Recognition of Specific Sequences in DNA by a Topoisomerase I Inhibitor Derived from the Antitumor Drug Rebeccamycin

CHRISTIAN BAILLY, PIERRE COLSON, CLAUDE HOUSSIER, ELISABETE RODRIGUES-PEREIRA, MICHELLE PRUDHOMME, and MICHAEL J. WARING

Laboratoire de Pharmacologie Moléculaire Antitumorale du Centre Oscar Lambret, Institut National de la Santé et de la Recherche Médicale Unité 124, 59045 Lille, France (C.B.), Laboratoire de Chimie Macromoléculaire et Chimie Physique, Université de Liège, Liège 4000, Belgium (P.C., C.H.), Synthèse, Electrolyse et Etude de Systèmes à Intérêt Biologique, Université Blaise Pascal, Centre National de la Recherche Scientifique UMR 6504, 63177 Aubière cedex, France (E.R.-P., M.P.) and University of Cambridge, Department of Pharmacology, Cambridge CB2 1QJ, UK (M.J.W.)

Received July 23, 1997; Accepted September 18, 1997

This paper is available online at http://www.molpharm.org

ABSTRACT

We investigated the interaction with DNA of two synthetic derivatives of the antitumor antibiotic rebeccamycin: R-3, which is a potent topoisomerase I inhibitor and contains a methoxyglucoside moiety appended to the indolocarbazole chromophore, and its aglycone, R-4. Spectroscopic measurements indicate that R-3 intercalates into DNA and that its carbohydrate domain contributes significantly to reinforce the affinity for DNA. Two complementary ligation assays concur that R-3, but not its aglycone counterpart, exerts a significant effect on the curvature and/or the flexibility of DNA. The sugar moiety may be responsible for preferential binding of R-3 to circular (or bent) DNA molecules as opposed to linear DNA fragments. The sequence selectivity of binding to DNA has been studied thoroughly by footprinting with DNase I and two other nucleases. The glycosylated compound is highly selective for nucleotide sequences containing GpT or TpG steps. The derivative lacking the sugar moiety on the indolocarbazole chromophore binds at essentially identical sites but with considerably lower affinity, so it seems that the chromophore rather than the carbohydrate is responsible for the preferential binding to sequences surrounding GpT and TpG steps. The influence of the exocyclic substituents present on the bases at the recognition sites (i.e., the 2-amino group of guanine and the 5-methyl group of thymine) was evaluated using two series of modified DNA molecules prepared by polymerase chain reaction containing inosine and/or 2,6-diaminopurine and uridine and/or 5-methylcytosine residues. The introduction of the amino group onto purine residues or the addition of a methyl group to pyrimidine residues suffices to create new drug binding sites. Therefore, unlike most DNA-binding small molecules, the rebeccamycin analogue seems to be highly sensitive to any modification of the exocyclic substituents on the bases in both the major and minor grooves of the double helix. The footprinting profiles with the different DNA fragments bear a remarkable resemblance to those determined for nogalamycin and bisnaphthalimide compounds known to recognize their preferred GpT and TpG sites via intercalation from the major groove. The unique DNA binding characteristics of the rebeccamycin analogue correlate well with its inhibitory effects on topoisomerase I.

The camptothecin derivatives irinotecan (CPT-11) and topotecan recently introduced in cancer chemotherapy are arguably the best characterized topoisomerase I poisons (Pommier and Tanizawa, 1993). These drugs interfere with the breakage-rejoining reaction by stabilizing a covalent topoisomerase I/DNA intermediate (usually referred to as cleavable complex), which can be detected by the appearance of DNA strand breaks on denaturation of the protein with a detergent. So far, relatively few other drugs have been shown to promote topoisomerase I-mediated DNA cleavage. Intoplicine, saintopins, and related naphthacene-dione antibiotics as well as alkaloids such as bulgarin and coraline (Gatto et al., 1996; Fujii et al., 1993, 1997; Leteurtre et al., 1994; Makhey et al., 1994; Nabiev et al., 1994) have been identified as inducers of the cleavable topoisomerase I/DNA complex. In the course of a screening program, Nakano and collaborators discovered that indolocarbazole derivatives could also promote DNA cleavage (Yamashita et al., 1994; Yoshinari et al., 1993). Subsequently, they synthesized water-soluble analogues such as compound NB-506 (Fig. 1), which proved to be not only a potent topoisomerase I poison but also a promising

ABBREVIATIONS: bp, base pair; l, inosine; U, uracil; DAP, 2,6-diaminopurine; PCR, polymerase chain reaction; 5MeC, 5-methylcytosine.
antitumor drug showing remarkable activity against colon and lung cancer xenografts as well as very low toxicity (Ara-kawa et al. 1995; Yoshinari et al., 1995).

In a recent study, we investigated the effects of a series of indolocarbazole derivatives structurally related to the antibiotic rebeccamycin (Bush et al., 1987), which is a cousin of NB-506 (Fig. 1). We found that the presence of chlorine on the indolocarbazole chromophore (as in rebeccamycin) significantly reduces the effect on topoisomerase I, whereas the substituents on the maleimido function and the functional group on the nonindolic moiety can be varied without loss of activity. In addition, we showed that the methoxyglucose residue attached to the chromophore plays a determinant role in facilitating interaction of the drug with DNA/topoisomerase I complexes (Bailly et al., 1997; Rodrigues-Pereira et al., 1996). However, the exact mechanism of interaction between these compounds and DNA with or without topoisomerase I remains unclear and poorly documented. To gain further insight into their binding to DNA and possible recognition of specific nucleotide sequences, we investigated the compound R-3 shown in Fig. 1, which is the most potent known topoisomerase I inhibitor among the indolocarbazole derivatives (Bailly et al., 1997). The interaction of this indolocarbazole derivative with DNA has been examined by a combination of spectroscopic and biochemical methods. Important information concerning the contribution of its sugar residue to the interaction with DNA has been obtained by comparing the binding properties of R-3 with those of its aglycone R-4.

**Materials and Methods**

**Drugs.** The syntheses of the indolocarbazole R-3 and its aglycone derivative, R-4, have been reported previously (Rodrigues-Pereira et al., 1996). In the dry state, the drugs were stored in a desiccator in the dark at 4°. Ligand concentrations were determined by direct weighing. The two drugs were dissolved in dimethylsulfoxide at 3 mg/ml and then diluted further with water. Fresh dilutions were prepared immediately before use. The final dimethylsulfoxide concentration never exceeded 0.3% (v/v) and were always made under conditions in which dimethylsulfoxide (also present in the controls) is known not to affect nuclease activity (Drew and Travers, 1984).

**Chemicals and biochemicals.** Calf thymus DNA and the double-stranded polymers poly(dA/dT)·poly(dA/dT) and poly(dG/dC)·poly(dG/dC) were from Sigma (La Verpillière, France). Their concentrations were determined through the application of molar extinction coefficients of 6600, 6600, and 8400 M⁻¹ cm⁻¹, respectively. Calf thymus DNA was deproteinized with sodium dodecyl sulfate (protein content, <0.2%), and all nucleic acids were dialyzed against 1 m[scap]m sodium cacodylate buffer, pH 6.5. Nucleoside triphosphates labeled with [α-³²P]dATP and [γ-³²P]ATP were obtained from Amersham International (Buckinghamshire, UK). Restriction endonucleases AvaI, EcoRI, and PvuII; alkaline phosphatase; T4 polynucleotide kinase; avian myeloblastosis virus reverse transcriptase, T4 DNA ligase; and exonuclease III were purchased from Boehringer-Mannheim Biochemica (Mannheim, Germany) and used according to the supplier’s recommended protocol in the activity buffer provided. All other chemicals were analytical grade reagents, and all solutions were prepared using doubly deionized, filtered water.

**Absorption spectroscopy and estimation of binding con- stants.** Absorption spectra were recorded on a Perkin-Elmer Cetus (Norwalk, CT) Lambda 5 spectrophotometer using a 10-mm optical pathlength. Titration of the drugs with DNA, covering a wide range of drug/DNA-phosphate ratios, were performed by the addition of aliquots of a concentrated DNA solution to a drug solution at a constant ligand concentration (20 μM). Binding parameters were determined using experimental spectrophotometric readings from absorbance titration experiments conducted at 320 nm. The apparent association constant K (M⁻¹) and number of sites per nucleotide (n) were estimated from Scatchard plots using two models: (1) the noncooperative overlapping binding site model of McGhee and von Hippel (1974) and (2) a two-site model that assumes the existence of two independent types of noncooperative binding sites. A better adjustment of the parameters to fit the experimental data could be obtained with the McGhee-von Hippel model than with the two-site model. The program Inplot 4 (Graphpad, San Diego, CA) (Leatherbarrow, 1990) was used to obtain the best fit of the data to each of these two models.

Circular dichroism measurements were recorded on a Jobin-Yvon CD6 dichrograph. Solutions of drugs and/or nucleic acids in 1 mM sodium cacodylate buffer, pH 6.5, were scanned in 1-cm quartz cuvettes. Measurements were made by progressive dilution of a drug/DNA complex at a high P/D ratio (P and D are nucleotide and drug concentrations, respectively) with a pure ligand solution to yield the desired drug/DNA ratios. Three scans were accumulated and averaged automatically.

Electric linear dichroism measurements were performed using a computerized optical measurement system built by C. Houssier (Houssier and O’Konasi, 1981). The procedures outlined previously were followed (Houssier, 1981). The optical setup incorporating a high-sensitivity T-jump instrument equipped with a Glan polarizer was used under the conditions of a bandwidth of 3 nm, sensitivity limit of 0.001 in ΔA/A, and response time of 3 msec. All experiments were conducted at 20° with a 10-mm-pathlength Kerr cell with 1.5-mm electrode separation in 1 mM sodium cacodylate buffer, pH 6.5.

![Fig. 1. Chemical structure of NB-506, rebeccamycin, and the two related compounds used in this study.](image-url)
DNA purification and labeling. Plasmids pBS (Stratagene, La Jolla, CA) and pKMp27 (Drew et al., 1985) were isolated from *Escherichia coli* by a standard sodium dodecyl sulfate-sodium hydroxide lysis procedure and purified by banding in CsCl-ethidium bromide gradients. Ethidium was removed through several isopropanol extractions followed by exhaustive dialysis against Tris-EDTA buffer. The purified plasmid then was precipitated and resuspended in appropriate buffer before digestion by the restriction enzymes. The two pBS DNA fragments were prepared by 3'- end labeling of the EcoRI/PvuII double digest of the plasmid using α-32P-dATP (6000 Ci/mmol) and avian myeloblastosis virus reverse transcriptase or by 5'-end labeling of the EcoRI/alkaline phosphatase-treated plasmid using γ-32P-ATP (6000 Ci/mmol) and T4 polynucleotide kinase followed by treatment with PvuII. Similarly, the tyrT fragment was prepared by 5'- or 3'- end labeling of the EcoRI/Aval digest of plasmid pKMp27. In each case, the digestion products were separated on a 6% polyacrylamide gel under native conditions in Tris/borate/EDTA buffer (89 mM Tris-borate, pH 8.3, 1 mM EDTA). After autoradiography, the band of DNA was excised, crushed, and soaked in elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate) overnight at 37°. This suspension was filtered through a 0.22-μm Millipore filter, and the DNA was precipitated with ethanol. After washing with 70% ethanol and vacuum-drying of the precipitate, the labeled DNA was resuspended in 10 mM Tris, adjusted to pH 7.0, containing 10 mM NaCl.

Preparation of DNA fragments containing modified bases (Baillie and Waring, 1995a). Molecules containing normal or modified bases (with I, DAP, 5-methylcytosine, or U in place of guanosine, adenine, cytosine, or thymine, respectively) were synthesized by the PCR using the primers 5'-AATTCCGGTTACCTTTATG and 5'-TCGG- GAAACCCCAACACGGG bearing a 5'-OH or 5'-NH₂ terminal group to permit 5'-phosphorylation of one strand only. The 160-bp tyrT(A93) fragment containing the *E. coli* tyrT promoter (Drew and Travers, 1984) was cut out of plasmid pKMp27 (Drew et al., 1985) by digestion with restriction enzymes EcoRI and Aval. This template bore a 5'-phosphate due to the action of EcoRI, and thus only the newly synthesized DNA (with normal or modified nucleotides) could be labeled by the kinase. Twenty amplification cycles were performed, with each cycle consisting of the following segments: (1) for normal, DAP-DNA, U-DNA, MeC-DNA, and U + MeC DNA: 94° for 1 min, 37° for 2 min, and 72° for 10 min; and (2) for I-DNA and I + DAP-DNA: 84° for 1 min, 30° for 2 min, and 62° for 10 min. The purified PCR products were 5'-end labeled with γ-32P-ATP in the presence of T4 polynucleotide kinase, and the labeled DNA was isolated by 6% polyacrylamide gel electrophoresis.

DNA circularization assay. Experiments were performed with either the 32P-labeled 160-bp tyrT(A93) fragment from plasmid pKMp27 or directly with the plasmid linearized with EcoRI. The experimental protocol has been described recently (Baillie et al., 1996c). Briefly, each sample consisted of 3 μl of DNA containing 0.5 μg of linearized plasmid DNA or ~0.3 ng (50 cps) of the 160-mer 32P-labeled EcoRI restriction fragment, 3 μl of water, 10 μl of drug at the desired concentration (or water in the controls), and 2 μl of 10X ligase buffer. After a 30-min incubation to ensure equilibration, 1 μl (5 units) of ligase was added to each tube, and the reaction was continued at room temperature (~20°) for 30 min. The ligase was denatured by heating at 65° for 5 min, and samples were electrophoresed immediately. To verify that the ligation products corresponded to circular DNA molecules, 1 μl of exonuclease III (50 units) was added to the ligase-treated sample (after the heat denaturation step) and incubated for 30 min at 37° before electrophoresis.

DNase I (EC 3.1.21.1) footprinting experiments were performed essentially as described previously (Baillie and Waring, 1995b). Briefly, reactions were conducted in a total volume of 10 μl. Samples (3 μl) of the labeled DNA fragment were incubated with 5 μl of the buffer solution containing the ligand at appropriate concentration. After a 30-min incubation at 37° to ensure equilibration of the binding reaction, the digestion was initiated by the addition of 2 μl of a DNase I solution whose concentration was adjusted to yield a final enzyme concentration of ~0.01 unit/ml in the reaction mixture. After 3 min, the reaction was stopped by freeze-drying. Samples were lyophilized and resuspended in 5 μl of an 80% formamide solution containing tracking dyes. The DNA samples then were heated at 90° for 4 min and chilled in ice for 4 min before electrophoresis.

Electrophoresis and quantification by storage phosphor imaging. DNA cleavage products were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3-mm-thick 8% acrylamide containing 8 x urea). After electrophoresis (~2.5 hr at 60 W, 1600 V in Tris/borate/EDTA buffer), gels were soaked in 10% acetic acid for 10 min, transferred to Whatman (Maidstone, UK) 3MM paper, and dried under vacuum at 80°. A Molecular Dynamics 425E PhosphorImager (Sunnyvale, CA) was used to collect data from the storage screens exposed to dried gels overnight at room temperature. Base-line-corrected scans were analyzed by integrating all the densities between two selected boundaries using ImageQuant software (v. 3.3; Sunnyvale, CA). Each resolved band was assigned to a particular bond within the DNA fragment by comparison of its position relative to sequencing standards generated by treatment of the DNA with dimethylsulfate (G) and/or formic acid (G + A) followed by piperidine-induced cleavage at the modified bases in DNA.

Results

DNA binding. Spectrophotometry reveals that the binding of the rebeccamycin derivative R-3 to DNA results in a large absorbance decrease in the 280–360-nm spectral region of the drug, whereas the addition of DNA to the sugar-free indolocarbazole analogue R-4 has little effect on its absorption spectrum (Fig. 2). A 28% hypochromism is observed in the 330-nm band of R-3 (at P/D = 30), whereas the hypochromism does not exceed 6% with R-4 precluding an accurate estimate of the binding affinity for this compound. The affinity of R-3 for calf thymus DNA was determined from the binding measurements using Scatchard analysis (Fig. 3). We could secure an acceptable fit to the data using the orthodox McGhee-von Hippel analysis (McGhee and von Hippel, 1974) but not with other models for a single class of binding sites. The binding constant calculated for R-3 is 1.75 ± 0.035 × 10^5 M⁻¹. It falls within the range of values reported for weak intercalating drugs, such as the acridine derivative amscarmine (Wilson et al., 1981). Although the strength of interaction between R-4 and DNA is too weak to measure accurately, we can estimate that the affinity constant for the aglycone must be ≥1 order of magnitude lower than that of R-3. There thus is no doubt that the carbohydrate moiety of R-3 contributes positively and significantly to the binding affinity.

Intercalation into DNA. Electric linear dichroism was used to compare the mode of binding of rebeccamycin (R-1), R-3, and R-4. As shown in Fig. 4A, the reduced dichroism measured at 320 nm with rebeccamycin or R-3 bound to DNA is slightly higher than that obtained with DNA alone at 260 nm, whereas that measured with R-4 in the presence of DNA is considerably weaker. These data indicate that in complexes formed with both drugs possessing the sugar residue, the chromophore is oriented parallel to the plane of the base pairs, as is the case with an intercalated drug. The results corroborate the initial hypothesis of Yamashita et al. (1992) and our recent measurements (Baillie et al., 1997) that compounds such as rebeccamycin and R-3 (which bear the same sugar) intercalate into DNA.

Circular dichroism measurements were carried out to com-
pare the binding of the drugs to poly(dA/dT) and poly(dG/dC). Unlike R-4, the ligand R-3 becomes optically active when it binds to DNA. The circular dichroism signal monitored at 320 nm in the presence of increasing DNA concentrations increases until a P/D value of 5 is reached and then remains more or less constant (Fig. 4B). The higher circular dichroism amplitude measured with poly(dG/dC) than with poly(dA/dT) suggests that the drug may prefer GC- to AT-rich sequences. Footprinting experiments confirm this prediction, as will be seen.

Effect on DNA structure and flexibility. For a first series of experiments, we resorted to the linear plasmid DNA ligation assay using T4 DNA ligase, which has been used to characterize the effect of intercalating agents, including adriamycin, amsacrine, and certain indolocarbazole derivatives (Yamashita et al., 1992). The linear pKmp27 DNA (cut with EcoRI) was treated with DNA ligase in the presence of increasing concentrations of R-3 and R-4. The gel in Fig. 5A...
shows unambiguously that the two indolocarbazole compounds have different effects on the rate of formation of circular DNA molecules and multimers. The ligation pattern observed with R-4, even at a high concentration, is indistinguishable from that seen in the control lanes with no drug present. In contrast, the addition of R-3 significantly changes the electrophoretic distribution of ligation products. With R-3, but not with R-4, a DNA species that comigrates with the supercoiled DNA (native plasmid) is formed. Moreover, R-3 is much more efficient than the aglycone at promoting the formation of supercoiled dimers. These data suggest that the carbohydrate moiety attached to the chromophore must increase the curvature of DNA and/or its intrinsic flexibility so as to facilitate the formation of circular products.

Next, to investigate this effect in greater detail, we applied a recently developed assay that also relies on the formation of DNA circles in the presence of T4 DNA ligase but uses a much shorter restriction fragment (Bailly et al., 1996c). Treatment of the 32P-labeled tyr T DNA fragment with ligase leads to the formation of 169-bp DNA circles identifiable by their complete resistance to digestion by exonuclease III. In the presence of rebeccamycin (R-1), the yield of circles remains more or less unchanged, although a slight decrease at 50 μM can be detected (Fig. 5B). Inhibition of the formation of DNA circles is much more pronounced with the aglycone R-4, which produces effects similar to those reported with intercalating drugs such as daunomycin and mitoxantrone (Bailly et al., 1996c). This could be attributed to an inhibition of the enzyme or, more likely, to a drug-induced stiffening of the helical rod. On the other hand, R-3 does not inhibit the ligase. Interestingly, we observe that with concentrations of R-3 of ≥10 μM, the electrophoretic mobility of the circular DNA species is considerably reduced, whereas that of the linear DNA remains unaffected. It seems that R-3 binds preferentially to the circular DNA species rather than the linear species.

The consistent observation that R-3 and R-4 interfere differently with the closure reaction confirms that they exert different effects on DNA structure. The data in Fig. 5, A and B, indicate that the methoxyglucose residue always affects the induction or recognition of circular DNA molecules by R-3, which implies that it contributes directly to processes connected with flexibility. Although the possible bending of DNA by R-3 remains speculative at present, it is quite conceivable that some such subtle effect on DNA structure could contribute to the mechanism by which the drug inhibits topoisomerase I.

**Sequence selective binding.** Footprinting studies were performed using the endonuclease DNase I, which has been used extensively in our laboratory over many years for mapping the DNA-binding sites of a large variety of drugs endowed with antimicrobial, antiviral, and antitumor properties (Waring and Bailly, 1994). Three different DNA restriction fragments, isolated from the plasmids pKMp27 and pBS and 3′- or 5′-end labeled on one or the other of the complementary strands, were used as substrates. A typical autoradiograph of a sequencing gel used to fractionate the products of partial digestion of the DNA complexed with R-3 is presented in Fig. 6. With the drug bound, the DNase I cleavage pattern differs significantly from that seen in the control lane. Numerous bands in the drug-containing lanes are weaker than the same bands in the drug-free lane, corresponding to attenuated cleavage, whereas others display relative enhancement of cutting. Many other gels (not shown)
cleavage at each internucleotide bond is affected by complexation with increasing concentrations of R-3. The intensity of each footprint (negative values) depends directly on the drug concentration; little or no preference can be detected at concentration of ≤15 μg/ml, and the footprinting profiles remain practically unchanged at concentrations of ≥60 μg/ml. The plots in Fig. 7B compare the footprinting data obtained with R-3 and its aglycone R-4 (both at 60 μg/ml). Although the footprints are much more pronounced with R-3 than with R-4, the two drugs bind at essentially identical sites. The weak footprints produced by R-4 surely correlate with its low affinity for DNA. Nevertheless, it seems that the sequence selectivity is driven principally by the indolocarbazole chromophore rather than by the carbohydrate moiety, which serves mainly to anchor the drug on the helix.

With both R-3 and its aglycone, the best binding sites correspond to sequences containing GC bp, often with a notable alternation of purine and pyrimidine nucleotides. In contrast, the susceptibility to DNase I cleavage seems substantially enhanced at AT-rich sequences, such as around nucleotide positions 30, 50, and 84. At first sight, it would seem that the binding of the indolocarbazole to GC sequences is favored over binding to AT or mixed sequences. However, closer inspection of the differential cleavage plots and the sequences protected from cleavage by R-3 in the 117- and 265-mer fragments (Fig. 7C) reveals that purely GC-rich sequences are not usually well protected from DNase I cleavage. A typical example can be seen in Fig. 7A, in which the sequence 5'-CGCGCCG from positions 73–79 corresponds to a region of enhanced cleavage. A search for the common denominator of the binding sites for R-3 suggests that with few exceptions, the rebeccamycin analogue is binding preferentially to 5'-YpGpT-ApCpR and 5'-RpTpG-CpApY sites but not to CpG or GpC steps. The strongest binding site in the 160-mer fragment encompasses the sequence TCTTACGTT, which contains three juxtaposed GpT/TpG steps. This GT-rich sequence was also protected from cleavage in experiments in which DNase II and micrococcal nuclease were used as footprinting probes (data not shown). With all three fragments tested, the sequences protected from DNase I cutting in the presence of R-3 contain at least one GpT or TpG site (Fig. 7C). In every experiment, the same sequences were found to be protected (weakly but significantly) by the addition of R-4, reinforcing the conclusion that the indolocarbazole chromophore, not the appended carbohydrate moiety, must be responsible to a large extent for the sequence selectivity.

Effects of the exocyclic substituents at TpG and GpT sites. In previous studies, we have shown that the exocyclic substituents of DNA bases, especially the 2-amino group of guanine, play a determinant role in specific recognition of DNA sequences by small molecules and proteins (Bailly et al., 1995, 1996c; Bailly and Waring, 1995a, 1995c). By analogy, we hypothesized that the exocyclic substituents at GpT and TpG sites must contribute to the selective binding of R-3 to TG/GT-containing sequences. This hypothesis was put to the test by synthesizing two series of DNA molecules in which the position of the exocyclic substituents, namely the guanine amino group and the thymine methyl group, are varied.

Together with the natural DNA fragment used as a control, the requisite series of different 160-bp DNA fragments containing modified bases were synthesized by PCR amplification with the aim of investigating the separate influence of using different concentrations of R-3 were run and used for the quantification.

To clarify the effect of R-3 on the rate of DNA cleavage by the nuclease, intensities from selected gel lanes were quantified by densitometry and converted into a set of differential cleavage plots (Fig. 7) that indicate the extent to which

![Figure 6](https://example.com/figure6.png)

**Fig. 6.** DNase I footprinting on the 160-, 117-, and 265-mer restriction fragments in the presence of R-3 at 20 μg/ml. In each case, the DNA was either 3'-end labeled at the EcoRI site with α-32P-dATP in the presence of avian myeloblastosis virus reverse transcriptase or 5'-end labeled at the EcoRI site with γ-32P-ATP in the presence of T4 polynucleotide kinase (3* or 5*). The products of nuclease digestion were resolved on an 8% polyacrylamide gel containing 7M urea. G, Guanine-specific sequence markers obtained by treatment of the DNA with dimethylsulfate followed by piperidine cleavage were run. Control tracks (Ct) contained no drug.
the 2-amino group of guanine lying in the minor groove and of the 5-methyl group of thymine exposed in the major groove (Fig. 8). In the first series, the substituted DNA contained either I residues in place of guanosines (G-to-I substitution) or DAP residues in place of adenines (A-to-DAP substitution) or both I and DAP residues. In the second series, the substituted DNA contained either U residues in place of thymines (T-to-U substitution) or 5-methyl cytidines in place of cytidines (C-to-MeC substitution) or both U and MeC residues. A Watson (upper strand) primer in which the 5'-terminal nucleotide residue bore a 5'-NH₂ terminal group instead of a conventional 5'-OH was used to enable selective labeling of the sense (Crick) strand in the PCR product. In each case, the DNA was incubated with R-3 for 20 min at 37° to establish equilibrium before exposure of the samples to the nicking activity of DNase I.

Fig. 7. Differential cleavage plots comparing the susceptibility of the 3'-labeled tyr T fragment to DNase I cutting in the presence of (A) increasing concentrations of R-3 and (B) R-3 and R-4 at 60 μg/ml. Negative values correspond to a ligand-protected site, and positive values represent enhanced cleavage. Vertical scale, units of ln(fc) ln(f_a), where f_a is the fractional cleavage at any bond in the presence of the drug, and f_c is the fractional cleavage of the same bond in the control (Bailly et al., 1997). C, Summary map representing the DNase I footprints detected with R-3 on both strands of the three DNA fragments studied: the 160-bp tyr T fragment from plasmid pKMp27 and the 117- and 265-bp EcoRI/PvuII restriction fragments from plasmid pBS. Only the region of the restriction fragments analyzed by densitometry is shown. Filled boxes, positions of inhibition of DNase I cutting in the presence of R-3 (presumptive binding sites inferred from differential cleavage plots). Data are compiled from quantitative analysis of several sequencing gels like the one shown in Fig. 6 and must be considered a set of averaged values.
Typical autoradiographs obtained with the normal and modified DNA fragments are shown in Figs. 9 and 10. Both types of substitutions markedly affect the recognition of specific sequences by R-3. In the DAP-substituted DNA, new footprints develop at sequences containing juxtaposed TD bp. For example, the pyrimidine sequence 5'-AAAG in normal DNA does not afford a receptor site for R-3, whereas the corresponding DAP-containing sequence, 5'-CTTT-DDDI, proves to be an excellent receptor for the drug. Similar trends appear at two other sites, DTD T and DTD C, indicated in Fig. 9. The footprinting patterns determined with the doubly substituted I + DAP DNA resemble those obtained with the DAP DNA, whereas those seen with the Inosine DNA do not differ greatly from those observed with normal DNA. The new binding sites at TpD or DpT sequences reinforce the belief that the drug prefers TpG or GpT steps in natural DNA. Indeed, viewed from the minor groove, TpG and GpT steps are equivalent to TpD and DpT steps, respectively, in terms of hydrogen-bond potential. Therefore, the data suggest that the purine 2-amino group normally present on guanine that occupies the minor groove of DNA acts as a positive element that drives the selective binding of R-3 to TpG or GpT sites.

Interestingly, the methyl group of thymine also seems to contribute significantly to the recognition process (Fig. 10). Removal of the methyl group apparently abolishes the preferential binding because no footprint can be detected with the uridine-containing DNA. In contrast, the C-to-MeC substitution results in the appearance of several strong binding sites at TpD or DpT in the first series and mCpG or GpmC in the second series, always is much more pronounced than that produced at the corresponding TpG or GpT sites in natural DNA. In conclusion, these footprinting experiments establish that both the guanine 2-amino group and the thymine methyl group, which project into the minor and major grooves of the double helix, respectively, contribute significantly to the recognition of GpT and TpG sites by the rebeccamycin analogue. This is very uncommon for a small molecule (see Discussion).

Discussion

The spectroscopic data reported here together with the previous results of others (Yamashita et al., 1992) as well as ourselves (Bailly et al., 1997) leave no room for doubt that rebeccamycin analogues such as R-3 intercalate into DNA. However, if the intercalation of the planar indolocarbazole chromophore is beyond dispute, one question remains concerning the location of the carbohydrate moiety when the drug is bound to DNA. Does the methoxyglucose residue of R-3 locate in the minor or the major groove of the double helix, or does it simply extend out from the helix? As far as the present data indicate, it must be acknowledged that the probable location of the sugar moiety has not been demonstrated directly, although it is certain that it contributes significantly to the interaction with nucleic acids. It would be hard to imagine that the carbohydrate residue plays no part in the interaction of the drug with its cellular targets, given the unequivocal evidence for the difference between R-3 and its aglycone R-4 regarding DNA-binding affinity but also in terms of topoisomerase I inhibition and cytotoxicity (Bailly et al., 1997; Rodrigues-Pereira et al., 1996).

At first sight, it seems logical to suggest that the carbohydrate occupies the minor groove when the drug chromophore is intercalated between consecutive base pairs. With the vast majority of DNA-binding small molecules equipped with carbohydrate groups, the sugar moiety nestles within the minor groove of the helix and participates positively in the DNA recognition process. This is the case for the well known anthracycline antibiotics (e.g., adriamycin) (Chaires, 1996) as well as the enediyne antibiotics (e.g., calicheamicin) (Kumar et al., 1997) and many other antitumor agents, such as mithramycin and chromomycin (Keniry et al., 1993). Moreover, the fact that addition, deletion, or relocation of the 2-amino group of guanine residues affects the sequence-selective binding of R-3 to DNA indicates that the drug must somehow sense the presence of a substituent on the edges of the bases in the minor groove. The exocyclic amino group is the only hydrogen-bond donor exposed in the minor groove, and we demonstrated previously that this group serves as a key element for minor groove recognition by small molecules as well as by proteins (Bailly et al., 1995, 1996c; Bailly and Waring, 1995a, 1995c). In this regard, the rebeccamycin analogue R-3 does not escape the rule because its interaction with TG/GT sites evidently depends on the position of the purine amino substituent in the minor groove. However, the totality of the situation cannot be so simple, and some involvement with the major groove cannot be excluded for several reasons. The chief finding is that as judged on the
basis of the footprinting data in Fig. 10, the methyl group of thymine must also contribute to the interaction between the drug and its preferred sites. With all the antibiotics and small molecules studied previously, relocation of the methyl group of thymine has had little or no influence on the sequence-selective interaction between the drug and DNA. For example, the methyl group can be eliminated completely without perturbing the AT-specific minor groove binding of netropsin and distamycin or the GC-specific intercalative binding of echinomycin and actinomycin, which also occupy the minor groove. So far as we are aware, this is the only known instance in which the transfer of the methyl group from thymines to cytosines (T-to-5MeC substitution) suffices to create new and strongly favored drug-binding sites. This key observation suggests that in one way or another, the drug must establish contacts with the DNA via the major groove of the double helix.

It is important to mention that the footprinting patterns determined here with R-3 are strongly reminiscent of those obtained previously with nogalamycin (Fox and Waring, 1986) and certain bisnaphthalimide derivatives (Bailly et al., 1996a). Nogalamycin is an intercalative antibiotic containing a planar anthracycline chromophore substituted with two bulky carbohydrate groups that come to lie simultaneously in the minor and major grooves of the double helix at DNA sequences containing GpT and/or TpG steps (Smith et al., 1995). Recently, we not only reported that a series of tumor active bisnaphthalimide derivatives exhibit a sharp selectivity for TG/GT-containing sequences but also presented strong evidence that these bisintercalating drugs recognize DNA sequences via major groove contacts (Bailly et al., 1996a). Our present footprinting studies show that R-3 shares with nogalamycin and the bisnaphthalimides the rare property of interacting selectively with TG/GT-containing sequences. The correspondence between the footprinting profiles obtained with R-3 and major groove binding drugs such as nogalamycin and the bisnaphthalimides, on the one hand, and the observation that the methyl group of thymine is a key element for the selective binding of R-3 to TG/GT sites, on the other, lead us to hypothesize that the sugar probably is located in the minor groove although the large planar chromophore extends well into the major groove of DNA. Although the elucidation of exactly how R-3 interacts with DNA must await more precise structural investigations (NMR and X-ray studies are in progress), we have adduced good reason to believe that the drug uses both exocyclic substituents of DNA exposed in the minor and the major grooves to recognize selectively TpG- and GpT-containing sequences. This unusual characteristic may be exploitable for targeting specific sites in genes.
the cleavage site (T at +5) is the only requirement for doxorubicin-stabilized cleavage of DNA by topoisomerase II (Capranico et al., 1990). Actinomycin D, which can inhibit both topoisomerase I and topoisomerase II, exhibits a sharp selectivity for GpC-containing sites (Bailly et al., 1994), but no specific sequence requirement for topoisomerase inhibition has been reported. Conversely, cleavage sites produced by topoisomerase I in response to camptothecin derivatives show a preponderance of G at the +1 position (Pommier et al., 1993), whereas the drug interacts loosely, if at all, with DNA in the absence of the enzyme. Thus, most studies have failed to find a correlation between the sequence selectivity of drug binding to (protein-free) DNA and effects on topoisomerases, but here the correlation is good. So far as we can tell, this is the first case in which the sequence selectivity of a drug coincides well with the topoisomerase I-mediated cleavage selectivity (Fig. 7A). It is refreshingy clear that the rebeccamycin analogue R-3 prefers TpG sequences, and we have recently shown that it stabilizes topoisomerase I preferentially at sites with a T and a G on the 5′ and 3′ side of the cleaved bond, respectively (Bailly et al., 1997). It would be a cruel coincidence if the preferred binding sites identified in our footprinting experiments were to play no part in determination of the effect of the drug on topoisomerase I cleavage reactions. Accordingly, the current results set the stage for the design of more effective rebeccamycin analogues.

Acknowledgments

C.B. and M.J.W. thank Chris Koncewicz for his helpful contribution to the footprinting experiments and Julie Morgan for expert technical assistance.

References


Fig. 10. DNase I footprinting of R-3 on the Crick strand of tyrT(A93) DNA containing the four natural nucleotides (normal) or uridine residues in place of thymidine (uridine), 5MeC residues in place of cytosine (5-MeC), or both U and methylcytosine residues in place of thymidine and cytosine (U + mC), respectively. For other details, see the legend to Fig. 9.

The last point to consider is the relationship between the drug-stimulated cleavage sites on DNA in the presence of topoisomerase I and the preferred drug-binding sites. In general, binding to DNA and topoisomerase inhibition are best viewed as two distinct molecular processes contributing separately to the cytotoxic activity because in nearly all cases, the known sequence selectivity of drug binding to DNA has little to do with the location of drug-induced topoisomerase II breaks. For example, the topoisomerase II inhibitor doxorubicin binds preferentially to ATGC and ATCG triplets (Chaires et al., 1990), whereas doxorubicin-induced topoisomerase II strand breaks can occur at many types of sites not necessarily encompassing the aforementioned triplets. The presence of an adenine residue at position −1 relative to
to 5'-A/T/C/G and 5'-A/T/G/C sequences revealed by footprinting titration experiments. *Biochemistry* **29**:6145–6153.


Send reprint requests to: Dr. Christian Bailly, IRCL–INSERM U124, Place de Verdin, 59045 Lille cedex, France. E-mail: bailly@lille.inserm.fr