Tetrahedron 68 (2012) 8908–8915

Contents lists available at SciVerse ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Alkali metal cations control over nucleophilic substitutions on aromatic fused pyrimidine-2,4-[1*H*,3*H*]-diones: towards new PNA monomers

Pengfa Li [a,b], Chuanlang Zhan [a,*], Shanlin Zhang [a], Xunlei Ding [c], Fengqi Guo [b,*], Shenggui He [c], Jiannian Yao [a,*]

[a] *Beijing National Laboratory of Molecular Science, Laboratory of Photochemistry, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, PR China*
[b] *Department of Chemistry, Zhengzhou University, Zhengzhou 450001, Henan Province, PR China*
[c] *Beijing National Laboratory for Molecular Sciences, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, PR China*

ARTICLE INFO

Article history:
Received 5 June 2012
Received in revised form 16 July 2012
Accepted 10 August 2012
Available online 16 August 2012

Keywords:
Nucleophilic substitution
Aromatic fused pyrimidine-2,4(3*H*,7*H*)-diones
Alkali metal cation
PNA monomer
Solvent effect

ABSTRACT

In this paper we report synthesis of a series of aromatic fused pyrimidine-2,4(3*H*)-dione-1-yl acetic acid and new PNA monomers containing these polycyclic nucleobase analogues. Introduction of a fused aromatic ring onto the 5,6-positions of the pyrimidine-2,4-[1*H*,3*H*]-diones brings about the steric effects and the charge delocalization, both weakening the nucleophilic substitutions on the 1- and 3-positions. We found that alkali metal cations play an important role in this alkylation reaction. LiOH brings out a much more efficient alkylation than NaOH does, while KOH nearly does not work on this reaction. Such influences from the alkali metal cations are probably due to that the charge-pairing interactions between the pyrimidine-2,4-dioxide anions and the alkali metal cations rearrange the charge distribution around the whole aromatic system and increase the negative charge distribution on the 1- and 3-nitrogen atoms, which then strengthens the nucleophilic reactivity on these positions.

1. Introduction

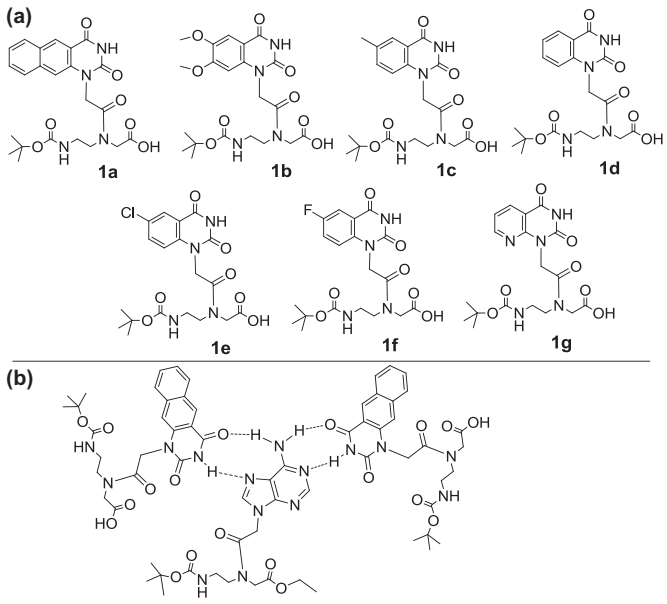
Peptide nucleic acid (PNA) is a typical class of structural mimics of nucleic acids,[1] in which the natural sugar-phosphate backbone is replaced by achiral *N*-(2-aminoethyl) glycine units.[2] PNA can bind to complementary sequence targets with high affinity and sequence specificity, forming either duplexes with mixed sequences,[3] triplexes with homopyrimidine/homopurine sequences,[4] or quadruplexes with G-rich sequences.[5] The PNA/DNA (RNA) complexes generally have a higher thermal stability than the corresponding DNA/DNA (RNA/RNA) complexes, and this encourages potential applications for PNA, for example, for antisense and antigene purposes.[6] The relative high binding affinity of PNA to natural oligonucleotides is determined by both the neutral backbone and conformational flexibility. The neutral backbone largely reduces the electrostatic repulsion formed between naturally DNA (RNA) strands and the conformational flexibility is compensated by the formation of internucleobase hydrogen-bonding and stacking as achieved between nucleic acid strands.[1,7] Attempts has been made to increase the binding affinity and selectivity of PNA. These include modifications of the PNA backbone (1) by using chiral amino acid to replace 2-aminoethyl, yielding chiral PNA backbone[8] or (2) by using achiral branched or *cyclo* alkyl units, achieving conformational constrained PNA backbone.[9]

Our purpose is to use polycyclic nucleobase analogues to tune the inter-nucleobase stacking and hydrogen-bonding interactions. Aromatic fused pyrimidine-2,4(1*H*,3*H*)-dione derivatives have a larger π-aromatic system and also possess hydrogen-bonding features similar to that of thymine, showing potential applications for the medicine or clinic purposes.[10,11] Also, these structural features endow the derivatives with possible utilizations as a typical class of polycyclic nucleobase analogues.[12] Encouraged by this, we aimed to convert the aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones as a proper substitute for the natural thymine nucleobase of PNA monomers (Scheme 1a), in which its different aromatic size and different electron nature will provide different internucleobase stacking and hydrogen-bonding interactions (Scheme 1b).[13]

Thymin-1-yl acetic acid, a key intermediate towards the thymine-based PNA monomers, is generally synthesized by following these two procedures: one is the alkylation of thymine-1*H* with 2-bromoacetate ester in the organic solvents, such as dimethylformamide (DMF) with K_2CO_3 as base and the followed

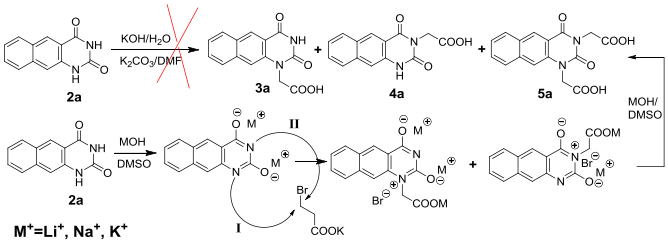
* Corresponding authors. Tel./fax: +86 10 82617312; e-mail addresses: clzhan@iccas.ac.cn (C. Zhan), jnyao@iccas.ac.cn (J. Yao).

0040-4020/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tet.2012.08.028



Scheme 1. (a) The chemical structures of the synthesized PNA monomers containing polycyclic nucleobase analogues of aromatic fused pyrimidine-2,4(3*H*)-diones. (b) The possible Watson–Crick and Hoogsteen H-bonding between the PNA monomers of **1a** and the natural T-monomer, for example.

hydrolysis of the resultant ester derivative to yield the target product;[14] the other is the alkylation of thymine-1*H* with 2-bromoacetic acid with water as the reaction medium and with KOH as the base.[8b,15] Both methods gave a typical yield of ~60%. However, introduction of the fused aromatic ring, for example, naphtha[2′,3′:5,6]pyrimidine-2,4-[1*H*,3*H*]-diones (**2a**) (Scheme 2), dramatically weakened the nucleophilic reactivity on the 1*H* and 3*H* positions under the conditions of the above two mentioned methods and there was only trace of 3-yl acetic acid product formed, whereas no 1-yl acetic acid product was yielded (Scheme 2). To resolve this problem, we have developed a straightforward synthetic procedure and report herein the synthesis of the aromatic fused pyrimidine-2,4(1*H*,3*H*)-dione-1-yl acetic acid (Scheme 2) and further the corresponding PNA monomers (Scheme 1a). In particular, we have observed that both the organic solvent and the alkali metal cations play an important role in the nucleophilic substitutions on the 1*H*- and 3*H*-positions.[16]

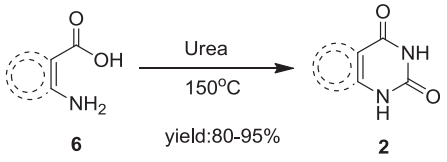


Scheme 2. Synthetic procedures of the naphtha[2′,3′:5,6]pyrimidine-2,4(3*H*)-dione-1-yl acetic acid (**3a**), naphtha[2′,3′:5,6]pyrimidine-2,4(1*H*)-dione-3-yl acetic acid (**4a**) and naphtha[2′,3′:5,6]pyrimidine-2,4-dione-1,3-diyl diacetic acid (**5a**). Detailed reaction conditions from **2a** to **3a**, **4a** and **5a** are shown in Table 2.

2. Results and discussion

2.1. Synthesis of aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones

Aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones were successfully synthesized by heating the mixture of the corresponding aromatic-3-amine-2-carboxylic acid (**6**, Scheme 3) with urea at 150 °C for 12 h. After workup, the desired product was obtained as a white or light yellow solid, typically in a yield of 80–95%.



Scheme 3. Synthesis of the aromatic fused pyrimidine-2,4(3*H*,7*H*)-diones. Reaction conditions: urea, 20 equiv; 150 °C, 12 h; yield: 80–95%.

2.2. Influences of the solvent on the nucleophilic substitutions of the 1- and 3-positions of the aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones

As we can see from Scheme 2 that introduction of the fused aromatic ring, such as naphthalene could produce steric effects, although slightly, for the nucleophilic reactivity on the 1*H* position of naphtha[2′,3′:5,6]pyrimidine-2,4(1*H*,3*H*)-diones. Again, the charge delocalization from the pyrimidine-2,4(1*H*,3*H*)-dioxide anions onto the naphthalene ring may weaken the nucleophilic ability of the 1- and 3-nitrogen atom. Quantum chemical calculations indicate that the negative charge distributed on the pyrimidine-2,4-dioxide anions, e.g., on the atoms of N1, C2, N3, C4, C5, C6, O7 and O8, is 96.2% (−1.923/−2) for 5-methylpyrimidine-2,4-dioxide anions (T^{2-}), whereas it is only 72.4% for naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide anions (A^{2-}) (Table 1 and Fig. 1). Synthetic experiments further support the weak nucleophilic reactivity on the N1 and N3 positions. For example, no desired product of **3a**, while only trace of product **4a** was obtained by following the traditional procedures using ethyl 2-bromoacetic ester (BrAcOEt) as the reactant, K_2CO_3 as the base and DMF as the reaction medium (Table 2, entry 1)[14] or using 2-bromoacetic acid (BrAcOH) as the reactant, KOH as the base and water as the reaction medium (Table 2, entry 2).[8b,15]

Table 1
Charge distributions on 5-methylpyrimidine-2,4-dioxide anions (T^{2-}), naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide anions (A^{2-}), naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide lithium (Li_2A), naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide sodium (Na_2A), and naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide potassium (K_2A), respectively

Compound	Σ (N1–O8)[a]	Σ (C9–C16)[b]	Σ ($2M^+$)[c]	N1	N3
T^{2-}	−1.923	−0.077[d]	—	−0.712	−0.701
A^{2-}	−1.449	−0.551	—	−0.694	−0.667
Li_2A	−1.762	0.016	1.746	−0.782	−0.786
Na_2A	−1.759	−0.052	1.811	−0.771	−0.773
K_2A	−1.772	−0.087	1.859	−0.776	−0.776

[a] Total Natural Bond Orbital (NBO) charge density distributed on the atoms of N1, C2, N3, C4, C5, C6, O7, and O8 (Fig. 1).
[b] Total NBO charge density distributed on the atoms of C9H, C10, C11, C12H, C13H, C14H, C15H, and C16H (Fig. 1).
[c] Total NBO charge density distributed on the two alkali metal cations of M17 and M18 (Fig. 1).
[d] For T^{2-} this represents the total NBO charge density distributed on the 5-substituted methyl (C9).

We then tried to change the reaction conditions. When we used 2-bromoacetic acid as the reactant and DMF as the reaction medium, while replaced K_2CO_3 with NaOH, desired products of **3a**, **4a**, and **5a** were isolated, respectively, in a yield of 3%, 6%, 4% (Table 2, entry 3). However, the insoluble intermediate of naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide sodium and the insoluble sodium salt of products **3a**, **4a**, and **5a** all mixed in the reaction system and this may prevent further conversion from the intermediate to the final

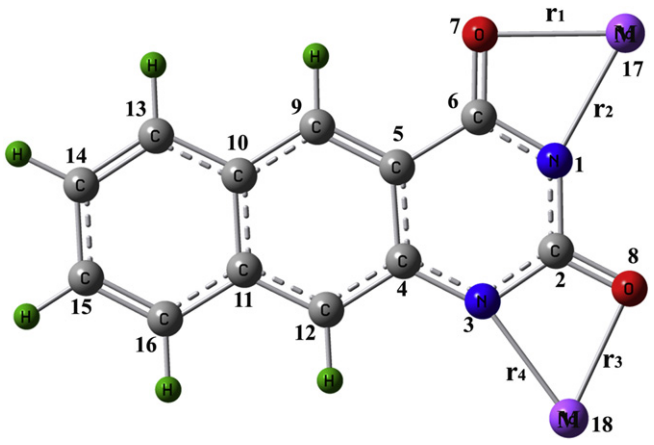


Fig. 1. Optimized geometry and molecular labelling numbers in naphtha[2′,3′:5,6] pyrimidine-2,4-dioxide salt (M_2A).

Table 2
Collections of the yields of products **3a**, **4a**, and **5a** under different reaction conditions[a]

Entries	Reactant	Base	Solvent	T (°C)	Yield of products (%)		
					3a	**4a**	**5a**
1	BrAcOEt	K_2CO_3	DMF	40	—	Trace	—
2	BrAcOH	KOH	H_2O	40	—	Trace	—
3	BrAcOH	NaOH	DMF	40	3	6	4
4	BrAcOH	NaOH	DMSO	40	30	20	10
5	BrAcOH	NaOH	DMSO	25	8	10	3
6	BrAcOH	NaOH	DMSO	30	16	15	10
7	BrAcOH	NaOH	DMSO	50	15	10	8
8	BrAcOH	LiOH	DMSO	40	55	30	5
9	BrAcOH	KOH	DMSO	40	10	8	3

[a] The reaction procedure and conditions are as follows: 1 equiv of naphtha [2′,3′:5,6]pyrimidine-2,4(1*H*,3*H*)-diones was dissolved or dispersed into a proper solvent (10 mL) and then 5 equiv of inorganic base was added. After the mixture was stirred at the selected temperature for 1 h, to which 3 equiv of reactant in 2 mL of this same solvent was dropwise added.

products, decreasing the yield of the desired products. We tried to solve this problem by adding water (in a volume ratio of 1–20%) and we observed that the addition of water could really dissolve the insoluble solids, but unfortunately, the yield of the desired product was also decreased obviously, only trace of desired product was detected by TLC. This may be due to that the strong hydrogen-bonding from water hinders naphtha[2′,3′:5,6]pyrimidine-2,4-dianions from attacking BrAcOH.

We then selected dimethyl sulfoxide (DMSO) as the reaction medium because it is aprotic and its dielectric constant (46.7) is larger than that of DMF (38.0). As expected, DMSO can dissolve the solids completely. As a result, the yield of products **3a**, **4a**, and **5a** was then significantly improved, up to 30%, 20%, and 10%, respectively (Table 2, entry 4). The solvent influences from water to DMF and DMSO imply that (1) a S_N2-type reaction may account for the alkylation on the 1- and 3-nitrogen atoms with 2-bromoacetic acid as the reactant and DMSO as the reaction medium[17] and (2) the balance between solubility and nucleophilicity is a key factor for selection of a reaction solvent.

Various reaction conditions, such as temperature, reaction time, and the relative equivalence of reactant and inorganic base as well as the solvent volume were checked out to optimize the reaction conditions. It was observed that the reaction temperature showed obvious influences on the yield (Table 2, entries 5–7). 40 °C was better for this reaction. Low or high reaction temperature, for example, 25 °C, 30 °C and 50 °C, all produced a low yield of the desired product.

2.3. Influences of the alkali metal cations on the nucleophilic substitutions of the 1- and 3-positions of the aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones

As shown in entries 4, 8 and 9 of Table 2, alkali metal cations influence the nucleophilic reactivity obviously: LiOH produced product **3a** in a yield of 55% and **4a** in a yield of 30% (entry 8), both are obviously increased as compared to that by using NaOH as the base (15% for **3a** and 10% for **4a**, entry 4), whereas KOH produced products of **3a** and **4a** both in a low yield of 10% and 8%, respectively (entry 9).

To get insightful information of the influences of the alkali metal cations on the nucleophilic reactivity, we carried out the DFT calculations. Table 1 shows that binding of Li^+ (from A^{2-} to Li_2A) significantly increases (1) the negative NBO charge density on the N1 and N3, going from −0.694 and −0.667 for A^{2-} to −0.782 and −0.786 for Li_2A, respectively, and (2) that on the part of pyrimidine-2,4-dioxide anions, e.g., on the atoms of N1, C2, N3, C4, C5, C6, O7, and O8 (Fig. 1), going from −1.449 for A^{2-} up to −1.762 for Li_2A. Both are comparable to those of T^{2-}. Therefore, the DFT calculations indicate that the rearrangement of the charge distributions after binding the alkali metal cations may account for the increase of the nucleophilic reactivity on the N1 and N3 positions.

Smaller alkali metal cation from K^+ to Na^+ and Li^+ can produces much stronger charge-pairing interactions between the metal cation and the oxide anion. As revealed by the DFT calculations, the binding energy between the alkali metal cations and the naphtha[2′,3′:5,6]pyrimidine-2,4-dioxide anions increases from −13.40 eV for K^+ to −15.06 eV for Na^+ and then −17.25 eV for Li^+, respectively (Table 3). This means that smaller alkali metal cation possesses stronger tendency to bind with the oxide anion in the DMSO solution and this contributes to the countering alkali metal cation effects on the nucleophilic reactivity, which decreases from Li^+ to Na^+ and then K^+, as observed from the experiments. Noted that Table 1 shows that the calculated NBO charge density on the pyrimidine-2,4-dioxide anions and on the 1- and 3-nitrogen atoms are both comparable to each other when binding with Li^+, Na^+, and K^+, respectively. This is because (1) delocalization of the partial negative charge from A^{2-} onto the $2M^+$ yields a decreases of total positive charge of the alkali metal cations, going from 1.746 for Li^+ to 1.811 for Na^+ and 1.859 for K^+, respectively (Table 1 and 2) the calculations are based on the gas models, and therefore, the solvated effects are excluded.

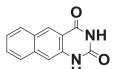
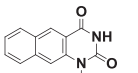
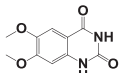
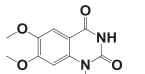
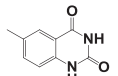
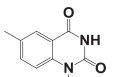
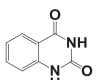
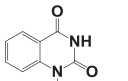
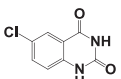
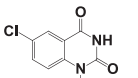
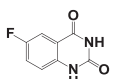
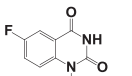
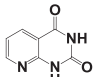
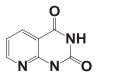
Table 3
Binding energy (E/eV) and distances (r_1, r_2, r_3 and r_4/Å) between the naphtha [2′,3′:5,6]pyrimidine-2,4-dioxide anions (A^{2-}) and alkali metal cations: E=$E(AM_2)$−$E(A^{2-})$−2*$E(M^+)$

	E (eV)	r_1 (Å)	r_2 (Å)	r_3 (Å)	r_4 (Å)
Li^+	−17.25	1.87	1.90	1.83	1.90
Na^+	−15.06	2.23	2.23	2.19	2.24
K^+	−13.40	2.54	2.56	2.48	2.58

The optimized geometry and distance of r_1, r_2, r_3 and r_4 are shown in Fig. 1.

By following this optimized reaction conditions, a series of aromatic fused pyrimidine-2,4(3*H*)-dione-1-yl acetic acid were successfully synthesized. As shown in Table 4, the yield of the desired product with the 1*H* position substituted by acetic acid was dependent on the utilization of the base, decreasing in the order of LiOH, NaOH and KOH for all these seven products.

Table 4
Synthesis of product **3** under the optimum reaction conditions[a]

Entries	Reactant	Product	Yield (%)		
			LiOH	NaOH	KOH
1	 2a	 3a	60–80	20–30	8–15
2	 2b	 3b	60–75	16–30	6–13
3	 2c	 3c	60–80	20–30	10–15
4	 2d	 3d	50–65	10–22	8–16
5	 2e	 3e	50–60	5–10	<5
6	 2f	 3f	40–50	10–20	10–15
7	 2g	 3g	40–50	12–21	8–14

[a] The reaction procedure and conditions are as follows: 1 equiv of aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones was dissolved or dispersed into DMSO (10 mL) and then 5 equiv of base was added. After the mixture was stirred at 40 °C for 1 h, to this solution 3 equiv of reactant in 2 mL of DMSO was dropwise added.

2.4. Discriminations of the aromatic fused pyrimidine-2,4(1*H*,3*H*)-dione-1-yl acetic acid (3) and aromatic fused pyrimidine-2,4(1*H*,3*H*)-dione-3-yl acetic acid (4)

In the ^{1}H NMR spectra of products **3c** and **4c** (Fig. 2b), seven sets of signals with the same integrals were observed in the region of 2–14 ppm. These signals include two singlets (protons a and k) in the region of 2–5 ppm, two doublets (protons c and d) and one singlet (proton g) in the region of 7–8 ppm, and two singlets (protons i and l) in the region of 11–14 ppm. These signals are in agreements well with the structure of the products of **3c** and **4c** (Fig. 2a). In the region of 4–5 ppm, the singlet was assigned to the protons of the methylene groups. For the 1-acetic acid derivatives (**3c**), they appear at the 4.78 ppm, whereas they occur at 4.55 ppm for the 3-acetic acid derivatives. This is the most useful feature in the chemical shifts to discriminate products **3c** and **4c**.

Assignments of the protons on **3c** and **4c** are confirmed by the H–H NOESY, HSQC and HMBC spectra (Fig. 3a–d and 4). In H–H NOESY spectrum of **3c** two obvious signals were observed between the protons of k and d (Fig. 3a), whereas no signals appeared between the protons of i and d. In contrast, two signals occurred between the protons of i and d and no signals appeared between the protons of k and d in the H–H NOESY spectrum of **4c** (Fig. 3b).

Similarly, in the ^{13}C NMR spectrum, the signals appearing in the region of 40–45 ppm are assigned to the carbons of the methylene groups (Fig. 1c). The methylene carbons in the 1-yl acetic acid derivative appear at 44.2 ppm, while those in the 3-yl acetic acid derivative occur at 41.7 ppm. The aromatic carbons were fully assigned further from the HSQC (Fig. 4) and HMBC spectra (Fig. 3c and d). In HMBC spectra, product **3c** shows three sets of strong signals between the protons k and the carbons e, h and l, respectively (Fig. 3c), whereas **4c** gives out three sets of signals between the protons k and the carbons h, j and l, respectively (Fig. 3d).

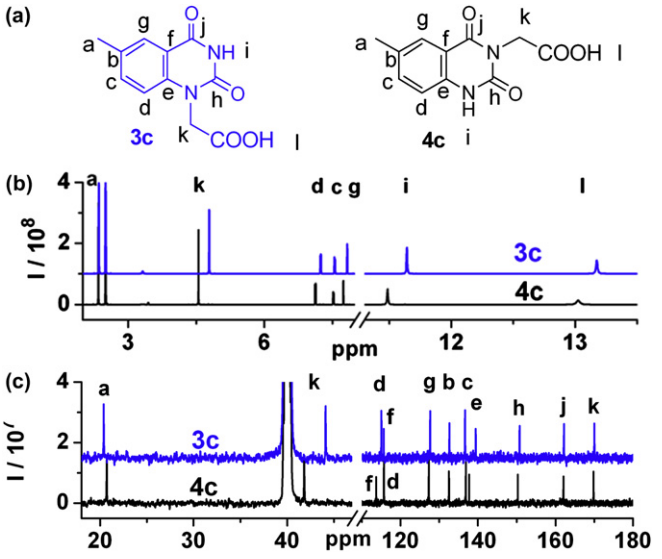


Fig. 2. Molecular structures (a), ^{1}H NMR (b) and ^{13}C NMR (c) spectra of products **3c** and **4c**, both in DMSO-d_6.

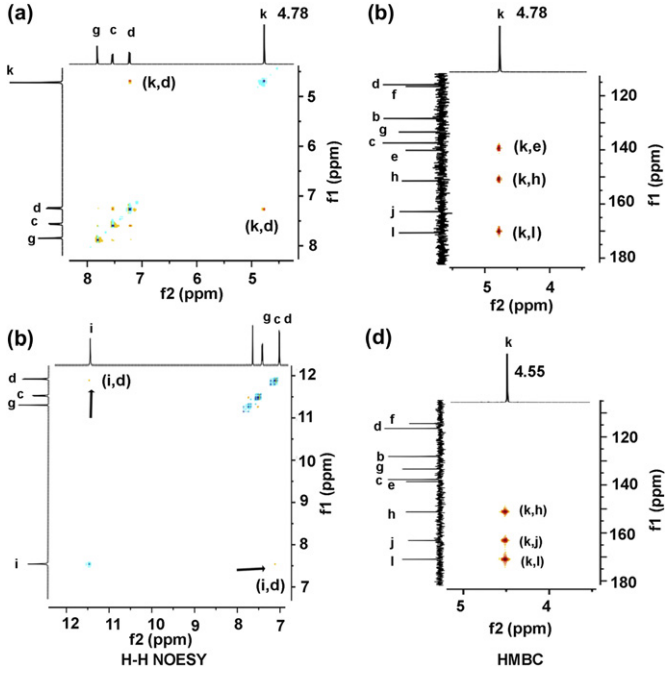


Fig. 3. Partial H–H NOESY (a and b) and HMBC (c and d) spectra of products **3c** (a and c) and **4c** (b and d), all in DMSO-d_6.

2.5. Synthesis of the PNA monomers with aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones as the base analogues

The Boc-protected PNA monomers were synthesized by following the procedure shown in Scheme 4, similar to that reported in the references.[14] In the condensation of **3** and **7** DCC

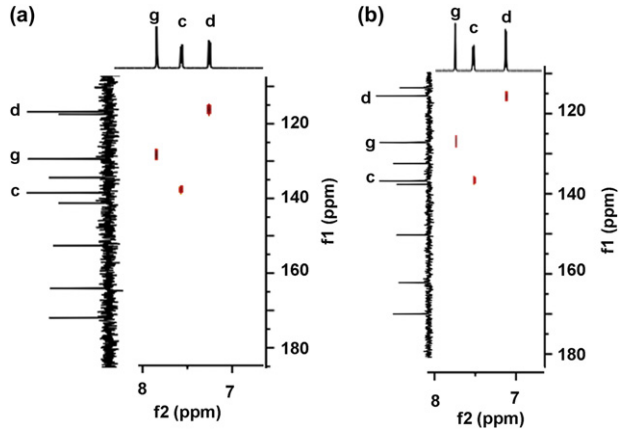
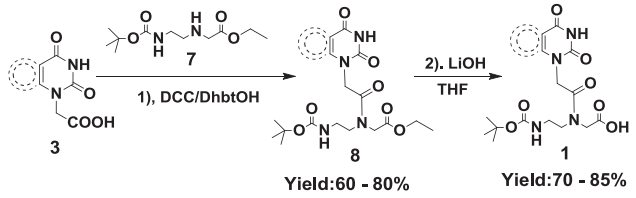


Fig. 4. Partial HSQC spectra of products **3c** (a) and **4c** (b), both in DMSO-d_6.



Scheme 4. Synthesis of the aromatic fused pyrimidine-2,4(3*H*,7*H*)-diones. Reaction conditions: (1), compound **7**, 1 mol; DhbtOH 1.1 mol; compound **3**, 1.1 mol; 1:1 DMF and DCM, 7.5 mL; 0 °C, 3 h; yield: 60–80%; (2), compound **8**, 1.0 mmol; THF, 5 mL; 1 M LiOH, 5 mL; rt, 45 min.

(dicyclohexylcarbodiimide) and DhbtOH (3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine) were both utilized to get a high yield of product **8**. In this procedure, trace of DCU (dicyclohexylurea) may mix with product **8** and they may be appeared as white precipitate during the hydrolysis of product **8**. The white precipitate should be removed by filtering the reaction mixture before the pH value of the filtration was adjusted to 2 by using 1 M hydrochloric acid aqueous solution.

3. Conclusions

We have developed a straightforward approach to the synthesis of aromatic fused pyrimidine-2,4(3*H*)-dione-1-yl acetic acid. We found that both a proper solvent of DMSO and a proper alkali metal hydroxide, such as LiOH yielded efficient alkylation of 2-bromoacetic acid on the 1- and 3-positions of the aromatic fused pyrimidine-2,4(3*H*)-dione derivatives. By using this approach we successfully synthesized a series of new PNA monomers containing these polycyclic base analogues. Currently, incorporation of these PNA monomers into PNA oligomers is undergoing.

4. Experimental part

4.1. Materials and methods

Starting materials are all commercially available reagents and solvents used as received except for statements. Tetrahydrofuran (THF) was distilled on sodium. All solvents were purified using the standard procedures. Reactions were monitored by thin-layer chromatography (TLC) on pre-coated silica gel plates (Yantai Shi Huagxue Gongye Yanjiusuo) and visualized using UV irradiation (254 or 365 nm). Flash chromatography was performed on silica gel H60 (Qingdao Haiyang Huagongchang).

^{1}H and ^{13}C NMR spectra were recorded on a Bruker AVANCE 400 or AVANCE 600 spectrometer and referenced to solvent signals. Chemical shifts were referenced to $Si(CH_4)_4$ signals. ESI mass spectrometric were obtained on LC-MS 2010 and BRUKER Apex IV FTMS and MALDI-TOF mass spectrometric measurements were performed on Bruker Biflex III MALDI-TOF spectrometer.

4.2. Quantum chemical calculations

Density functional theory (DFT) calculations were performed using the Gaussian 03 program[18] with the B3LYP exchange-correlation functional.[19] All-electron triple-ξ valence basis sets with diffusion and polarization functions (6-311++G**)[20] are used for all atoms. Geometry optimizations were performed with full relaxation of all atoms. Calculations were performed in gas phase without solvent effects. Vibrational frequency calculations were performed to check that the stable structures had no imaginary frequency. Charge distribution of the molecules was calculated by natural population analysis.[21]

4.3. Synthetic procedures

4.3.1. Synthesis of aromatic fused pyrimidine-2,4(1H,3H)-diones. The typical procedure to aromatic fused pyrimidine-2,4(1*H*,3*H*)-diones (**2**) is as follows. Aromatic-3-amine-2-carboxylic acid (0.1 mol) and urea (2 mol) were both mixed in a flask and then heated at 150 °C for 12 h. When the mixture was cooled down to 100 °C, 40 mL of water was added and the resultant mixture was stirred for another 10 min to dissolve the unreacted urea. After the reaction mixture was cooled down to room temperature, it was filtered and 400 mL, 1 mol/L NaOH aqueous solution was then added. Another 1 h later 55 mL of acetic acid was dropwise added for acidification and the resultant light yellow or white precipitates were filtered and dried to give the desired product **2** in a yield of 80–95%.

For **2a**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.31 (s, 1H, –NH), 11.20 (s, 1H, –NH), 8.64 (s, 1H, –naph–H), 8.10 (d, 1H, J=8.4 Hz, –naph–H), 7.90 (d, 1H, J=8.4 Hz, –naph–H), 7.61 (t, 1H, J=8.0 Hz, –naph–H), 7.52 (s, 1H, –naph–H), 7.47 (t, 1H, J=8.0 Hz, –naph–H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 162.9, 150.3, 136.5, 136.3, 129.5, 129.2, 129.0, 128.5, 126.7, 124.9, 115.2, 110.2; Anal. Calcd: C, 67.92; H, 3.80; N, 13.20; found: C, 68.07; H, 3.74; N, 13.17; GCT-MS m/z (%): 212.0 (M^+, 100%).

For **2b**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.36 (s, 1H, –NH), 11.21 (s, 1H, –*NH*), 7.29 (s, 1H, –ph–H), 6.71 (s, 1H, –ph–H), 3.85 (s, 3H, $-CH_3$), 3.80 (s, 3H, $-CH_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 161.6, 149.8, 137.2, 136.4, 132.1, 126.8, 115.3, 113.2, 56.2, 54.3; Anal. Calcd: C, 54.05; H, 4.54; N, 12.61; found: C, 53.97; H, 4.70; N, 12.57; GCT-MS m/z (%): 222.2 (M^+, 100%).

For **2c**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.21 (s, 1H, –NH), 11.05 (s, 1H, –NH), 7.69 (s, 1H, –phH), 7.47 (d, 1H, J=7.2 Hz, –phH), 7.08 (d, 1H, J=7.2 Hz, –phH), 2.32 (s, 3H, $-CH_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 162.8, 150.3, 138.7, 135.9, 131.5, 126.5, 115.3, 114.2, 20.2; Anal. Calcd: C, 61.36; H, 4.58; N, 15.90; found: C, 61.11; H, 4.74; N, 16.07; GCT-MS m/z (%): 176.0 (M^+, 100%).

For **2d**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.21 (s, 1H, –NH), 11.10 (s, 1H, –NH), 8.00 (d, 1H, J=7.6 Hz, –phH), 7.71 (t, 1H, J=7.6 and 7.6 Hz, –phH), 7.32 (d, 1H, J=8.4 Hz, –phH), 7.30 (t, 1H, J=7.6 and 7.5 Hz, –phH); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 162.6, 150.1, 138.6, 135.8, 131.4, 126.5, 115.2, 114.1; Anal. Calcd: C, 59.26; H, 3.73; N, 17.28; Found: C, 59.17; H, 3.74; N, 17.07; GCT-MS m/z (%): 162.0 (M^+, 100%).

For **2e**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.44 (s, 1H, –NH), 11.27 (s, 1H, –NH), 7.82 (s, 1H, –phH), 7.70 (d, 1H, J=8.4 Hz, –phH), 7.19 (d, 1H, J=8.4 Hz, –phH); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 161.8, 150.1, 139.7, 134.8, 126.3, 115.9, 117.5, 115.8; Anal. Calcd: C, 48.88; H, 2.56; N, 14.25; Found: C, 48.97; H, 2.70; N, 14.27; GCT-MS m/z (%): 196.0 (M^+, 100%).

For **2f**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.29 (s, broad, 2H, −NH), 7.60 (m, 2H, −phH), 7.21 (s, 1H, −phH); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 162.1, 158.5, 156.1, 150.1, 137.5, 123.0, 122.8, 117.6, 117.5, 115.4, 115.3, 112.1, 111.8; Anal. Calcd: C, 53.34; H, 2.80; N, 15.55; found: C, 53.17; H, 2.74; N, 15.37; GCT-MS *m*/*z* (%): 180.0 (M^+, 100%).

For **2g**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 11.64 (s, 1H, −NH), 11.48 (s, 1H, −NH), 8.60 (d, 1H, *J*=8.0 Hz, −pyH), 8.26 (d, 1H, *J*=8.0 Hz, −pyH), 7.26 (q, 1H, *J*=4.8 and 6.8 Hz, −pyH); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 162.2, 154.8, 152.4, 150.4, 136.7, 118.9, 109.9; Anal. Calcd: C, 51.54; H, 3.09; N, 25.76; found: C, 51.37; H, 2.99; N, 25.67; GCT-MS *m*/*z* (%): 163.0 (M^+, 100%).

4.3.2. Synthesis of the aromatic fused pyrimidine-2,4(3H)-dione-1-yl acetic acid. The typical procedure is as follows: **2** (10 mmol) and LiOH (25 mmol) was dissolved into DMSO (60 mL) and stirred at 40 °C for 2 h, to which a 5 mL DMSO solution of $BrCH_2COOH$ (10 mmol) was dropwise added inside within 5 min. The reaction was continued at 40 °C for 12 h. Then it was poured into 500 mL ethyl acetate. The resultant white or light yellow precipitates were collected and re-dissolved into 100 mL water. The water solution was adjusted to pH=6 using 4 M hydrochloric acid and then was put at 3–6 °C for 2 h. The precipitates (**2**) were then filtered out and the solution was again adjusted to pH=2 and put at 3–6 °C for another 2 h. The product **3** was filtered out in a typical yield shown in Tables 1 and 2, respectively. The yield for each compound of **3a**–**g** was listed in Table 2.

For **3a**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.21 (s, 1H, −COOH), 11.76 (s, 1H, −NH), 8.76 (s, 1H, −naph−H), 8.15 (d, 1H, *J*=8.4 Hz, −naph−H), 7.99 (d, 1H, *J*=8.4 Hz, −naph−H), 7.81 (s, 1H, −naph−H), 7.64 (t, 1H, *J*=8.0 and 8.0 Hz, −naph−H), 7.52 (t, 1H, *J*=8.0 and 8.0 Hz, −naph−H), 4.90 (s, 2H, $-CH_2$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 169.6, 161.8, 150.3, 136.7, 136.5, 129.5, 129.2, 128.2, 127.3, 125.6, 115.7, 110.5, 56.0; Anal. Calcd: C, 62.22; H, 3.73; N, 10.37; found: C, 62.42; H, 3.60; N, 10.17; ESI-MS *m*/*z* (%): 269.0 (M^-, $-H^+$, 100%).

For **3b**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.08 (s, 1H, −COOH), 11.51 (s, 1H, −NH), 7.40 (s, 1H, −ph−H), 6.82 (s, 1H, −ph−H), 4.85 (s, 2H, $-CH_2$), 3.88 (s, 3H, $-CH_3$), 3.81 (s, 3H, $-CH_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 169.8, 161.2, 155.0, 150.5, 145.0, 136.9, 108.0, 107.4, 98.0, 56.4, 55.8, 44.0. Anal. Calcd: C, 51.43; H, 4.32; N, 10.00; found: C, 51.23; H, 4.42; N, 10.20; ESI-MS *m*/*z* (%): 279.1 (M^-, $-H^+$, 100%).

For **3c**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.16 (s, 1H, −COOH), 11.64 (s, 1H, −NH), 7.83 (s, 1H, −phH), 7.56 (d, 1H, *J*=8.2 Hz, −phH), 7.25 (d, 1H, *J*=8.4 Hz, −phH), 4.79 (s, 2H, $-CH_2$), 2.36 (s, 3H, $-CH_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.2, 162.4, 150.7, 139.5, 136.7, 132.5, 127.5, 115.9, 115.0, 44.4, 20.4; Anal. Calcd: C, 56.41; H, 4.30; N, 11.96; found: C, 56.21; H, 4.35; N, 12.06; ESI-MS *m*/*z* (%): 233.1 (M^-, $-H^+$, 100%).

For **3d**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.17 (s, 1H, −COOH), 11.71 (s, 1H, −NH), 8.03 (d, 1H, *J*=7.6 Hz, −phH), 7.75 (t, 1H, *J*=7.6 and 7.6 Hz, −phH), 7.35 (d, 1H, *J*=8.4 Hz, −phH), 7.29 (t, 1H, *J*=7.6 and 7.5 Hz, −phH), 4.81 (s, 2H, $-CH_2$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 169.6, 161.7, 150.3, 141.1, 135.4, 127.6, 122.8, 115.4, 114.6, 43.7; Anal. Calcd: C, 54.55; H, 3.66; N, 12.72; found: C, 54.35; H, 3.78; N, 12.85; ESI-MS *m*/*z* (%): 219.0 (M^-, $-H^+$, 100%).

For **3e**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.23 (s, 1H, −COOH), 11.89 (s, 1H, −NH), 7.95 (s, 1H, −phH), 7.79 (d, 1H, *J*=8.4 Hz, −phH), 7.42 (d, 1H, *J*=8.8 Hz, −phH), 4.81 (s, 2H, $-CH_2$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 169.4, 160.7, 150.1, 140.1, 135.0, 127.2, 126.5, 117.1, 114.9, 44.0; Anal. Calcd: C, 47.17; H, 2.77; N, 11.00; found: C, 47.08; H, 2.85; N, 11.15; ESI-MS *m*/*z* (%): 253.0 (M^-, $-H^+$, 100%).

For **3f**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.22 (s, 1H, −COOH), 11.91 (s, 1H, −NH), 7.85 (s, 1H, −phH), 7.85 (d, 1H, *J*=8.4 Hz, −phH), 7.43 (d, 1H, *J*=8.8 Hz, −phH), 4.80 (s, 2H, $-CH_2$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 169.5, 160.8, 150.0, 140.2, 135.1, 127.3, 126.6, 117.2, 114.9, 112.3, 44.1; Anal. Calcd: C, 50.43; H, 2.96; N, 11.76; found: C, 50.25; H, 3.09; N, 11.58; ESI-MS *m*/*z* (%): 237.0 (M^-, $-H^+$, 100%).

For **3g**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.27 (s, 1H, −COOH), 11.98 (s, 1H, −NH), 8.60 (d, 1H, *J*=8.0 Hz, −pyH), 8.27 (d, 1H, *J*=8.0 Hz, −pyH), 7.23 (q, 1H, *J*=6.4, 6.4 and 6.8 Hz, −pyH), 4.68 (s, 2H, $-CH_2$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.5, 161.6, 154.4, 151.9, 151.3, 136.7, 118.4, 109.5, 44.3; Anal. Calcd: C, 48.87; H, 3.19; N, 19.00; found: C, 48.68; H, 3.07; N, 19.14; ESI-MS *m*/*z* (%): 220.0 (M^-, $-H^+$, 100%).

4.3.3. Synthesis of the PNA monomers containing aromatic fused pyrimidine-2,4(3H)-diones. Briefly, ethyl *N*-[2-Boc-aminoethyl]glycinate (**7**, 1.0 mmol), DhbtOH (3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine, 1.1 mmol), and aromatic fused pyrimidine-2,4(13*H*)-dione-1-yl acetic acid (**3**, 1.1 mmol) were dissolved in 1:1 mixed DMF and DCM (7.5 mL). After the solution was cooled to 0 °C, DCC (6.6 mmol) was added and the reaction was allowed to react for 3 h at this temperature. After removal of DCU by filtration, the organic layer was washed with dilute aqueous $NaHCO_3$ and brine aqueous solution, respectively. After the solvent was removed under vacuum, the resultant oil residue was dissolved in DCM (3 mL) and the precipitate was filtered out. The product was precipitated by the addition of petroleum ether (3 mL) at 0 °C. The precipitate was re-dissolved in DCM and precipitated once more with petroleum ether, affording **8**. The yield is 60–80%.

Compound **8** (1.0 mmol) was suspended in THF (5 mL) and 1 M LiOH (aqueous, 5 mL) was added. The mixture was stirred for 45 min at room temperature and the white precipitate of DCU was filtered out. The filtrate was then extracted with DCM and water for several times until no white precipitate appeared. The water layer was collected and cooled to 0 °C and its pH was adjusted to 2 by the dropwise addition of 1 M HCl. The water layer was extracted with EtOAc. After removal of EtOAc, the residue was re-dissolved in methanol and the desired product **1** was re-precipitated by using cyclohexane as a colourless solid. The yield for each PNA monomer is in the range of 70–85%.

For **1a**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.09 (s, 0.2H, −COOH) and 12.61 (s, 0.8H, −COOH), 11.67 (s, 1H, −NH), 8.71 (s, 1H, −naph−H), 8.15 (d, 1H, *J*=8.4 Hz, −naph−H), 7.95 (d, 1H, *J*=8.4 Hz, −naph−H), 7.63 (t, 1H, *J*=8.0 Hz, −naph−H), 7.59 (s, 1H, −naph−H), 7.48 (t, 1H, *J*=8.0 Hz, −naph−H), 6.93 (s, 0.8H, −COOH) and 6.74 (s, 0.2H, −COOH), 4.90 (s, 1.2H, $-CH_2$) and 4.70 (s, 0.8H, $-CH_2$), 4.28 (s, 0.8H, $-CH_2$) and 3.97 (s, 1.2H, $-CH_2$), 3.47 (t, 1.2H, *J*=6.0 and 6.4 Hz, $-CH_2$) and 3.32 (d, 0.8H, *J*=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, *J*=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, *J*=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.9, 170.6, 166.8, 166.5, 161.7, 155.7, 149.8, 141.3, 136.4, 135.0, 129.6, 129.4, 128.5, 126.8, 125.0, 114.3, 110.2, 78.1, 77.1, 49.0, 47.4, 44.8, 44.4, 38.1, 37.7, 28.2; Anal. Calcd: C, 58.72; H, 5.57; N, 11.91; found: C, 58.48; H, 5.49; N, 12.16; TOF-MS *m*/*z* (%): 493.2 ($M+Na^+$, 100%).

For **1b**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.05 (s, 0.2H, −COOH) and 12.62 (s, 0.8H, −COOH), 11.50 (s, 1H, −NH), 7.29 (s, 1H, −ph−H), 6.94 (s, 0.8H, −COOH) and 6.73 (s, 0.2H, −COOH), 6.73 (s, 1H, −ph−H), 4.81 (s, 1.2H, $-CH_2$) and 4.70 (s, 0.8H, $-CH_2$), 4.28 (s, 0.8H, $-CH_2$) and 3.97 (s, 1.2H, $-CH_2$), 3.88 (s, 3H, $-CH_3$), 3.81 (s, 3H, $-CH_3$), 3.47 (t, 1.2H, *J*=6.0 and 6.4 Hz, $-CH_2$) and 3.32 (d, 0.8H, *J*=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, *J*=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, *J*=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.9, 170.6, 167.3, 167.3, 167.0, 161.1, 155.7, 155.1, 150.0, 145.2, 135.2, 107.5, 105.5, 97.6, 78.0, 77.7, 55.9, 55.7, 49.1, 47.0, 44.8, 44.4, 38.1, 37.7, 28.2; Anal. Calcd: C, 52.50; H, 5.87; N, 11.66; found: C, 52.38; H, 5.73; N, 11.74; TOF-MS *m*/*z* (%): 503.2 ($M+Na^+$, 100%).

For **1c**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.07 (s, 0.2H, −COOH) and 12.64 (s, 0.8H, −COOH), 11.64 (s, 1H, −NH), 7.82 (s, 1H, −phH), 7.54 (d, 1H, *J*=8.0 Hz, −phH), 7.24 (d, 1H, *J*=8.0 Hz, −phH), 6.93 (s, 0.8H, −COOH) and 6.74 (s, 0.2H, −COOH), 4.80 (s, 1.2H, $-CH_2$) and

4.71 (s, 0.8H, $-CH_2$), 4.29 (s, 0.8H, $-CH_2$) and 3.98 (s, 1.2H, $-CH_2$), 3.47 (t, 1.2H, J=6.0 and 6.4 Hz, $-CH_2$) and 3.33 (d, 0.8H, J=6.4 and 6.4 Hz, $-CH_2$), 3.21 (d, 1.2H, J=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, J=6.4 Hz, $-CH_2$), 2.38 (s, 3H, $-CH_3$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.9, 170.7, 167.2, 166.8, 161.7, 155.8, 149.9, 137.4, 136.2, 131.9, 126.8, 115.2, 113.4, 78.1, 77.1, 49.1, 46.7, 44.8, 44.4, 38.1, 37.7, 28.2, 20.3; Anal. Calcd: C, 55.29; H, 6.03; N, 12.90; found: C, 55.36; H, 6.07; N, 12.80; TOF-MS m/z (%): 457.3 (M+Na^+, 100%).

For **1d**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.06 (s, 0.2H, −COOH) and 12.63 (s, 0.8H, −COOH), 11.64 (s, 1H, −NH), 8.03 (d, 1H, J=7.6 Hz, −phH), 7.75 (t, 1H, J=7.6 and 7.6 Hz, −phH), 7.30 (m, 2H, J=8.4, 7.6 and 7.5 Hz, −phH), 6.93 (s, 0.8H, −COOH) and 6.74 (s, 0.2H, −COOH), 4.81 (s, 1.2H, $-CH_2$) and 4.71 (s, 0.8H, $-CH_2$), 4.29 (s, 0.8H, $-CH_2$) and 3.98 (s, 1.2H, $-CH_2$), 3.47 (t, 1.2H, J=6.0 and 6.4 Hz, $-CH_2$) and 3.34 (d, 0.8H, J=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, J=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, J=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.9, 170.6, 167.1, 166.8, 161.6, 155.7, 149.9, 139.4, 135.2, 127.4, 122.6, 115.2, 113.5, 78.0, 77.7, 49.0, 46.9, 44.8, 44.4, 38.1, 37.7, 28.2; Anal. Calcd: C, 54.28; H, 5.75; N, 13.33;found: C, 54.43; H, 5.70; N, 13.43; TOF-MS m/z (%): 443.2 (M+Na^+, 100%).

For **1e**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.04 (s, 0.2H, −COOH) and 12.62 (s, 0.8H, −COOH), 11.67 (s, 1H, −NH), 7.93 (s, 1H, −phH), 7.82 (d, 1H, J=8.4 Hz, −phH), 7.46 (d, 1H, J=8.8 Hz, −phH), 6.93 (s, 0.8H, −COOH) and 6.74 (s, 0.2H, −COOH), 4.81 (s, 1.2H, $-CH_2$) and 4.70 (s, 0.8H, $-CH_2$), 4.28 (s, 0.8H, $-CH_2$) and 3.97 (s, 1.2H, $-CH_2$), 3.47 (t, 1.2H, J=6.0 and 6.4 Hz, $-CH_2$) and 3.32 (d, 0.8H, J=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, J=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, J=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.8, 170.5, 166.9, 166.5, 160.6, 155.7, 149.6, 138.4, 135.2, 126.6, 126.3, 117.5, 114.9, 78.0, 77.7, 48.9, 46.9, 44.7, 44.2, 38.1, 37.7, 28.2; Anal. Calcd: C, 50.17; H, 5.10; N, 12.32; found: C, 50.23; H, 5.00; N, 12.32; TOF-MS m/z (%): 467.3 (M+Na^+, 100%).

For **1f**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.07 (s, 0.2H, −COOH) and 12.64 (s, 0.8H, −COOH), 11.66 (d, 1H, −NH), 7.64 (m, 2H, J=8.8 and 8.4 Hz, −ph−H), 7.27 (q, 1H, J=4.0, 4.0, 4.4 Hz, −ph−H), 6.94 (s, 0.8H, −COOH) and 6.74 (s, 0.2H, −COOH), 4.81 (s, 1.0H, $-CH_2$) and 4.70 (s, 1.0H, $-CH_2$), 4.28 (s, 0.8H, $-CH_2$) and 3.97 (s, 1.2H, $-CH_2$), 3.47 (t, 1.2H, J=6.0 and 6.4 Hz, $-CH_2$) and 3.32 (d, 0.8H, J=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, J=6.4 Hz, $-CH_2$) and 3.02 (d, 0.8H, J=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 170.7, 170.6, 166.9, 166.6, 160.5, 155.7, 149.6, 138.5, 135.3, 126.6, 126.4, 117.6, 115.0, 78.1, 77.7, 49.0, 46.8, 44.7, 44.2, 38.1, 37.7, 28.2; Anal. Calcd: C, 52.05; H, 5.29; N, 12.78; found: C, 52.16; H, 5.22; N, 12.89; TOF-MS m/z (%): 461.3 (M+Na^+, 100%).

For **1g**: ^{1}H NMR (400 MHz, DMSO-d_6) δ: 13.06 (s, 0.2H, −COOH) and 12.63 (s, 0.8H, −COOH), 12.07 (d, 1H, −NH), 8.66 (d, 1H, J=4.8 Hz, −py−H), 8.31 (d, 1H, J=8.0 Hz, −py−H), 7.32 (q, 1H, J=5.6, 7.2, 6.8 Hz, −py−H), 6.96–6.70 (s, broad, 1H, −COOH), 4.81 (s, 1.0H, $-CH_2$) and 4.68 (s, 1.0H, $-CH_2$), 4.27 (s, 0.8H, $-CH_2$) and 3.96 (s, 1.2H, $-CH_2$), 3.49 (t, 1.2H, J=6.4 and 6.4 Hz, $-CH_2$), 3.32 (t, 0.8H, J=6.4 and 6.4 Hz, $-CH_2$), 3.20 (d, 1.2H, J=6.4 Hz, $-CH_2$), 3.02 (d, 0.8H, J=6.4 Hz, $-CH_2$), 1.37 (s, 9H, −Boc−H); ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 171.7, 171.5, 166.9, 166.3, 161.3, 155.7, 155.0, 151.0, 150.1, 136.9, 119.3, 109.0, 78.0, 77.6, 51.0, 48.1, 44.8, 44.3, 38.1, 37.7, 28.6; Anal. Calcd: C, 51.30; H, 5.50; N, 16.62; found: C, 51.13; H, 5.58; N, 16.42; TOF-MS m/z (%): 444.2 (M+Na^+, 100%).

Acknowledgements

This work was financially supported by NSFC (Nos. 20973182 and 21173233), the Chinese Academy of Sciences, Projects 973 (2011CB808400) and 863 (2009AA03Z323).

Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.tet.2012.08.028. These data include MOL files and InChIKeys of the most important compounds described in this article.

References and notes

1. Nielsen, P. E. *Peptide Nucleic Acids, Protocols and Applications*, 2nd ed.; Horizon Bioscience: Copenhagen, 2004.
2. (a) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Nordén, B. *Nature* **1994**, *368*, 561–563; (b) Böhler, C.; Nielsen, P. E.; Orgel, L. E. *Nature* **1995**, *376*, 578–581; (c) Petersson, B.; Nielsen, B. B.; Rasmussen, H.; Larsen, I. K.; Gajhede, M.; Nielsen, P. E.; Kastrup, J. S. *J. Am. Chem. Soc.* **2005**, *127*, 1424–1430.
3. (a) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566–568; (b) Baker, E. S.; Hong, J. W.; Gaylord, B. S.; Bazan, G. C.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 8484–8492.
4. (a) Bentin, T.; Nielsen, P. E. *J. Am. Chem. Soc.* **2003**, *125*, 6378–6379; (b) Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E. A.; Sprankle, K. G.; Griffey, R. H.; Kiely, J. S.; Freier, S. M. *J. Am. Chem. Soc.* **1995**, *117*, 831–832; (c) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, *42*, 13987–13995; (d) Kaihatsu, K.; Braasch, D. A.; Cansizoglu, A.; Corey, D. R. *Biochemistry* **2002**, *41*, 11118–11125; (e) Hansen, M. E.; Bentin, T.; Nielsen, P. E. *Nucleic Acids Res.* **2009**, *37*, 4498–4507; (f) Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, *132*, 8676–8681.
5. (a) Marin, V. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 8032–8033; (b) Petraccone, L.; Pagano, B.; Esposito, V.; Randazzo, A.; Piccialli, G.; Barone, G.; Mattia, C. A.; Giancola, C. *J. Am. Chem. Soc.* **2005**, *127*, 16215–16223.
6. (a) Veselkov, A. G.; Demidov, V. V.; Frank-Kamenetskii, M. D.; Nielsen, P. E. *Nature* **1996**, *379*, 214; (b) Shiraishi, T.; Pankratova, S.; Nielsen, P. E. *Chem. Biol.* **2005**, *12*, 923–929; (c) Tyagi, P.; Banerjee, R.; Basu, S.; Yoshimura, N.; Chancellor, M.; Huang, L. *Mol. Pharmacol.* **2006**, *3*, 398–406; (d) Janowski, B. A.; Kaihatsu, K.; Huffman, K. E.; Schwartz, J. C.; Ram, R.; Hardy, D.; Mendelson, C. R.; Corey, D. R. *Nat. Chem. Biol.* **2005**, *1*, 210–215; (e) Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996**, *24*, 494–500; (f) Pieck, J. C.; Kuch, D.; Grolle, F.; Linne, U.; Haas, C.; Ca, T. *J. Am. Chem. Soc.* **2006**, *128*, 1404–1405.
7. (a) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404–412; (b) Bendifallah, N.; Rasmussen, F. W.; Zachar, V.; Ebbesen, P.; Nielsen, P. E.; Koppelhus, U. *Bioconjugate Chem.* **2006**, *17*, 750–758; (c) Dragulescu-Andrasi, A.; Rapireddy, S.; He, G.; Bhattacharya, B.; Hyldig-Nielsen, J. J.; Zon, G.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16104–16112.
8. (a) Lusvarghi, S.; Murphy, C. T.; Roy, S.; Tanious, F. A.; Sacui, I.; Wilson, W. D.; Ly, D. H.; Armitage, B. A. *J. Am. Chem. Soc.* **2009**, *131*, 18415–18424; (b) Katritzky, A. R.; Narindoshvili, T. *Org. Biomol. Chem.* **2008**, *6*, 3171–3176; (c) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939–1942; (d) Dilek, I.; Madrid, M.; Singh, R.; Urrea, C. P.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 3339–3345; (e) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258–10267.
9. (a) Shibata, N.; Das, B. K.; Honjo, H.; Takeuchi, Y. *J. Chem. Soc., Perkin Trans. 1* **2001**, 1605–1611; (b) Slaitas, A.; Yeheskiely, E. *Eur. J. Org. Chem.* **2002**, 2391–2399; (c) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Am. Chem. Soc.* **2005**, *127*, 4144–4145; (d) Vilaivan, T.; Srisuwannaket, C. *Org. Lett.* **2006**, *8*, 1897–1900.
10. (a) Etson, S. R.; Mattson, R. J.; Sowell, J. W., Sr.. *J. Heterocycl. Chem.* **1979**, *16*, 929–933; (b) Takahashi, M.; Nagaoka, H.; Inoue, K. *J. Heterocycl. Chem.* **2004**, *41*, 525–530; (c) Powers, D. L.; Sowell, J. W., Sr.; Freeman, J. J.; Kosh, J. W. *J. Pharm. Sci.* **1980**, *69*, 473–475; (d) Sasaki, S.; Cho, N.; Nara, Y. *J. Med. Chem.* **2003**, *46*, 113–124.
11. (a) Abdel-Razik, H. H. *J. Chin. Chem. Soc.* **2005**, *52*, 141–148; (b) Nikpour, F.; Paibast, T.; Green, A. *Chem. Lett.* **2005**, *34*, 1438–1439; (c) Manfred, S.; Siegfried, J. *Monatsh. Chem.* **1987**, *118*, 71–80.
12. (a) Godde, F.; Toulmé, J.-J.; Moreau, S. *Biochemistry* **1998**, *37*, 13765–13775; (b) Mizuno, T.; Ishino, Y. *Tetrahedron* **2002**, *58*, 3155–3158; (c) Patil, Y. P.; Tambade, P. J.; Deshmukh, K. M.; Bhanage, B. M. *Catal. Today* **2009**, *133*, 201–208; (d) Michel, J.; Toulmé, J.-J.; Vercauteren, J.; Moreau, S. *Nucleic Acids Res.* **1996**, *24*, 1127–1135; (e) Skibo, E. B. *J. Org. Chem.* **1985**, *50*, 4861–4865; (f) Michel, J.; Gueguen, G.; Vercauteren, I. J.; Moreau, S. *Tetrahedron* **1997**, *53*, 8457–8478.
13. (a) Ausín, C.; Ortega, J.-A.; Robles, J.; Grandas, A.; Pedroso, E. *Org. Lett.* **2002**, *4*, 4073–4075; (b) Sanjayan, G. J.; Pedireddi, V. R.; Ganesh, K. N. *Org. Lett.* **2000**, *2*, 2825–2828.
14. Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767–5773.
15. Kospkina, L. W.; Wang, T. W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173–5176.
16. Yao, J. N.; Zhan, C. L.; Li, P. F. Chinese Patents, Application No. CN 201010224200.7, CN 201010224196.4, CN 201010224267.0, CN 201010224258.1.
17. (a) Zhang, X.; Pang, S. F.; Zhang, Z. G.; Ding, X. L.; Zhang, S. L.; He, S. G.; Zhan, C. L. *Tetrahedron Lett.* **2012**, *53*, 1094–1097; (b) Ke, D. M.; Zhan, C. L.; Li, X.; Li, A. D. Q.; Yao, J. N. *Synlett* **2009**, 1506–1510; (c) Ke, D. M.; Zhan, C. L.; Li, X.; Wang, Y. B.; Li, A. D. Q.; Yao, J. N. *Tetrahedron Lett.* **2009**, *50*, 3926–3928; (d) Ke, D. M.; Zhan, C. L.; Li, X.; Wang, Y. B.; Li, A. D. Q.; Yao, J. N. *Tetrahedron* **2009**, *65*, 8269–8276.

18. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian: Wallingford CT, 2004.
19. (a) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B: Condens. Matter* **1988**, *37*, 785–789; (b) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
20. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
21. Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. *NBO Version 3.1*; 1995.