Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 18 (2008) 2073–2077
Bioorganic & Medicinal Chemistry Letters

An asymmetric C8/C8′-tripyrrole-linked sequence-selective pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) dimer DNA interstrand cross-linking agent spanning 11 DNA base pairs

Arnaud C. Tiberghien,[a] David A. Evans,[a] Konstantinos Kiakos,[b] Christopher R. H. Martin,[a] John A. Hartley,[a,b] David E. Thurston[a,c,*] and Philip W. Howard[a,c]

[a]*Spirogen Ltd, 29/39 Brunswick Square, London WC1N 1AX, UK*
[b]*Cancer Research UK Drug–DNA Interactions Research Group, UCL Cancer Institute, Paul O'Gorman Building, Huntley Street, London WC1E 6BT, UK*
[c]*CR-UK Gene Targeting Drug Design Research Group, School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, UK*

Received 3 December 2007; revised 23 January 2008; accepted 24 January 2008
Available online 30 January 2008

Abstract—A novel sequence-selective extended PBD dimer **4** has been synthesized that binds with high affinity to an interstrand cross-linking site spanning 11 DNA base pairs. Despite its molecular weight (984.07) and length, the molecule has significant DNA interstrand cross-linking potency (~100-fold greater than the clinically used agent melphalan) and sub-micromolar cytotoxicity in a number of tumour cell lines, suggesting that it readily penetrates cellular and nuclear membranes to reach its DNA target.

There is growing interest in small molecules that recognize DNA sequence,[1] and examples are known that interact with DNA non-covalently (e.g., the hairpin polyamides[2]) or covalently through monoalkylation (e.g., the pyrrolobenzodiazepines[3]) or cross-linking (e.g., bizelesin[4]) mechanisms. SJG-136 (**1**, Fig. 1) is a potent, sequence-selective interstrand DNA cross-linking agent containing pyrrolobenzodiazepine units that is presently in Phase I evaluation in the clinic.[5–7] The propensity of SJG-136 to preferentially target 5′-Pu-GATC-Py DNA sequences is thought to contribute to its antitumour activity.[8] In 2001, Bando and co-workers reported the structure of an efficient CPI-based cross-linking agent (**2**), the polypyrrole-imidazole core of which allows sequence-specific cross-linking.[9] More recently, syntheses of the first examples of heterocyclic-linked PBD dimers have been independently reported; Kumar and Lown have described molecules of type **3**, although DNA cross-linking and sequence-selectivity data were not reported,[10] and similar extended PBD dimers have been reported in the patent literature by our group.[11]

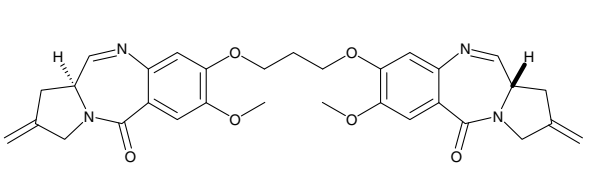
In order to systematically study the effect of length and composition of the polyheterocyclic linker on DNA sequence-selectivity, binding affinity and cross-linking efficiency of C8/C8′-linked PBD dimers of type **3**, we have synthesized different families of dimers containing a variety of heterocyclic linkers and with different regiochemical arrangements of constituent components. These heterocyclic units are known from the work of Dervan and co-workers to recognize sequences in the minor groove.[2]

We report here the synthesis and evaluation of the extended asymmetrically-linked tripyrrole-containing PBD dimer **4** (AT-235) which spans 11 DNA base pairs with a sequence-selectivity that can be rationalized based on both its covalent and non-covalent interactions with DNA base pairs in the minor groove.

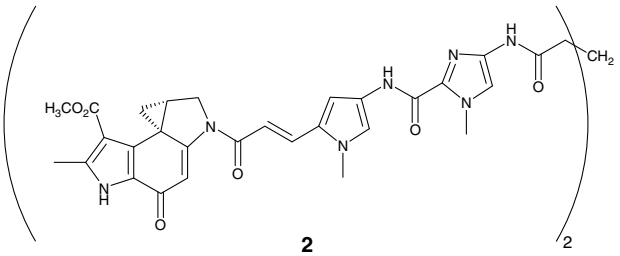
The synthesis of **4** was achieved through sequential amide coupling between the three main components **5**, **6** (to give **7**) and **8** as shown in Scheme 1.

Keywords: DNA-binding; Sequence-selective; Interstrand; Cross-linking; Anticancer agent; Pyrrolobenzodiazepine; PBD; PBD dimer.
* Corresponding author. Tel.: +44 2077535932; fax: +44 2077535964; e-mail: david.thurston@pharmacy.ac.uk

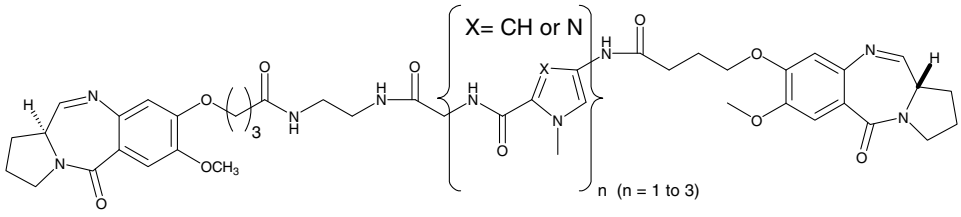
0960-894X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2008.01.096



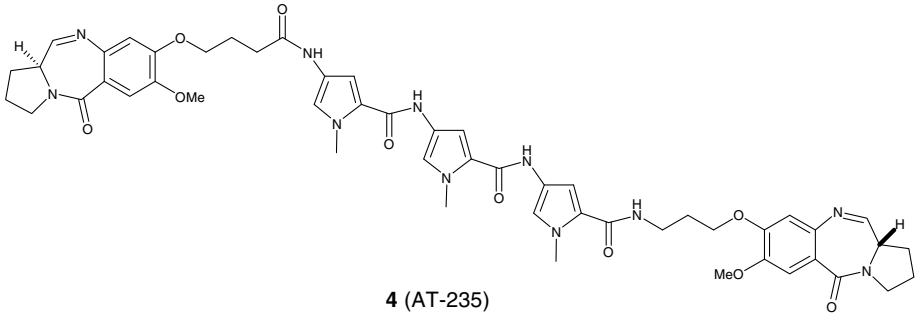
1 (SJG-136)



2

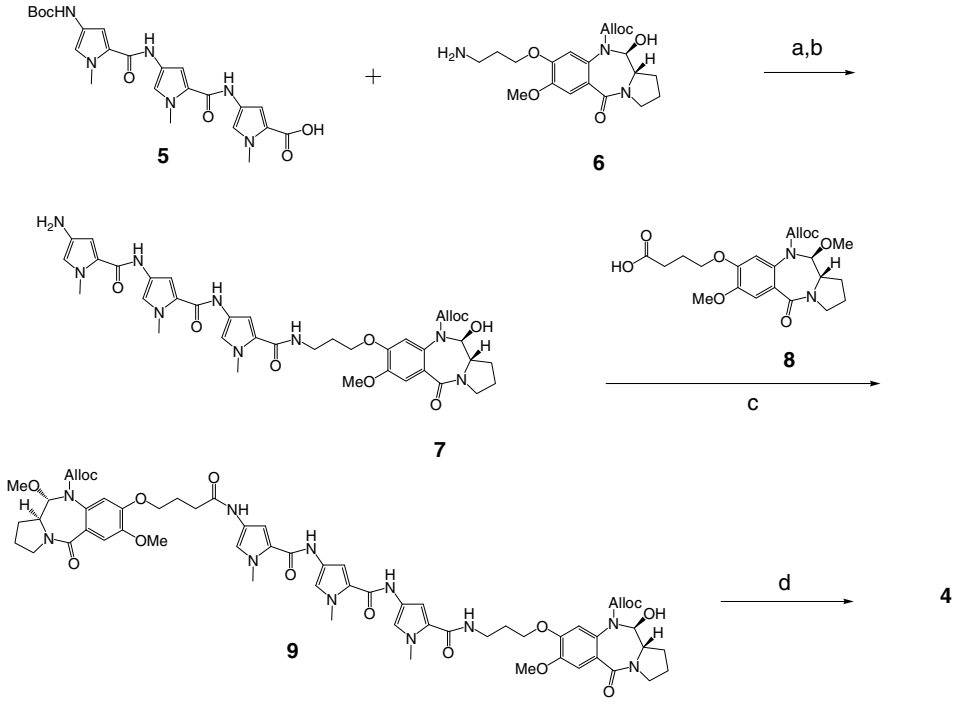


3



4 (AT-235)

Figure 1. Structures of the DNA interstrand cross-linking agents **1** (SJG-136), the CPI-based dimer **2**, the previously reported PBD dimer **3** and the new asymmetric PBD dimer **4** (AT-235).



Scheme 1. Reagents and condition: (a) EDCI, HOBt, DMF, 71%; (b) 4 N HCl in dioxane, quant; (c) oxalyl chloride, THF, DMF cat, then **7** and DIPEA, 44%; (d) $Pd(PPh_3)_4$, pyrrolidine, $CHCl_3$, 85%.

The known Boc-tripyrrole acid **5** was prepared by an improvement of a method of Boger and co-workers[12] that required no complex work-up or purification procedures (Scheme 2).

The novel chirally pure C8-aminopropyl-N10-alloc-protected PBD **6** was synthesized in 7 steps from the previously reported 4-hydroxy-5-methoxy-2-nitrobenzaldehyde starting material (**14**)[13] (Scheme 3). The aliphatic Boc amino side chain was introduced by Mitsunobu phenol etherification and the aldehyde oxidized with potassium permanganate to yield the nitro acid **15**. Coupling to *S*-pyrrolidinemethanol gave the nitro alcohol **16** which was hydrogenated followed by alloc protection of the resulting amine to give alcohol **17**. Exposure to TEMPO/BAIB[14] led to smooth oxidative ring closure to give **18**. The best conditions for the final Boc deprotection were found to be TFA/DCM/water 47/47/6 which afforded the capping unit **6** in an overall yield of 17% over 7 steps. This intermediate required only one chromatographic purification and was chirally pure. Further protection of the C11-hydroxy functionality before the next coupling step was found to be unnecessary due to the superior nucleophilicity of the amino functionality. The C11-methyl ether (**8**) of the PBD acid capping unit described by Wells and co-workers[15] was employed in order to improve coupling efficiency and avoid side-reactions. Finally, cascade removal of all protective groups of **9** under Deziel conditions[16] yielded the extended PBD dimer **4**[17] in 85% yield.

PBD dimer **4** was shown to effectively interstrand cross-link linear plasmid pUC18 DNA with an XL_{50} value of 0.23 μM (Fig. 2 and Table 1),[18] a potency 4-fold less than SJG-136 (0.06 μM) but approximately 100-fold greater than the clinically used nitrogen mustard melphalan (20.0 μM). As the molecule is designed to span

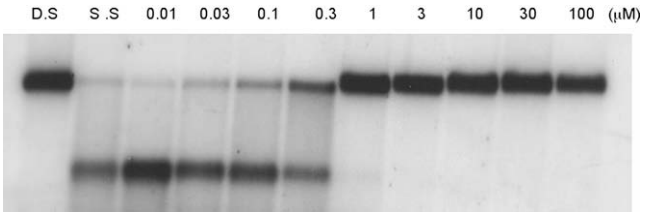


Figure 2. Cross-linking gel[18] for extended PBD dimer **4** (AT-235) using linearized pUC18 DNA. Controls: DS, double stranded DNA; SS, single stranded DNA. The XL_{50} was determined to be 0.23 μM.

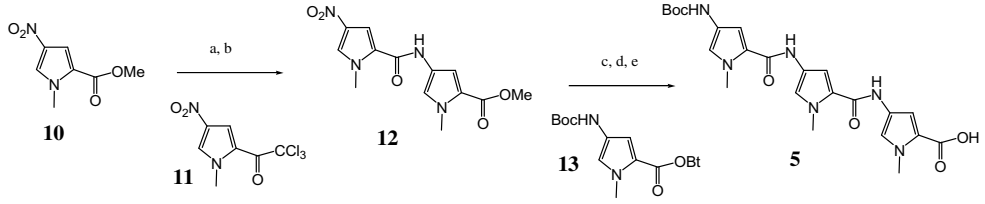
Table 1. Average cytotoxicity values from the NCI 60-cell panel and results from a K562 line, and DNA cross-linking data for **1** (SJG-136) and extended dimer **4**

Compound	NCI (μM)			K562 GI_{50}[b] (μM)	XL_{50}[a] (μM)
	GI_{50}[b]	TGI[c]	LC_{50}[d]		
1	0.007	0.087	0.562	0.008	0.060
4	0.014	0.49	22.9	0.037	0.23

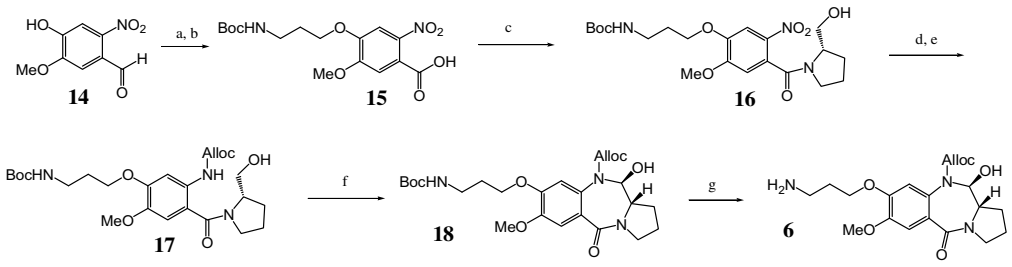
[a] XL_{50}: Dose providing 50% cross-linking of DS pUC18 plasmid DNA.[18]
[b] GI_{50}: Dose inhibiting 50% cell growth.
[c] TGI: Dose inhibiting 100% cell growth.
[d] LC_{50}: Dose killing 50% of cells.

11 DNA base pairs with a degree of selectivity within this span, the probability of the perfectly matched site(s) appearing within the pUC18 sequence is lower than that for SJG-136 which spans 6 bp, and this could be one possible explanation for the lower observed potency of **4** in this assay compared to **1**.

PBD dimer **4** was also shown to have significant cytotoxicity in both the NCI 60-cell line panel and in the K562 leukaemia cell line (Table 1).



Scheme 2. Reagents and conditions: (a) 10% Pd/C, H_2, DMF, quant; (b) **11**, DMF, 50%; (c) 10% Pd/C, H_2, DMF, quant; (d) **13**, DMF, 94%; (e) NaOH, MeOH, water, quant.



Scheme 3. Reagents and conditions: (a) PPh_3, DEAD, (3-hydroxy-propyl)carbamic acid *tert*-butyl ester, THF; (b) $KMnO_4$, water, acetone, 32% over 2 steps; (c) *S*-pyrrolidinemethanol, EDCI, HOBt, DMF, 93%; (d) 10% Pd/C, H_2, EtOAc, quant; (e) allyl chloroformate, pyridine, DCM, 86%; (f) TEMPO, BAIB, DCM, 72%; (g) TFA, DCM, water (47/47/6), quant.

The sub-micromolar GI_{50} and TGI cytotoxicity values suggest that, despite its molecular weight (984.07) and length (spanning 11 bp), **4** is still capable of traversing cellular and nuclear membranes to reach the genome of these cells.

DNA footprinting[8] of **4** on a 543 bp fragment of BCL-2 DNA[19] showed that it binds with high affinity (i.e., 10 nM) to a 5′-G**C**TTATAAT**G**G-3′ sequence (Fig. 3). This may be rationalized through simulation studies (Table 2) involving **4** bound to a range of potential binding sites which predict a preference for interstrand rather than intrastrand cross-linking with the covalently modified guanines on opposite strands separated by 7 bp (i.e., sequence B). The predicted binding energy at the main footprint site ($\Delta\Delta E_{bind}$ = 42.6. kcal mol^{-1}) was higher than those values calculated for the hypothetical cross-linking motifs shown in Table 2. However, as sequences A, B and C are not present in the footprinted BCL-2 fragment, it is possible that higher-affinity binding sites may exist in longer DNA sequences. Figure 4 shows a molecular model in which **4** is bound to the footprinted 5′-G**C**TTATAAT**G**G-3′ sequence. The two PBD moieties are bound to the 3′-C**G**A (opposite 5′-GCT) and 5′-T**G**G triplet sequences (covalent binding

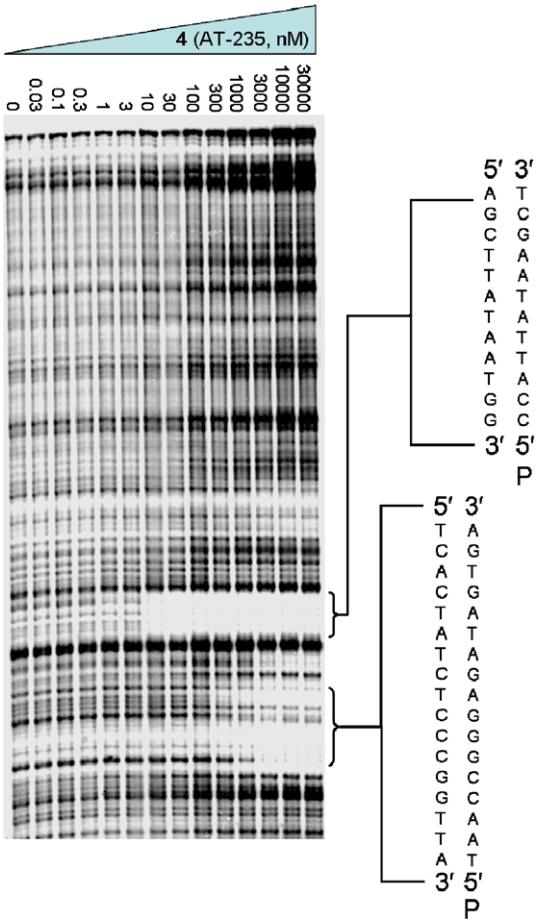


Figure 3. DNA footprinting gel of **4** bound to a 543 bp fragment of BCL-2 DNA[19] showing a well-defined footprint at a 5′-AG**C**TTA TAAT**G**G-3′ site at a concentration of 10 nM or above after 18 h incubation.

Table 2. Energy calculations using duplex DNA sequences **A**–**F** to predict preferences for interstrand (**A**–**C**) versus intrastrand (**D**–**F**) cross-linking and the optimal number of base pairs (6–8) separating covalently modified guanines

	Simulated binding sequence (5′ → 3′)	$\Delta\Delta E_{bind}$[a]/kcal mol^{-1}
A	CGCA**G**AAATTT**C**TGCG	13.6 ± 0.6
B	CGCA**G**AAAATTT**C**TGCG	0
C	CGCA**G**AAAATTTT**C**TGCG	7.4 ± 0.6
D	CGCA**G**AAATTT**G**TGCG	10.1 ± 0.6
E	CGCA**G**AAAATTT**G**TGCG	17.6 ± 0.6
F	CGCA**G**AAAATTTT**G**TGCG	23.5 ± 0.6

The results predict that sequence **B** (5′-A**G**AAAATTT**C**T-3′) has the lowest binding energy, suggesting that **4** is more likely to form an interstrand rather than an intrastrand cross-link with a spacing of 7 bp between the covalently-linked guanine residues.

[a] $\Delta\Delta E_{bind}$ is the MM-PBSA[22,23] binding energy relative to the lowest energy binding sequence which was **B** (thus appearing as 0 kcal mol^{-1}). The error estimate is the standard error in the mean over all included trajectory snapshots.

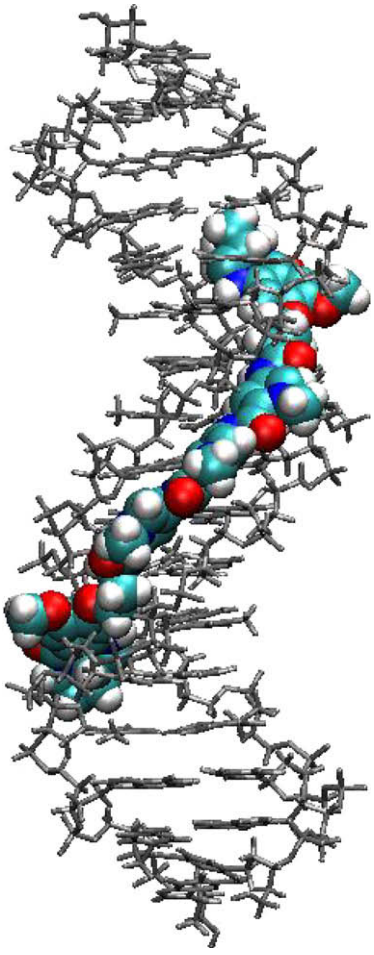


Figure 4. Representative snapshot from a molecular dynamics simulation (Sander program in Amber8[20]) showing **4** bound to a 5′-G**C**TTATAAT**G**G-3′ sequence with the two PBD units covalently bound to guanines on opposite strands (positions underlined).

base pair positions underlined) with the heterocyclic linker interacting with the central 5′-TATAA base pairs as predicted by the literature.[2,15] According to this model,

the dimer takes up the curvature of the DNA minor groove with little distortion of the helix.

In Figure 3 an additional weaker footprint was observed at a 5′-TCACTATCTCCCGGTTA-3′ sequence at a concentration of approximately 1 μM. This sequence does not contain a predicted cross-linking site for **4**, but does contain potential sites of monoalkylation for one PBD unit such as 3′-GGG, a preferred monomeric PBD binding sequence.[21] This could explain the lower preference (by 2 orders of magnitude) of **4** for this site.

Through a greater understanding of SAR for extended PBD dimers of this type, we hope to provide the basis of a novel approach to DNA sequence recognition through an interstrand cross-linking mechanism. This is being further explored by the synthesis and evaluation of analogues of **4** with differently-structured linkers which will be reported in due course.

Acknowledgments

Marissa Coffils and Dr. Robert Schultz are thanked for providing the K562 and NCI 60-cell line data, respectively, and Laura Calderon is acknowledged for initially synthesizing the PBD intermediate **6**.

References and notes

1. Thurston, D. E. *Chemistry and Pharmacology of Anticancer Drugs*; CRC Press (Taylor & Francis): Boca Raton, Florida, USA, 2006.
2. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111.
3. Thurston, D. E. *Br. J. Cancer* **1999**, *80*, 65.
4. Schwartz, G. H.; Patnaik, A.; Hammond, L. A.; Rizzo, J.; Berg, K.; Von Hoff, D. D.; Rowinsky, E. K. *Ann. Oncol.* **2003**, *14*, 775.
5. Alley, M. C.; Hollingshead, M. G.; Pacula-Cox, C. M.; Waud, W. R.; Hartley, J. A.; Howard, P. W.; Gregson, S. J.; Thurston, D. E.; Sausville, E. A. *Cancer Res.* **2004**, *64*, 6700.
6. Gregson, S. J.; Howard, P. W.; Hartley, J. A.; Brooks, N. A.; Adams, L. J.; Jenkins, T. C.; Kelland, L. R.; Thurston, D. E. *J. Med. Chem.* **2001**, *44*, 737.
7. Hartley, J. A.; Spanswick, V. J.; Brooks, N.; Clingen, P. H.; McHugh, P. J.; Hochhauser, D.; Pedley, R. B.; Kelland, L. R.; Alley, M. C.; Schultz, R.; Hollingshead, M. G.; Schweikart, K. M.; Tomaszewski, J. E.; Sausville, E. A.; Gregson, S. J.; Howard, P. W.; Thurston, D. E. *Cancer Res.* **2004**, *64*, 6693.
8. Martin, C.; Ellis, T.; McGurk, C. J.; Jenkins, T. C.; Hartley, J. A.; Waring, M. J.; Thurston, D. E. *Biochemistry* **2005**, *44*, 4135.
9. Bando, T.; Iida, H.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.* **2001**, *123*, 5158.
10. Kumar, R.; Lown, J. W. *Eur. J. Med. Chem.* **2005**, *40*, 641.
11. Howard, P. W.; Gregson, S. J.; Tiberghien, A.C. WO 2005/085250 A1 Sep 15th, 2005.
12. Boger, D. L.; Fink, B. E.; Hedrick, M. P. *J. Am. Chem. Soc.* **2000**, *122*, 6382.
13. Fukuyama, T.; Lin, S. C.; Li, L. P. *J. Am. Chem. Soc.* **1990**, *112*, 7050.
14. DeMico, A.; Margarita, R.; Parlanti, L.; Vescovi, A.; Piancatelli, G. *J. Org. Chem.* **1997**, *62*, 6974.
15. Wells, G.; Martin, C. R. H.; Howard, P. W.; Sands, Z. A.; Laughton, C. A.; Tiberghien, A.; Woo, C. K.; Masterson, L. A.; Stephenson, M. J.; Hartley, J. A.; Jenkins, T. C.; Shnyder, S. D.; Loadman, P. M.; Waring, M. J.; Thurston, D. E. *J. Med. Chem.* **2006**, *49*, 5442.
16. Deziel, R. *Tetrahedron Lett.* **1987**, *28*, 4371.
17. Data for **4**. ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.90 (s, 3H), 8.10 (m, 1H), 7.79 (d, 2H, J = 4.36 Hz), 7.35 (s, 2H), 7.25 (s, 1H), 7.20 (2s, 2H), 7.05 (s, 1H), 6.90 (s, 2H), 6.84 (s, 2H), 4.21–4.01 (m, 4H), 3.85–3.81 (m, 15H), 3.61 (m, 2H), 3.46–3.35 (m, 6H), 2.45 (m, 2H), 2.38–2.17 (m, 4H), 2.13–1.89 (m, 8H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 168.7, 164.2, 158.3, 150.2, 150.1, 146.8, 140.5, 122.7, 122.0, 121.9, 119.7, 118.4, 118.4, 117.8, 114.8, 111.2, 110.0, 104.6, 104.2, 103.9, 67.7, 66.4, 55.8, 55.5, 53.6, 53.3, 48.5, 46.3, 36.0, 35.8, 35.5, 30.2, 28.9, 28.9, 28.7, 24.7, 23.6, 22.3; IR (cm^{-1}): ν 3306, 2949, 1638, 1597, 1555, 1508, 1465, 1435, 1405, 1262, 1216, 1090, 1064. MS (ES^+) *m/z* (relative intensity): 984.4 ($[M+H]^{+\cdot}$, 100); HRMS (TOF MS ES^+): Calcd for $C_{51}H_{57}N_{11}O_{10}$ ($[M+H]^+$): 984.4363. Found: 984.4359.
18. Hartley, J. A.; Berardini, M. D.; Souhami, R. L. *Anal. Biochem.* **1991**, *193*, 131.
19. Accession No. NC_000018.8 (Chromosome 18; DNA Bases 59136813–59136823).
20. Amber Molecular Dynamics Package: http://amber.scripps.edu.
21. Thurston, D. E.. In *Molecular Aspects of Anticancer Drug–DNA Interactions*; Neidle, S., Waring, M. J., Eds.; The Macmillan Press Ltd.: London, UK, 1993; Vol. 1, p 54.
22. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889.
23. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401.