BIOCHEMISTRY

 Volume 44, Number 11 March 22, 2005

*Articles*

# Sequence-Selective Interaction of the Minor-Groove Interstrand Cross-Linking Agent SJG-136 with Naked and Cellular DNA: Footprinting and Enzyme Inhibition Studies†

Chris Martin,‡ Tom Ellis,‡ Claire J. McGurk,§ Terence C. Jenkins,|| John A. Hartley,§ Michael J. Waring,‡ and David E. Thurston*,⊥

*Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, United Kingdom, Cancer Research UK Drug−DNA Interactions Research Group, Department of Oncology, Royal Free and University College Medical School, UCL, 91 Riding House Street, London W1W 7BS, United Kingdom, School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom, and Cancer Research UK Gene Targeted Drug Design Research Group, The School of Pharmacy, University of London, 29/39 Brunswick Square, London, WC1N 1AX, United Kingdom*

*Received September 19, 2004; Revised Manuscript Received November 29, 2004*

ABSTRACT: SJG-136 (**3**) is a novel pyrrolobenzodiazepine (PBD) dimer that is predicted from molecular models to bind in the minor groove of DNA and to form sequence-selective interstrand cross-links at 5′-Pu-GATC-Py-3′ (Pu = purine; Py = pyrimidine) sites through covalent bonding between each PBD unit and guanines on opposing strands. Footprinting studies have confirmed that high-affinity adducts do form at 5′-G-GATC-C-3′ sequences and that these can inhibit RNA polymerase in a sequence-selective manner. At higher concentrations of SJG-136, bands that migrate more slowly than one of the 5′-G-GATC-C-3′ footprint sites show significantly reduced intensity, concomitant with the appearance of higher molecular weight material near the gel origin. This phenomenon is attributed to interstrand cross-linking at the 5′-G-GATC-C-3′ site and is the first report of DNA footprinting being used to detect interstrand cross-linked adducts. The control dimer GD113 (**4**), of similar structure to SJG-136 but unable to cross-link DNA due to its C7/C7′-linkage rather than C8/C8′-linkage, neither produces footprints with the same DNA sequence nor blocks transcription at comparable concentrations. In addition to the two high-affinity 5′-G-GATC-C-3′ footprints on the MS2 DNA sequence, other SJG-136 adducts of lower affinity are observed that can still block transcription but with lower efficiency. All these sites contain the 5′-GXXC-3′ motif (where XX includes AG, TA, GC, CT, TT, GG, and TC) and represent less-favored cross-link sites. In time-course experiments, SJG-136 blocks transcription if incubated with a double-stranded DNA template before the transcription components are added; addition after transcription is initiated fails to elicit blockage. Single-strand ligation PCR studies on a sequence from the c-*jun* gene show that SJG-136 binds to 5′-GAAC-3′/5′-GTTC-3′ (preferred) or 5′-GAGC-3′/5′-GCTC-3′ sequences. Significantly, adducts are obtained at the same sequences following extraction of DNA from drug-treated K562 cells, confirming that the agent reaches the cellular genome and interacts with the DNA in a sequence-selective fashion. Finally, SJG-136 efficiently inhibits the action of restriction endonuclease *Bgl*II, which has a 5′-A-GATC-T-3′ motif at its cleavage site.

Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD)[1] dimers are synthetic sequence-selective interstrand DNA minor-groove cross-linking agents developed from the monomeric anthramycin family of naturally occurring antitumor agents (*1*). The tricyclic PBD subunit incorporates a chiral center at

† This work was partly supported by Cancer Research U.K. (C180/A1060, SP1938/0402, SP1938/0201, SP1938/0301, and SP2000/0402 to D.E.T. and C2259/A3083 to J.A.H.) and Yorkshire Cancer Research (T.C.J.).

* Corresponding author. (D.E.T.) Phone: +44-(0)207 753 5932. Fax: +44 (0)207-753-5935. E-mail: david.thurston@ulsop.ac.uk.

‡ University of Cambridge.
§ Cancer Research UK Drug−DNA Interactions Research Group.
|| University of Manchester.
⊥ Cancer Research UK Gene Targeted Drug Design Research Group.

10.1021/bi0479813 CCC: $30.25 
Published on Web 02/19/2005

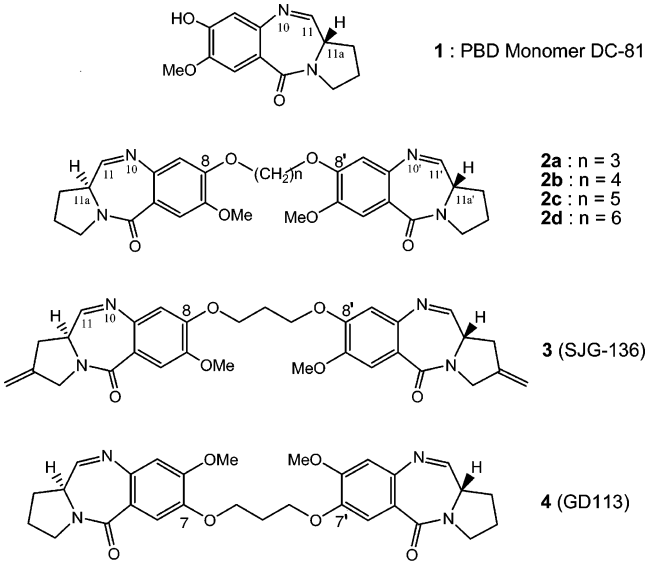


FIGURE 1: Structures of a PBD monomer (**1**, DC-81), the homologous C8/C8′-linked synthetic DSB-120 PBD dimer series (**2a**−**d**), the C2-exo-methylene PBD dimer SJG-136 (**3**), and the control C7/C7′-linked PBD dimer GD113 (**4**).

the C11a position and a DNA-reactive imine moiety at N10−C11 [e.g., DC-81 (**1**) in Figure 1]. The molecules adopt the appropriate 3-D shape to fit into the minor groove of double-stranded (duplex) DNA, such that the imine can effect subsequent alkylation at the exocyclic N2-position of a guanine base. PBD monomers typically span ∼3 base pairs (bp) of DNA with an alkylating preference for 5′-A**G**A-3′ sequences. Furthermore, PBD monomers have been shown to inhibit both endonuclease (*2*) and RNA polymerase (i.e., transcription) (*3*) enzymes in a sequence-dependent manner. A number of versatile synthetic routes to PBDs have been developed (*4*, *5*), and their structure−activity relationships (SARs) are now well-understood (*6*−*8*). Such agents are currently being used as part of a gene-targeting program designed to exploit their inherent reactivity toward embedded G-containing DNA sequences of biological significance (*9*).

The first C8/C8′-linked PBD dimer (DSB-120; **2a** in Figure 1) was reported in 1992 and represents two monomeric DC-81 subunits joined through the C8-positions of their aromatic A-rings through an inert propyldioxy (i.e., -O-$(CH_2)_3$-O- diether) linker (*10*). Dimeric homologues **2b**−**d** (-O-$(CH_2)_n$-O-, where $n = 4-6$) were later synthesized (*11*), and their relative cytotoxic potencies and DNA cross-linking efficiencies were reported (*12*), together with their cellular pharmacology (*13*). Extensive dynamic molecular modeling and an NMR structural study showed that PBD dimer **2a** spans a ∼6 bp tract in a targeted DNA duplex (Figure 2; Z = H,H; X = purine, Y = pyrimidine), with a cross-linking reactivity preference for 5′-Pu-GATC−Py-3′ sequences (*14*−*16*). However, while **2a** shows potent cytotoxicity toward a panel of human cancer cell lines in vitro (*10*, *12*), it was found to lack significant antitumor activity in vivo (*17*). Derivatives with either different C2/C2′ substitution and/or unsaturation in the pyrrolidine C-rings have since been evaluated, and a firm basis for the observed SAR has been established (*18*). From these studies, a C2/C2′ exo-unsaturated analogue SJG-136 (**3**) with an $n = 3$ diether linkage was identified that has significant activity both in vitro (*19*−*23*) and in vivo (*24*), and this compound is presently being evaluated in clinical trials.

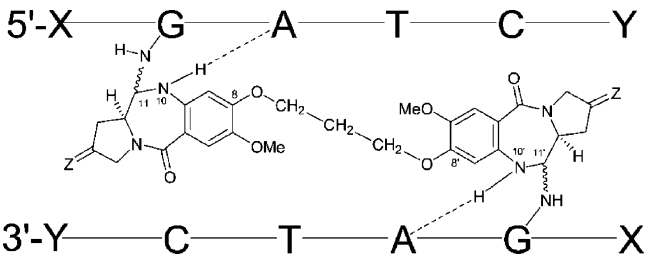


FIGURE 2: Proposed mechanism for the interstrand cross-linking of DNA by the PBD dimers [**2a** (DSB-120: Z = $H_2$) or **3** (SJG-136: **Z** = $CH_2$)]. Sequence selectivity is due in part to (i) covalent bonds between the C11/C11′ positions of the PBD moieties and the exocyclic N2 groups of guanines on opposite strands (shown as ∼∼∼) and (ii) hydrogen bonds (shown as - - -) formed between the N10/N10′ protons of the PBD moieties and the ring nitrogen N3 acceptors of the adenines adjacent to the covalently modified guanines. X and Y denote flanking bases (X = purine and Y = pyrimidine preferred).

One possible explanation for the biological activity of PBD dimers is that they bind sequence-selectively to critical genomic regions containing an over-abundance of relevant binding sites (*25*, *26*), thereby blocking transcription of key genes essential for tumor growth and/or inhibiting DNA replication. This concept has been used to explain the activity of other sequence-selective DNA-binding agents (*27*−*29*). An alternative explanation is that, compared to normal cells, tumor cells have an inherently reduced ability to repair such stable and complex interstrand cross-linked adducts across the entire genome and so enter an apoptotic pathway.

We now describe experiments that probe the sequence-selective binding of SJG-136 to both naked and cellular DNA using footprinting and enzyme inhibition techniques. Quantitative methods have been used to reveal both high-affinity preferred sites for DNA interstrand cross-linking and a population of lower-affinity sites. Furthermore, single-strand ligation PCR studies show that SJG-136 can reach the nucleus of cells and interact with genomic DNA with a similar sequence-selectivity to that occurring with naked DNA in vitro.

## MATERIALS AND METHODS

### *PBD Compounds and Solutions.*

SJG-136 (**3**) was synthesized as described previously (*19*, *20*). The structurally related PBD dimer control molecule, GD113 (**4**), was obtained from Spirogen Ltd. Drug solutions were prepared in HPLC-grade $CH_3OH$ prior to use; the $CH_3OH$ concentration in working DNA−drug mixtures was <0.5% v/v.

### *Footprinting.*

MS2-forward (5′-CAGGAAACAGCTATGAC-3′) or MS2-reverse (5′-GTAAAACGACGGCCAGT-3′) primer was 5′-end-labeled with $^{32}P$ using T4 polynucleotide kinase. Briefly, the labeling mixture contained the DNA primer (5 pmol; Sigma-Genosys), 10x kinase buffer (1 μL; Promega),

[1] Abbreviations: bp, base pair; PBD, pyrrolo[2,1-*c*][1,4]benzodiazepine; SAR, structure−activity relationship; Pu, purine; Py, pyrimidine.

($\gamma$-$^{32}$P)-ATP (2 $\mu$L of 10 $\mu$Ci/$\mu$L, 6000 Ci/mmol, Perkin-Elmer Life Sciences), and T4 polynucleotide kinase (10 U; Promega) in a 10 $\mu$L final volume. The labeling mixture was incubated at 37 °C for 30 min followed by inactivation of the kinase at 70 °C for 10 min. A 262 bp DNA fragment (MS2F or MS2R) was amplified from the MS2 plasmid using PCR with either a labeled forward or a reverse primer and the unlabeled primer as required. Reactions were comprised of a labeled primer mixture (10 $\mu$L), 10x PCR buffer (5 $\mu$L; Sigma), dNTPs (5 $\mu$L; 2.5 mM each of A, C, G, and T, Amersham), unlabeled primer (5 pmol) as required, MS2 plasmid template DNA (10 ng), and Taq DNA polymerase (5 U; Sigma) in a 50 $\mu$L final volume. Radiolabeled 262-mer was purified by non-denaturing polyacrylamide gel electrophoresis, elution, filtration, and EtOH precipitation and was then resuspended in a solution of Tris-HCl (10 mM) pH 7.4/NaCl (10 mM) to yield an activity of 200 counts/s per 5 $\mu$L when held up, in a pipet tip, to a capped Morgan series 900 minimonitor. The same protocol was used to produce a 320 bp DNA fragment from the pGL3-C plasmid (Promega). This fragment, chosen to allow footprinting of the plasmid multiple cloning site (MCS), was produced with a labeled forward primer GL3-forward (5′-AGTGCAGGTGCCAGAACATTTC-3′) and unlabeled reverse primer GL3-reverse (5′-TTTTGGCGTCTTCCATGGTG-3′).

For DNase I footprinting, the radiolabeled DNA solution (2 $\mu$L) was incubated overnight (16−18 h) with either a drug solution (388 $\mu$L) or an aqueous drug-free buffer [HEPES (20 mM) pH 7.9, NaCl (50 mM), $MgCl_2$ (1 mM), DTT (5 mM), 10% glycerol] for control reactions. Enzymatic cleavage was initiated by the addition of a DNase I solution (10 $\mu$L of 0.05 U/mL; Sigma) in aqueous NaCl (200 mM), $MgCl_2$ (20 mM), and $MnCl_2$ (20 mM) and halted with stop solution [40 $\mu$L of NaCl (2.25 M) and EDTA (150 mM) pH 8.0 containing glycogen (0.57 $\mu$g/$\mu$L; Roche), and poly(dI-dC)·poly(dI-dC) DNA (19.3 ng/$\mu$L Amersham Pharmacia)]. From this mixture, digested fragments were prepared for electrophoresis using the protocol of Trauger and Dervan (*30*) and loaded on to a preheated 8% denaturing polyacrylamide gel. Electrophoresis in 1x TBE buffer was allowed to proceed for 100 min at 70 W (1600−2000 V). The gel was then fixed in acetic acid (10% v/v) and methanol (10% v/v) for 15 min, followed by being blotted onto Whatmann 3MM paper and vacuum-dried at 80 °C for 45 min.

Dried gels were stored at room temperature in phosphor storage screens (Molecular Dynamics) for a minimum period of 12 h. Gel images were collected from exposed screens using a Molecular Dynamics 425E phosphorimager and transferred to ImageQuant v1.2 software (Molecular Dynamics) for analysis. Data were expressed as the differential cleavage between control and sample lanes (see **SI-1** and **SI-4**, Supporting Information). Peaks were integrated, and differential cleavage was calculated as $\ln(f_d) - \ln(f_c)$, where $f_d$ and $f_c$ are the fractional cleavage at any particular bond in the presence and absence (control) of a drug, respectively (*31*). The G + A markers prepared according to the usual protocol (*32*) were run alongside the samples to allow sequence identification. Estimates of the apparent binding affinity ($K_{app}$) of SJG-136 toward specific sites were calculated from gel images as the inverse of the concentration where DNase I digestion is half-maximal (*30*).

*In Vitro Stop Assay.*

An extended 282 bp DNA fragment, containing the 262 bp sequence used in the footprinting assays attached to a 5′ T7 promoter, was amplified from the MS2 plasmid by PCR using the T7-MS2-forward primer (5′-TAATACGACTCACTATAGGGCAGGAAACAGCTATGAC-3′) and the MS2-reverse primer. The 282-mer product was purified by phenol-$CHCl_3$ extraction and EtOH precipitation before being re-suspended in nuclease-free water to a fixed concentration (1 $\mu$g/$\mu$L).

A solution of the 282-mer (1 $\mu$L) was incubated overnight (16−18 h) with either 1.25× drug solution or water (as a control) (4 $\mu$L). The transcription mix [5 $\mu$L volume: 5× transcription buffer (2 $\mu$L), RNasin (0.25 $\mu$L, 40 U/$\mu$L), T7 RNA polymerase (0.25 $\mu$L, 20 U/$\mu$L), 100 mM DTT (0.25 $\mu$L; all purchased from Promega), NTPs (0.5 $\mu$L; 25 mM each of A, C, G, and T, Amersham), ($\alpha$-$^{32}$P)-CTP (0.25 $\mu$L of 10 $\mu$Ci/$\mu$L, 3000 Ci/mmol, Perkin-Elmer Life Sciences), and nuclease-free water (1.25 $\mu$L) was added, and synthesis of RNA transcripts was allowed to proceed at 37 °C for 90 min. After the addition of loading dye (10 $\mu$L), samples were heated to 94 °C for 4 min, briefly cooled on ice, and then loaded onto a preheated 0.5× TBE 8% denaturing polyacrylamide gel. Electrophoresis was performed for 80 min in 0.5× TBE buffer at 60 W (2000−2250 V), then gel blotted onto Whatmann 3MM paper and dried under vacuum at 80 °C for 45 min. Exposure to phosphor storage screens and image collection was performed as described for the footprinting gels.

For the temperature-dependence assays, DNA and drug solutions were cooled or heated to the appropriate temperature before being mixed and then maintained at that temperature after being mixed for the overnight period of incubation. In the assessment of time-dependence, DNA and the drug were incubated for the indicated time period prior to addition of transcription mix (positive time), after addition of the mix (negative time), or when DNA was added to a combination of transcription mix and drug (i.e., time $t = 0$).

Marker lanes were created by restriction endonuclease digestion of the 282-mer. Ten different restriction endonucleases, each with a single cutting site present in the selected 282-mer sequence, were employed to cleave the template 282-mer DNA in separate reactions according to the manufacturer's guidelines. Digested DNA was pooled and purified by phenol-$CHCl_3$ extraction and EtOH precipitation. The DNA was re-suspended in nuclease-free water to a fixed concentration (1 $\mu$g/$\mu$L). Marker DNA solution (1 $\mu$L) was used as a template for transcription as described previously, yielding an RNA ladder corresponding to transcription termination at the 10 enzyme cutting sites. The lengths of attenuated transcripts produced in the presence of drug were calculated by a graphical method. The electrophoresed distance of each marker band was accurately measured, and a plot of transcript length against this distance was curve-fitted. The electrophoresed distance of the attenuated transcripts was then used to estimate their approximate length and thus determine the position of transcription stop sites (T-stops) on the DNA template.

An analysis of the intensity of each T-stop was carried out by first measuring the raw intensity of each band using ImageQuant v1.2 (**SI-2**, Supporting Information). Then,

assuming a 100% efficiency of incorporation of radiolabeled cytidine into each fragment, normalized intensities were obtained by dividing the raw intensities of each band by the number of cytidine residues in the corresponding transcript (taken from the sequence).

*Single Strand Ligation PCR Assay.*

*Cell Culture.* The K562 human myeloid leukaemic cell line was grown in RPMI-1640 medium, supplemented with 10% foetal bovine serum and glutamine (2 mM), in a humidified incubator with 5% $CO_2$ at 37 °C.

*Oligonucleotides.* The oligonucleotide primer sets specific for the human c-*jun* promoter are as follows: J(Y)-1.B, 5′-biotin-CCCCCGACGGTCTCTC-3′; J(Y)-2, 5′-TTC-AAAATGTTTGCAACTGCTGCGTTAG-3′; and J(Y)-3, 5′-CTGCGTTAGCATGAGTTGGCACCC-3′.

All oligonucleotides were synthesized by MWG, Ebersberg, Germany for use in these studies. The double-stranded oligonucleotide linker $\gamma$ was as described previously (*33*) and was comprised of complementary upper and lower oligonucleotide strands where three additional cytosines are included at the 3′ terminus for the upper oligonucleotide. The lower oligonucleotide strand had a 3′-amine and a 5′-phosphate group.

*Drug Treatment of Isolated DNA.* Genomic DNA was isolated as follows. Aliquots (3 $\mu$g) were incubated at 37 °C with or without SJG-136 for 1 h in aqueous TEOA buffer [triethanolamine (25 mM) and EDTA (1 mM) pH 7.2] in a total volume of 50 $\mu$L. DNA was ethanol precipitated with 0.1 volumes of aqueous NaOAc (3 M) and 3 volumes of EtOH, washed twice with 70% v/v EtOH, then dried and finally re-suspended in water (10 $\mu$L) before immunoprecipitation.

*Drug Treatment of Cells.* Cells in serum-free RPMI ($3 \times 10^6$ cells in 1.5 mL) were added to 6-well tissue culture plates. An equal volume of RPMI containing the PBD dimer was then added to the cell suspension to give a final volume of 3 mL containing SJG-136 (0.03 $\mu$M). The drug-treated cells were finally incubated at 37 °C for 2 h.

*DNA Isolation from Cells.* DNA was isolated using the Wizard Genomic DNA Purification Kit (Promega). Restriction enzymes (40 U in a 120 $\mu$L volume of 1× buffer at 37 °C for 18 h) were used to cut the DNA at sites both immediately 3′ and 200−400 bases 5′ of the primer binding sites for each region studied. The DNA was then precipitated and dissolved in water, and the concentration determined with a fluorimeter (Perkin-Elmer, LS-2B) using Hoechst 33258 dye (Sigma-Aldrich) (*34*). The DNA was re-suspended in sterile water to an adjusted final concentration of 0.3 $\mu$g/$\mu$L and then stored at −20 °C.

*Measurement of Sequence Specificity of Covalent Adduct Formation.* The methodology used has recently been described (*35*). Briefly, linear PCR was performed on extracted DNA samples using biotinylated primer J(Y)-1.B. The biotinylated PCR products were captured and purified using streptavidin M-280 Dynabeads (Dynal). Terminal deoxynucleotidyl transferase was used to add a ribo-tail to the 3′-ends of the PCR products. DNA ligase was then used to add an oligonucleotide DNA linker to the tailed end. A second round of exponential PCR was carried out using primer J(Y)-2, followed by a third round of linear PCR using a 5′-$^{32}$P-end-labeled primer J(Y)-3. The samples were then run on polyacrylamide sequencing gels and visualized by autoradiography.

*DNA Sequencing.* Di-deoxy sequencing was carried out using the T7 Sequenase version 2.0 kit (Amersham, UK) using $^{32}$P labeled primers. DNA fragments for sequencing were generated by amplifying undamaged DNA. The sequencing products were resolved in parallel to the damaged DNA processed by the PCR-based technique outlined previously to determine the sites of nucleotide damage. The sequencing products run 27 nucleotides faster than the corresponding PCR products due to the addition of three guanines and the use of a linker in the damage detection assay.

*BglII Inhibition Assay and Footprinting on the Multiple Cloning Site (MCS) Region of pGL3-C.*

*BglII Inhibition Assay.* SJG-136 was incubated overnight with pGL3-C (20 ng, Promega) at a range of concentrations. The plasmid DNA was then linearized with *Bgl*II or *Mlu*I (Roche) according to the manufacturer's guidelines, and the digestion products were separated by 0.7% agarose gel electrophoresis. The ethidium bromide-stained gel was imaged under UV light, and the band intensities for linear products were calculated using NIHimage (NIH) software. Intensities were normalized as a percentage of control bands, and KaleidaGraph (Synergy Software, Reading, PA) was used to fit the results to a sigmoidal curve and to determine values for both $IC_{50}$ and the Hill slope.

*Footprinting on the MCS Region of pGL3-C.* A 320 base pair fragment of 5′-end-labeled $^{32}$P DNA (MCS and SV40 promoter region of pGL3-C) was incubated overnight (16−18 h) at room temperature with increasing concentrations of SJG-136 (0.01, 0.03, 0.1, 1.0, 3.0, and 10.0 $\mu$M) before being footprinted with DNase I (Figure 9B). The resulting products were separated by denaturing polyacrylamide gel electrophoresis as described for MS2 DNA in the previous section. The DNA fragment was produced by PCR from pGL3-C using $^{32}$P 5′-end-labeled forward primer DKGL3F (5′-AGTGCAGGTGCCAGAACATTTC-3′) and cold reverse primer DKGL3R (5′-TTTTGGCGTCT-TCCATGGTG-3′). The DNA fragment, around 320 bp in length, was labeled on the Watson strand to give resolution of footprints in the MCS region.

## RESULTS

*MS2 Sequence Analysis.*

The 262-bp sequence amplified from the MS2 plasmid was developed as a tool for footprinting studies as this DNA duplex contains all of the possible combinations of tetranucleotide sequences (*36*). This sequence was analyzed for occurrence of the SJG-136 preferred cross-linking site (5′-Pu-GATC-Py-3′) as predicted by molecular modeling studies (*19*) (see Figure 2) and also for regions containing the tetranucleotide 5′-GXXC-3′ motif (X = any base) that may provide less-favored binding sites. Figure 3 shows that the 262-mer contains two preferred sequences at 5′-$^{48}$GGATCC$^{53}$-3′ and 5′-$^{208}$GGATCC$^{213}$-3′ (indicated by double underlining), together with 17 other 5′-GXXC-3′ core motifs (indicated by single underlining). The site at 5′-$^{39}$GGTACC$^{44}$-3′ resembles the preferred 5′-GGATCC-3′

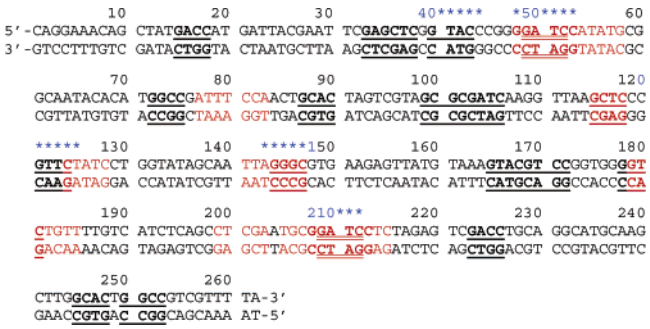

FIGURE 3: Analysis of the MS2 T7 sequence showing potential SJG-136 cross-linked binding sites and the results of the footprinting and in vitro stop assays: Single underline = all potential 5′-GXXC-3′ cross-linking sites where X is any base; double underline = preferred cross-linking sites [5′-Pu-GATC-Py-3′] based on molecular modeling; red bases = observed footprint sites; and overhead blue asterisks = observed T-stops.

Table 1: Analysis of the Observed Distribution of Footprints and T-Stops for the Interaction of SJG-136 (**3**) with the MS2 T7 Sequence (See Figure 3)[a]

| 5′-X-GXXC-X-3′ site | Footprint | T-stop[b] |
|---|---|---|
| 5′-$^{39}$G-GTAC-C$^{44}$ | not visible | yes |
| 5′-$^{48}$G-GATC-C$^{53}$ | 5′-$^{48}$GGATCCATATG$^{58}$-3′ (highest intensity)[c] | yes |
| not present[d] | 5′-$^{77}$ATTTCCA$^{83}$-3′ | no |
| 5′-$^{114}$A-GCTC-C$^{119}$ | 5′-$^{115}$GCTC$^{118}$ | no |
| 5′-$^{120}$C-GTTC-T$^{125}$ | 5′-$^{124}$CTATC$^{128}$ | yes |
| 5′-$^{143}$A-GGGC-G$^{148}$ | 5′-$^{141}$TTAGGGC$^{148}$-3′ | yes |
| 5′-$^{177}$G-GGTC-T$^{182}$ | 5′-$^{179}$GTCTGTT$^{185}$ | no |
| not present[e] | 5′-$^{199}$CTCGA$^{203}$-3′ | no |
| 5′-$^{208}$G-GATC-C$^{213}$ | 5′-$^{205}$TGCGGATCCTC$^{215}$-3′[f] | yes |

[a] Column 1 indicates the potential 5′-X-GXXC-X-3′ interstrand cross-linking sites associated with the corresponding footprints (column 2) and T-stops (column 3). [b] Two additional very low intensity T-stops were observed at the potential cross-linking sites 5′-$^{71}$T-GGCC-G$^{76}$-3′ and 5′-$^{86}$T-GCAC-T$^{91}$-3′ in the time-dependence experiment (see Figure 6). [c] Apparent affinity = $1.6 \times 10^5$ M$^{-1}$. [d] Footprint could be due to mono-adduct formation at 5′-$^{80}$TCC$^{82}$-3 (5′-GGA lower strand). [e] Footprint could be due to mono-adduct formation at 5′-$^{201}$CGA$^{203}$-3′ (or 5′-$^{202}$CGA$^{200}$-3′ lower strand). [f] Apparent affinity = $6.4 \times 10^4$ M$^{-1}$.

motif but differs in that the two central AT base pairs are reversed. The bases corresponding to observed footprint sites and T-stops (see Table 1) have been highlighted in red and blue (above the bases as blue asterisks), respectively.

*DNase I Footprinting.*

SJG-136 (**3**) was assessed for sequence-selective DNA binding by DNase I footprinting followed by subsequent data analysis using a computational method. Both forward-(MS2F) and reverse-(MS2R) labeled 262-mers (MS2F and MS2R, respectively) were used to allow complete coverage of the internal 200 bp sequence. Apparent binding constants for each site on the host DNA were calculated using the method of Dervan and co-workers (*37*) by direct comparison of quantitative experimental data. It should be stressed that covalent binding of a PBD can be reversible for a monoalkylation event, but cross-linking by a PBD dimer requires synchronization of two distinct alkylation events. Cross-linked sites of low affinity are inherently less stable than their high-affinity counterparts, and this will be manifest in a lower apparent binding constant in this assay. Furthermore, the assay method used revealed a time- and concentration-dependent behavior for the covalent DNA modifications. A full analysis of the binding profile in terms of association and dissociation kinetics, which is not possible from the results of this assay, would be required to obtain the equilibrium binding constants.

Figures 4 shows that SJG-136 binds to several sites on both MS2F and MS2R in the low micromolar range. A summary of the distribution of footprints is given in Table 1 (column 2), together with the closest 5′-GXXC-3′ cross-linking site likely to be responsible for each footprint (column 1). The footprints observed are also highlighted in red in Figure 3. It is noteworthy that the two predicted most energetically-favored binding sites, 5′-$^{48}$GGATCC$^{53}$-3′ and 5′-$^{208}$GGATCC$^{213}$-3′, are footprinted with high magnitude as indicated by the differential cleavage plot (**SI-1**, Supporting Information). The sequence 5′-$^{48}$GGATCC$^{53}$-3′ (the strongest site) has an apparent affinity of $1.6 \times 10^5$ M$^{-1}$ for SJG-136, whereas the second preferred site 5′-$^{208}$GGATCC$^{213}$-3′ has an affinity of $6.4 \times 10^4$ M$^{-1}$ (Table 1). Four of the six other footprints appear to be associated with sites containing the 5′-GXXC-3′ tetranucleotide core sequence, although the footprinting assay is unable to distinguish directly between cross-linked or mono-alkylated sites. Some of these lower-affinity sites contain either 5′-GGG-3′ or 5′-GGA-3′ triplets (on either the top or bottom strands) in close proximity to the proposed 5′-GXXC-3′ cross-linking sequence. On this basis, it is not possible to eliminate the formation of covalent mono-adducts at these sites where only one-half of the SJG-136 molecule might have interacted covalently with the second PBD residue unreacted. The footprints at 5′-$^{77}$ATTTCCA$^{83}$-3′ and 5′-$^{199}$CTCGA$^{203}$-3′ do not appear to be associated with a 5′-GXXC-3′ cross-linking motif and may be due to monomeric PBD binding sites at 5′-$^{80}$TCC$^{82}$-3′ (i.e., 5′-$^{82}$GGA$^{80}$-3′ on lower strand), or 5′-$^{201}$CGA$^{203}$-3′ (top strand), or 5′-$^{202}$CGA$^{200}$-3′ (lower strand), respectively.

In a control experiment, the PBD dimer GD113 (**4**), a C7/C7′-linked analogue of DSB-120 (**2a**), which does not have the correct 3-D shape to allow interstrand DNA cross-linking, was footprinted on the MS2 sequence (data not shown). As anticipated, no footprints were visible on any part of the sequence at concentrations up to 100 $\mu$M.

*In Vitro Transcription (T-Stop) Assay.*

The 282-mer DNA (20-bp T7 promoter and 262-bp MS2 sequence) was used as a template for in vitro transcription in the presence or absence of SJG-136. The dependence of the DNA−drug interaction on concentration, temperature, and time factors was examined.

*Concentration-Dependence.* Alterations in the concentration of SJG-136 revealed the presence of significant transcriptional stop sites (T-stops) considered to be associated with the potential 5′-$^{39}$GGTACC$^{44}$-3′, 5′-$^{48}$GGATCC$^{53}$-3′, 5′-$^{120}$CGTTCT$^{125}$-3′, 5′-$^{143}$AGGGCG$^{148}$-3′ and 5′-$^{208}$GGATCC$^{213}$-3′ cross-linking sites (Figure 5). The findings are summarized in Table 1 (columns 1 and 3) and are correlated with the corresponding footprints (column 2). Further weak concentration-dependent T-stops are also evident in Figure 5 (e.g., between markers 72−97). Other T-stops are found at the origin of the gel but are not properly resolved (e.g., just below the 222/210 marker).

The overall pattern of appearance for the T-stops appears to mirror the process of single-hit kinetics in that as stop

sites near the start of the sequence develop and become more prominent, then production of longer transcripts (including full-length product) becomes increasingly blocked. For example, as the preferred cross-linked adduct at 5′-$^{48}$GGATCC$^{53}$-3′ develops, longer transcripts are gradually inhibited until the full-length transcript at the origin disappears by 1 $\mu$M drug dose. T-stops at 5′-$^{48}$GGATCC$^{53}$-3′, 5′-$^{120}$CGTTCT$^{125}$-3′, and 5′-$^{143}$AGGGCG$^{148}$-3′ are clearly present at the lowest concentration tested (0.05 $\mu$M). Conversely, the T-stop at 5′-$^{39}$GGTACC$^{44}$-3′ does not reach full intensity until ~2 $\mu$M.

With the exception of 5′-$^{39}$GGTACC$^{44}$-3′, all of the T-stops coincide convincingly with footprints (see Table 1). However, as Figure 5 makes clear, the 5′-$^{39}$GGTACC$^{44}$-3′ sequence must be one of the lower affinity sites as the T-stop does not appear to reach maximum intensity until an elevated drug dose of ~2 $\mu$M is reached, and the adduct formed appears insufficiently stable to provide a significant footprint. Conversely, the 5′-$^{48}$GGATCC$^{53}$-3′ sequence, which gives rise to a T-stop at the lowest concentration (0.05 $\mu$M), provides the strongest footprint.

*Time-Dependence.* Increases in drug−DNA incubation time for more than 15 min prior to initiating transcription have no effect on the intensity of the T-stops produced (see Figure 6). This behavior was seen in independent experiments at three different SJG-136 concentrations of 0.25 and 0.5 $\mu$M (data not shown) and 1 $\mu$M. However, addition of drug at either $t = 0$ or after transcription initiation (i.e., $t = -30$ or $-60$ min) produced no T-stops and had no effect on the production of full-length transcript. In these experiments, two additional weak T-stops became apparent that may correspond to potential cross-linking sites at 5′-$^{71}$TGGCCG$^{76}$-3′ and 5′-$^{86}$TGCACT$^{91}$-3′ (not included in Table 1).

*Temperature-Dependence.* At 1 $\mu$M SJG-136 concentration, the T-stop produced at 5′-$^{39}$GGTACC$^{44}$-3′ increases in intensity in a temperature-dependent manner (Figure 7). At 4 °C, it represents the least intense T-stop, whereas at 37 °C, it becomes the most prominent such that production of full-length transcript is completely blocked. The intensities of all other T-stops, and synthesis of full-length product, diminish as the temperature is raised.

*Concentration-Dependence of GD113.* An identical method was adopted to assess the ability of GD113 (**4**) to inhibit transcription by T7 RNA polymerase (see **SI-3**, Supporting Information). Between the 5−50 $\mu$M concentration levels, GD113 produced a concentration-dependent T-stop at 5′-$^{127}$TCC$^{129}$-3′ (i.e., 5′-GGA-3′ on the opposite strand) with a weaker T-stop at 5′-$^{192}$TCT$^{194}$-3′ (i.e., 5′-AGA-3′ on the opposite strand) at the 50 $\mu$M drug level. It is noteworthy that significant T-stops are not seen below 15 $\mu$M, a concentration at least 300-fold higher than that required for transcription attenuation by SJG-136 (**3**).

*Single Strand Ligation PCR Assay.*

The sequence selectivity of binding of SJG-136 to human genomic DNA has been evaluated using a single-strand ligation PCR (sslig-PCR/TD-PCR) combined technique (*35*). Several binding sites within a region of the non-transcribed strand of the human c-*jun* gene were observed following treatment of naked genomic DNA with SJG-136 for 1 h at 3 nM (Figure 8A). All of the binding sites observed correspond to potential 5′-GXXC-3′ interstrand cross-linking sites for SJG-136. Although there are no preferred 5′-GATC-3′ sequences within this fragment of the c-*jun* gene, the three most intense sites correspond to 5′-GAAC-3′/5′-GTTC-3′. The other two weaker sites observed correspond to 5′-GAGC-3′/5′-GCTC-3′ sequences. This pattern of binding observed in naked genomic DNA was compared with that obtained in intact cells after treatment with SJG-136 (Figure 8B). The binding in cells was found to be similar to naked DNA in that the 5′-GAAC-3′/5′-GTTC-3′ preference was preserved.

*Restriction Endonuclease Inhibition.*

SJG-136 (**3**) was found to impair restriction endonuclease digestion of the plasmid pGL3-C MCS in a sequence-selective manner. Inhibition of cleavage at the *Bgl*II site occurred at a lower concentration (~10-fold less) than at the *Mlu*I site (~3 vs 30 $\mu$M). A plot of quantified intensities of digestion products is shown in Figure 9A. DNase I footprinting of pGL3-C MCS after treatment with SJG-136 (Figure 9B) confirmed that a high-affinity adduct forms at the 5′-$^{35}$AGATCT$^{40}$-3′ motif found at the *Bgl*II cleavage site, whereas no significant SJG-136 binding is observed at the *Mlu*I site 5′-$^{14}$TGCGCA$^{19}$-3′ (see analysis of footprinting gel in **SI-4**, Supporting Information). Interestingly, the footprinting gel (Figure 9B) shows similar features to that for SJG-136-treated MS2 (Figure 4) in that, at SJG-136 concentrations of 1 $\mu$M and above, high molecular weight material appears in a concentration-dependent manner near the origin of the gel accompanied by a concentration-dependent decrease in intensity of all bands of higher molecular weight at positions beyond the preferred 5′-GATC-3′ binding site. This is interpreted as the formation of an interstrand cross-link at the energetically favored 5′-GATC-3′ site with subsequent formation of duplex DNA. There are less-favored potential interstrand cross-linking sequences such as 5′-AGCTCT-3′ and 5′-CGTGCT-3′ flanking the *Mlu*I site, which might be responsible for the observed inhibition of this enzyme. The other footprints observed in Figure 9B are most likely due to the formation of mono adducts at sites such as 5′-AGA-3′/5′-TCT-3′ (−11 to −9), 5′-TGA-3′/5′-TCA-3′ (59−61), or 5′-AGT-3′/5′-ACT-3′ (57−59) (see **SI-4**, Supporting Information).

## DISCUSSION

*DNase I Footprinting.* The footprinting experiments performed here corroborate the results of dynamic computer modeling studies of covalent interactions between SJG-136 (**3**) and DNA duplexes of defined sequence (*19*, *38*). This also demonstrates, for the first time, the usefulness of DNA footprinting assays to study sequence-selective cross-linking agents. As anticipated, two of the highest affinity binding sites present in the footprinted 262-mer sequence contain 5′-GGATCC-3′ motifs. However, the data reveal that SJG-136 binds to a number of lower affinity sites as well. With the exception of the footprints at 5′-$^{77}$ATTTCCA$^{83}$-3′ and 5′-$^{199}$CTCGA$^{203}$-3′, which may represent mono-adduct formation, these sites contain 5′-GXXC-3′ motifs where the spanned central XX doublets are TA, GC, CT, TT, GG, or GT rather than the 5′-AT-3′ predicted to be the most

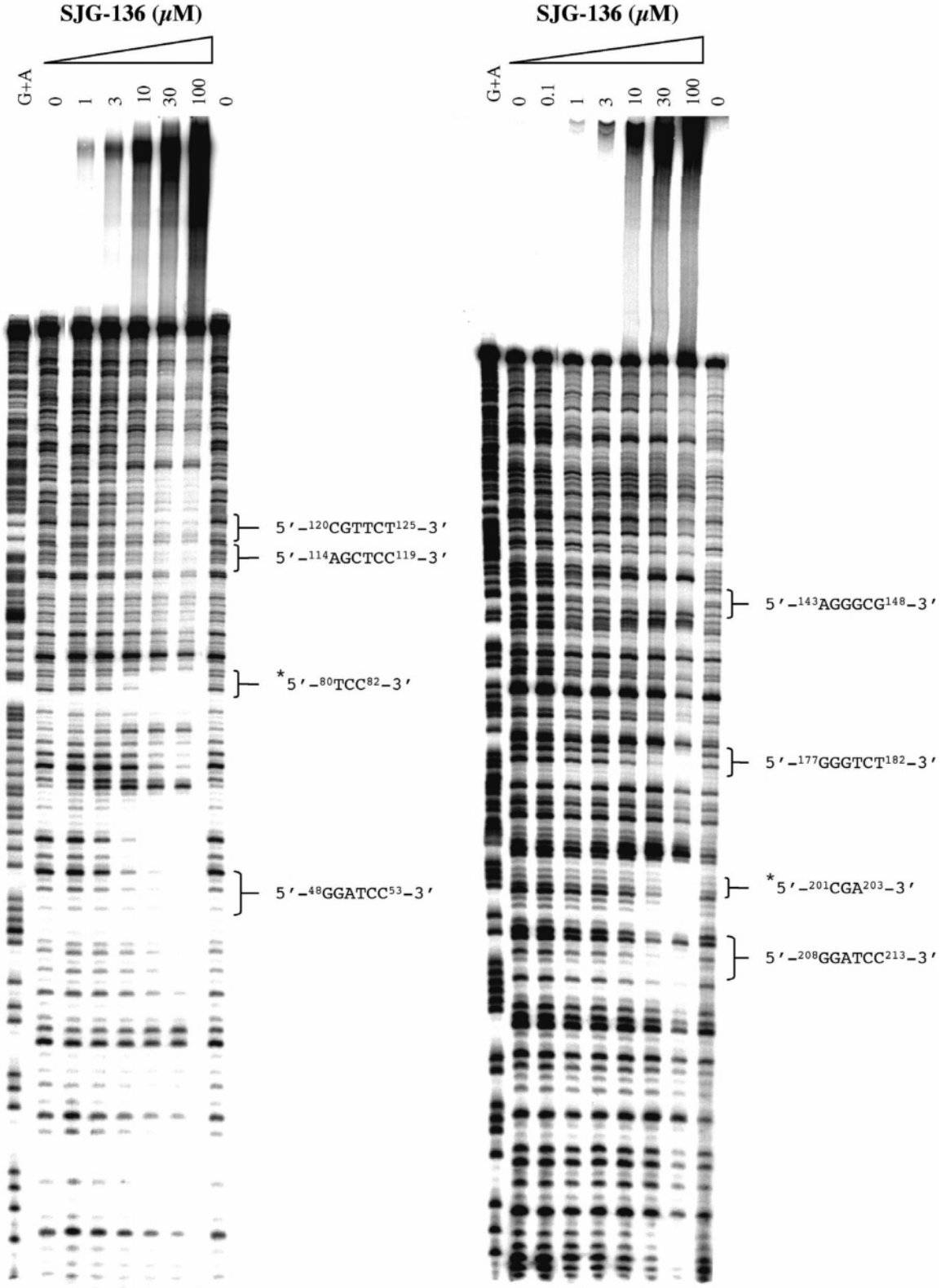


FIGURE 4: Footprinting gels for the interaction of SJG-136 (**3**) with the MS2 sequence shown in Figure 3 at drug concentrations of 0.1, 1, 3, 10, 30, and 100 $\mu$M (left panel = top strand [MS2-F]; right panel = bottom strand [MS2-R]). The labels to the right of each gel correspond to the potential cross-linking sites listed in column 1 of Table 1 with the exception of 5′-$^{80}$TCC$^{82}$-3 (left-hand gel) and 5′-$^{201}$CGA$^{203}$-3 (right-hand gel) (asterisked), which are thought to be mono-covalent binding sites (see footnotes d and e in Table 1). An analysis of footprint intensities for the 3, 10, and 30 $\mu$M dose levels is provided in Supporting Information (**SI-1**).

energetically favored based on models. This contradicts a previously held belief that a guanine (and opposing cytosine) residue might not be well-tolerated between the cross-linked guanines due to the bulky nature of the C2-exocyclic amino group of a guanine projecting into the minor groove. One possible explanation is that these are not strictly cross-linking sites but may reflect either mono-alkylation events at these sites (i.e., where only a single PBD unit from one end of the dimer is attached) or possibly adducts generated as a result of non-covalent processes. For example, the

footprints relating to the 5′-$^{143}$AGGGCG$^{148}$-3′ and 5′-$^{114}$AGCTCC$^{119}$-3′ potential cross-linking sites (see Figure 3 and Table 1) all contain known PBD monomeric binding sites (i.e., 5′-Pu-G-Pu-3′ on the top and bottom strands, respectively [indicated by underlining]) (*3*).

It is worth noting that the gels shown in Figures 4 and 9B illustrate a novel feature of the DNase footprinting of SJG-136. At concentrations of 1 $\mu$M and above, high molecular weight material appears at the top of the gel and increases in intensity in a concentration-dependent manner. This is accompanied by a corresponding loss of intensity for all bands migrating slower than the footprint site thought to correspond to an interstrand cross-link. The higher molecular weight material is thought to represent duplex DNA species that cannot melt to single strands before and/or during electrophoresis due to the presence of one or more SJG-136-induced interstrand cross-links and is thus retarded in the gel appearing as a smear. A useful consequence of this is that interstrand DNA cross-link positions can be identified by noting the footprint site from which the band wipe-out extends. This phenomenon is visible in the footprinting gels shown in both Figures 4 and 9B and in the latter case clearly identifies the strong interstrand cross-link at the *Bgl*II site.

*In Vitro Transcription.* Table 1 reveals that four of the five major T-stops coincide with accompanying footprints. Furthermore, all the T-stops appear to be associated with 5′-GXXC-3′ sequences, although the assay is unable to differentiate between interstrand cross-linked (i.e., bifunctional alkylation) or mono-alkylated adducts. However, it is generally accepted that covalently bound adducts are required to produce significant T-stops in this assay (*3*), and so non-covalent adducts can probably be excluded. Interestingly, the T-stop produced at 5′-$^{39}$GGTACC$^{44}$-3′ did not produce a significant footprint under the conditions of the experiment. This may be attributable to the relatively low affinity of SJG-136 for this site as indicated by the elevated concentrations required to achieve a T-stop.

It might be anticipated that cross-linked duplex sites should provide the most stable blockage to transcription whereas mono-alkylated species may provide adequate but less robust inhibition. On first inspection of the gel in Figure 5, it appears that the T-stops at 5′-$^{120}$CGTTCT$^{125}$-3′ and 5′-$^{143}$AGGGCG$^{148}$-3′ are similar in intensity to that produced at 5′-$^{48}$GGATCC$^{53}$-3′, which is one of the proposed interstrand cross-linked sites. However, when band intensity is normalized to radiolabel content (i.e., longer RNA transcripts contain more radiolabeled bases), then the T-stop at 5′-$^{48}$GGATCC$^{53}$-3′ is some 60% more intense than other T-stops (see **SI-2**, Supporting Information), suggesting that 5′-$^{48}$GGATCC$^{53}$-3′ may be the principal cross-linking site. In support of the previous assignment, this T-stop also corresponds to the strongest footprint.

*Kinetic Behavior.* The time-scale for interaction between SJG-136 and DNA duplex can be estimated from the kinetic data shown in Figure 6. At the concentrations of SJG-136 and DNA used, it appears that incubation for only 15 min is required to obtain the full spectrum and intensity of transcription blocks. However, when transcription is initiated before addition of SJG-136 (or simultaneously), then no inhibition occurs, suggesting that the adducts need to be fully formed before they can block transcription.

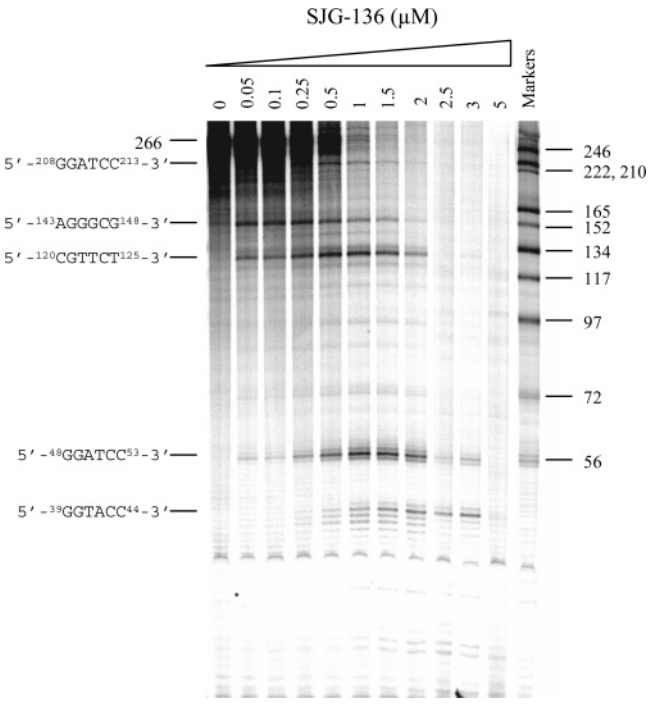


FIGURE 5: T-stop assay showing the effect of SJG-136 (**3**) concentration on the inhibition of T7 RNA polymerase at concentrations ranging from 0.05 to 5.0 $\mu$M. Stop sites, including the most-favored interstrand cross-link site (5′-$^{48}$GGATCC$^{53}$-3′), can be seen developing at the lowest concentration studied (0.05 $\mu$M). Marker labels (right of gel) refer to the length of transcript produced and not the corresponding base positions as depicted in Figure 3. Thus, all markers contain an extra four 5′-nucleotides transcribed from the T7 promoter (e.g., marker 56 refers to a T-stop at position 52).

It is known that T7 RNA polymerase produces RNA transcripts at a rate of 20−50 nucleotides per second (*39*). Thus, in the experiments leading to Figure 6, each full-length transcript (266 nt) should be synthesized in 5−13 s. It follows that formation of the covalent bonds between SJG-136 and DNA required to block the transcribing enzyme must require longer than this time period to form. It is likely that SJG-136 is initially held in the minor groove by non-covalent interactions that are fully reversible. During these early stages of adduct formation, the molecule may either dissociate from the DNA if the local sequence is unfavorable or be expelled from the groove by the advancing polymerase as the duplex unwinds. Presumably, it is only when SJG-136 finds itself in an energetically favored sequence (i.e., 5′-Pu-GXXC-Py-3′, and preferably, 5′-Pu-GATC-Py-3′) that a covalent adduct forms capable of blocking transcription. This kinetic control process presumably requires >5−13 s for demonstrable efficiency.

*Effects of Temperature.* The effect of temperature on the inhibition of transcription (Figure 7) provides an insight into the energetics of covalent binding to each DNA reaction site. It is clear that an input of heat energy is able to increase the probability of drug binding at all sites. This behavior confirms the bimolecular nature of the covalent drug−DNA interaction and the inherent requirement for activation processes. It is likely that each sequence motif targeted by SJG-136 is associated with a specific activation energy for covalent bond formation. At the preferred 5′-Pu-GATC-Py-3′ site (i.e., 5′-$^{48}$G-GATC-C$^{53}$-3′), this activation energy must be sufficiently low that even at 4 °C

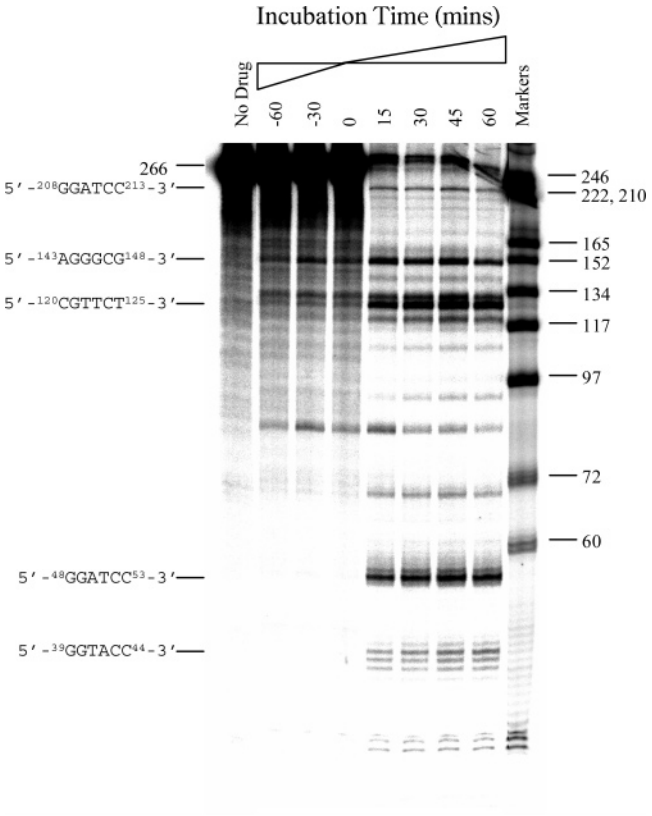


FIGURE 6: T-stop assay showing the effect of incubation time on the ability of SJG-136 (**3**) at 1.0 $\mu$M to inhibit transcription. The right-hand side of the panel shows that most stop sites are already visible after only 15 min incubation of drug and DNA duplex prior to the addition of transcription buffer, and that there is little significant change up to 60 min incubation. The left-hand side of the panel shows that once transcription has started ($t = 0$), there is little effect upon adding SJG-136 at either 30 or 60 min time points (labeled as −30 and −60 min, respectively).

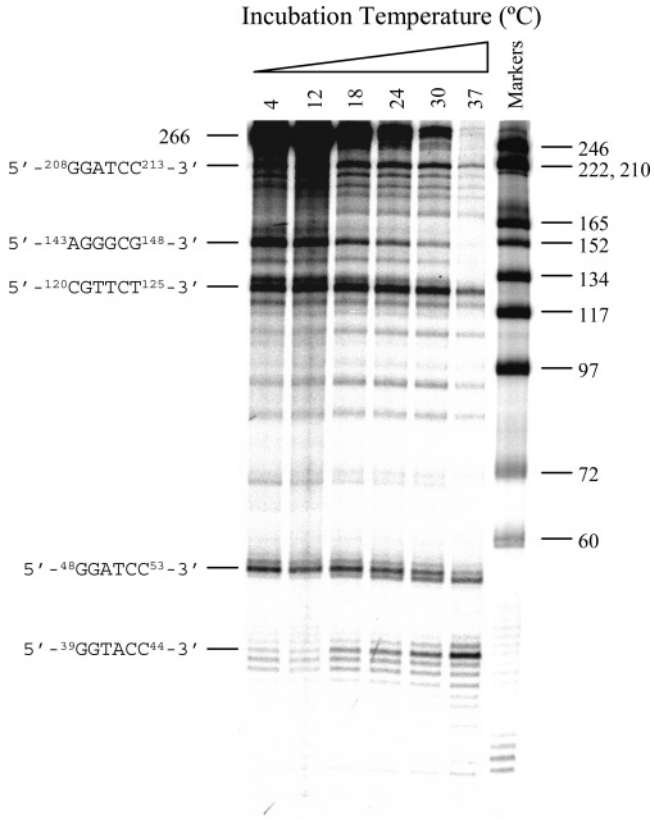


FIGURE 7: T-stop assay showing the effect of temperature on the ability of SJG-136 (**3**) to inhibit transcription. For the preferred interstrand cross-linked sequence (5′-$^{48}$GGATCC$^{53}$-3′), a strong stop site is evident at 4 °C with an intensity that is not changed significantly at 12, 18, 24, 30, or 37 °C. A stop for the less-favored cross-linked adduct (5′-$^{39}$GGTACC$^{44}$-3′) is much less intense at 4 °C but then increases in intensity proportionately with temperature until by 37 °C it is similar in intensity to that of the 5′-$^{48}$GGATCC$^{53}$-3′ stop site at 4 °C.

adduct formation is adequate enough to provide significant attenuation of transcription. Conversely, binding must be negligible at the less-favored 5′-$^{39}$GGTACC$^{44}$-3′ site at 4 °C, although the degree of binding increases significantly as the temperature is raised. The argument of a larger activation energy being required for the 5′-$^{39}$G-GTAC-C$^{44}$-3′ sequence as compared to 5′-$^{48}$G-GATC-C$^{53}$-3′ is consistent with the slower rate of reaction of SJG-136 at the former site as clearly shown in Figure 6.

*Control Molecule GD113.* The structure of GD113 (**4**) is similar to that of DSB-120 (**2a**) and SJG-136 (**3**), although the two PBD units are connected through their A-ring C7/C7′-positions rather than their C8/C8′-positions. Both molecular modeling and physical studies suggest that PBD dimers with this configuration are unable to adopt the necessary isohelical three-dimensional shape to successfully form interstrand cross-links within the DNA minor groove (*38*, *40*). T-stop data (see **SI-3**, Supporting Information) clearly demonstrate that transcription blockage by GD113 occurs only at concentrations of 15 $\mu$M and above, and only becomes significant at concentrations approaching 50 $\mu$M. The two transcription blocks observed (5′-$^{127}$TCC$^{129}$-3′ and 5′-$^{192}$TCT$^{194}$-3′) appear to correspond to mono-alkylation sites (i.e., 5′-GGA-3′ and 5′-AGA-3′, respectively, on the opposite strand) rather than 5′-GXXC-3′ potential interstrand DNA cross-linking sites. These data are consistent with the lack of footprints for GD113 at concentrations of up to 100 $\mu$M and confirm the greater sensitivity of the in vitro transcription assay as compared to DNase I footprinting as a biophysical tool to study mono-alkylated PBD-DNA adducts.

*Single-Strand Ligation PCR Assay.* Figure 8A indicates the presence of five adducts on this 220-bp fragment of naked c-*jun* DNA treated with SJG-136 (**3**). The three most intense sites all occur at 5′-GAAC-3′ sequences (or 5′-GTTC-3′ on the lower strand) with two less intense ones at 5′-GCTC-3′ (or 5′-GAGC-3′ on the lower strand). These sites coincide with the footprints and T-stops observed at 5′-$^{120}$CGTTCT$^{125}$-3′ and 5′-$^{114}$AGCTCC$^{119}$-3′, respectively, on the MS2 DNA fragment (see Figures 4−7 and Table 1). Although this c-*jun* fragment does not contain a preferred 5′-Pu-GATC-Py-3′ sequence for SJG-136, the observed binding pattern supports the notion that it targets 5′-GXXC-3′ sites.

The data shown in Figure 8B demonstrate that, in cells treated with SJG-136, the agent successfully traverses both the cellular and the nuclear membranes to interact with the same fragment of DNA at the genomic level with the same degree of sequence selectivity.

*Restriction Endonuclease Inhibition.* The data from the DNase I footprinting, in vitro transcription, and single-strand ligation PCR assays all concur that SJG-136 is capable of sequence-selective interaction with linear ∼250-bp DNA fragments both in naked DNA and (in the case of the latter

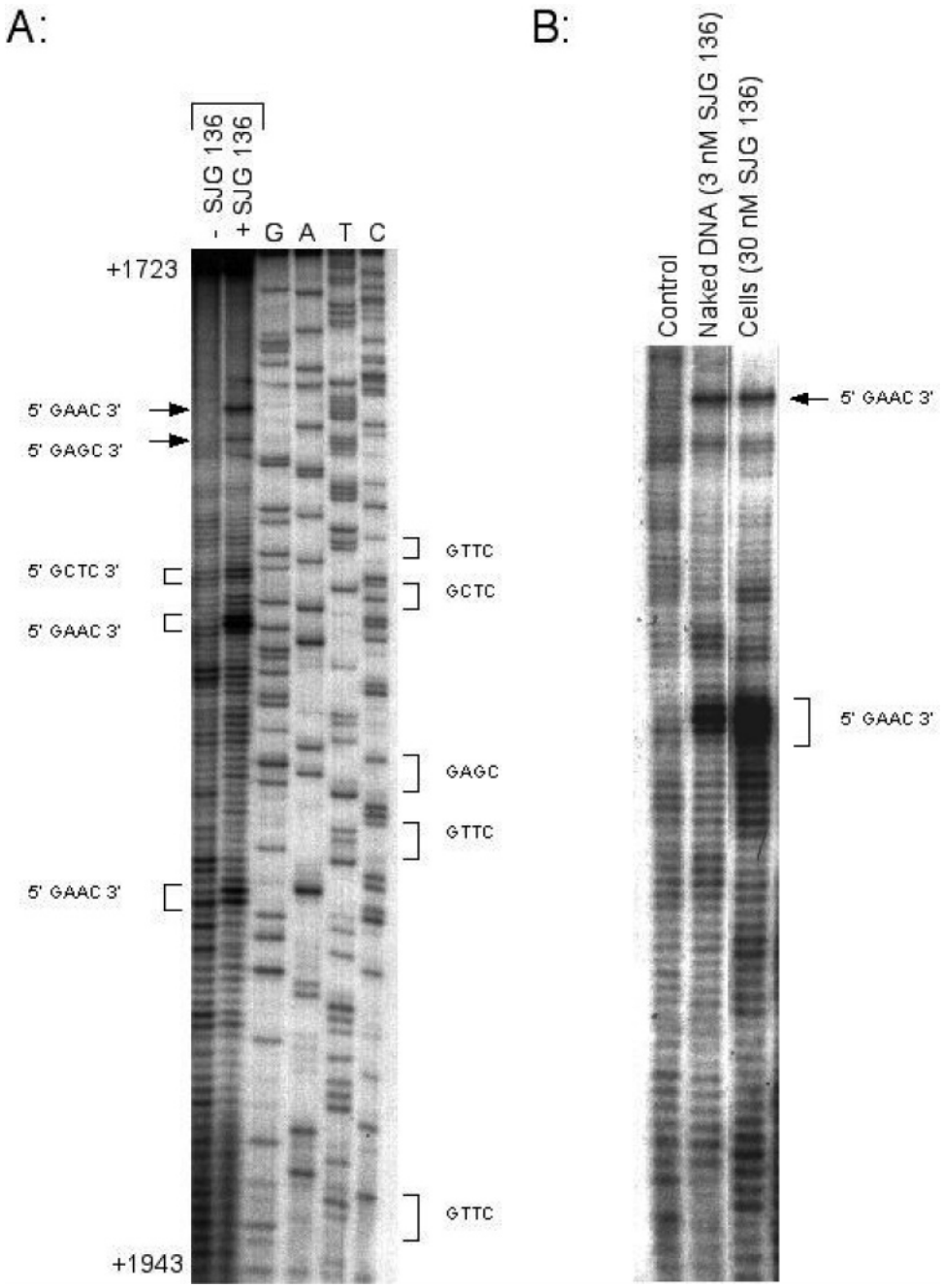


FIGURE 8: Detection of SJG-136 (**3**) adducts on the non-transcribed strand of the human c-*jun* gene using the sslig-PCR/TD-PCR combined technique: (A) human genomic DNA extracted from K562 cells treated with SJG-136 at 3 nM for 1 h. Numbers on the left of the gel indicate the nucleotide positions relative to the major transcription start site. The main drug binding sites are labeled on the left-hand side of the gel, and di-deoxy sequencing ladders are indicated as G, A, T, and C on the right. The bands in the sequencing lanes are 27 nucleotides shorter than the corresponding sslig-PCR/TD-PCR generated products. (B) Comparison of the sequence-specific binding of SJG-136 to naked DNA (treated at 3 nM for 1 h) with DNA extracted from K562 cells treated with SJG-136 at 30 nM for 2 h. The two major 5′-GAAC-3′ binding sites are indicated.

assay) within cells. The restriction endonuclease inhibition study demonstrates that the same sequence-selective interactions occur with circular plasmid DNA of ~5 kb in length. The *Bgl*II restriction site contains a preferred 5′-AGATCT-3′ cross-linking site for SJG-136, and an adduct can be shown to form here by DNA footprinting (Figure 9B and **SI-4** in Supporting Information), which presumably blocks access to the restriction endonuclease. Access is not blocked to the same degree at the *Mlu*I site, and a significant degree of cutting still occurs. However, the small reduction in cutting at this site compared to control could reflect binding of SJG-136 to neighboring lower affinity sites. For example, there are less-favored potential cross-linking sequences such as 5′-AGCTCT-3′ and 5′-CGTGCT-3′ flanking the *Mlu*I site, along with some potential mono-covalent sites, although none of these are of sufficiently high affinity to become apparent by footprinting.

## CONCLUSIONS

The results of the DNase I footprinting, in vitro transcription, and endonuclease restriction experiments reported here confirm that, as predicted from molecular modeling studies,

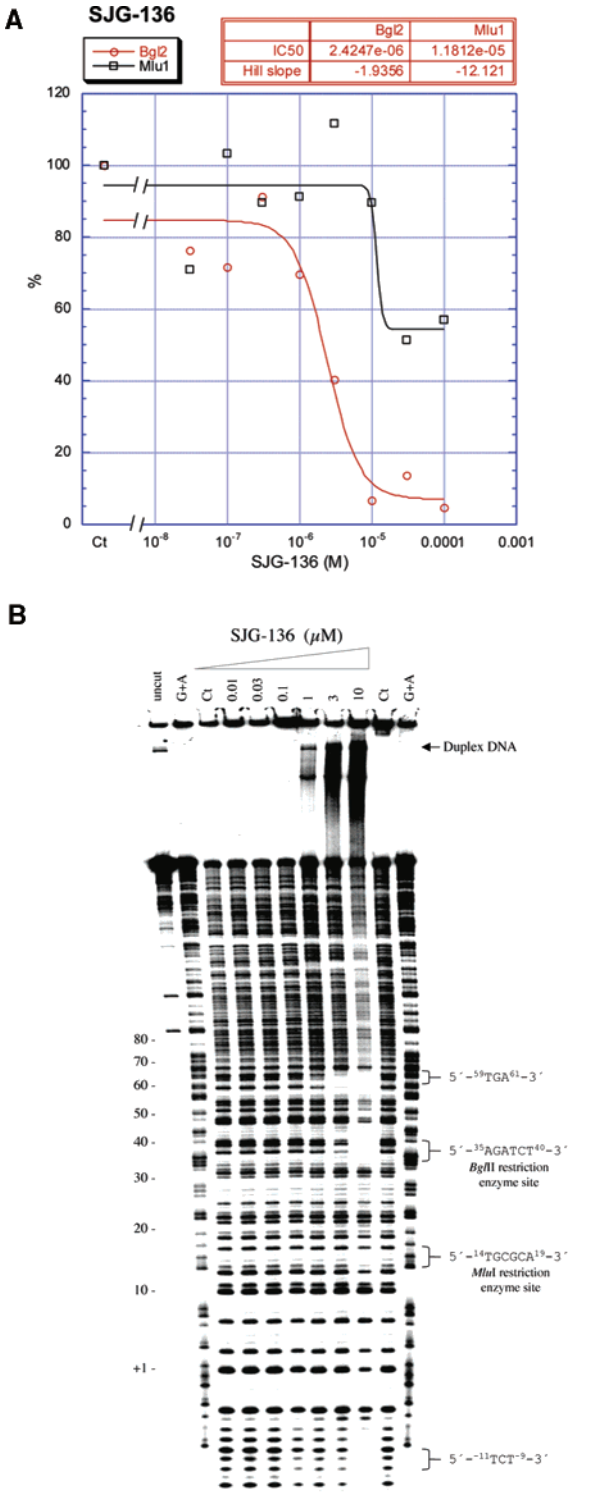


FIGURE 9: (A) Inhibition of cleavage of the pGL3-C plasmid by site-specific restriction endonucleases *Bgl*II and *Mlu*I by SJG-136 (**3**). The plasmid was linearized with *Bgl*II and *Mlu*I after incubation with increasing concentrations of SJG-136, and the products were separated by 0.7% agarose gel electrophoresis, stained with ethidium bromide, and then viewed under UV light. Band intensities of linearized products were measured and plotted against SJG-136 concentration. (B) Footprinting gel showing the binding sites of SJG-136 on a 320 base pair fragment of 5′-end-labeled $^{32}$P DNA (MCS and SV40 promoter region of pGL3-C) after incubation overnight (16−18 h) at room temperature with increasing concentrations of SJG-136 (0.01, 0.03, 0.1, 1.0, 3.0, and 10.0 $\mu$M) before being treated with DNase I. The resulting products were separated by denaturing polyacrylamide gel electrophoresis as described in the text for MS2 DNA.

SJG-136 forms high-affinity adducts at interstrand cross-linking 5′-Pu-GATC-Py-3′ sites. However, these data, along with the single-strand ligation PCR assay results, demonstrate that SJG-136 also binds to a number of other 5′-GXXC-3′ potential interstrand cross-linking sites of lower affinity. There is also evidence that it forms lower-affinity monocovalent adducts at sequences such as 5′-Pu-G-Pu-3′ or 5′-Pu-G-Py-3′. Of the other nine possible 5′-GXXC-3′ combinations (i.e., 5′-AG-3′/5′-CT-3′, AC/GT, AA/TT, TG/CA, TC/GA, TA/TA, GG/CC, CG/CG, and GC/GC), six of these (TA/TA, GC/GC, TC/GA, AA/TT, GG/CC, and AC/GT) are represented in the gel-based data (i.e., footprints or T-stops, or both), but three (TG/CA, TC/GA, and CG/CG) are not. Significantly, two of the latter sites (5′-TG/5′-CA and 5′-CG/5′-CG) contain the least preferred pyrimidine bases in their first 5′-X positions on both the top and bottom strands, and the other has a pyrimidine (T) in one of the 5′-X positions. With one exception, all of the bound sites have at least one and sometimes both of the 5′-flanking X positions (top and bottom strands) occupied by a purine (and preferably adenine) according to the footprinting data. The one exception is the 5′-GTAC-3′ sequence, which has a thymine at each of the 5′-X positions. However, this adduct gave some of the weakest footprints and T-stops observed. Therefore, the overall conclusion is that the 5′-Pu-GATC-Py-3′ motif provides the highest affinity binding site for SJG-136, which is likely to be the most biologically relevant. Furthermore, the single-strand ligation PCR experiment in cells demonstrates that SJG-136 is able to cross both the cellular and nuclear membranes to reach its genomic target.

It has been hypothesized that the remarkable cytotoxicity of SJG-136 across a number of in vitro cell lines (*22*), as well as its broad-spectrum antitumor activity in numerous human tumor xenograft models (*24*), may be due to the formation of high-affinity interstrand cross-links at 5′-Pu-GATC-Py-3′ sequences as suggested by the results reported here. Such lesions are presumably difficult for tumor cells to repair, thereby leading to apoptosis. However, the data reported here also raise the possibility that a number of other lower-affinity interstrand cross-linked and monocovalent sites may play a role in the mechanism of action of this agent.

The knowledge gained from these studies will inform the design of next-generation agents aimed at producing lethal non-repairable interstrand cross-links in longer DNA targets. This strategy differs from previous small-molecule approaches to gene targeting and DNA recognition (e.g., hairpin polyamides or information-reading lexitropsins) that rely upon non-covalent interaction with DNA, and offers the potential to modulate gene expression by blocking transcription in the coding region (*9*).

## ACKNOWLEDGMENTS

Drs. Gyoung-Dong Kang, Stephen J. Gregson, and Philip W. Howard of Spirogen Ltd. are acknowledged for synthesizing and supplying SJG-136 and the control dimer GD113 (**4**). Ipsen and Spirogen Ltd. are thanked for permission to use SJG-136 for these experiments.

## SUPPORTING INFORMATION AVAILABLE

Further information is available relating to the analysis of intensities of footprints in Figure 4 (**SI-1**), the method of analysis of the intensity of T-stops in Figures 5−7 (**SI-2**), the results (gel) of a T-stop assay showing the concentration-dependent effect of GD113 (**4**) on the inhibition of T7 polymerase at concentrations ranging from 0.5 to 50 $\mu$M (**SI-3**), and an analysis of intensities of footprints in Figure 9B (**SI-4**). This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Thurston, D. E. (1993) in *Molecular Aspects of Anticancer Drug−DNA Interactions* (Neidle, S., and Waring, M. J., Eds.) pp 54−88, The Macmillan Press Ltd., London.
2. Puvvada, M. S., Hartley, J. A., Jenkins, T. C., and Thurston, D. E. (1993) A quantitative assay to measure the relative DNA-binding affinity of pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) antitumor antibiotics based on the inhibition of restriction-endonuclease *Bam*H1, *Nucleic Acids Res. 21*, 3671−3675.
3. Puvvada, M. S., Forrow, S. A., Hartley, J. A., Stephenson, P., Gibson, I., Jenkins, T. C., and Thurston, D. E. (1997) Inhibition of bacteriophage T7 RNA polymerase in vitro transcription by DNA-binding pyrrolo[2,1-*c*][1,4]benzodiazepines, *Biochemistry 36*, 2478−2484.
4. Thurston, D. E., and Bose, D. S. (1994) Synthesis of DNA-interactive pyrrolo[2,1-*c*][1,4]benzodiazepines, *Chem. Rev. 94*, 433−465.
5. Kamal, A., Rao, M. V., Laxman, N., Ramesh, G., and Reddy, G. S. K. (2002) Recent developments in the design, synthesis, and structure−activity relationship studies of pyrrolo[2,1-*c*][1,4]-benzodiazepines as DNA-interactive antitumour antibiotics, *Curr. Med. Chem.−Anti-Cancer Agents 2*, 215−254.
6. Thurston, D. E., Bose, D. S., Howard, P. W., Jenkins, T. C., Leoni, A., Baraldi, P. G., Guiotto, A., Cacciari, B., Kelland, L. R., Foloppe, M. P., and Rault, S. (1999) Effect of A-ring modifications on the DNA-binding behavior and cytotoxicity of pyrrolo[2,1-*c*]-[1,4]benzodiazepines, *J. Med. Chem. 42*, 1951−1964.
7. Gregson, S. J., Howard, P. W., Barcella, S., Nakamya, A., Jenkins, T. C., Kelland, L. R., and Thurston, D. E. (2000) Effect of C2/C3-endo unsaturation on the cytotoxicity and DNA-binding reactivity of pyrrolo[2,1-*c*][1,4]benzodiazepines, *Bioorg. Med. Chem. Lett. 10*, 1849−1851.
8. Gregson, S. J., Howard, P. W., Corcoran, K. E., Barcella, S., Yasin, M. M., Hurst, A. A., Jenkins, T. C., Kelland, L. R., and Thurston, D. E. (2000) Effect of C2-exo unsaturation on the cytotoxicity and DNA-binding reactivity of pyrrolo[2,1-*c*][1,4]benzodiazepines, *Bioorg. Med. Chem. Lett. 10*, 1845−1847.
9. Thurston, D. E. (1999) Nucleic acid targeting: Therapeutic strategies for the 21st century, *Br. J. Cancer 80*, 65−85.
10. Bose, D. S., Thompson, A. S., Ching, J. S., Hartley, J. A., Berardini, M. D., Jenkins, T. C., Neidle, S., Hurley, L. H., and Thurston, D. E. (1992) Rational design of a highly efficient irreversible DNA interstrand cross-linking agent based on the pyrrolobenzodiazepine ring-system, *J. Am. Chem. Soc. 114*, 4939−4941.
11. Thurston, D. E., Bose, D. S., Thompson, A. S., Howard, P. W., Leoni, A., Croker, S. J., Jenkins, T. C., Neidle, S., Hartley, J. A., and Hurley, L. H. (1996) Synthesis of sequence-selective C8-linked pyrrolo[2,1-*c*][1,4]benzodiazepine DNA interstrand cross-linking agents, *J. Org. Chem. 61*, 8141−8147.
12. Bose, D. S., Thompson, A. S., Smellie, M., Berardini, M. D., Hartley, J. A., Jenkins, T. C., Neidle, S., and Thurston, D. E. (1992) Effect of linker length on DNA-binding affinity, cross-linking efficiency, and cytotoxicity of C8-linked pyrrolobenzodiazepine dimers, *Chem. Commun.* 1518−1520.
13. Smellie, M., Kelland, L. R., Thurston, D. E., Souhami, R. L., and Hartley, J. A. (1994) Cellular pharmacology of novel C8-linked anthramycin-based sequence-selective DNA minor-groove cross-linking agents, *Br. J. Cancer 70*, 48−53.
14. Mountzouris, J. A., Wang, J. J., Thurston, D., and Hurley, L. H. (1994) Comparison of a DSB-120 DNA interstrand cross-linked adduct with the corresponding bis-tomaymycin adduct−an example of a successful template-directed approach to drug design based upon the monoalkylating compound tomaymycin, *J. Med. Chem. 37*, 3132−3140.
15. Jenkins, T. C., Hurley, L. H., Neidle, S., and Thurston, D. E. (1994) Structure of a covalent DNA minor-groove adduct with a pyrrolobenzodiazepine dimer−evidence for sequence-specific interstrand cross-linking, *J. Med. Chem. 37*, 4529−4537.
16. Adams, L. J., Jenkins, T. C., Banting, L., and Thurston, D. E. (1999) Molecular modelling of a sequence-specific DNA-binding agent based on the pyrrolo[2,1-*c*][1,4]benzodiazepines, *Pharm. Pharmacol. Commun. 5*, 555−560.
17. Walton, M. I., Goddard, P., Kelland, L. R., Thurston, D. E., and Harrap, K. R. (1996) Preclinical pharmacology and antitumor activity of the novel sequence-selective DNA minor-groove cross-linking agent DSB-120, *Cancer Chemother. Pharmacol. 38*, 431−438.
18. Gregson, S. J., Howard, P. W., Corcoran, K. E., Jenkins, T. C., Kelland, L. R., and Thurston, D. E. (2001) Synthesis of the first example of a C2−C3/C2′-C3′-endo-unsaturated pyrrolo[2,1-*c*]-[1,4]benzodiazepine dimer, *Bioorg. Med. Chem. Lett. 11*, 2859−2862.
19. Gregson, S. J., Howard, P. W., Hartley, J. A., Brooks, N. A., Adams, L. J., Jenkins, T. C., Kelland, L. R., and Thurston, D. E. (2001) Design, synthesis, and evaluation of a novel pyrrolobenzodiazepine DNA-interactive agent with highly efficient cross-linking ability and potent cytotoxicity, *J. Med. Chem. 44*, 737−748.
20. Gregson, S. J., Howard, P. W., Jenkins, T. C., Kelland, L. R., and Thurston, D. E. (1999) Synthesis of a novel C2/C2′-exo unsaturated pyrrolobenzodiazepine cross-linking agent with remarkable DNA binding affinity and cytotoxicity, *Chem. Commun.* 797−798.
21. Smellie, M., Bose, D. S., Thompson, A. S., Jenkins, T. C., Hartley, J. A., and Thurston, D. E. (2003) Sequence-selective recognition of duplex DNA through covalent interstrand cross-linking: Kinetic and molecular modeling studies with pyrrolobenzodiazepine dimers, *Biochemistry 42*, 8232−8239.
22. 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., and Thurston, D. E. (2004) SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity: Part 1: Cellular pharmacology, in vitro and initial in vivo antitumor activity, *Cancer Res. 64*, 6693−6699.
23. Pepper, C. J., Hambly, R. M., Fegan, C. D., Delavault, P., and Thurston, D. E. (2004) The novel sequence-specific DNA cross-linking agent SJG-136 (NSC 694501) has potent and selective in vitro cytotoxicity in human B-cell chronic lymphocytic leukemia cells with evidence of a p53-independent mechanism of cell kill, *Cancer Res. 64*, 6750−6755.
24. 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., and Sausville, E. A. (2004) SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumour activity: Part 2: Efficacy evaluations, *Cancer Res. 64*, 6693−6699.
25. Neidle, S., Puvvada, M. S., and Thurston, D. E. (1994) The relevance of drug−DNA sequence specificity to antitumor activity, *Eur. J. Cancer 30A*, 567−568.
26. Neidle, S., and Thurston, D. E. (1994) in *New Targets for Cancer Chemotherapy* (Kerr, D. J., and Workman, P., Eds.) pp 159−175, CRC Press Ltd, Boca Raton, FL.
27. Woynarowski, J. M., Herzig, M. C., Trevino, A. V., Rodriguez, K. A., Hurley, L., and Hardies, S. C. (2001) Identification of cellular targets for DNA-reactive antitumor drugs by a combination of bioinformatics and molecular pharmacology, *Clin. Cancer Res. 7*, 705.
28. Woynarowski, J. M. (2002) Targeting critical regions in genomic DNA with AT-specific anticancer drugs, *Biochim. Biophys. Acta 1587*, 300−308.
29. Woynarowski, J. M., Hardies, S. C., Trevino, A. V., and Arnett, B. M. (1998) Targeting of AT-rich islands in genomic DNA by bizelesin, an AT-specific DNA-reactive antitumor drug, *Annals Oncol. 9*, 528.
30. Trauger, J. W., and Dervan, P. B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes, in *Drug-Nucleic Acid Interactions*, pp 450−466, Academic Press Inc., San Diego.

31. Low, C. M. L., Drew, H. R., and Waring, M. J. (1984) Sequence-specific binding of echinomycin to DNA—evidence for conformational changes affecting flanking sequences, *Nucleic Acids Res. 12*, 4865−4879.
32. Maxam, A. M., and Gilbert, W. (1977) A new method for sequencing DNA, *Proc. Natl. Acad. Sci. U.S.A. 74*, 560−564.
33. Komura, J., and Riggs, A. D. (1998) Terminal transferase-dependent PCR: A versatile and sensitive method for in vivo footprinting and detection of DNA adducts, *Nucleic Acids Res. 26*, 1807−1811.
34. Cesarone, C. F., Bolognesi, C., and Santi, L. (1979) Improved microfluorometric DNA determination in biological material using 33259 Hoechst, *Anal. Biochem. 100*, 188−197.
35. McGurk, C. J., McHugh, P. J., Tilby, M. J., Grimaldi, K. A., and Hartley, J. A. (2001) Measurement of covalent drug−DNA interactions at the nucleotide level in cells at pharmacologically relevant doses, in *Methods in Enzymology*, pp 358−376, Academic Press Inc., San Diego.
36. Lavesa, M., and Fox, K. R. (2001) Preferred binding sites for [*N*-MeCys(3),*N*-MeCys(7)]TANDEM determined using a universal footprinting substrate, *Anal. Biochem. 293*, 246−250.
37. Trauger, J. W., Baird, E. E., Mrksich, M., and Dervan, P. B. (1996) Extension of sequence-specific recognition in the minor groove of DNA by pyrrole-imidazole polyamides to 9−13 base pairs, *J. Am. Chem. Soc. 118*, 6160−6166.
38. Gregson, S. J., Howard, P. W., Gullick, D. R., Hamaguchi, A., Corcoran, K. E., Brooks, N. A., Hartley, J. A., Jenkins, T. C., Patel, S., Guille, M. J., and Thurston, D. E. (2004) Linker length modulates DNA cross-linking reactivity and cytotoxic potency of C8/C8′ ether-linked C2-exo-unsaturated pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) dimers, *J. Med. Chem. 47*, 1161−1174.
39. Skinner, G. M., Baumann, C. G., Quinn, D. M., Molloy, J. E., and Hoggett, J. G. (2004) Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase—a single-molecule view of the transcription cycle, *J. Biol. Chem. 279*, 3239−3244.
40. Farmer, J. D., Rudnicki, S. M., and Suggs, J. W. (1988) Synthesis and DNA cross-linking ability of a dimeric anthramycin analogue, *Tetrahedron Lett. 29*, 5105−5108.

BI0479813