Sequence requirements for mammalian topoisomerase II mediated DNA cleavage stimulated by an ellipticine derivative

Philippe Fossé, Brigitte René, Marc Le Bret, Claude Paoletti and Jean-Marie Saucier
Laboratoire de Biochimie-Enzymologie, CNRS URA 147, INSERM U 140, Institut Gustave Roussy,
94805 Villejuif cedex, France

Received March 20, 1991; Revised and Accepted April 30, 1991

ABSTRACT

Various antitumor drugs stabilize DNA topoisomerase II-DNA transient covalent complexes. The complexes distribution along pBR322 DNA was shown previously to depend upon the nature of the drug (Tewey et al. (1984) *Science* 226, 466 – 468). The position in pBR322 of DNA cleavage by calf DNA topoisomerase II for 115 such sites stabilized by an ellipticine derivative and the relative frequency of cleavage at most of these sites were determined. The nucleotide sequence surrounding the 25 strongest sites was analyzed and the following ellipticine specific consensus sequence was deduced: 5'-ANCNT(A/G)T^ˆNN(G/C)N(A/G)-3' where cleavage occurs at the indicated mark. A thymine is always present at the 3' end of at least one strand of the strong cleavage sites, and the dinucleotide AT or GT at the 3' end of the break plays a major role in the complex stabilisation. The predictive value of cleavage of the consensus was tested for two regions of SV40 DNA and cleavage was indeed detected at the majority of the sites matching the consensus. Some complexes stabilized by ellipticine are resistant to salt dissociation and this property seems to be correlated with the presence of symmetrical sequences in the cleavage site with a center of symmetry staggered relatively to the center of symmetry of cleavage.

INTRODUCTION

Several antitumor drugs such as anthracyclines, epipodophyllotoxins, *m*-AMSA and ellipticine derivatives interfere with the reaction catalyzed by DNA topoisomerase II, and evidence has accumulated to suggest that this interaction is directly related to their cytotoxicity (1, 2). DNA topoisomerase II is a nuclear enzyme which modifies DNA topology by breaking and resealing both DNA strands in concert, allowing the passage through the transient break of independent DNA segments. The catalytic reaction involves the formation of a transient covalent bond between a tyrosine residue of each enzyme subunit and the

5'- end created by cleavage of the DNA strands (3). *In vitro* studies have demonstrated that the inhibitors mentioned above stabilize this covalent complex termed cleavable complex (4–6). Previous studies have shown that the cleavable complexes distribution along a DNA molecule is not random, but exhibits a sequence specificity determined by the nature of the drug stabilizing the complexes (6–8). Since the frequency of DNA breaks required to induce a given cytotoxic effect varies for drugs belonging to different chemical series, one may speculate that the toxic effect of a cleavable complex varies according to its genomic location (1, 5, 9). Recently, some characteristic features of the sequence dependence of DNA cleavage induced by doxorubicin have been determined (8). In the present study, we analyzed the sequence specificity of DNA cleavage by calf DNA topoisomerase II in the presence of another antitumor DNA intercalating drug belonging to the ellipticine series. The alignment of sequences of 25 strong DNA cleavage sites stimulated by this ellipticine derivative enabled us to derive the consensus sequence: 5'-ANCNT(A/G)T^ˆNN(G/C)N(A/G)-3' (where the mark ^ˆ indicates the cleavage site and N is any base). The predictive value of this consensus sequence was assessed using independent sequences from SV40 DNA.

MATERIALS AND METHODS

Drugs, enzymes and materials

EPC (2, 11-dimethyl-5-ethyl-9-hydroxy-6H-pyrido[4,3-b]carbazolium) was provided by the SANOFI Co. (Sisteron, France) and dissolved in distilled water at a concentration of 2 mM just before use. Etoposide (VP-16) was kindly given by Dr. Kiechel (laboratories Sandoz, Rueil-Malmaison, France). *m*-AMSA (4'-(9-acridinylamino)-methanesulfon-*m*-anisidide) was provided by National Cancer Institute (Bethesda, MD). Stock solutions were made in dimethylsulfoxide at 10 mM and kept frozen at –20°C. Genistein was purchased from Extrasynthese laboratories (France) and dilutions were prepared as described previously (7). Restriction endonucleases, T4-polynucleotide kinase were purchased from New England Biolabs (Beverly, MA). Calf intestine alkaline phosphatase, *E. coli* DNA polymerase I (Klenow fragment) and proteinase K were

purchased from Boehringer Mannheim (Hamburg, FRG). DNA topoisomerase II was purified from calf thymus using previously published procedures (10, 11). Circular pBR322 DNA was purified from *E. coli* (*hsdR*, *hsdM*, *recA13*) according to published methods (12). SV40 DNA was purchased from Bethesda Research Laboratories (Gaithersburg, MD). Reagents for Maxam-Gilbert sequencing were from Dupont De Nemours Inc., NEN Products (Boston, MA). Adenosine [γ - 32 P] triphosphate and [α - 32 P] dNTP were purchased from Amersham (Bucks, UK). Autoradiographies were performed with R-X films (Fuji-Photo Film Co., Japan).

End-labeling and isolation of DNA restriction fragments

For preparation of 3'-end-labeled pBR322 fragments, circular pBR322 DNA was cut with *EcoR* I, *Hind* III or *Nde* I restriction endonucleases and then labeled with [α - 32 P]dATP using the large fragment of *E. coli* DNA polymerase I and unlabeled dTTP at 20°C for 20 min. The labeled DNA was purified by two cycles of ethanol precipitation and was further cut with a second enzyme: *Hind* III, *EcoR* I or *Pvu* II in the case of *EcoR* I, *Hind* III or *Nde* I labeling, respectively. This procedure generates in each case two fragments, a large and a small one which were both labeled at one end. These two DNA fragments were present in the reaction of topoisomerase II mediated DNA cleavage. For preparation of 5'-end-labeled DNA fragments, pBR322 DNA and SV40 DNA were first cut with a restriction endonuclease; then the 5'-DNA termini were dephosphorylated with calf intestine phosphatase and labeled with [γ - 32 P]ATP and T4 polynucleotide kinase. The labeled fragments were then subjected to a second enzyme digestion. This procedure generated DNA fragments uniquely 5'-end-labeled which were then isolated by electrophoresis on a 6% polyacrylamide gel followed by electroelution and ethanol precipitation.

Topoisomerase II mediated DNA cleavage reactions

End-labeled DNA restriction fragments (10^4 - 10^5 cpm) were incubated with topoisomerase II (1.1 μ g/ml), in the presence or absence of drug, in the cleavage buffer (40 mM Tris-HCl¹, pH 7.5, 0.5 mM DTT, 100 mM KCl, 10 mM MgCl₂, 0.5 mM EDTA, 1 mM ATP and 30 μ g/ml bovine serum albumin) for 15 min. at 37°C. The cleavage reaction was terminated by the addition of SDS followed by proteinase K to final concentrations of 0.4% and 0.1 mg/mL, respectively and samples were incubated for an additional 30 min. at 50°C. When DNA was loaded on agarose gels, cleavage reactions were made in 15 μ l and after proteinase K digestion, 5 μ l of loading buffer (0.05% bromophenol blue, 50 mM EDTA, 50% sucrose) was added to samples. The products of cleavage reactions were analyzed on 1% agarose gels for ca 15 h (2.5 V/cm) in TBE (89 mM Tris

Base, 89 mM boric acid, 2.5 mM EDTA, pH 8.3) at 11°C. After electrophoresis, gels were dried on a 3 MM paper sheet and autoradiographed with Fuji X-Ray films. When samples were loaded on sequencing gel, cleavage reactions were made in 60 μ l reaction volume and after proteinase K digestion, extraction with phenol-chloroform and ethanol precipitation were performed. The samples were resuspended into 3 μ l of loading buffer [80% (v/v) formamide, 20 mM EDTA, 0.1% bromophenol blue and 0.1% xylene cyanol] and were heated at 100°C for 3 min. before loading into sequencing gels. DNA sequencing reactions (13, 14) and the products of cleavage reactions were analyzed on 6, 8 or 10% polyacrylamide sequencing gels [29:1 acrylamide: bis (acrylamide) ratio] containing 7 M urea, run for 1-4 h (45 V/cm) in TBE. Gels were then dried and autoradiographed as above.

Quantification of DNA cleavage stimulated by drugs

To quantify the amount of topoisomerase II mediated DNA cleavage at one site, the bands on the autoradiograms, which represented unique cleavage sites, were scanned using a double beam recording microdensitometer MKIII C Joyce, Loeb and Co. Ltd. (U.K.). In order to determine the linear range of both the autoradiogram and the microdensitometer, different known amounts of [32 P] labeled pBR322 DNA fragments and the products of cleavage reactions were loaded on the same gel. Several exposures of each gel, ranging from 7 hours to 6 days, were performed and linear ranges were determined by plotting the cpm loaded versus the absorbance. By this procedure, quantification and comparison of bands of weak and strong intensities could be performed.

Figure 2. DNA double strand break cleavage patterns induced by calf thymus topoisomerase II in the presence of EPC. pBR322 DNA labeled with [32 P] at its *EcoR* I ends (A) or *Hind* III ends (B) was incubated in the presence of topoisomerase II and 6 μ M EPC as described in Materials and Methods. Size markers in base pairs are in lane M. C) *Ava* I-*Bam* H I pBR322 DNA fragment (3314 bp) 5'-end-labeled at the *Ava* I site was incubated with topoisomerase II and 6 μ M EPC for 15 min. (lanes 1 to 4). After incubation, NaCl at various concentrations (lanes 2 to 4) or cleavage buffer (lane 1) were added and the samples were incubated 30 min. at 37°C before SDS-proteinase K treatment. Final NaCl concentrations of samples in lanes 2 to 4 were 0.12, 0.25 and 0.50 M respectively. Control without NaCl, lane 1; topoisomerase II without drugs, lane 5; control without drugs and topoisomerase II, lane 6. Reaction products were analyzed on a 1% agarose gel and autoradiography was performed. The arrows, marked with Roman numerals, indicate some regions where the DNA cleavage is stimulated by EPC.

Figure 1. Chemical structure of EPC.

Data analysis

The position of topoisomerase II cleavage sites was determined at the nucleotide level by using sequencing gels. The [^{32}P] labeled DNA fragments generated by topoisomerase II cleavage have 3'-hydroxyl ends, and migrate slightly slower than the fragments with 3'-phosphoryl ends obtained by chemical cleavage (13, 15). Taking this fact into account, the phosphodiester bond cleaved by topoisomerase II could be determined unambiguously. Nucleotides surrounding the cleavage site were numbered in such a way that the 3'-terminal nucleotide is at position -1 and the 5'-terminal nucleotide is at position +1. Since in most of the cleavage sites no clear symmetry appears and the relative orientation of the sites is not known, both DNA strands have to be compared. Forty nucleotides long sequences were aligned, and the matrix of frequency with which each base occurs at a given position was determined. The information theory (16) was used to evaluate the significance of each position within a combination of aligned sequences. This method is applicable for genomes in which all bases are equally represented. In pBR322 DNA the frequencies of bases are close to 0.25 ($f(\text{A})=f(\text{T})=0.2314$ and $f(\text{C})=f(\text{G})=0.2686$). The 'information content' (called R_{sequence}) at each position L derived from the formula (16):

$R_{\text{sequence}}(L) = H_g - H_s(L)$ where H are the uncertainties and

$$H_g = - \sum_{B=A}^T P(B) \log_2 P(B); H_g = 2 \text{ when } P(B) = 0.25$$

(The frequencies of bases are used as estimates for probabilities $P(B)$).

$$H_s(L) = - \sum_{B=A}^T f(B,L) \log_2 f(B,L); f(B,L) \text{ is the frequency}$$

of a base in aligned sequences. $R_{\text{sequence}}(L)$ values are between zero (random) and two (total conservation of a single base). The sum of $R_{\text{sequence}}(L)$ was calculated for all the possible

Figure 3. Map of cleavage regions stimulated by EPC in pBR322. Strong, medium and weak cleavage regions are respectively indicated by large, medium and small arrows. Shaded areas represent sequenced regions where the position of cleavage sites was exactly determined.

combinations of one strand of 25 sites and the combination giving the higher score was considered as the best alignment for the series of sites studied. Calculations were carried out on a Silicon Graphics IRIS 4D 70/G T workstation, using a program developed in our laboratory.

RESULTS

Mapping of the ellipticine-stimulated cleavage regions in pBR322

A previous study (17) has shown that ellipticine and 2-Methyl-9-hydroxyellipticinium stimulate topoisomerase II mediated pBR322 DNA cleavage with significant differences between the ellipticine and m-AMSA stimulated cleavage pattern. In the present work, the site specificity of cleavage of a 5-ethyl substituted ellipticine derivative termed EPC (Fig. 1) which displays a high potency in stimulating DNA cleavage over a broad concentration range with the same pattern as those observed with other ellipticine derivatives (manuscript in preparation) was studied. In order to map accurately all the ellipticine-stimulated double strand DNA cleavage sites, the reaction was carried out at a concentration of 6 μM of EPC with various end-labeled DNA fragments covering the whole pBR322 molecule (*Hind* III-*Eco* R I fragment labeled at its *Eco* R I or *Hind* III end; *Ava* I-*Bam* H I fragment labeled at its *Ava* I end; Fig. 2). In these experiments,

Figure 4. Sequence analysis of cleavage sites stimulated by EPC in the regions I, V, VI and VII. Topoisomerase II cleavage assay was carried out as described in Materials and Methods. Reaction products were analyzed on a 10% (A) or 8% (B) denaturing polyacrylamide gel. The pBR322 DNA fragments used were *Eco* R I-*Rsa* I (163 bp) (A) and *Taq* I-*Bgl* I (908bp) (B) which were 5'-end-labeled at the *Eco* R I and *Taq* I sites, respectively. At the top of the gels are indicated the drugs studied and Maxam-Gilbert reactions (C and G+A). Topo II: topoisomerase II without drugs, control: without drugs and enzyme. Drugs concentrations were 6 μM EPC, 20 μM m-AMSA, 100 μM VP-16 and 100 μM genistein. The arrows point out the cleavage sites stimulated by EPC. The position of cleavage sites is numbered according to Peden (18).

the enzyme concentration was such that the average number of cleavages per molecule was smaller than one. Therefore, multiple cleavage of one molecule was very unfrequent and the intensity of the bands shown in Figure 2 is directly related to the probability of cleavage in the corresponding region, independently of its distance from the labeled end (for example, compare band III in Figures 2A and 2B). From the data of Figure 2 together with data not shown obtained with a *Nde* I–*Pvu* II fragment labeled at its *Nde* I end, the location of the regions of cleavage was deduced and the extent of cleavage in each region was estimated from the intensity of the bands. Strong and medium cleavage regions are those which appear after ≤ 24 hours autoradiography whereas weak regions appear after > 24 hours autoradiography. A map of the EPC-stimulated cleavage regions is shown in Figure 3. Due to the limited resolution of agarose gel electrophoresis, a minimum distance of 45 bp between two regions is necessary to consider them as distinct regions.

Sequencing of the EPC-stimulated cleavage sites in pBR322 DNA

In order to determine the exact position of a large sample of strong EPC-stimulated cleavage sites, various restriction fragments of

pBR322 covering the main regions of cleavage described above and [32 P] labeled at one end, were used as substrate for topoisomerase II mediated DNA cleavage. The cleavage products were analysed by polyacrylamide sequencing gel electrophoresis and, by using appropriate markers, the position of the phosphodiester bond cleaved was determined. The cleavage sites positions were designated by the number of the 5'-terminal nucleotide of the pBR322 upper strand, taking into account the 4 bp separation between bonds cleaved on each strand. Figure 4A shows as an example the cleavage pattern of a 163 bp fragment (*Eco*R I–*Rsa* I) labeled at its *Eco*R I end. This fragment allows to study cleavage in region I and the data indicate that cleavage indeed occurs at site 22 in this region with a high yield. The region V of strong cleavage was studied using a 908 bp fragment (*Taq* I–*Bgl* I) labeled at its *Taq* I end: a strong EPC-stimulated cleavage site was observed in this region at position 2750 (Fig. 4B). In the region II, cleavage occurs at sites 86, 102 and 142 (Fig. 5). Cleavage was analyzed on the opposite strands of some sites (22, 86, 102 and 142) and the 4 bp separation characteristic of topoisomerase II double strand cleavage was observed (data not shown). Using other restriction fragments, the position of 115 EPC-stimulated cleavage sites was exactly

Figure 5. Determination of the strength of DNA cleavage sites stimulated by EPC. Panel A, pBR322 DNA linearized (4363 bp) and 5'-ends-labeled at the *Nhe* I site (a) and the *Nhe* I–*Ava* I pBR322 DNA fragment (1192 bp) 5'-end-labeled at the *Nhe* I site (b) were incubated with topoisomerase II as described in Materials and Methods. Reaction products were analyzed on a 8% denaturing polyacrylamide gel. EPC was used at 6 μ M concentration. Salt reversibility (lane 3) was assayed as described in the legend of Figure 2. Maxam-Gilbert G+A reaction was performed with the *Nhe* I–*Ava* I pBR322 DNA fragment (lane 6). The cleavage positions are indicated by arrows. Panel B, lanes 1, 2, 4, 5, 7 and 8 of the autoradiogram were scanned. The gel was exposed for 6 days in order to see the weak DNA cleavage sites, but with this exposure the sites 22 and 142 were off the linear range of the film. On the right side of the scans are shown the positions of DNA cleavage sites within the linearized DNA (4363 bp) and the DNA fragment (1192 bp). The stars indicate the labeled ends. The arrow indicates the electrophoresis direction.

determined in the regions indicated in Figure 3 covering 53% of the pBR322 entire genome.

The frequency of cleavage at the different sites was variable, and in order to analyse the degree of sequence specificity of cleavage, the relative cleavage efficiency at each individual site had to be determined at the same DNA, enzyme and drug concentration. For this purpose, linear pBR322 DNA molecules were labeled at both ends, and the equal labeling of the ends was checked (not shown). After topoisomerase II mediated DNA cleavage, the fragments were analyzed by polyacrylamide gel

electrophoresis, and the relative yield of cleavage at a particular site was deduced from the peak height measured by scanning the autoradiogram. In order to improve the resolution of the gel, electrophoresis was carried out for different time intervals, allowing to examine two adjacent regions of up to 700 bp in a single gel. Figure 5 shows as an example the cleavage pattern of linear pBR322 labeled at its *Nhe* I ends. To determine, relatively to the *Nhe* I restriction site, the position of the cleavage sites corresponding to each band observed in the polyacrylamide gel, a *Nhe* I-Ava I DNA fragment labeled at its *Nhe*I end was

Figure 6. Sequences alignment used to derive the consensus sequence. (A) The 25 cleavage sites stimulated by EPC are ranked according to their relative strength. 100% of cleavage is arbitrarily set as the cleavage intensity measured at site 22. Each value (indicated at the right side) is the mean of two or three determinations. At the left side are indicated the positions of the EPC-stimulated cleavage sites within pBR322 DNA. The sequences are aligned at the point of the cut (between -1 and +1) and given in the 5' to 3' direction. The bases in the sequences that adhere to the consensus sequence are printed in capital letters. (B) Base matrix frequency in the EPC-stimulated cleavage sites. Consensus sequence was derived from the underlined bases frequencies. (C) Information content profile ($R_{\text{sequence}}(L)$) for the 25 EPC-stimulated cleavage sites. The aligned sequences (see above) are analyzed for 20 nucleotides in each direction from the point of cleavage. Following Schneider *et al.* representation (16) the zero without sampling error correction (continuous line) and the zero with the correction made (broken line) are shown.

subjected to topoisomerase II mediated DNA cleavage, and the cleavage products were analysed in the same gel. This analysis showed that in some instances, two bands were overlapping such as in the case of peaks 22/437 and peaks 86/375. However, the intensity of bands 437 and 375 was very low compared to that of bands 22 and 86 and did not affect significantly the height estimation of peaks 22 and 86. When the strength of the EPC-stimulated cleavage sites was compared, peak 22 appeared as the strongest. Therefore, the relative yield of cleavage at a particular site was expressed relatively to site 22 set arbitrarily at 100%. Figure 6A lists the positions of 25 sites with a relative yield of cleavage $> 2\%$ with their sequences. Two of these sites (2019 and 2431) are cleaved in the absence of drug with a relative yield of cleavage of about 4%. For the other sites of Figure 6A, cleavage efficiency in the absence of EPC is either undetectable or $< 1\%$. Ninety EPC-stimulated sites (not presented) with a relative yield of cleavage $< 2\%$ are referred to as weak sites.

Sequence analysis of the EPC-stimulated cleavage sites

Topoisomerase II mediated DNA cleavage occurs at sites determined by the structure of the drug stimulating the cleavage (6). This is illustrated by data in Figure 4A showing that *m*-AMSA and genistein at concentrations where they exhibit high cleavage efficiencies (7), do not induce pBR322 DNA cleavage in the region of site 22. Similarly, Figure 4A indicates that site 23 is VP-16 specific since cleavage is not stimulated by either *m*-AMSA, genistein or EPC at this site. In order to determine the sequences characteristics of the EPC-stimulated cleavage sites,

Figure 7. Sequencing analysis of cleavage sites stimulated by EPC in two SV40 DNA regions. *Bgl* I-*Ban* I fragment (296 bp) (A) and *Alw* I-*Alw* I fragment (241 bp) (B) 5'-end-labeled at the *Ban* I and *Alw* I sites, respectively, were incubated with topoisomerase II and 6 μ M EPC. Products of DNA cleavage were analyzed on a 8% denaturing polyacrylamide gel. The arrow heads point out the EPC-stimulated cleavage sites discussed in the text. The cleavage sites positions are designated by the number of the 5'-terminal nucleotide of the strand which has the same sense as the late mRNA of SV40 DNA (22).

the nucleotide sequences of the 25 main cleavage sites were analyzed. For these sites, the strand combination corresponding to the highest sum of $R_{\text{sequence}}(L)$ was determined and each site was represented in Figure 6A by the strand selected according to this criterion. The information content (16) of the sites aligned by this procedure is plotted in Figure 6C showing a high value at position -1 and -2 and a value significantly higher than background at positions -7 , -5 , -3 , $+3$, $+5$. The consensus sequence was derived from these results considering exclusively positions with $R_{\text{sequence}}(L) > 0.2$. For these positions, one base was retained if its frequency was ≥ 0.60 , or two bases if their cumulated frequencies were ≥ 0.80 (Fig. 6B). Thymine is highly conserved at position -1 since it is observed with a frequency of 1 in the 25 strongest sites. None of the 90 weak EPC-stimulated sites had a purine simultaneously at the position -1 and at corresponding position on the opposite strand. Among the 90 weak cleavage sites, 77 (86%) had a thymine at position -1 of at least one member of each pair of strand constituting a double strand break. Cleavage at position 102 in the absence of drug has been extensively studied (15, 19). This site has neither T nor C at position -1 on either strand and is not stimulated by EPC (Fig. 5). At positions -2 , C and T are excluded and the matrix frequency indicates a slight preference for A over G (Fig. 6B). In fact, the first three strongest sites have an adenine at this position.

The consensus sequence derived from our results should predict EPC-stimulated cleavage sites in a DNA molecule provided its sequence is known. In the 53% of pBR322 which were analyzed by sequencing gel, excluding the 25 strong sites used to derive the consensus sequence, one site at position 2894 matches exactly the consensus sequence, and EPC-stimulated cleavage was indeed observed at this site. With one mismatch at positions either -7 , -5 , $+3$ or $+5$, where the information content is not very high,

Figure 8. Strong DNA cleavage sites stimulated by EPC in two regions of SV40 DNA. The bases in the sequences that adhere to the consensus sequence are printed in capital letters. The arrow heads point out the cut which occurs between -1 and $+1$.

Figure 9. Symmetrical structures of salt-resistant cleavage sites. Elements of sequence symmetry are indicated by boxes. The heptanucleotide CTTATCG is printed in bold type. The arrow heads point out the cleavage sites.

cleavage is predicted at 16 sites, and actually observed at 10 (62%) of these. Accepting two mismatches at these positions, 68 cleavage sites are predicted, and 30 (44%) are actually observed in our experiments. Sixty five (72%) of the 90 weak sites match the consensus sequence on at least four out of seven positions on one strand.

We attempted to predict EPC-stimulated cleavage sites in two regions of the SV40 DNA molecule where *in vivo* cleavage occurs preferentially in the presence of VM-26 and *m*-AMSA (20,21). One region, comprised between nucleotides 100 and 300, is the transcription enhancer region and the other is the region between nucleotides 675 and 860. In the enhancer region, characterized by the presence of a 72 bp direct repeat, the consensus sequence is present, altered at two positions (either -7, -5, +3 or +5), at two sites (242 and 256). Using the *Ban* I-*Bgl* I fragment of SV40 [³²P] labeled at its *Ban* I site, EPC-stimulated DNA cleavage was observed at both sites (data not shown). Figure 7A shows EPC-stimulated cleavage at sites 112, 172 and 184 which match the consensus sequence at five of out seven positions but contain a purine at position -3 and at sites 135, 170 and 207 which do not match the consensus sequence, but however contain the motif T(A/G)T. The criteria used above did not allow to predict two EPC strong cleavage sites (Fig. 7A) which are within the 72 bp repeat and have identical sequences. These sites (122 and 194) match the consensus sequence at four out of seven positions on both strands (Fig. 8). Furthermore, the dinucleotide AT is present on both strands at the 3' ends of the cleavage. In the region 675-860, the consensus sequence is present, altered at two positions (either -7, -5, +3 or +5) at 10 sites (696, 711, 735, 765, 773, 786, 792, 822, 831 and 850), and cleavage is observed at 8 of these sites (696, 735, 765, 773, 786, 792 shown in Figure 7B and 831 and 850 not shown). Site 786, the strongest EPC cleavage site in this region matches the consensus sequence at five of seven positions on both strands (Fig.8). Most of the 61 EPC-stimulated cleavage sites detected in the two SV40 regions studied are weak. Forty two (69%) of these sites match the consensus sequence on at least four out seven positions on one strand. Furthermore, 48 (79%) of the 61 EPC-stimulated cleavage sites contain a thymine at position -1 on at least one of the DNA strands. Only one weak DNA cleavage site (785) is devoid of a pyrimidine at position -1 on both strands.

Sequences of salt-resistant sites

It has been observed previously that cleavable complexes stabilized by 2-Methyl-9-hydroxyellipticinium at different sites in various regions of pBR322 DNA are resistant to dissociation by 0.5 M NaCl, whereas cleavable complexes formed in the presence of *m*-AMSA dissociated quickly in these conditions (17). In agreement with this observation, we have observed that, when NaCl was added at various concentrations to preformed EPC-stabilized complexes, cleavage persisted after 30 min. in some regions of pBR322 DNA (I and VI in Figure 2C and IV, data

not shown). By analysis of these cleavage regions in sequencing gels, it was determined that three sites were resistant to salt dissociation: 22 (Fig. 5), 1529 and 2829 (not shown). Site 22 displays an 8 bp symmetry centered on base 24 whereas the center of symmetry of cleavage lies between bases 23 and 24 (Fig. 9). In a similar fashion, sites 1529 and 2829 exhibit a 4 bp symmetry centered on bases 1531 and 2830, respectively, whereas the center of symmetry of cleavage lies between bases 1530/1531 and 2830/2831 (Fig. 9). Sites sensitive to salt dissociation exhibit either no sequence symmetry or a symmetry with the same center of symmetry as the cleavage site. Furthermore, the heptanucleotide CTTATCG is present (with zero or one mismatch) at the cleavage site on one strand of the three salt-resistant sites.

DISCUSSION

The DNA sequence specificity of cleavage was studied previously for DNA topoisomerase II from *Drosophila melanogaster* in the absence of drugs (23) and chicken erythrocytes in the presence of *m*-AMSA (24), and consensus sequences for the cleavage sites were derived from these studies. Antitumor drugs belonging to different series stimulate DNA cleavage by mammalian DNA topoisomerase II at sites which are determined by the structure of the drug involved (6). Since DNA sequence specificity of mammalian topoisomerase II mediated cleavage was studied previously for inhibitors such as doxorubicin (8) and a 5-substituted 2-nitroimidazole (25), it was of interest to analyse the sequence specificity of mammalian topoisomerase II mediated DNA cleavage stimulated by ellipticine derivatives. EPC was selected for this study because it is a potent inducer of topoisomerase II mediated DNA cleavage and it yields cleavage patterns identical to those observed with all other derivatives in the series (manuscript in preparation). The position of 115 EPC-stimulated cleavage sites was determined at the nucleotide level in pBR322 and a 12 bp consensus sequence covering the cleavage site was deduced from a statistical analysis of the 25 sites displaying the highest cleavage frequency. The consensus size is within the minimal duplex length of 16 bp required for cleavage activity (19) and compatible with the size of *ca* 25 bp observed for the protected DNA region in footprinting experiments (26, 27).

Our results show clearly that bases at positions -1 and -2 of the cleavage site play a major role in the cleavable complex stabilization by ellipticines. The dinucleotides AT and GT are observed at the 3' end of the 25 strong cleavage sites analyzed in pBR322 and the 3 SV40 strong sites. Furthermore, the presence of an adenine at position -2 of the strongest cleavage sites 22, 2750 and 142 of pBR322 suggests a preference for an A at this position. The DNA sequence specificity of cleavage by mouse DNA topoisomerase II in the presence of doxorubicin was studied recently (8). Since topoisomerases II from mouse and calf thymus exhibit the same DNA sequence specificity (28), it is reasonable to assume that the differences in sequence specificity of the doxorubicin-stabilized and EPC-stabilized sites are due exclusively to the drugs. All the doxorubicin-stabilized cleavage sites have an adenine at the 3' terminus of at least one of the two single strand breaks whereas in the absence of drug, a pyrimidine was preferentially observed at position -1 of the cleavage site (8). However, cleavage has been observed in the absence of drugs at sites devoid of a pyrimidine at position -1 (25, 29). Similarly, stabilization by VP-16 (site 102 of pBR322

Figure 10. Comparison between ellipticine sequence (A) and doxorubicin sequence (B) taken from ref. (8). N indicates no base preference; in doxorubicin sequence, T at +4 is weakly preferred.

(not shown) and (30)) and by a 5-substituted 2-nitroimidazole (25) occurs at sites devoid of pyrimidine at position -1. In contrast with the doxorubicin sites, the ellipticine-stimulated cleavage sites in pBR322 and SV40 have a thymine at the 3' terminus of at least one of the single strand breaks. In a few weak sites, a cytosine was observed instead of a thymine, but two purines were never observed simultaneously at the 3' termini of both single strand breaks of an EPC-stimulated cleavage site with one exception in SV40 (site 785). This weak site is distant of 1 bp from the strong site 786, and may result of a mispositioning of the enzyme on the strong site. The thymine requirement at position -1 for ellipticine is a characteristic feature of this drug, and suggests a direct interaction of the drug, within the ternary complex, with the AT base pair adjacent to the cleavage. Figure 10 which compares the ellipticine consensus sequence with the doxorubicin sequence (8) shows that, in addition to position -1, positions -2 and +5 are mutually exclusive. Ellipticine sites exhibit a more stringent requirement than doxorubicin sites at positions -7, -5 and +3, whereas the opposite is observed at sites +1, +4, +6 and +7.

As mentioned in the Results section, the proposed consensus sequence allows to predict accurately a large number of pBR322 and SV40 cleavage sites. However, two strong EPC cleavage sites within the 72 bp repeat of SV40 match the consensus sequence at only 4 of 7 positions if one strand is considered. If both strands were examined concurrently, cleavage at these sites could indeed be predicted. If homology with the consensus is scored on both strands in the regions of pBR322 and SV40 which have been studied, 8 or more matches with the consensus sequence (of 14 possible) are observed for 78% of the strong sites and 31% of the weak sites. A hypothetical site matching exactly the consensus on both strands would exhibit a partial sequence symmetry with the same center of symmetry as the cleavage site. The sequence of such a site: ANCNT(A/G)TN(G/C)(G/C)NA(T/C)ANGNT is observed with one mismatch at site 142 which is indeed a strong pBR322 cleavage site. More interestingly, the strongest and salt-resistant EPC site 22 has a remarkable sequence symmetry with a center which does not coincide with the center of symmetry of cleavage. An identical pattern is found in two other cleavage sites which are also resistant to salt dissociation, and one can assume such a staggered symmetry induces conformational constraint in the enzyme subunits bound to DNA and to ellipticines which results in a slower rate of the transesterification reaction restoring the phosphodiester backbone, thus increasing the cleavable complex concentration. Another less likely hypothesis is that the salt resistance is due to the presence of the heptanucleotide CTTATCG which is unfavorable for the transesterification reaction in the presence of bound ellipticine. For other complexes, such as this formed at site 142 (Fig.5), which are readily dissociated by salts and are devoid of staggered symmetry, the increased level of cleavable complex might result from a stimulation of their rate of formation as suggested previously for etoposide (31). Alternatively, one may argue that these sites are stable in the standard conditions, and become unstable upon addition of salt.

Finally, our data, together with previously published observations, indicate that the extent of cleavage by DNA topoisomerase II in the presence of an antitumor drug, and the resistance of the complex to dissociation by salt at high concentrations is determined by the base sequence of the DNA

engaged in the complex. Single base substitution experiments should provide precise informations on the role of each individual nucleotide in the stability of this complex.

ACKNOWLEDGEMENTS

We thank Chantal Buffenoir and Monique Charra for expert technical assistance. This work was supported by Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Rhône-Poulenc Santé and Fondation pour la Recherche Médicale.

REFERENCES

- Kohn, K.W., Pommier, Y., Kerrigan, D., Markovits, J., and Covey, J.M. (1987) *NCI Monographs* **4**, 61-71.
- D'Arpa, P. and Liu, L.F. (1989) *Biochim. Biophys. Acta* **989**, 163-177.
- Wang, J.C. (1985) *Annu. Rev. Biochem.* **54**, 665-697.
- Chen, G.L., Yang, L., Rowe, T.C., Halligan, B.D., Tewey, K.M. and Liu, L.F. (1984) *J. Biol. Chem.* **259**, 13560-13566.
- Rowe, T.C., Chen, G.L., Hsiang, Y.H., and Liu, L.F. (1986) *Cancer Res.* **46**, 2021-2026.
- Tewey, K.M., Rowe, T.C., Yang, L., Halligan, B.D., and Liu, L.F. (1984) *Science* (Washington, D.C.) **226**, 466-468.
- Markovits, J., Linossier, C., Fossé, P., Couprie, J., Pierre, J., Jacquemin-Sablon, A., Saucier, J.-M., Le Pecq, J.-B., and Larsen, A.K. (1989) *Cancer Res.* **49**, 5111-5117.
- Capranico, G., Kohn, K.W. and Pommier, Y. (1990) *Nucleic Acids Res.* **18**, 6611-6619.
- Zwelling, L.A., Michaels, S., Kerrigan, D., Pommier, Y., and Kohn, K.W. (1982) *Biochem. Pharmacol.* **31**, 3261-3267.
- Miller, K.G., Liu, L.F., and Englund, P.T. (1981) *J. Biol. Chem.* **256**, 9334-9339.
- Halligan, B.D., Edwards, K.A., and Liu, L.F. (1985) *J. Biol. Chem.* **260**, 2475-2482.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) in *Molecular Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Maxam, A.M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
- Bencini, D.A., O'Donovan, G.A., and Wild, J.R. (1984) *BioTechniques* **4**, 4-5.
- Sander, M., and Hsieh, T.-S. (1983) *J. Biol. Chem.* **258**, 8421-8428.
- Schneider, T.D., Stormo, G.D., Gold, L., and Ehrenfeucht, A. (1986) *J. Mol. Biol.* **188**, 415-431.
- Tewey, K.M., Chen, G.L., Nelson, E.M., and Liu, L.F. (1984) *J. Biol. Chem.* **259**, 9182-9187.
- Peden, K.W.C. (1983) *Gene* **22**, 277-280.
- Lund, K., Andersen, A.H., Christiansen, K., Svejstrup, J.Q. and Westergaard, D. (1990) *J. Biol. Chem.* **265**, 13856-13863.
- Yang, L., Rowe, T.C., Nelson, E.M. and Liu, L.F. (1985) *Cell* **41**, 127-132.
- Yang, L., Rowe, T.C. and Liu, L.F. (1985) *Cancer Res.* **45**, 5872-5876.
- Toozé, T. (1981) Editor of *DNA Tumor Viruses*, part 2, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sander, M., and Hsieh, T.-S. (1985) *Nucleic Acids Res.* **13**, 1057-1072.
- Spitzner, J.R., and Muller, M.T. (1988) *Nucl. Acids. Res.* **16**, 5533-5556.
- Sorensen, B.S., Jensen, P.S., Andersen, A.H., Christiansen, K., Alsner, J., Thomsen, B. and Westergaard, O. (1990) *Biochemistry* **29**, 9507-9515.
- Lee, M.P., Sander, M. and Hsieh, T.-S. (1989) *J. Biol. Chem.* **264**, 21779-21787.
- Thomsen, B., Bendixen, C., Lund, K., Andersen, A.H., Sorensen, B.S. and Westergaard, O. (1990) *J. Mol. Biol.* **215**, 237-244.
- Sperry, A.O., Blasquez, V.C. and Garrard, W.T. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 5497-5501.
- Liu, L.F., Rowe, T.C., Yang, L., Tewey, K.M., and Chen, G.L. (1983) *J. Biol. Chem.* **258**, 15365-15370.
- Darby, M.K., Herrera, R.F., Vosberg, H.-D. and Nordheim, A. (1986). *EMBO J.* **5**, 2257-2265.
- Osheroff, N. (1989) *Biochemistry* **28**, 6157-6160.