Synthesis of a novel GC-specific covalent-binding DNA affinity-cleavage agent based on pyrrolobenzodiazepines (PBDs)

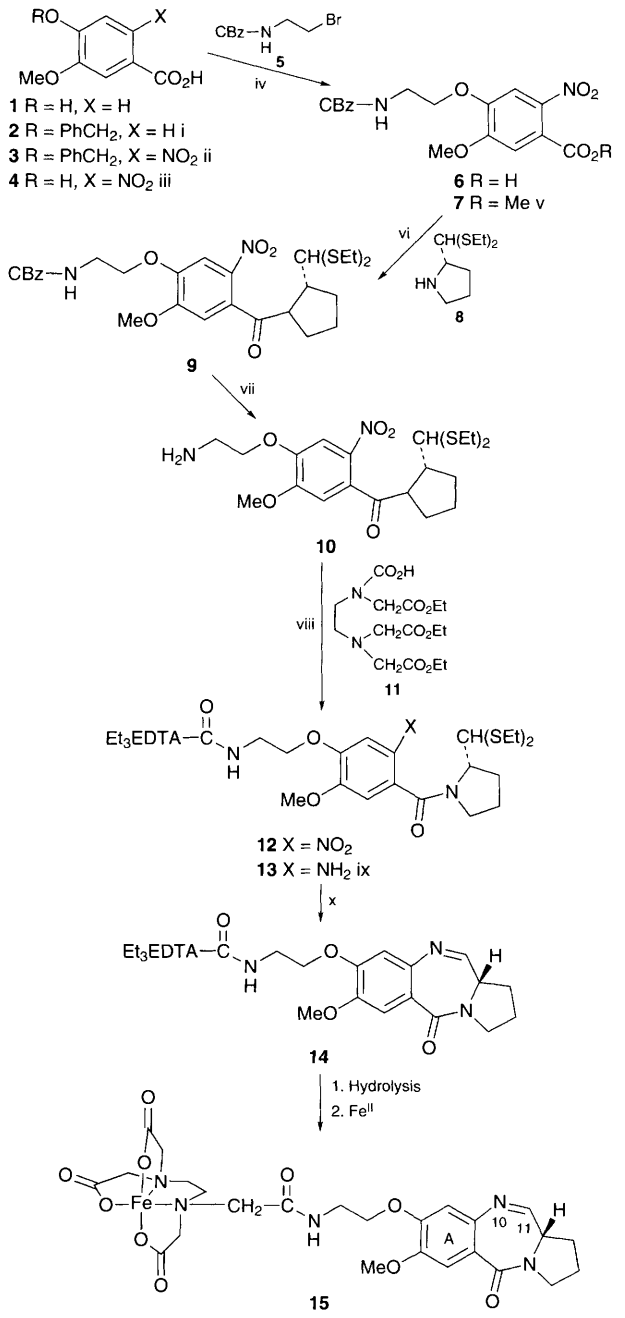
David E. Thurston,*[a] Steven J. Morris[a] and John A. Hartley[b]

[a] *School of Pharmacy and Biomedical Science, University of Portsmouth, King Henry I Street, Portsmouth, UK PO1 2DZ*
[b] *CRC Drug-DNA Interactions Group, Department of Oncology, University College London Medical School, 91 Riding House Street, London, UK W1P 8BT*

A novel GC-selective pyrrolobenzodiazepine–EDTA conjugate is synthesized that covalently binds to DNA at 5′-PuGPu sequences leading to site-specific cleavage.

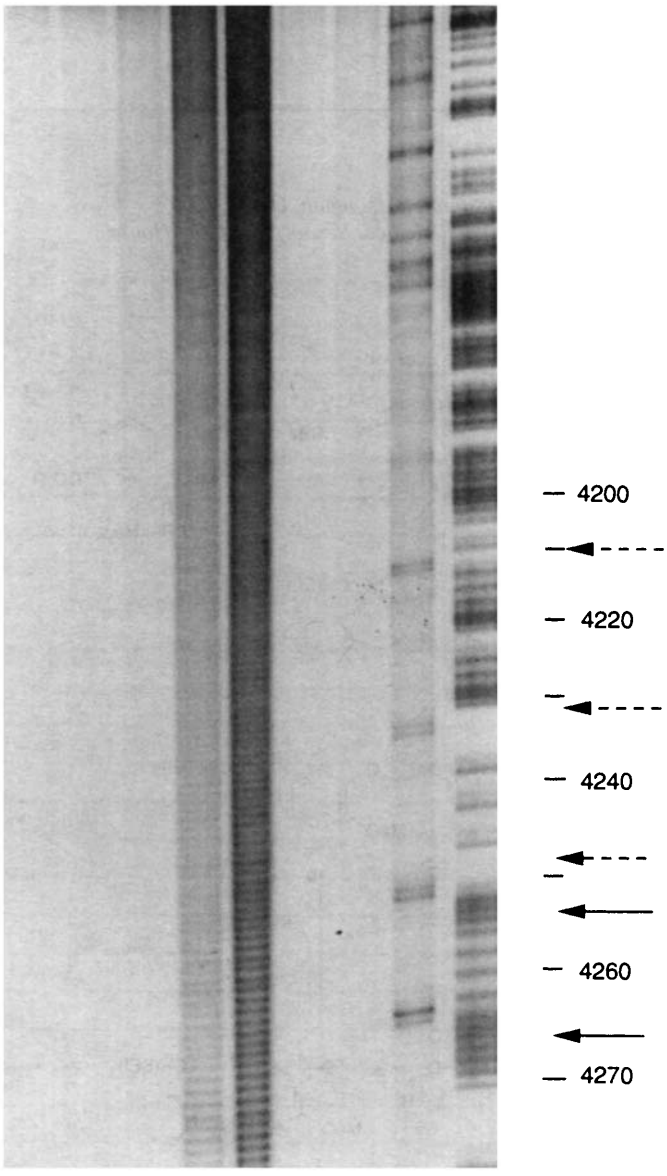
There is presently interest in developing simple low molecular weight molecules with sequence-selective DNA-interactive properties as tools for molecular biology and as possible therapeutic agents to inactivate particular genes.[1] One approach to this involves the synthesis of molecules with predictable affinity–cleavage properties that might function as artificial restriction enzymes. Efforts in this area have so far included, for example, the attachment of cleaving moieties such as ethylenediamine tetraacetic acid (EDTA) or metal-chelating porphyrins to intercalating molecules, non-covalent minor-groove binders, triple-helix forming oligonucleotides and DNA-binding proteins.[2] With the exception of compounds such as dynemicin A[3] and bleomycin,[4] most of the known low molecular weight DNA-cleaving agents have a sequence-selectivity for AT-rich sites. However, as GC-rich sequences appear to be the major site of action of a number of clinically-useful antitumour drugs such as the nitrogen mustards, mitomycin and many intercalating agents, the development of GC-specific affinity-cleavage agents is of interest.[1] We report here, attachment of an EDTA moiety to DC-81,[5] a member of the guanine(N2)-specific pyrrolobenzodiazepine (PBD) family of antitumour antibiotics,[6] to produce the first example of a covalent-binding GC-specific DNA-cleaving agent with a selectivity for 5′-PuGPu sequences (Pu = purine; G = guanine).

Factors considered in the design of this molecule included attachment of the EDTA moiety to the PBD skeleton prior to formation of the relatively unstable electrophilic N10–C11 imine, and at a position that would not compromise DNA-binding.[6] The natural product DC-81 was chosen as the PBD nucleus as it is well characterized, relatively cytotoxic and has a clear sequence-selectivity for PuGPu motifs according to footprinting.[5–7] The convergent synthesis shown in Scheme 1 was based on the 'thioacetal' route previously developed in this laboratory.[8,9] It started with *O*-benzylation (**2**, 76%), nitration (**3**, 70%) and then debenzylation of vanillic acid **1** to afford 4-hydroxy-5-methoxy-2-nitrobenzoic acid **4** in yields of >90%. The linker component, *N*-(benzyloxycarbonyl)-2-bromoethylamine **5**, prepared by treatment of 2-bromoethylamine with benzylchloroformate, was then attached to **4** through an ether linkage using Me_2SO–aq.NaOH at 80 °C. The crude product, formed in yields of up to 65%, was converted into the methyl ester **7** (>90%), purified by chromatography, and then hydrolysed (aq. NaOH, THF) to afford pure 4-[*N*-(benzoxycarbonyl)-2-aminoethoxy]-5-methoxy-2-nitrobenzoic acid **6**. This was linked to (2*S*)-pyrrolidine-2-carboxaldehyde diethyl thioacetal **8**, prepared in six steps from L-proline,[9] to provide the amide **9** in 40% yield. An alternative approach, which involved the direct coupling of crude **6** to **8**, followed by chromatographic purification, gave a marginally improved yield over the three steps (**6** → **9**). Deprotection of **9** with Me_3SiI (**10**, 60%) was followed by



Scheme 1 *Reagents and conditions*: i, $PhCH_2Cl$, THF, H_2O, NaOH, 76%; ii, $SnCl_4$, HNO_3, CH_2Cl_2, −20 °C, 70%; iii, C_2H_5SH, $BF_3.OEt_2$, room temp., 75% or HBr–AcOH 30%, >90%; iv, Me_2SO, aq. NaOH, 80 °C, 3 h, 65%; v, $COCl_2$, THF, MeOH, >90%; vi, THF, $COCl_2$, Et_3N, H_2O, 40%; vii, Me_3SiI, CH_2Cl_2, 60%; viii, carbonyldiimidazole (CDI), DMF, 84%; ix, H_2, 10% Pd–C, MeOH; x, $HgCl_2$, $CaCO_3$, MeCN : H_2O (4 : 1).

(*a*)



(*b*)

Fig. 1 (*a*) Autoradiograph of approximately the first two hundred base pairs of an acrylamide gel showing the affinity cleavage of a linear 5′-^{32}P-end-labelled 4330 base-pair restriction fragment of pBR322 DNA after incubation with Et_3EDTA–DC-81 **14** for 24 h at room temp. (so as to allow hydrolysis of the EDTA esters, and sufficient time for the drug to bind to DNA), followed by the addition of $Fe(NH_4)_2(SO_4)_2$ (to give an equimolar drug : Fe ratio) and then DTT to initiate cleavage. The lanes are: a, control; b–e = 1, 10, 100, 1000 μmol dm^{-3} of EDTA; f–h = 1, 10, 100 μmol dm^{-3} Et_3EDTA–DC-81 **14**; i = purine marker lane. The arrows indicate PuGPu sites either on the labelled strand (solid arrows) or opposite strand (dotted arrows). The numbers correspond to the sequence of pBR322. (*b*) DNA sequence of part of the resolved portion of the gel shown in (*a*). The arrows indicate the observed cleavage sites.

attachment of the EDTA triester[2a] **11** through its free carboxylic group using CDI to afford the nitro amide **12** in yields of up to 84%. Reduction of **12** was achieved through catalytic hydrogenation to provide the amine **13** which was immediately cyclized using $HgCl_2$–$CaCO_3$,[8,9] and the resulting crude product purified by chromatography to afford Et_3EDTA–DC-81 **14** as a viscous yellow oil (12% yield from **12** → **14**). The structure of **14** was confirmed by ^{1}H and ^{13}C NMR, and mass spectrometry.† In particular, the diagnostic PBD imine (H11) signal could be observed as a doublet in the ^{1}H NMR spectrum at δ 7.66 (*J* 4.4 Hz).[8] Treatment of **14** with esterase (Porcine liver: EC 3.1.1.1) in Tris buffer (40 mmol dm^{-3} Tris-HCl at pH 7.8/5 mmol dm^{-3} NaOAc) and Me_2SO (4% *v/v* to aid solubility) for 5 min lead to hydrolysis of the EDTA esters as judged by UV spectroscopic (loss of intense ester absorption bands at *ca.* 236 nm) and TLC (MeOH–$CHCl_3$, 1 : 3, loss of higher R_f fluorescent component) studies in comparison with a control sample which excluded esterase.[10]

The DNA cleavage potential of **14** was initially assessed by monitoring[11] the conversion of supercoiled pBR322 plasmid DNA (form I) into the open circular (form II) and linear forms (form III) by electrophoresis (data not shown). Et_3EDTA–DC-81 **14** was incubated with esterase, complexed in a 1 : 1 ratio with Fe^{II} to form **15**, and then added to supercoiled pBR322. Dithiothreitol (DTT) was added as reducing agent, and the mixture incubated at 37 °C for 1 h. Both single- and double-stranded cleavage of DNA was observed at Et_3EDTA–DC-81 concentrations down to 2.5 μmol dm^{-3} and, at higher concentrations, the extent of double-stranded cleavage could be increased with incubation time. Cleavage was enhanced 10-fold by pre-incubation of esterase-treated **14** with DNA for 8 h prior to the addition of Fe^{II} and DTT, consistent with the usual time-dependent covalent-binding process associated with the PBDs.[6] No cleavage was observed in the absence of either Fe^{II} or DTT, and no double-stranded cleavage was observed for EDTA alone at this concentration. Interestingly, some cleavage was also observed in the absence of esterase suggesting that either spontaneous hydrolysis of one or more of the EDTA esters had occurred under the conditions of the experiment, or that Et_3EDTA–DC-81 itself is capable of chelating iron and then cleaving DNA. Further UV and TLC studies in Tris buffer–Me_2SO (*vide supra*) indicated that hydrolysis of **14** can occur in the absence of esterase but over a longer time period (*e.g.* loss of 20% of **14** in 40 min compared to complete hydrolysis in < 5 min in the presence of esterase). This spontaneous hydrolysis may be associated with the buffer composition and pH.

In order to study the sequence-selectivity of **14**, it was incubated for 24 h at room temperature with a 4330 base-pair 5′-^{32}P-end-labelled restriction fragment of pBR322 DNA, followed by the addition of $Fe(NH_4)_2(SO_4)_2$ (to provide an equimolar drug : Fe ratio) and then DTT to initiate cleavage. Analysis by gel electrophoresis allowed the first two to three hundred base pairs to be accurately resolved to the nucleotide level [Figs. 1(*a*), (*b*)]. A range of concentrations of both **14** (1–100 μmol dm^{-3}) and EDTA (1–1000 μmol dm^{-3}) were studied, and a sample of formic acid treated (depurinated) DNA was included as a 'sequence-marker'. In contrast to the EDTA lanes (b–e), non-random cleavage patterns were observed at highly localised sites at an (EDTA–DC-81)Fe^{II} concentration of 100 μmol dm^{-3} (lane h), indicative of sequence-specific cleavage. Ten cleavage sites are visible in the first 200 base-

pairs, all adjacent to PuGPu sequences, consistent with the known sequence-selectivity of the PBDs.[6,12] Detailed analysis of the cleavage sites by laser densitometry suggested that the results are best explained by two major modes of binding/cleaving for (EDTA–DC-81)Fe^{II} (Fig. 2). In one mode, the molecule is attached to the 5′-labelled strand with the A-ring orientated towards the 3′-end [Fig. 2(*a*)]. In this case, cleavage appears to occur in the 3′-direction (to different extents) at the 3rd and 4th base pairs from the covalently-modified guanine, resulting in a fragment of lower electrophoretic mobility. In the other mode [Fig. 2(*b*)], the PBD is bound to the unlabelled strand in the same orientation; the cleavage sites are now observed on the labelled top strand at the 1st and 2nd base pairs 3′- from the covalently-modified guanine, resulting in a fragment of higher electrophoretic mobility. This unsymmetrical cleavage pattern, and the fact that two adjacent base pairs appear to be cleaved to different extents, are most likely due to the helical structure of right-handed DNA and the diffusability of the OH· radical, respectively.[2a] These results are generally consistent with the preferred orientation of binding of the PBDs as indicated by fluorimetry, molecular modelling and NMR studies.[6]

In summary, it has been possible to design and synthesise a sequence-specific affinity-cleavage agent, (EDTA–DC-81)Fe^{II}, that spans approximately 4–5 base pairs and cleaves specifically at sites adjacent to 5′-PuGPu sequences in the 3′-direction. This approach can provide useful information about the mode of binding of PBDs and, in this case, suggests that DC-81 prefers to orientate with its A-ring towards the 3′-end of the covalently-modified strand. Having recently extended the length of DNA recognition of PBD-type molecules by synthesizing interstrand cross-linking dimers (*e.g.* DSB-120) capable of binding and spanning up to six or seven base pairs with a selectivity for 5′-PuGATCPy or 5′-PyGATCPu sequences (Py = pyrimidine),[13] the basis now exists to attach a similar EDTA-linker moiety to PBD dimers in order to effect site-specific affinity-cleavage on spans of 6 base-pairs or more. Such molecules should approach the level of usefulness of many restriction enzymes.

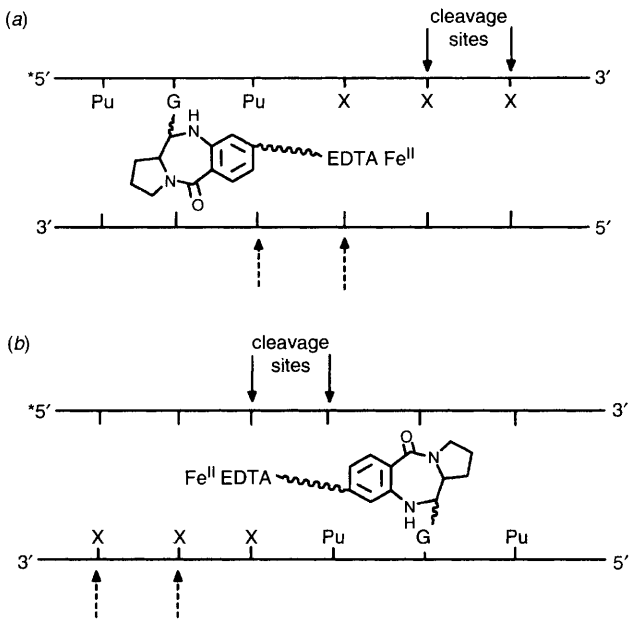


Fig. 2 Two possible models that explain the observed cleavage pattern on the 5′-end-labelled strand upon covalent-binding of (EDTA–DC-81)Fe^{II} to pBR322 DNA: (*a*) PBD bound to labelled (top) strand with 3′-orientation of the A-ring; (*b*) PBD bound to unlabelled (bottom) strand with 3′-orientation of the A-ring. The solid arrows indicate cleavage sites detected on the labelled strand, and the dotted arrows indicate predicted cleavage sites on the opposite unlabelled strand.

The authors are indebted to Professor L. H. Hurley (University of Texas at Austin, USA) for initiating this work, and also thank Dr Eric Benghiat for carrying out some preliminary synthetic studies. The PCFC (UK) and the Cancer Research Campaign (UK) are gratefully acknowledged for their financial support.

Footnote

† *Spectroscopic data* for **14**: ^1H NMR ($CDCl_3$): δ_H 1.21–1.26 (m, 9 H), 2.06–2.10 (m, 2 H), 2.28–2.36 (m, 2 H), 2.85 (br s, 4 H), 3.43–3.87 (m, 7 H; inc. H11a), 3.43 (s, 2 H), 3.48 (s, 2H), 3.59 (s, 4 H), 3.92 (s, 3 H), 4.09–4.17 (m, 6 H), 6.85 (s, 1H; H6 or H9), 7.50 (s, 1 H, H6 or H9) and 7.66 (d, 1 H, *J* 4.4 Hz, H11); ^{13}C NMR ($CDCl_3$): δ_C 13.2, 23.2, 28.6, 37.5, 45.6, 50.8, 51.5, 52.7, 53.6, 54.2, 54.9, 55.0, 57.1, 59.5, 59.6, 66.5, 109.8, 110.5, 119.1, 139.7, 149.4, 161.4, 163.5 and 170.1; MS (EI) *m/z* (relative intensity) 647 (M^+, 0.4%), 419 (6), 401 (8), 346 (16), 331 (21), 272 (14), 245 (12), 217 (44), 202 (100), 200 (20), 171 (54), 130 (53) and 116 (78); MS (FAB) (VG analytical model ZAB-E Mass spectrometer): = 648 (MH^+); MS (high resolution): Calc. for 648.3245 (C_{31} $H_{46}N_5O_{10}$ MH^+), found 648.3245 (MH^+).

References

1 For example see: S. Neidle, M. S. Puvvada and D. E. Thurston, *Eur. J. Cancer*, 1994, **30A**, 567; J. A. Hartley and R. L. Souhami, '*DNA Sequence Specificity of Anticancer Agents*', in Cancer Chemotherapy; Frontiers in Pharmacology and Therapeutics Series, ed J. A. Hickman and T. Tritton, Blackwell Scientific Ltd, 1993, pp. 251–280; Y. Hashimoto, *J. Pharm. Soc. Jpn*, 1994, **114**, 357.

2 For example see: (*a*) J. P. Sluka, J. H. Griffin, D. P. Mack and P. B. Dervan, *J. Am. Chem. Soc.*, 1990, **112**, 6369; (*b*) G. Mehta, T. Sambaiah, B. G. Maiya, M. Sirish and D. Chatterjee, *J. Chem. Soc., Perkin Trans. 1*, 1993, **22**, 2667; (*c*) P. E. Nielson, M. Egholm and O. Buchardt, *Bioconjugate Chemistry*, 1994, **5**, 1043; (*d*) Y. W. Ebright, Y. Chen, P. S. Pendergrast and R. H. Ebright, *Biochemistry*, 1992, **31**, 10664.

3 M. D. Lee, G. A. Ellestrad and D. B. Borders, *Acc. Chem. Res.*, 1991, **24**, 235.

4 L. L. Guan, J. Kuwahara and Y. Sugiura, *Biochemistry*, 1993, **32**, 6141.

5 D. S. Bose, G. B. Jones and D. E. Thurston, *Tetrahedron*, 1992, **48**, 751.

6 D. E. Thurston, 'Advances in the Study of Pyrrolo[2,1-c][1,4]benzodiazepine (PBD) Antitumour Antibiotics', in 'Molecular Aspects of Anticancer Drug–DNA Interactions', Vol. 1, Topics in Molecular and Structural Biology, ed. S. Neidle and M. J. Waring, Macmillan Press, 1993, pp. 54–88.

7 G. B. Jones, PhD Thesis, 1991, University of Portsmouth, UK.

8 D. E. Thurston and D. S. Bose, *Chem. Rev.*, 1994, **94**, 433.

9 D. R. Langley and D. E. Thurston, *J. Org. Chem.*, 1987, **52**, 91.

10 S. J. Morris, PhD Thesis, 1992, University of Portsmouth, UK.

11 P. H. Johnson and L. I. Grossman, *Biochemistry*, 1977, **16**, 4217.

12 R. P. Hertzberg, S. M. Hecht, V. L. Reynolds, I. J. Molineux and L. H. Hurley, *Biochemistry*, 1986, **25**, 1249; L. H. Hurley, T. Reck, D. E. Thurston, D. R. Langley, K. G. Holden, R. P. Hertzberg, J. R. E. Hoover, G. Gallagher, L. F. Faucette Jnr, S.-M. Mong and R. K. Johnson, *Chem. Res. Toxicol.*, 1988, **1**, 258.

13 D. S. Bose, A. S. Thompson, J. Ching, J. A. Hartley, M. D. Berardini, T. C. Jenkins, S. Neidle, L. H. Hurley and D. E. Thurston, *J. Am. Chem. Soc.*, 1992, **114**, 4939; D. S. Bose, A. S. Thompson, M. Smellie, M. D. Berardini, J. A. Hartley, T. C. Jenkins, S. Neidle and D. E. Thurston, *J. Chem. Soc., Chem. Commun.*, 1992, 1518.

Received, 7th September 1995; Com. 5/05928C