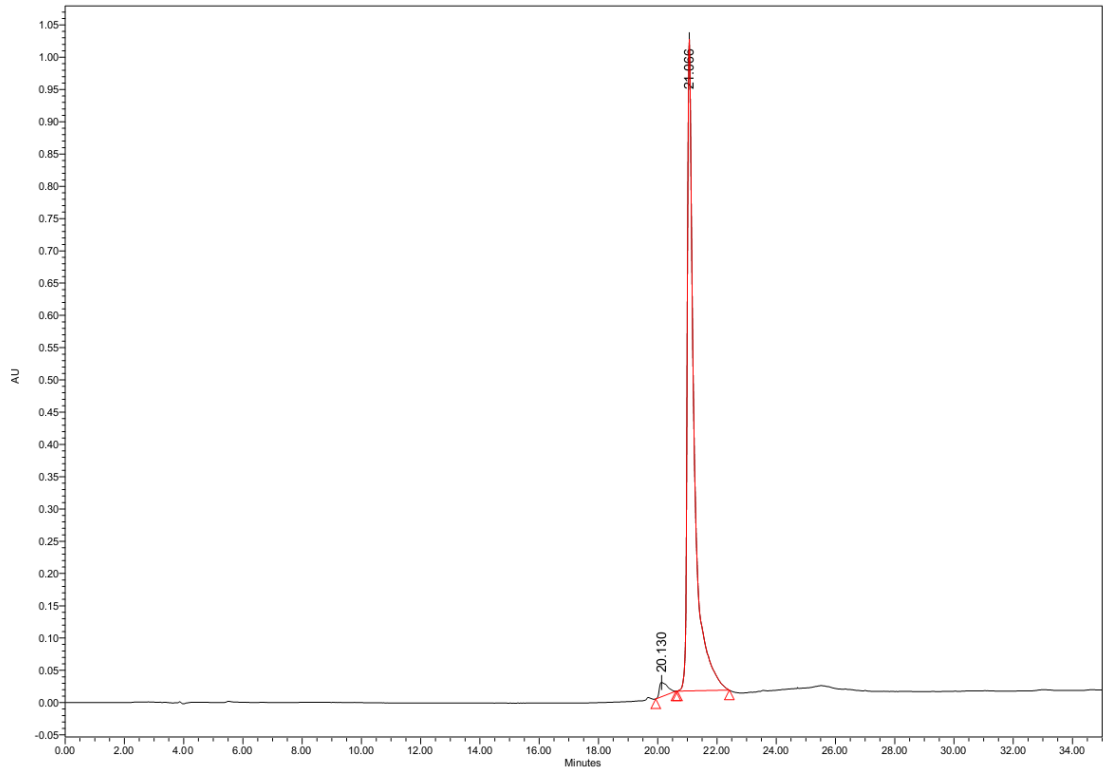


(a)

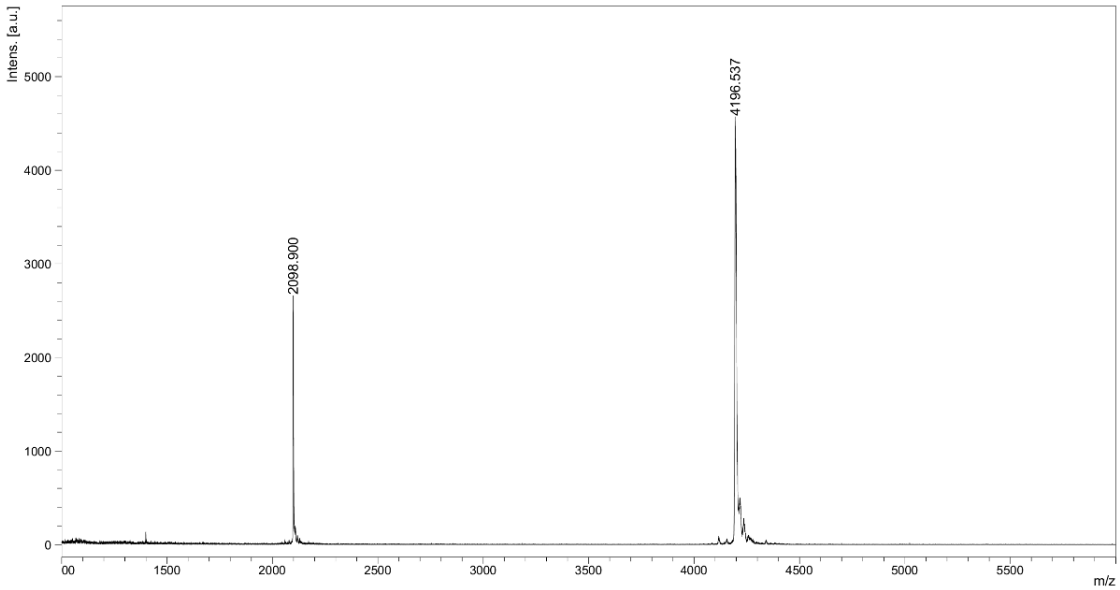


(b)

Fig. S14 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA10** (Flu-O-ICTGCTTCACT-LysNH_2) (calcd. *m/z* 4178.4)



(a)



(b)

Fig. S15 HPLC chromatogram (a) and MALDI-TOF mass spectrum (b) of **PNA11** (Flu-O-GCTGCTTCACT-LysNH_2) (calcd. *m/z* 4193.4)

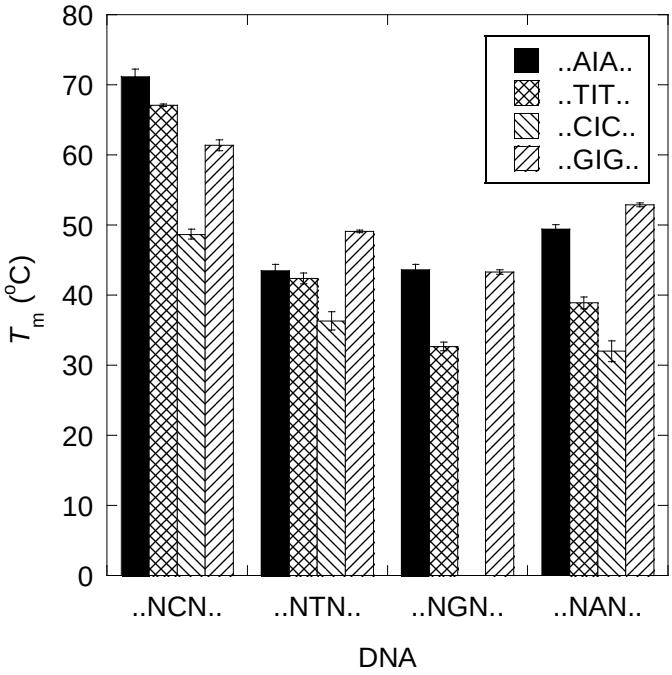


Fig. S16 Melting temperatures (T_m) of DNA hybrids of hypoxanthine-containing acpcPNA **PNA2** (AIA: ■), **PNA3** (TIT: ⊠), **PNA4** (CIC: ▨) and **PNA5** (GIG: ▨). The symbols ..NXN.. in the x-axis denote the three bases in the middle position of the DNA sequences. The identity of N depends on the sequence of PNA, eg. ..TCT.. for PNA ..AIA.. Conditions: 1.0 μM PNA, 1.0 μM DNA, 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.

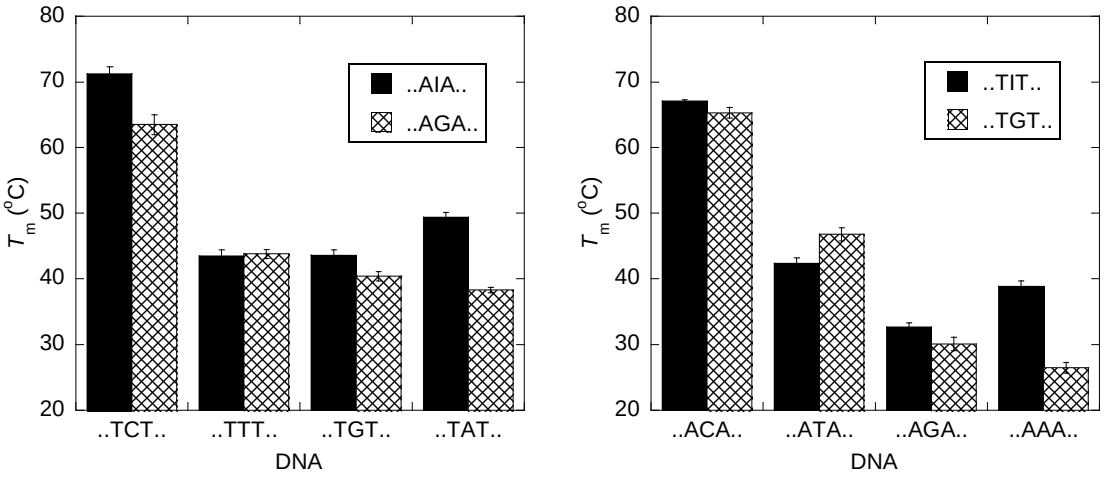


Fig. S17 Comparison of melting temperatures (T_m) of DNA hybrids of hypoxanthine-containing acpcPNA **PNA2** (AIA: ■) and guanine-containing acpcPNA **PNA6** (AGA: ⊠) (left) and hypoxanthine-containing acpcPNA **PNA3** (TIT: ■) and guanine-containing acpcPNA **PNA7** (TGT: ⊠) (right). The symbols ..NNN.. in the x-axis denote the three base in the middle position of the DNA sequences. Conditions: 1.0 μM PNA, 1.0 μM DNA, 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.