Local sequence requirements for DNA cleavage by mammalian topoisomerase II in the presence of doxorubicin

Giovanni Capranico+, Kurt W.Kohn and Yves Pommier*
Laboratory of Molecular Pharmacology, Division of Cancer Treatment, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received July 16, 1990; Revised and Accepted October 15, 1990

ABSTRACT

Doxorubicin, a DNA-intercalator, is one of several anti-cancer drugs that have been found to stabilizes topoisomerase II cleavage complexes at drug-specific DNA sites. The distribution and DNA sequence environments of doxorubicin-stabilized sites were determined in the SV40 genome. The sites were found to be most concentrated in the major nuclear matrix-associated region and nearly absent in the vicinity of the replication origin including the enhancer sequences in the 21-bp and 72-bp tandem repeats. Among 97 doxorubicin-stabilized sites that were localized at the DNA sequence level, none coincided with any of the 90 topoisomerase II cleavage sites detected in the same regions in the absence of drug. Cleavage at the 90 enzyme-only sites was inhibited by doxorubicin and never stimulated even at low drug concentrations. All of the doxorubicin-stabilized sites had an A at the 3' terminus of at least one member of each pair of strand breaks that would constitute a topoisomerase II double-strand scission. Conversely, none of the enzyme-only sites had an A simultaneously at the corresponding positions on opposite strands. The 3'-A requirement for doxorubicin-stabilized cleavage is therefore incompatible with enzyme-only cleavage and explains the mutual exclusivity of the two classes of sites.

INTRODUCTION

Eukaryotic DNA topoisomerase II is a nuclear enzyme involved in DNA replication, transcription and recombination, as well as in chromatin organization (1, 2). The enzyme catalyzes the interconversion of DNA topoisomers by making a transient double-strand break in a DNA segment through which another DNA duplex passes before resealing of the DNA break (1, 2). Mammalian topoisomerase II has been identified as the cellular target of some of the most active anticancer drugs (3, 4). Drugs such as doxorubicin, etoposide, and *m*-AMSA inhibit the enzyme by stabilizing DNA cleavage complexes which may correspond to the normal DNA cleavage intermediates in the topoisomerase DNA strand-passing reaction. Although it has been established that each drug induces topoisomerase II-mediated DNA cleavage at specific sites (5), the molecular basis for the sequence selectivity of drug action remains unexplained.

We have recently shown that in the presence of purified topoisomerase II, the most active anthracyclines induce prominent cleavage in a small region of the entire SV40 DNA, between nucleotides 4200 and 4300 (6). This SV40 region also corresponds to a cluster of major enzyme sites in the absence of drug and to a nuclear matrix attachment region (MAR) (7). In the present study, the DNA sequence of topoisomerase II cleavage sites induced in the presence and absence of doxorubicin was analyzed in several segment of the SV40 genome including the MAR in order to investigate the molecular basis for the sequence-selectivity of doxorubicin action upon topoisomerase II. The analysis of doxorubicin-induced sites was limited to the effects of low doxorubicin concentrations (never exceeding 0.5 uM) since doxorubicin induces maximum cleavage at relatively low drug concentrations (between 0.5 and 2 uM) and suppresses DNA cleavage at higher concentrations (6).

MATERIALS AND METHODS

Materials and Enzymes

Doxorubicin and 4-demethoxydaunorubicin were obtained from Farmitalia-Carlo Erba, Milan, Italy. Stock solutions were made in deionized water at 0.2 mM and aliquots were kept frozen at −20°C. Etoposide (VM-26) and *m*-AMSA were obtained from the Drug Synthesis and Chemistry Branch, National Cancer Institute, Bethesda, MD. Stock solution were made in dimethylsulfoxide at 10 mM immediately before use. SV40 DNA, restriction endonucleases, T4 polynucleotide kinase, and agarose were purchased from Bethesda Research Laboratories (Gaithersburg, MD). Calf intestine phosphatase was purchased from New England Biolabs (Beverly, MA) and [gamma-^{32}P]ATP from New England Nuclear Research Products (Boston, MA). DNA topoisomerase II was purified from

* To whom correspondence should be addressed

+ Present address: Division of Experimental Oncology B, Istituto Nazionale Tumori, v. Venezian 1, 20133 Milan, Italy

mouse leukemia L1210 cell nuclei as described previously (8). The purified enzyme yielded a single 170 kDa band after silver staining of SDS-polyacrylamide gels.

Topoisomerase II-induced DNA cleavage reactions

SV40 DNA fragments were uniquely 5′ end-labeled as already described (6, 7, 9) and then purified by electroelution and ethanol precipitation. DNA fragments were reacted with 40−70 ng of topoisomerase II with or without doxorubicin (0.1 to 0.5 μM final concentration) in 0.01 M Tris-HCl, pH 7.5, 0.05 M KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM ATP and 15 μg/ml bovine serum albumin for 20 min at 37°C. VM-26, *m*-AMSA and 4-demethoxydaunorubicin were used at the indicated concentrations. Topoisomerase II was added to 20 μl final reaction volume in 3 μl of 40% (v/v) glycerol, 0.35 M NaCl, 5 mM $MgCl_2$, 1 mM EGTA, 1 mM KH_2PO_4, 0.2 mM dithiothreitol and 0.1 mM phenylmethanesulfonyl fluoride, pH 6.4. Reactions were stopped by adding sodium dodecyl sulfate (SDS), EDTA and proteinase K (1%, 10 mM and 250 μg/ml, respectively) and samples incubated for 1 h at 42°C. For agarose gel analysis 3 μl (10×) loading buffer (0.3% bromophenol blue, 16% Ficoll, 0.01 M Na_2HPO_4) was added to each sample which was then heated at 65°C for 1−2 min before being loaded into an agarose gel made in (1×)TBE (89 mM Tris, 89 mM Boric acid, 2mM EDTA, pH 8) containing 0.1% SDS in order to remove DNA-bound drug molecules which otherwise retard the electrophoretic migration of DNA fragments (6, 10). For DNA sequence analysis, samples were ethanol-precipitated and then resuspended in 2.5 μl loading buffer (80% formamide, 10 mM NaOH, 1 mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue). They were heated at 90°C before loading into DNA sequencing gels (8% polyacrylamide; 29:1, acrylamide:bis) containing 7 M urea in 1 × TBE buffer.

Genomic mapping of DNA breaks

Doxorubicin-induced topoisomerase II DNA cleavage sites were mapped by means of densitometer and computer analysis of agarose gel autoradiographies, as described (7).

Statistical tests

The chi-square one-sample test was used to determine the deviation from the random distribution of bases at each position of the aligned sequences. The expected occurrence of each nucleotide was calculated either relative to the base frequencies averaged over the entire SV40 DNA (A=T=0.296; C=G=0.204) or averaged over the local regions within 35 bases of each cutting site. These two methods of calculation showed no significant difference.

Probabilities of deviations from expectation were calculated as follows. Let p be the frequency of occurrence of a given base (p=0.296 for A or T, p=0.204 for C or G). The expected number of sites having a given base at any particular location is pn, were n is the total number of sites analyzed. Let m be the observed number of sites which have the given base at that location. If $m > pn$, then the probability, P, of the chance occurrence of m or more instances was calculated as

$$P = \sum_{i=m}^{n} \frac{n!}{i!(n-i)!} p^i(1-p)^{n-i}$$

If $m < pn$, then the chance occurrence of m or fewer instances was calculated as

$$P = \sum_{i=o}^{m} \frac{n!}{i!(n-i)!} p^i(1-p)^{n-i}$$

Factorials were computed to 10-digit precision as their logarithms using the Lanczos approximation (11).

RESULTS

DNA breaks induced by topoisomerase II in the presence of doxorubicin

Because doxorubicin-induced topo II cleavage complexes are relatively rare in most regions of DNA (6), the entire SV40 genome was surveyed in order to locate regions where the cleavage sites tend to be most frequent. Neutral agarose gel electrophoresis was used to disclose regions containing double-strand breaks (Fig. 1). In agreement with our previous study of anthracycline derivatives (6), DNA cleavage induced by 0.5 uM doxorubicin was most prominent between positions 4100 and 4300 (Fig. 1, upper panel). This region was also preferentially cleaved by topoisomerase II in the absence of drug (Fig. 1, lower panel) (7) and corresponds to the major nuclear matrix attachment region (MAR) of SV40 DNA (7, 12). It should be noted that only minor cleavage was found around the replication origin including the enhancer sequences in the 21-bp and 72-bp tandem repeats (Fig. 1).

Topoisomerase II cleavage sites induced in the presence and absence of doxorubicin were sequenced in the nuclear matrix attachment region in order to determine the position and the intensity of DNA breaks on each strand. Figure 2A shows that, within the 90 bp region containing the major doxorubicin-induced cleavage sites, all sites on one strand were paired with a corresponding site, staggered with the expected 5′ overhang of four bases, on the complementary strand. A more complete analysis indicated that only 3 out of 25 doxorubicin-induced cleavage sites did not have a corresponding site on the complementary strand. These results are consistent with those of figure 1 and previous observations in cultured cells showing that anthracyclines, in contrast to epipodophyllotoxins (VP-16 and VM-26) and *m*-AMSA, predominantly induce DNA double-strand breaks (13). The possibility however is not excluded that many of the breaks detected at corresponding sites on opposite strands in the current work actually exist as single-strand breaks in different DNA molecules.

DNA cleavage sites induced by topoisomerase II alone and sites stimulated by doxorubicin are mutually exclusive

Topoisomerase II inhibition by doxorubicin was analyzed further by sequencing the DNA cleavage sites induced by topoisomerase II in the presence and in the absence of the drug in three different regions of the SV40 DNA (arrows in Fig. 1). In no case was a doxorubicin-induced cleavage site detectable without drug (Fig. 2A−B). Conversely, the topoisomerase II cleavage sites detected in the absence of drug were never enhanced (even at drug concentrations as low as 0.1 μM) but instead were suppressed by doxorubicin (Fig. 2A−B). Similar results were obtained with other anthracyclines. For instance, 4-demethoxydaunorubicin, although more potent than doxorubicin on a concentration basis (6), stimulated cleavage at the same sites and usually with the same relative intensities (figure 2B). In contrast, other classes of topoisomerase II inhibitors, including VM-26 and *m*-AMSA, did not suppress any of the sites detected in the absence of drug

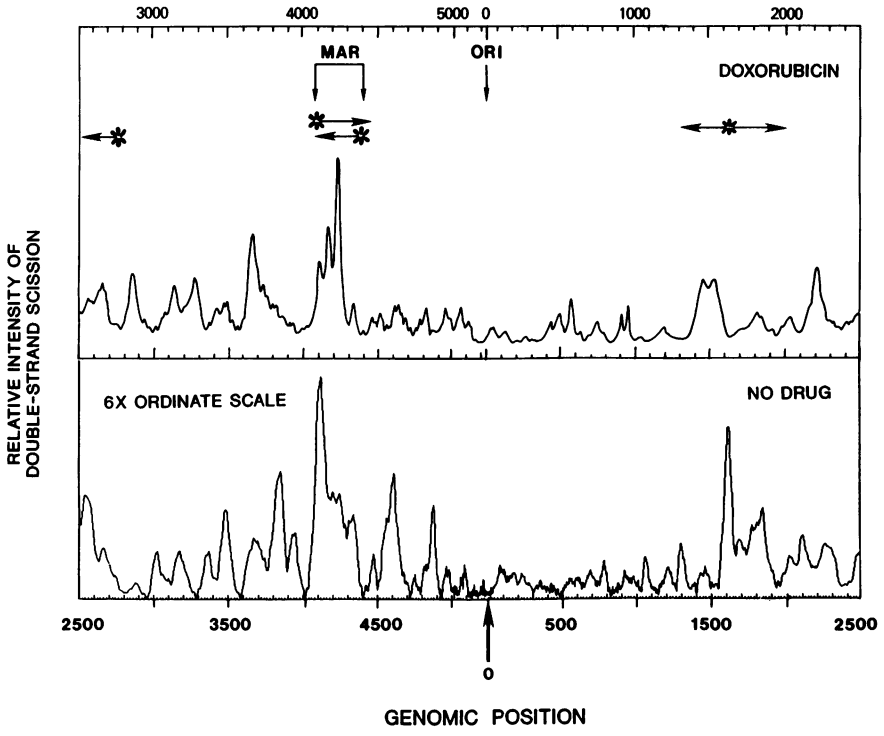

Figure 1. DNA double strand scission induced by murine topoisomerase II in SV40 DNA. Upper panel: cleavage in the presence of 0.5 μM doxorubicin; lower panel: cleavage in the absence of drug (The scale ordinate is amplified 6-fold relative to the upper panel). Densitometer scanning curves of DNA cleavage patterns analyzed by means of agarose gels are corrected to give a linear measure of cleavage frequency per base pair. Stars indicate the labeled 5'-end termini and arrows the sequenced fragments (see Materials and Methods); ORI, replication origin; MAR, major nuclear-matrix associated region (7, 12).

and often enhanced cleavage at these sites (Figs. 2A-B). Thus, all topoisomerase II DNA cleavage sites could be assigned to one of two types: 1) 'no-drug sites', that were cleaved by topoisomerase II in the absence of drug, and suppressed by doxorubicin; 2) 'doxorubicin sites', that were cleaved in the presence of doxorubicin, and that were not detectable in the absence of drug.

Requirement of an adenine at the site of doxorubicin-induced topoisomerase II DNA cleavage

The base sequence of a total of 187 topoisomerase II DNA cleavage sites (97 'doxorubicin sites' and 90 'no-drug sites') was analyzed. DNA sequences were aligned at the point of measured phosphodiester bond cleavage in the 5′→ 3′ orientation and the bases immediately 5′ and 3′ to the analyzed break point were numbered −1 and +1, respectively. Positions −1 and +5 are equivalent relative to topoisomerase II-mediated DNA double-strand breaks since they are 5′ adjacent to each site of cleavage (see Fig. 6). When the deviation of the base distribution from the global SV40 DNA base frequencies was evaluated at each position by chi-square analysis (Fig. 3), a core region of non-random base composition was found between positions −3 and +7 for both groups of cleavage sites. The most significantly biased positions were −1 and +5, while the center of the core region (positions +2 and +3) showed little or no preferential base selection (Fig. 3).

Base frequency and probability of base occurrence were determined at each position for the 'doxorubicin sites' and 'no-drug sites' (Table 1 and Fig. 4). Positions −1 and +5 exhibited the most striking differences. An adenine and a thymine were strongly preferred at positions −1 and +5, respectively, in the case of 'doxorubicin sites': 71% had an A at position −1 and 67% had a T at position +5 (Table 1). On the contrary, in the case of the 'no-drug sites', at position −1, adenine frequency was markedly low (3%) and pyrimidines were frequent (80%), and at position +5, thymine frequency was significantly lower than elsewhere (10%) (Table 1 and Fig. 4).

The 'doxorubicin sites', contrary to consensus sequences previously reported for other topoisomerase inhibitors, conformed to a universal rule: 'doxorubicin sites' invariably had either an A at position −1 or a T at position +5. The rule actually addresses a single requirement, because the two positions, A(−1) and T(+5), are symmetrically equivalent with respect to double-strand cleavage by topoisomerase II. The 'no-drug sites', on the other hand, usually obeyed the converse of the above rule: 87% had neither an A at −1 nor a T at +5, and the remaining 13% never had both an A at −1 and a T at +5. Therefore either the A at position −1 or the T at position +5 must be present for doxorubicin-induced DNA cleavage, and these same features tend to exclude cleavage by topoisomerase II in the absence of drug.

The relationship between DNA cleavage intensity and the presence of an adenine at the 5′ side of the break site in one or both strands was examined (Table 3). Among the 22 strong 'doxorubicin sites' (cleavage intensity = 3), 11 sites (50%) had both an A at position −1 and a T at +5, 8 (37%) had the A and lacked the T, and only 3 sites (13%) lacked the A and had the T. Thus, high cleavage intensity was more often observed when an adenine was 5′ adjacent to the two cleavage sites than when only a T was found at position +5. The major effect appeared to be that sites lacking an A at −1 tended to be weak.

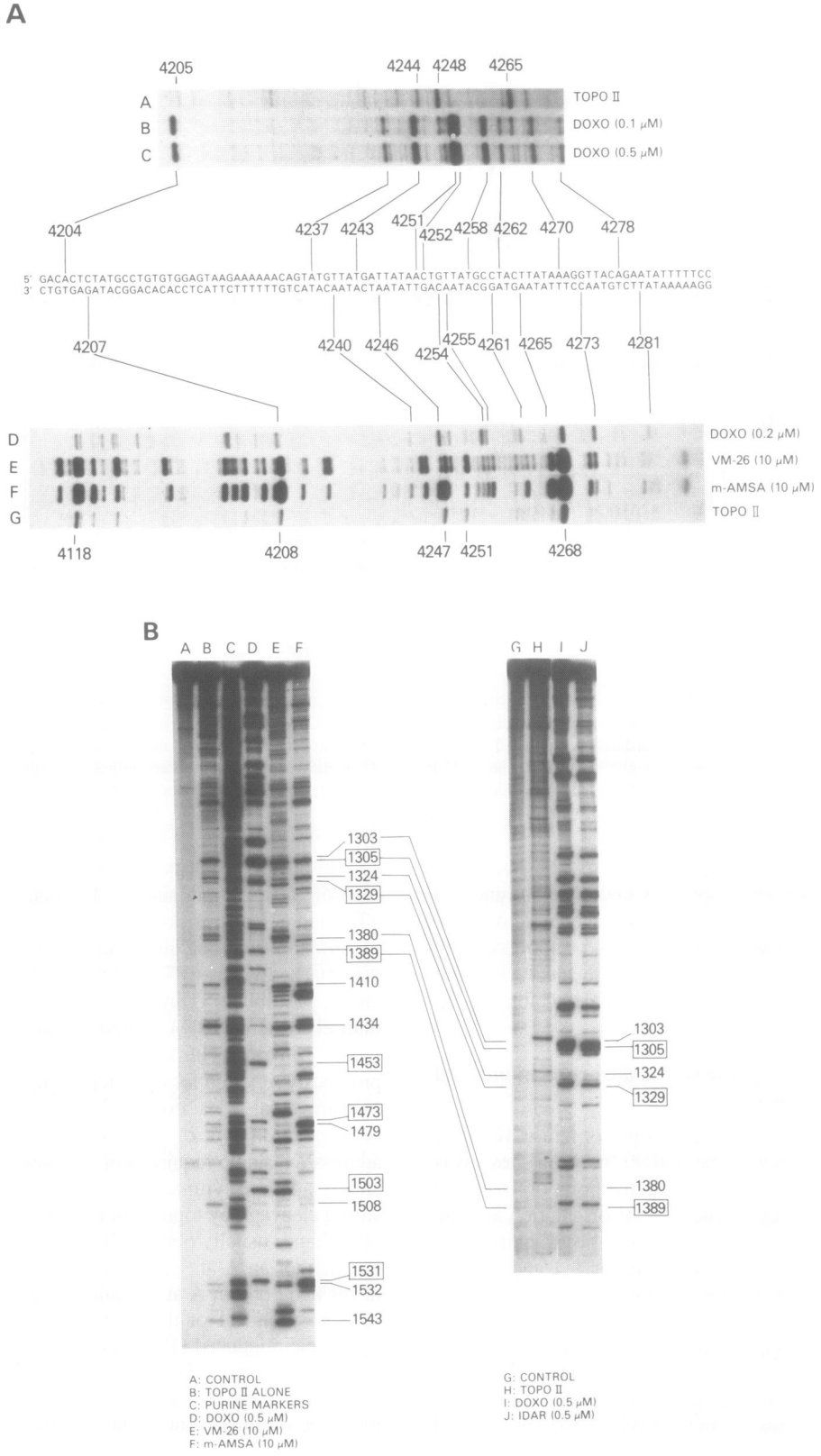

Figure 2. A) Sequencing analysis of topoisomerase II cleavage sites in the presence or absence of doxorubicin in the MAR of SV40. Upper panel: the late template strand labeled at 4099 genomic position; lane a, topoisomerase II without drugs; lanes b and c, with doxorubicin at 0.1 and 0.5 uM, respectively. Lower panel: the late coding strand was labeled at 1631 genomic position; lane d, 0.2 uM doxorubicin; lane e, 10 uM VM-26; lane f, 10 uM *m*-AMSA; lane g, topoisomerase II without drugs. Numbers mark the nucleotide covalently bound to the enzyme at the break sites determined by comparison with purine markers run in the same gel. B) Sequencing analysis of topoisomerase II cleavage sites induce by doxorubicin and 4-demethoxydaunorubicin in SV40 DNA. The early template strand was labeled at 1631 genomic position. Lanes a and g, control DNA; lanes b and h, topoisomerase II without drugs; lane c, purine markers; lane d and i, 0.5 uM doxorubicin; lane e, 10 uM VM-26; lane f, 10 uM *m*-AMSA; lane j, 0.5 uM 4-demethoxydaunorubicin. Numbers mark the nucleotide covalently bound to the enzyme at the break sites (numbers in boxes mark selected anthracycline-induced sites).

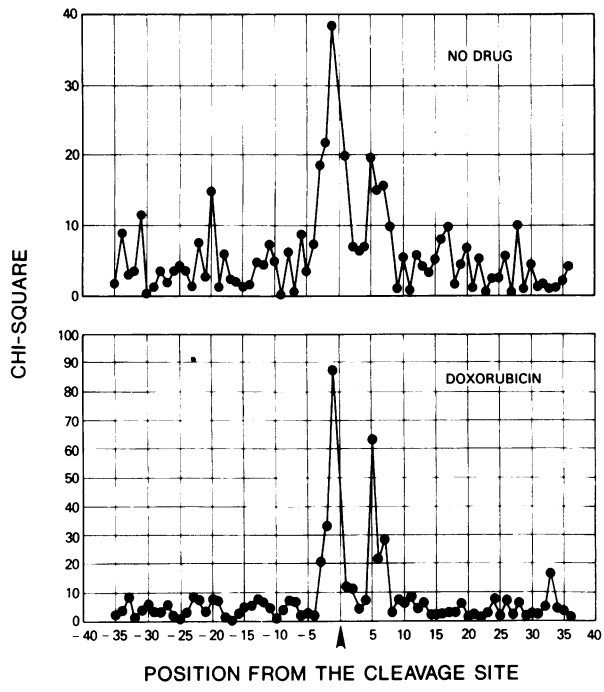

Figure 3. Chi-square values of the nucleotide distribution at each position of the cleavage site for the 'no-drug sites' (upper) and 'doxorubicin sites' (lower). The chi-square values for p = 0.05 and p = 0.01 are 5.99 and 11.34, respectively (3 degree of freedom). The 'no-drug sites' (with the convention that L and E indicate the strand which has the same sense as the late and early mRNA of SV40 DNA, respectively) were: E4309, E4311, E4307, E4272, E4268, E4266, E4251, E4247, E4343, E4339, E4334, E4331, E4208, E4203, E4190, E4197, E4169, E4143, E4129, E4167, E4214, L4126, L4139, L4140, L4164, L4187, L4194, L4205, L4211, L4234, L4240, L4244, L4248, L4265, L4269, L4304, L4306, L4308, L4263, L4200, L1672, L1681, L1684, L1689, L1700, L1706, L1743, L1758, L1769, L1812, L1820, L1860, E2747, E2742, E2721, E2718, E2709, E2686, E2678, E2652, E2647, E2644, E2629, E2605, E1606, E1588, E1581, E1569, E1567, E1566, E1543, E1540, E1534, E1532, E1508, E1485, E1482, E1479, E1474, E1472, E1466, E1460, E1458, E1438, E1434, E1431, E1416, E1410, E1448, E1449. The 'doxorubicin sites' were: E4252, E4273, E4261, E4254, E4255, E4248, E4246, E4281, E4271, E4265, E4240, E4207, E4201, E4195, E4192, E4191, E4160, E4141, E4134, E4338, E4162, E4154, E4297, L4243, L4270, L4251, L4204, L4132, L4151, L4157, L4237, L4252, L4258, L4262, L4278, L4282, L4294, L4198, L4192, L4189, L4188, L4186, L4159, L4138, L4131, L4249, L4245, L4268, L1703, L1707, L1728, L1736, L1752, L1765, L1782, L1785, L1725, L1729, L1821, L1827, L1842, L1848, L1854, L1866, L1874, E2736, E2728, E2697, E2691, E2682, E2679, E2676, E2673, E2667, E2666, E2665, E2661, E2654, E2648, E2646, E2639, E2633, E2628, E2604, E1589, E1531, E1503, E1497, E1489, E1478, E1453, E1484, E1481, E1601, E1564, E1557, E1520.

Base preferences around doxorubicin-induced DNA cleavage sites

An alternative method to gauge base preferences was to calculate the probability of the observed deviation from expectation for each base at each position. Expectations were calculated relative to the overall base frequencies in SV40: 29.6% for A or T, 20.4% for G or C. The number of occurrences of each base at a given position were compared with expectation, and the probability (P) of a deviation as great or greater than that observed was calculated (see Materials and Methods) and plotted as $-\log P$ (Fig. 4). This method discloses the most significant deviation, either above or below expectation, among all four bases considered independently.

For the 97 doxorubicin sites, the greatest deviations from expectation are again an excess of A at −1 (69 instances compared to 29 expected out of a total of 97, $-\log P=16.1$) and an excess of T at +5 (65 instances, $-\log P=13.2$). Moreover, T at −1 was deficient relative to C, and A at +5 was deficient relative to G (Fig. 4), showing that the base preferences at −1 and +5 have in all detail the dyadic symmetry of topoisomerase II double-strand cuts.

Going to the next base towards the 5′ end on each strand, we come to positions −2 and +6. Here preferences can be seen for T at −2 (55 instances, $-\log P=7.3$) and for A at +6 (48 instances, $-\log P=4.2$). At the next positions towards the 5′ ends, the largest $-\log P$ values are for a deficiency of C at −3 (4 instances compared to 20 expected out of a total of 97, $-\log P=5.1$) and for a deficiency of G at +7 (2 instances, $-\log P=6.8$). All of these base preferences and deficiencies are consistent with the dyadic symmetry of topoisomerase II double-strand cuts.

The above 3 dyadic pairs (the 3 bases immediately 5′ to the cutting site on each strand) are the only positions where substantial deviations from expectation were seen when all 97 doxorubicin sites were included. However, additional information was obtained when subsets of these sites were analyzed. The doxorubicin-induced cleavage sites can be separated into 3 mutually exclusive subsets, depending upon whether there is (i) A(−1) and not-T(+5), (ii) T(+5) and not-A(−1), or (iii) both A(−1) and T(+5) (table 2). There were no instances where both A(−1) and T(+5) were lacking. Base preferences for each subset are shown in Fig. 5 (top, subset (i); middle, subset (ii); bottom, subset (iii)). It can be seen that the base preferences near the cleavage site are asymmetric for subset (i) and for subset (ii). The 2 subsets relative to each other however show the expected dyadic symmetry at several positions near the cleavage site, corresponding to the following base preference pattern:

position: (5′→ 3′)	−3	−2	−1	+1	+2	+3	+4	+5	+6	+7
	not-C	T	(A)	not-C	.	A	.	(not-T)	.	not-G
(i)	2.9	7.4	...	2.9	.	1.0	.	...	.	1.9
(ii)	2.5	4.9	...	2.3	.	2.9	.	...	.	2.5

The bottom 2 lines are $-\log P$ values for subset (i) (32 sites) and subset (ii) (28 sites). Positions −1 and +5 are marked with parentheses to indicate that the subsets were selected according to these base requirements. In order to demonstrate the symmetry, subset (ii) has been subjected to a diadic transformation (180° rotation) of the DNA duplex about the midpoint between positions +2 and +3 (e.g., the base at position +4 is replaced by the complementary base at position +1). Although the strongest bias at position +1 is 'not-C', there is also an almost equally strong symmetrical preference for A (fig. 5).

A possible asymmetry (not shown above) occurred at position +4, where subset (ii) showed a preference for T ($-\log P=2.9$), while subset (i) had no discernable preference at the corresponding position. Another deviation from dyadic symmetry was noted at position +10, where subset (i) had a highly significant preponderance of pyrimidines (28 out of a total of 32 sites in this subset), the most significant single base deviation from expectation being an excess of C (16 occurrences out of 32, $-\log P=3.4$) (fig. 5). The corresponding position +6 in subset (ii) showed no significant deviations from expectation.

For subset (iii), base pattern preferences were seen at positions −2 and +6, where the dyadic symmetry suggests a preference for T at the second position towards the 5′end from the cleavage site on each strand (fig 5, bottom). Position +7 showed a deficiency of G (1 instance out of 37, $-\log P=2.3$), while the

Table 1. Base frequencies at the DNA cleavage sites induced by murine DNA topoisomerase II in the presence or in the absence of doxorubicin[a]

Position 5′→ 3′	DOXORUBICIN SITES (97 sites)					NO DRUG SITES (90 sites)				
	A	C	G	T	X^2	A	C	G	T	X^2
−6	25	26	20	29	1.8	43	14	14	29	8.7
−5	29	16	18	37	2.7	25	29	18	28	3.4
−4	33	16	22	29	1.7	23	20	31	26	7.5
−3	37	4	16	43	21.0	43	14	6	37	18.7
−2	22	11	10	57	33.0	39	15	33	13	22.0
−1	71	15	9	5	87.0	3	39	17	41	38.0
+1	*38*	9	16	37	11.6	50	16	11	23	20.0
+2	34	13	12	41	11.0	32	15	29	24	6.9
+3	35	15	24	26	4.0	30	30	12	28	6.3
+4	26	19	13	42	7.3	28	14	16	42	7.0
+5	10	9	14	67	63.0	36	23	31	10	20.0
+6	50	12	12	26	21.0	20	28	9	43	15.0
+7	37	15	2	46	28.0	43	11	10	36	15.8
+8	30	18	15	37	3.0	21	33	15	31	9.8
+9	36	12	26	26	7.5	31	22	21	26	0.9
+10	23	24	12	24	6.4	24	15	21	40	5.4

[a] Base frequencies are expressed as percentage; in the whole SV40 DNA they are: A=T=29.6%, G=C=20.4%. Preferred nucleotides (17) are in italic and were considered as follows: for a single base, frequency greater or equal 50%; for any pair of bases, frequency above or equal 70% (if no single base is above 50%).

dyad corresponding position, −3, showed a non-significant deficiency of C (4 instances out of 37, −log P=0.7).

position: (5′→ 3′)	−3	−2	−1	+1	+2	+3	+4	+5	+6	+7
	.	T	(A)	.	.	.	.	(T)	A	not-G
(iii)	.	3.5	...	.	.	.	.	...	2.1	2.3

The bottom line contains the -log P values for subset (iii) (37 sites).

The base preference pattern for subset (iii) is generally consistent with those of subsets (i) and (ii) when the effects of symmetry are taken into account. However a possible additional dyadic pair appeared at positions −9 (18 instances of T out of 37, −log P=1.7) and +13 (19 instances of A, −log P=2.1) (see fig. 5, bottom).

Base preferences around topoisomerase II DNA cleavage sites in the absence of drug

Similar probability tests were employed to analyze the base preferences of the 90 'no-drug sites' (Fig. 4, top; Table 4). The largest deviations were not-A at −1 (3 instances compared to an expected 26.6 out of a total of 90, −log P=9.4) and not-T at +5 (9 instances, −log P=4.8). The positions flanking the −1 cleavage site showed preference for A at +1 (45 instances, −log P=4.1), not-T at −2 (12 instances, −log P=3.3), not-G at −3 (5 instances compared to 18.4 expected, −log P=3.8, and G at −4 (28 instances, −log P=1.66). Significant deviations, although not as large as at the −1 position were also observed around the +5 position with preference for T at +4 (38 instances compared to 26.6 expected out of 90, −log P=1.8), not-G at +6 (8 instances compared to 18.4 expected, −log P=2.2), A at +7 (39 instances, −log P=2.1), and C at +8 (30 instances, −log P=2.3). Thus the 'no-drug sites' show the expected dyadic symmetry, although the base sequence preferences differ from and show the already mentioned relationship to the 'doxorubicin sites' (Table 4).

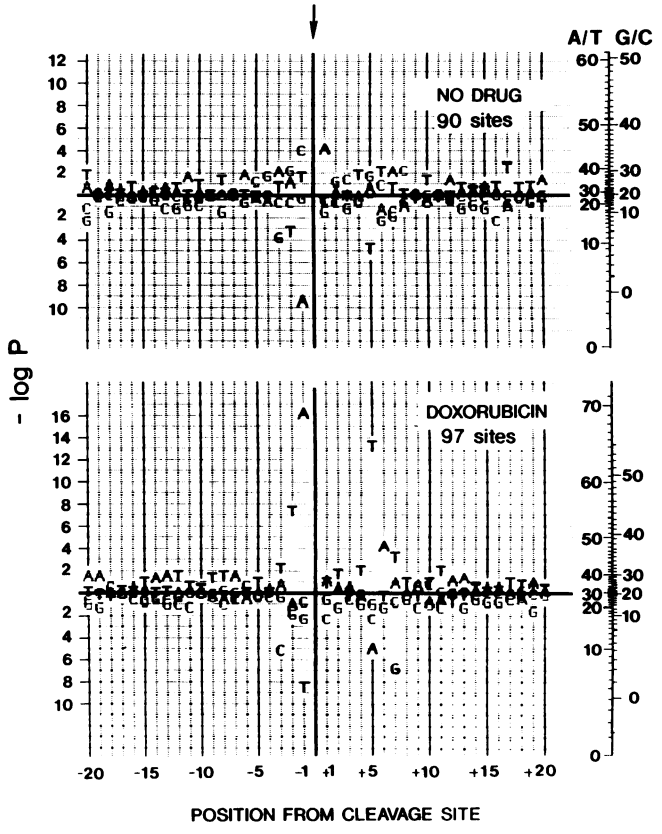

Figure 4. Probabilities of the observed base frequency deviations from expectation at each position of 'no-drug sites' (top) and 'doxorubicin sites' (bottom). In the y-axis, P is the probability of observing that deviation or more, either as excess (above base line) or deficiency (below base line) relative to the expected frequency of each individual base. The expected frequencies were based on the overall base composition of SV40. The number of occurrences of A/T and G/C corresponding to the −log P scale are given on separate scales on the right.

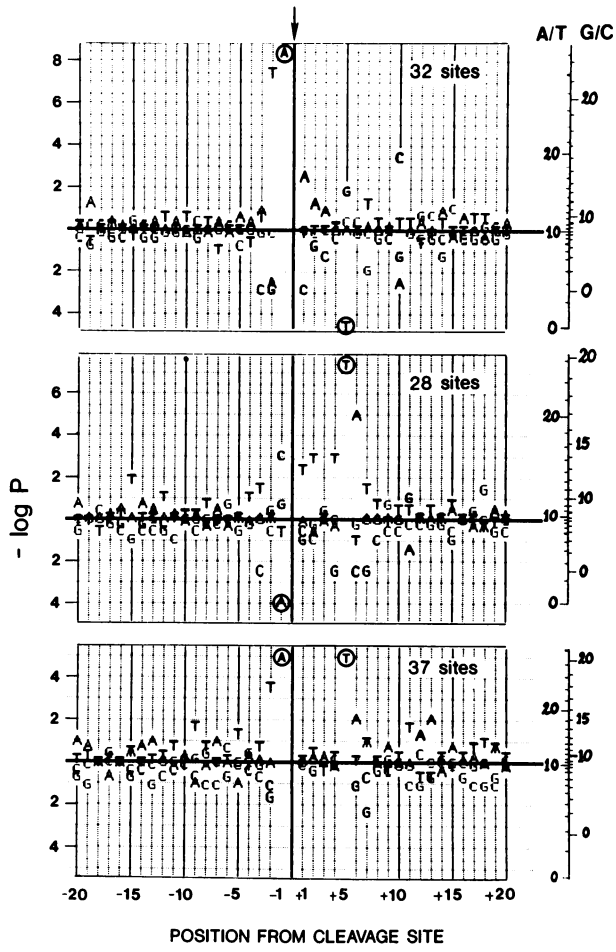

Figure 5. Probability calculation of base occurrence at each position for the three subsets of 'doxorubicin sites' (see Table 2): top, sites having the A at position −1, but lacking the T at position +5; middle, sites having the T, but lacking the A; bottom, sites having both the A and the T. Circle letters represent selected nucleotides at positions −1 and +5.

Table 2. DNA cleavage sites induced by topoisomerase II with or without doxorubicin with respect to the presence of an adenine at position −1 and/or a thymine at position +5

Position −1	Position +5	No-drug sites (%)		Doxorubicin sites(%)	
A	T	0	(0)	37	(38)
A	not-T	3	(3)	32	(33)
not-A	T	9	(10)	28	(29)
not-A	not-T	78	(87)	0	(0)
Total		90	(100)	97	(100)

DISCUSSION

The coincidence between doxorubicin-induced DNA cleavage and nuclear matrix attachment in the same region, taken together with the observation that doxorubicin-induced cytotoxicity occurs at low level of DNA cleavage in cells (13, 14), raises the possibility that the high antitumor activity of doxorubicin may be related to its selective action at critical sites of topoisomerase II activity, such as nuclear matrix attachment regions.

The current work sought to determine the base sequence dependence for the formation of DNA cleavage complexes by topoisomerase II in the presence or absence of doxorubicin. The strongest dependence noted had to do with the presence or absence of an A immediately 5′ to a potential cleavage site. Topoisomerase II produces double-strand cuts containing 4-base pair 5′ overhangs (15−18); thus, relative to an observed cut between positions −1 and +1 on one strand, there may also be a cut between positions +4 and +5 on the opposite strand (fig. 6). The doxorubicin-induced cleavage sites analyzed in the current work showed an absolute requirement for an A at −1 or a T at +5. The latter corresponds to an A immediately 5′ to the expected cleavage on the opposite strand, in agreement with the dyadic symmetry of double-strand cuts. The requirement for doxorubicin-induced cuts was the converse of what was found in the absence of drug: cleavage sites in the absence of drug never had both an A at −1 and a T at +5 (table 2).

Table 3 Relationship between cleavage intensity and presence of an adenine and/or a thymine at positions −1 and +5 of the 'Doxorubicin-Induced Sites'.

Position −1	Position +5	Cleavage Intensity[a] 1	2	3	Total Number of sites
A	T	12	14	11	37
A	not-T	11	13	8	32
not-A	T	16	9	3	28
Total		39	36	22	97

[a] 1, 2 and 3 correspond to weak, intermediate and strong cleavage intensities, respectively.

The enzyme appears to be positioned by the local base sequence at a 'primary site' where the strand may be cut. Once positioned, the enzyme may also cut at a 'secondary site' on the opposite strand where the local base sequence requirement is less stringent. An A immediately 5′ to a potential site seems to prevent primary cutting at that site in the absence of drug, but is necessary for primary cutting in the presence of doxorubicin.

For the doxorubicin-induced sites, a T at +5 was nearly as effective in facilitating cleavage as an A at −1 (table 2). Thus the primary cut could be on either strand (fig. 6). Of a total of 32 + 37 = 69 doxorubicin-induced sites that had an A at −1, the number of T's at +5 would be expected by chance, assuming independence of the bases at the 2 locations, to be $69 \times 0.296 = 20.4$ (29.6% A in SV40), whereas the number found, 37, was significantly greater ($P < 0.0001$). Similarly, there were a total of 65 sites that had a T at +5, of which 37, rather than the expected 19.2, had an A at −1 ($P < 0.00001$). Thus the simultaneous presence of A at −1 and T at +5 increases the likelihood of cleavage; it also tends to increase the average intensity of cleavage (table 3). Clearly however there are additional base sequence requirements, not yet completely identified, for site recognition and cleavage. These calculations take into account the high A+T content of SV40 DNA. It is possible that high A+T content favors topoisomerase II cleavage sites since GC-rich regions that we have examined, both in SV40 and in the c-myc gene have shown very few doxorubicin sites.

We have tried to develop a model to account for the striking requirement of an A on the 5′ side of primary cleavage sites induced by doxorubicin and for the prohibition of cleavage at such sites in the absence of drug (figs. 6 and 7). Nearly all of the known topoisomerase inhibitors either are DNA intercalators or have planar ring systems that might intercalate between base pairs when additional stabilization is conferred in the enzyme-DNA complex. We assume therefore that doxorubicin is

Table 4. Preferred Bases around Topoisomerase II Cleavage Sites

	NO DRUG SITES			DOXORUBICIN SITES
Position 5′→ 3′	Present study (90 sites) (*Mouse*)	Sander & Hsieh (Ref. 17) (16 sites) (*Drosophila*)	Spitzner & Muller (Ref. 18) (71 sites) (*Chicken*)	Present study (97 sites) (*Mouse*)
−4	(G)	.	.	.
−3	not-G (not-C)	A/T	.	not-C
−2	not-T (not-C)	A	G	T
−1	not-A (not-G)	C/T	C/T	A(not-T)
+1	A	A	.	(not-C)
+2	.	T	G	.
+3	.	T	G/T	.
+4	(T)	.	T	(T)
+5	not-T	A	.	T(not-A)
+6	(not-G)(not-A)	T	C/T	A
+7	(A)(T)	.	.	not-G
+8	(C)	.	C/T	.
+9	.	G	G	.

A dot indicates no base preference; parentheses indicate weak bias.

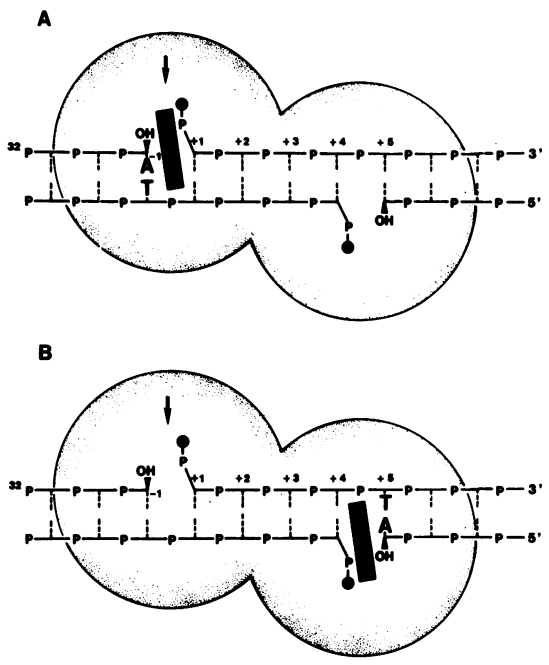

Figure 6. Schematic representation of the ternary complex of topoisomerase II, doxorubicin and DNA, showing intercalation of doxorubicin (black rectangle) at (**A**) the observed cleavage site (arrow), or at (**B**) the expected cleavage site on the opposite strand.

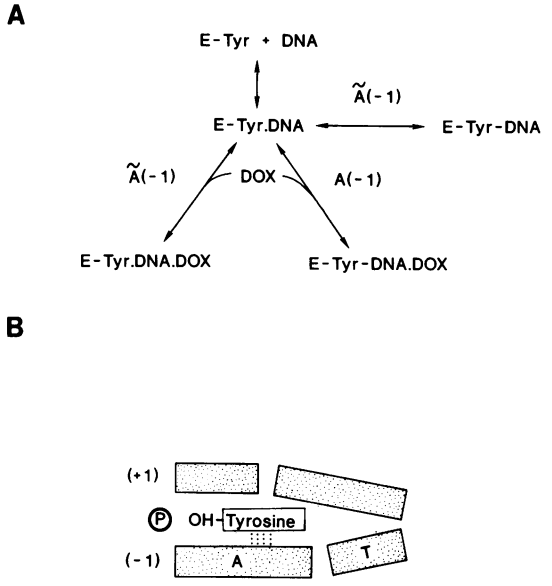

Figure 7. A: Reactions at a primary topoisomerase II DNA cleavage site in the presence or absence of doxorubicin (DOX), showing the dependence on the base (whether it is an A or not an A) immediately 5′ (−1 position) to the cleavage site. Topoisomerase II (E-Tyr) first forms a non-covalent complex (E-Tyr.DNA) at a potential cleavage site. If the base at −1 is not an A, this complex readily converts to a (transient) covalent cleavage complex (E-Tyr-DNA); however an A at −1 usually prevents the conversion. Doxorubicin can bind to the initial non-covalent complex, forming a ternary complex that can be a covalent DNA cleavage complex only if the −1 base is an A (right branch), but otherwise would be a non-covalent complex (left branch). **B**: Proposed configuration of the initial non-covalent complex between topoisomerase II and DNA with partial intercalation of the tyrosine residue. If the base at the −1 position is an A, then a specific stacking interaction between the tyrosine and adenine rings is assumed to reduce access of the tyrosine OH group to the phosphorus atom of the DNA backbone. A potential space adjacent to the tyrosine ring is assumed to provide a site where doxorubicin or other intercalating ring systems can enter with high affinity. When doxorubicin binds in the complex, it displaces the intercalated tyrosine ring, either bound (DNA cleavage complex) or not bound (hypothetical non-covalent complex) to the DNA phosphorus, depending on whether or not the base at −1 is an A.

intercalated between base pairs in the enzyme-DNA complex. We suggest further that the intercalation is at the cleavage site. The intercalated doxorubicin could then interact with the A on the 5′ side by stacking in some special way that allows the formation and stabilization of a DNA cleavage complex.

How might this come about, and how could one account for the converse effect of the same A in the absence of drug? To form a cleavage complex, the OH group of the enzyme's reactive tyrosine must have access to the phosphorus atom in the DNA backbone. We propose that the enzyme forces the tyrosine ring into a partial intercalation complex from which the backbone phosphorus atom is accessible (fig 7, bottom). When there is an A on the 5′ side, however, the interaction between that A and the tyrosine ring may be such that access to the phosphorus atom is prevented. Since the tyrosine ring is much smaller than a base pair, there would be a potential space which could serve as a

high affinity entry point for a variety of intercalator ring systems, including doxorubicin. The entry of an intercalator ring system would displace the tyrosine ring. There are 2 displacement paths that the tyrosine could follow: (1) it could leave by the same path by which it was forced in, leading to a non-covalent complex without DNA cleavage, or (2) it could be pushed towards the phosphorus atom and lead to a covalent cleavage complex. The path chosen would depend on the interaction geometry between the stacked intercalator and the base-pairs. When the 5′ base is an A and the incoming intercalator is doxorubicin, the tyrosine would be displaced from its stacking with the A and would, we assume, be pushed towards the phosphorus atom to produce a cleavage complex. According to this model, doxorubicin would also bind to sites that do not have a 5′ A, but in this case the displacement path of the tyrosine would be such that the resulting complex would be non-covalent and would not involve DNA cleavage.

Although topoisomerase inhibitors differ widely in the sites where they stabilize cleavage complexes, doxorubicin and related anthracyclines are unique in that sites cleaved by topoisomerase II in the absence of drug are invariably suppressed and never stimulated by these drugs. This could perhaps be related to the perpendicular orientation of intercalation of these drugs (19, 20) and to a required phenolic group (6) on the anthracycline ring which might interact analogously to a tyrosine residue. In general, the cleavage site selection for different inhibitors would depend on the electric dipole interactions between the intercalated ring and the adjacent base pairs. Depending upon the exact orientations of these forces, an intercalator could stabilize either a cleavage complex or a non-covalent complex (fig. 7, top).

Consensus sequences for cleavage sites have been reported for topoisomerase II from Drosophila (17) and from chicken erythrocytes (18). Our findings with the mouse enzyme are generally consistent with these previous reports, although we express our conclusions in terms of negative as well as positive bias (Table 4). For example, position −2 is reported as consensus A by Sander & Hsieh and G by Spitzner & Muller, while our finding of not-T and secondarily not-C is consistent with both. The main discrepancy is at position +9 where both previous reports found consensus G, while we did not find any consistent bias. Differences due to species of origin of the enzymes are not excluded. Another possible source of variation is that the data of Spitzner & Muller included some sites observed in the presence of *m*-AMSA or VM-26. In any case, our findings deviate from the previous consensus sequences no more, and perhaps less than the deviation between the previous reports.

In fact, the relatively low DNA sequence selectivity of mammalian topoisomerase II suggests that different classes of topoisomerase II cleavage sites may exist in order for the enzyme to act at multiple sites of the genome. This would be consistent with the pleiotropic roles of DNA topoisomerase II *in vivo*, including regulation of DNA topology during replication and RNA transcription (21−24), nuclear matrix organization (7, 25−27), and chromosome structure (28−30). Doxorubicin might have revealed one extreme case of cleavage sites, which are not cut *in vitro* but that may be used for topoisomerase II-mediated DNA strand passage and topoisomerization *in vivo*.

ACKNOWLEDGEMENTS

G.C. was supported by awards from the Associazione Italiana per la Ricerca sul Cancro and from the EORTC−NCI Research Training Program.

REFERENCES

1. Wang, J.C. (1985) *Annu. Rev. Biochem.* **54**, 665−697.
2. Wang, J.C. (1987) *Biochim. Biophys. Acta* **909**, 1−9.
3. Liu, L.F. (1989) *Annu. Rev. Biochem.* **58**, 351−375.
4. Pommier, Y., and Kohn, K.W., (1989) In Glazer, R.I. (ed.), Developments in Cancer Chemotherapy. CRC Press, Inc., Boca Raton, FA, vol 2, pp. 175−196.
5. Tewey, K.M., Rowe, T.C., Yang, L., Halligan, B.D., and Liu, L.F. (1984) *Science* **226**, 466−468.
6. Capranico, G., Zunino, F., Kohn, K.W., and Pommier, Y. (1990) *Biochemistry* **29**, 562−569.
7. Pommier, Y., Cockerill, P.N., Kohn, K.W., and Garrard, W. (1990) *J. Virol.* **64**, 419−423.
8. Minford, J., Pommier, Y., Filipski, J., Kohn, K.W., Kerrigan, D., Mattern, M., Michaels, S., Schwartz, R., and Zwelling, L.A. (1986) *Biochemistry* **25**, 9−16.
9. Fesen, M., and Pommier, Y. (1989) *J. Biol. Chem.* **19**, 11354−11359.
10. Pommier, Y., Covey, J.M., Kerrigan, D., Markovits, J., and Pham, R. (1987) *Nucleic. Acids. Res.* **15**, 6713−6731.
11. Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., (1986) In The art of scientific computing. Cambridge University Press, pp. 156−160.
12. Prives, C., Covey, L., Michael, T., Lewis, E.D., and Manley, J.L., (1986) In Botchan, M., Grodzicker, T., and Sharp, P. (eds.), DNA tumor virus. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., vol 4, pp. 125−135.
13. Zwelling, L.A., Michaels, S., Erickson, L.C., Ungerleider, R.S., Nichols, M., and Kohn, K.W. (1981) *Biochemistry* **20**, 6553−6563.
14. Ross, W.E., Glaubiger, D., and Kohn, K.W. (1979) *Biochim. Biophys. Acta* **562**, 41−50.
15. Sander, M., and Hsieh, T. (1983) *J. Biol. Chem.* **258**, 8421−8428.
16. Liu, L.F., Rowe, T.C., Yang, L., Tewey, K.M., and Chen, G.L. (1983) *J. Biol. Chem.* **258**, 15365−15370.
17. Sander, M., and Hsieh, T.S. (1985) *Nucleic. Acids. Res.* **13**, 1057−1072.
18. Spitzner, J.R., and Muller, M.T. (1988) *Nucleic. Acids. Res.* **16**, 5533−5556.
19. Quigley, G.J., Wang, A.H., Ughetto, G., Van der Marel, G., Van Boom, J.H., and Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A* **77**, 7204−7208.
20. Frederick, C.A., Williams, L.D., Ughetto, G., Van der Marel, G.A., Van Bloom, J.H., Rich, A., and Wang, A.H.J. (1990) *Biochemistry* **29**, 2538−2549.
21. Liu, L.F., and Wang, J.C. (1987) *Proc. Natl. Acad. Sci. U.S.A* **84**, 7024−7027.
22. Tsao, Y.P., Wu, H.Y., and Liu, L.F. (1989) *Cell.* **56**, 111−118.
23. Yang, L., Wold, M.S., Li, J.J., Kelly, T.J., and Liu, L.F. (1987) *Proc. Natl. Acad. Sci. U.S.A* **84**, 950−954.
24. Snapka, R.M. (1986) *Mol. Cell. Biol.* **6**, 4221−4227.
25. Berrios, M., Osheroff, N., and Fisher, P.A. (1985) *Proc. Natl. Acad. Sci. U.S.A* **82**, 4142−4146.
26. Izaurralde, E., Mirkovitch, J., and Laemmli, U.K. (1988) *J. Mol. Biol.* **200**, 111−125.
27. Cockerill, P.N., and Garrard, W.T. (1986) *FEBS. Lett.* **204**, 5−7.
28. Earnshaw, W.C., Halligan, B., Cooke, C.A., Heck, M.M., and Liu, L.F. (1985) *J. Cell. Biol.* **100**, 1706−1715.
29. Earnshaw, W.C., Halligan, B., Cooke, C.A., Heck, M.M., and Liu, L.F. (1985) *J. Cell. Biol.* **100**, 1706−1715.
30. Boy de la Tour, E., and Laemmli, U.K. (1988) *Cell.* **55**, 937−944.