Lin, S. Y., & Grollman, A. P. (1981) *Biochemistry 20*, 7589–7598.
Lo, K.-M., Jones, S. S., Hackett, N. R., & Khorana, H. G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2285–2289.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
Nelson, J. W., & Tinoco, I., Jr. (1985) *Biochemistry 24*, 6416–6421.
Povirk, L. F., & Goldberg, I. H. (1983) *J. Biol. Chem. 258*, 11763–11767.
Povirk, L. F., Dattagupta, N., Warf, B. C., & Goldberg, I. H. (1981) *Biochemistry 20*, 4007–4014.
Sausville, E. A., Stein, R. W., Peisach, J., & Horwitz, S. B. (1978) *Biochemistry 17*, 2746–2754.
Schultz, P. G., & Dervan, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6834–6837.
Sekiya, T., Besmer, P., Takeya, T., & Khorana, H. G. (1976) *J. Biol. Chem. 251*, 634–641.
Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.
Takeshita, M., Grollman, A. P., Ohtsubo, E., & Ohtsubo, H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5983–5987.
Takeshita, M., Kappen, L. S., Grollman, A. P., Eisenberg, M., & Goldberg, I. H. (1981) *Biochemistry 20*, 7599–7606.
Van Dyke, M. W., & Dervan, P. B. (1982) *Cold Spring Harbor Symp. Quant. Biol. 47*, 347–353.
White, S. A., & Draper, D. E. (1987) *Nucleic Acids Res. 15*, 4049–4064.
Wu, J. C., Kozarich, J. W., & Stubbe, J. (1985) *Biochemistry 24*, 7562–7568.

Sequence Specificity in Photoreaction of Various Psoralen Derivatives with DNA: Role in Biological Activity†

Véronique Boyer, Ethel Moustacchi, and Evelyne Sage*
Section de Biologie, Institut Curie, 26 rue d'Ulm, 75231 Paris Cedex 05, France
Received June 22, 1987; Revised Manuscript Received December 11, 1987

ABSTRACT: The sequence specificity in the photoreaction of various psoralen derivatives with DNA is investigated by using DNA sequencing methodology. The 3′–5′ exonuclease activity associated with T_4 DNA polymerase serves as a probe to map the psoralens' photoaddition (monoadducts plus biadducts) on DNA fragments of defined sequence. This approach has already allowed us to demonstrate a strong sequence context effect on the 8-methoxypsoralen photobinding to DNA [Sage, E., & Moustacchi, E. (1987) *Biochemistry 26*, 3307–3314]. The psoralens studied include bifunctional derivatives [8-methoxypsoralen, 5-methoxypsoralen, and 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen] and monofunctional derivatives (angelicin, 3-carbethoxypsoralen, and three pyridopsoralens). Maps of photochemical binding on two DNA fragments of the *lacI* gene of *Escherichia coli* are established for all the derivatives. These maps demonstrate the following general qualitative rules in the photoreaction of the furocoumarins with DNA: thymine residues in a GC environment are cold, adjacent thymines are better targets, 5′-TpA sites are strongly preferred versus 5′-ApT, and alternating $(AT)_n$ sequences are hot spots for photoaddition. Depending on the chemical structure of the derivatives and on their affinity for DNA, some minor differences in the binding spectrum are detected. A most interesting example is 3-carbethoxypsoralen, which specifically reacts with $(AT)_n$ sites. Our observations lead us to define two types of target sites: the "strong sites", which are preferential targets for all psoralen derivatives, and the "weak sites", which are targets only for derivatives having a high affinity for DNA. The frequency of DNA lesions is much higher in the former sites. The specific photochemical binding of each derivative is discussed in conformational terms and related to the repair and mutagenesis induced by the studied furocoumarins.

Psoralens are a class of compounds of particular interest. In addition to their clinical use in the treatment of certain skin diseases, they have proven to be excellent probes for studying chromatin and nucleic acid structure. Moreover, psoralen–DNA photoadducts provide an attractive model to study the influence of DNA sequence in relation to mutagenicity and to analyze basic DNA repair mechanisms.

The photoreaction of psoralens with DNA is a multistep process [see for review Song and Tapley (1979)]. First, the psoralen forms an intercalative noncovalent complex with DNA. Upon exposure to UVA radiation (320–400 nm), cycloaddition occurs to the 5,6 double bond of pyrimidine bases (mainly thymine) through the pyrone or the furan ring of the psoralen. In the case of bifunctional psoralens, the thymine–furan side monoadduct can absorb a second photon of near-UV light and, when properly located, be converted to a diadduct that yields an interstrand cross-link. The effect of adducts on the local structure of the DNA depends on the psoralen derivatives. Typically, 8-methoxypsoralen (8-MOP)[1] photoadducts provoke an unwinding of the double helix (Wiesehahn & Hearst, 1978).

In addition to natural furocoumarins, several families of psoralen compounds have been synthesized. The studies on

† This work was financially supported by the CNRS, the CEA (Saclay, France), the ARC (Villejuif, France), and Contracts 852017 (INSERM) and B16-151-F (CCE, Bruxelles).
* Author to whom correspondence should be addressed.

[1] Abbreviations: bp, base pair(s); EDTA, ethylenediaminetetraacetic acid; ATP, adenosine triphosphate; 8-MOP, 8-methoxypsoralen; 5-MOP, 5-methoxypsoralen; 3-CPs, 3-carbethoxypsoralen; MePyPs, 7-methylpyrido[3,4-*c*]psoralen; BCH 394, 11-methylpyrido[3,4-*h*]psoralen; HMT, 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride.

0006-2960/88/0427-3011$01.50/0

the genotoxic effects of some of these derivatives exhibit quite extensive differences, if one compares to each other bifunctional compounds, monofunctional derivatives, or monofunctional versus bifunctional compounds (Ben-Hur & Song, 1984). Their physical constants have been described [Hearst, 1981; see Gasparro (1987) for review], and their photochemistry is often known. For some of them, the photoadducts have been identified. Nevertheless, the biological difference observed cannot be explained only by the above physical data.

In order to get a better understanding of the differential genotoxic effects of various psoralen derivatives, we have studied the photobinding of several furocoumarins to DNA at the nucleotide level. The hope was to detect subtle differences in the sequence specificity of the photoreaction of these compounds. Indeed, using DNA sequencing methodology, we have been able to study the sequence context effect on the photobinding of 8-MOP to DNA fragments of defined sequence (Sage & Moustacchi, 1987). Using the same approach, we now show a general rule in the photobinding of psoralens to DNA. In addition, we observe some limited variations in the photoreaction. These observations lead us to define "hot sites" that are preferentially hit by any furocoumarins. The last part of this paper is an attempt to correlate our results together with physicochemical data obtained by others to the reparability and mutagenicity of the photolesions induced by the studied furocoumarins. This work defines in qualitative terms the sequence specificity of the photoreaction of diverse compounds known to differ in their genotoxic activity.

Experimental Procedures

Chemicals and Enzymes. 8-Methoxypsoralen (8-MOP) was a commercial product from Sigma, and 5-methoxypsoralen (5-MOP) was from Sarsyntex-Interchim (Montluçon, France). 3-Carbethoxypsoralen (3-CPs), pyrido[3,4-*c*]psoralen (PyPs), 7-methylpyrido[3,4-*c*]psoralen (MePyPs), 11-methylpyrido-[3,4-*h*]psoralen (BCH 394) were synthesized in Dr. E. Bisagni group (Institut Curie, Orsay, France). Angelicin was synthesized by Dr. R. Royer (Institut Curie, Paris). 4′-(Hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT), a product from HRI Associates Inc. (Emeryville, CA), was a gift from Dr. J. Hearst. The structure of molecules used is given in Figure 1. The psoralen derivatives were in solution at 5×10^{-3} M in dimethyl sulfoxide. [γ-^{32}P]ATP (specific activity 3000 Ci/mmol) was obtained from either Amersham or New England Nuclear. Restriction enzymes, T_4 polynucleotide kinase, and T_4 DNA polymerase were purchased from Bethesda Research Laboratories. Calf intestine alkaline phosphatase came from P-L Biochemicals and proteinase K from Boehringer-Mannheim.

DNA Preparations. Phage M13 mp8 *lacI* 935 resulted from cloning of the 935 bp *Hin*cII fragment of the *lacI* gene of *Escherichia coli* at the *Hin*dIII site of the polylinker of M13 mp8 phage. Replicative form (RF) DNA of phages M13 mp8 and M13 mp8 *lac*I 935 were isolated by standard procedure from infected *E. coli* JM 109 or JM 101.

Preparation of 5′-^{32}P-End-Labeled DNA Fragments. End-labeled fragments HN2 (49 bp) and HN3 (78 bp) were prepared from M13 mp8 *lac*I 935 RF DNA as described in Sage and Moustacchi (1987). EP1 and EP2 fragments of 123 and 178 bp, respectively, were produced by digestion of M13 mp8 RF DNA with *Eco*RI, 5′ end labeling, and further hydrolysis by *Pvu*II.

Photobinding of Psoralen Derivatives to DNA. The protocol is described in Sage and Moustacchi (1987). Briefly, aliquots of 5′-end-labeled DNA were incubated in 20 μL of TE buffer

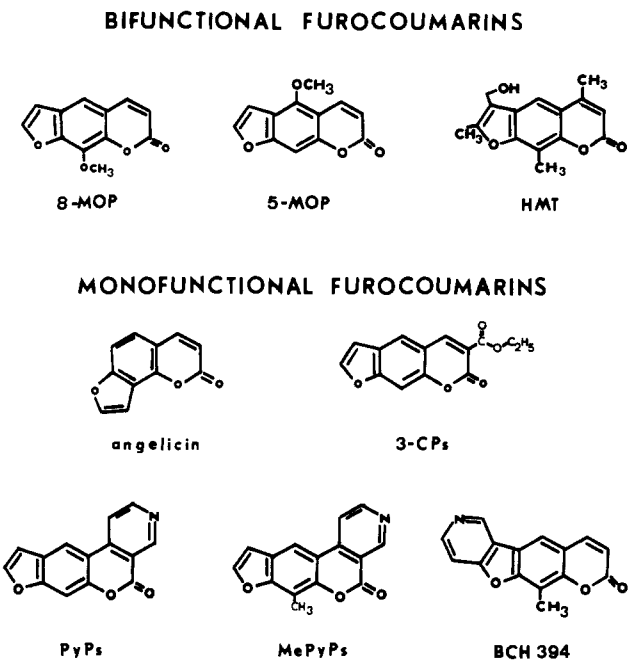

FIGURE 1: Molecular structure of furocoumarins used.

(10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and 10–20% dimethyl sulfoxide, in the presence of psoralen derivative (concentration given in text). Whatever were the affinity constants of the psoralen derivatives for the DNA, there was a large excess of psoralen in the reaction mixture. The mixture received various doses of UVA light (365 nm). Unbound psoralen molecules were eliminated by four chloroform–isoamyl alcohol (19:1 v/v) extractions followed by ethanol precipitation. Controls included DNA samples treated with psoralen and no UVA or with UVA and no psoralen.

Digestion of Psoralen-Modified DNA with T_4 DNA Polymerase 3′–5′ Exonuclease. Treated samples were digested with 3.5 units of T_4 DNA polymerase 3′–5′ exonuclease in the absence of dNTP, as indicated in Sage and Moustacchi (1987).

Analysis of Damaged Sites by Sequencing Gel Electrophoresis. Before electrophoresis of the DNA fragments modified with bifunctional derivatives (8-MOP, 5-MOP, HMT) and digested with T_4 DNA polymerase exonuclease, interstrand cross-links were photoreversed by irradiation at 254 nm UV at a dose of 6 kJ m^{-2} (Sage & Moustacchi, 1987).

Digestion products were analyzed on sequencing gels as described by Maxam and Gilbert (1980). The polyacrylamide concentrations of the gels varied between 12% and 20% according to the length of the DNA fragments to be analyzed. The digested samples were coelectrophoresed with undigested samples and with the four Maxam and Gilbert sequencing reaction products of unmodified DNA. The T_4 DNA polymerase associated exonuclease termination products were quantified by cutting bands out of the gel and measuring by Cerenkov counting the radioactivity relative to the total radioactivity in the lane. To account for more than one photoproduct per DNA strand, a correction factor was applied as described in Gordon and Haseltine (1980).

Results and Discussion

General Sequence Specificity in the Photoreaction of Several Psoralen Derivatives. We have previously shown that the 3′–5′ exonuclease associated with the DNA polymerase of phage T_4 is blocked by 8-MOP monoadducts, as well as biadducts, and have concluded that it is a suitable tool for mapping psoralen DNA adducts (Sage & Moustacchi, 1987). We report here a qualitative study of the photoreaction of

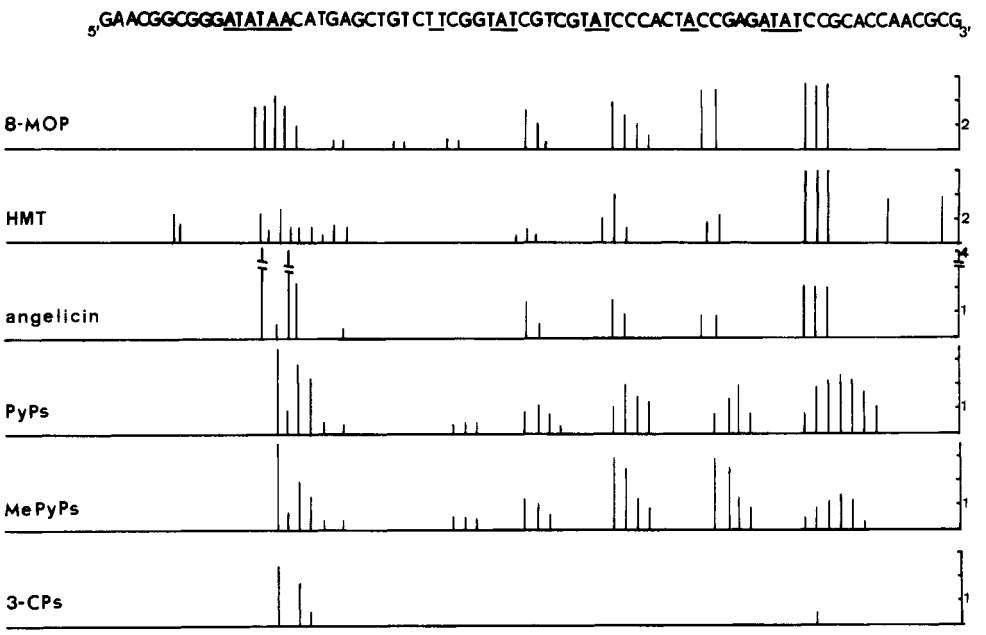

FIGURE 2: Distribution of photoadducts induced in HN3 DNA fragment by several psoralen derivatives. The height of bars reflects the intensity of corresponding bands on autoradiograms. A bar expresses the percentage of digestion product migrating at a particular position over the total material in the lane, corrected for multiple hits in the same strand [see Sage and Moustacchi (1987)]. The conditions of photoreaction are (2.5–5) × 10^{-4} M furocoumarin and a UVA irradiation dose of 6 kJ m^{-2} (8-MOP and HMT), 36 kJ m^{-2} (angelicin and 3-CPs), or 72 kJ m^{-2} (PyPs and MePyPs).

various psoralen derivatives at different sites in DNA fragments of defined sequence. In our definition a "site" is a group of adjacent nucleosides that may include one or several thymidines (or cytidines).

Two 5′-end-labeled DNA fragments HN2 and HN3 were incubated in presence of the bifunctional psoralen derivatives, 8-MOP, 5-MOP, or HMT, or the monofunctional derivatives, angelicin, PyPs, MePyPs, BCH 394, or 3-CPs, at a concentration of (2.5–5) × 10^{-4} M and irradiated with various doses of 365-nm UVA light. Reacted DNAs were then subjected to hydrolysis by T_4 DNA polymerase 3′–5′ exonuclease, and termination sites were analyzed by high-resolution gel electrophoresis alongside Maxam and Gilbert sequence ladders. Undigested controls and modified DNA fragments were run in parallel. Photoreversion of cross-links after enzymatic hydrolysis was systematically done when bifunctional compounds were the damaging agents (see Experimental Procedures). In that case, the termination products appearing on autoradiograms correspond to a blockage of the exonuclease at monoadducts and biadducts as well. An example of a sequence band pattern representing the exonuclease termination sites at photoadducts is given in Figure 4 and will be described later on. Mapping and quantitation are described in Sage and Moustacchi (1987). Indeed, this method allows the determination of total amount of interstrand cross-links and also the quantification of total DNA adducts. This last determination was not the purpose of the actual work and consequently was not performed. However, in some instances the total amount of biadducts was quantified from undigested DNA fragments in the absence of photoreversion. In the case where denaturation in untreated samples was not complete, a correction factor was applied.

The length of the exonuclease termination products allows us to determine the sites of photoaddition. In order to account for the difference in the 3′ termini of digested and chemically cleaved DNA fragments, a correction of 1.5–2 nucleotides should be applied. Furthermore, the presence of the psoralen moiety delays the migration of a given termination product (in the absence of photoreversion). The delay in migration

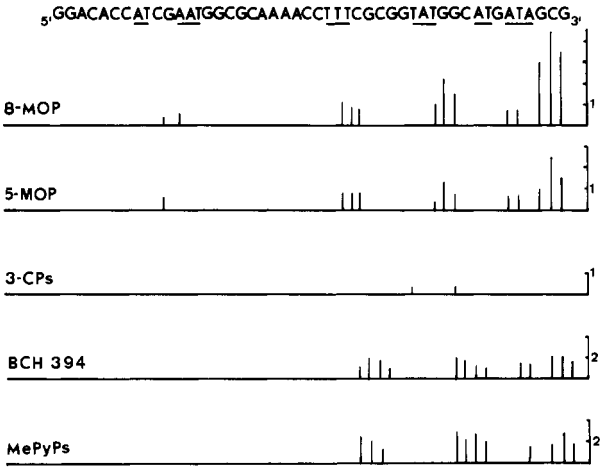

FIGURE 3: Distribution of photoadducts induced in HN2 DNA fragment by several psoralen derivatives. The conditions of photoreaction are (2.5–5) × 10^{-4} M furocoumarin and a UVA irradiation dose of 24 kJ m^{-2} (8-MOP), 36 kJ m^{-2} (5-MOP, BCH 394, and MePyPs), and 72 kJ m^{-2} (3-CPs).

will depend on the structure of the furocoumarin derivative. The psoralen moiety is eliminated by photoreversion (Cimino et al., 1986; Sage & Moustacchi, 1987). Figures 2 and 3 show the mapping of the exonuclease termination products observed on the HN3 and HN2 DNA fragments treated by several psoralen derivatives. Each bar corresponds to a band on autoradiograms (see Figure 4). The presence of several termination products relative to a single potential target (i.e., 5′-TpA) is remarkable and requires some comments. Cytosine monoadducts are unlikely to contribute much in the numerous bands observed, in view of their low frequency for most derivatives with the exception of HMT (Kanne et al., 1982; Calvin & Hanawalt, 1987). In this last case a strong sequence specificity is observed (see Sequence Context Effect in the Formation of HMT–Cytosine Photoadducts). Psoralen derivatives form, in different proportions, several isomeric photoadducts (Kanne et al., 1982; Straub et al., 1981; Moysan, 1987), and some of them are likely to be resolved on gels

(Cimino et al., 1986). The termination site of the enzyme and the number of termination products may be influenced by both the configuration of the different isomeric photoadducts and the strength of the stacking interaction of the psoralen moiety with adjacent bases. These factors as well as the local conformation at the modified sites influenced by the sequence context would surely affect the kinetics of hydrolysis, but this problem is circumvented in the conditions of our assay [2 h of incubation, and see Sage and Moustacchi (1987)]. Although there is a certain ambiguity about the origin of the multiple bands, the sites of photoaddition, as defined earlier, are quite clear and can be compared in their reactivity toward various psoralen derivatives, at least on a qualitative basis.

The most striking feature revealed by Figures 2 and 3 is that a similar photobinding spectrum appears for the different derivatives. The same sites are hit with relatively similar frequencies, by most derivatives. This observation demonstrates a general rule in the photoreaction of the studied psoralen derivatives with DNA and is likely to be applicable to all furocoumarins. The sequence specificity in the photobinding of 8-MOP to DNA was studied in detail in Sage and Moustacchi (1987). The sequence selectivity can be summarized as the following: weak reactivity of thymine in a GC environment, better reactivity of adjacent thymines, strong preference for 5′-TpA sites versus 5′-ApT, and repeated $(AT)_n$ sequences as hot spots for the photofixation of psoralens. The 49 bp DNA fragment possesses three 5′-ApT sites, and two of them are almost unreactive, whatever is the furocoumarin used (Figure 3). In addition, this fragment has a run of three thymines that are quite reactive with most of the derivatives. The conditions used for the photoreaction are such that there is a maximal amount of intercalated drug in DNA prior to irradiation (Sage & Moustacchi, 1987). The fraction of occupied sites in DNA depends on the affinity constant of the drug for DNA and varies from one compound to another (Hearst, 1981). Furthermore, a UVA dose response was established for all derivatives (data not shown). Indeed, we are able to detect photoadditions at all potential sites. Premature stops of 3′–5′ exonuclease activity, bypass of lesions, and possible differential kinetics according to lesion are avoided in the extreme conditions of digestion used in our experiments (see above). As discussed below, local differences can be observed from one compound to another, i.e., number of bands per site and reactivity of sites more or less important for the different derivatives.

Sequence Context Effect in the Formation of HMT–Cytosine Photoadducts. This compound is known to form substantial photoadducts at cytosine moieties (Straub et al., 1981). Indeed some termination products related to cytosine photoadducts appear on Figure 2. They correspond to cross-linkable sites 5′-CpA and/or 5′-ApC. All such sites are reactive, except one in the sequence CCCAC. The run of cytosines may confer a greater stability and a lesser flexibility to the double helix, preventing intercalation. In contrast to the photoaddition at thymine positions, runs of cytosines or cytosine residues that are not bordered by an adenine moiety on one side are fully unreactive. This is the first demonstration of sequence specificity in the formation of the cytosine photoadducts. The site ACCA on the 3′ end of the DNA fragment is barely cross-linked since the frequency of photoadducts at this site increases very slightly with increasing UVA doses, in contrast to the site ATAT (not shown). 5′-TpA remains the major cross-linkable site. The ratio of thymine photoadducts to cytosine photoadducts is about 3.5. No cytosine photoadducts are detected in our assay, in the photoreaction of the other derivatives. These observations are in agreement with Straub et al. (1981) and Kanne et al. (1982). The total amount of interstrand cross-links was determined and revealed that HMT forms 5 times more cross-links in DNA than 8-MOP, at a UVA dose of 6 kJ m^{-2}.

HMT–cytosine photoadducts are probably formed in vivo. Piette et al. (1985) have localized mutations photoinduced by HMT in the promoter region of the *lac* operon in phage M13 mp10. Most of the base substitutions occur in AT-rich sequences. Nevertheless, two of them occur at thymine in GTG/CAC sites. These thymines are not reactive toward angelicin (unpublished results). These mutations are probably targeted at HMT cross-links between the thymine and cytosine residues. This is in good agreement with the sequence specificity that we observe in the formation of HMT–cytosine photoadducts.

Peculiarity of 3-CPs. Surprisingly, in the HN3 76 bp DNA fragment (Figure 2), 3-CPs photoreacts only with the hottest site ATATAA. A barely detectable band appears at ATATCC. All the other sites are unreactive. 3-CPs does not photoreact with any of the potential sites of the HN2 DNA fragment (Figure 3). Because 3-CPs has a low affinity for DNA and is highly photodegradable (Gaboriau et al., 1981), we tried to increase photoadduct formation by adding 3-CPs during irradiation and by increasing the UVA doses (up to 10^{-3} M 3-CPs and 144 kJ m^{-2}). No new sites showed up. This demonstrates a strong preference of 3-CPs for repeated $(AT)_n$ regions. In the EPI 123 bp DNA fragment, only the site TTAAT is slightly reactive (1% of termination products when DNA is treated with 5×10^{-4} M 3-CPs plus 72 kJ m^{-2} of UVA light, to be compared to results in Table I). A study of the photoreaction of 3-CPs with DNA reveals that the major adduct is the cis–syn furan-side thymine monoadduct (Gaboriau et al., 1987). Two diastereoisomers have been found in about equal proportion (52% in 5′-TpX sites and 48% in 5′-XpT sites; Moysan, 1987). Fluorescence anisotropy measurements and fluorescence quenching shows that the bulky carbethoxy group does not prevent an internal location of the psoralen ring in the double helix, in agreement with a cis–syn configuration of the adduct, but the secondary DNA structure is locally disrupted (Gaboriau et al., 1987). This slight local denaturation in adducted DNA is sensitive to the S1 endonuclease (Gaboriau, personal communication). The mean angle between the psoralen moiety and the normal to the plane of the nucleic acid bases is slightly lower for 3-CPs (68°–75°) than for 8-MOP (76°–79°) (Vigny et al., 1987). The unwinding angle found for 3-CPs is higher than expected for an average monoadduct (27° versus 18°; Isaacs et al., 1984).

These data show that the polar and bulky carbethoxy group on the carbon 3 of the furocoumarin prevents a "proper" intercalation that would permit a maximum overlapping between reactive atoms of the furocoumarin and the thymine residue. The photoreaction will occur only in a region of great flexibility and instability, i.e., the repeated $(AT)_n$ sequences, allowing structural rearrangement. The low affinity of 3-CPs for DNA may reflect the small number of sites available for the photoreaction of 3-CPs. Our results are in agreement with the low in vitro photoreactivity of 3-CPs observed by Isaacs et al. (1984). In addition to this strong sequence selectivity in the photoreaction of 3-CPs, we have been able to detect a small amount of interstrand cross-links photoinduced by 3-CPs at high UVA doses (unpublished data). A low level of 3-CPs-induced cross-links has also been found in vivo (Gruenert et al., 1985; E. Cundari and D. Averbeck, personal communication). This derivative is not purely monofunctional.

Comparison in the Photoreaction of Pyridopsoralens on DNA. Pyridopsoralens were found to be pure monofunctional derivatives both in vitro and in vivo (Blais et al., 1984; Magaña-Schwencke & Moustacchi, 1985; E. Sage, unpublished data). These compounds have a fused pyridine ring on the 3,4 (PyPs and MePyPs) or the 4′,5′ site (BCH 394) of the furocoumarin (Figure 1). PyPs and MePyPs form only furan-side monoadducts, whereas BCH 394 is able to photoreact with DNA only through its pyrone side. It was of interest to study the photoreactivity of these three derivatives on defined sequences. Indeed, no study on the sequence specificity in the formation of the pyrone-side photoadduct has been reported.

Figure 2 shows that sites of formation of the furan-side photoadducts of PyPs and MePyPs do not differ from the photoadditions of 8-MOP or angelicin. It appears that the two adjacent thymine residues in the fragment are better targets than for 8-MOP or angelicin. It has to be noticed that, at site ATATAA, the exonuclease stops one nucleotide earlier when it is compared to termination sites due to angelicin photoadducts. Figure 3 shows that the mapping of MePyPs and of BCH 394 photoadducts is very similar. It implies that both types of monoadducts, furan side and pyrone side, constitute a block for the T_4 DNA polymerase 3′–5′ exonuclease activity. It demonstrates the same sequence specificity in the formation of furan-side and pyrone-side monoadducts. Again, 5′-ApT sites are poorly reactive and 5′-TpA sites are strongly preferred. Nevertheless, the map obtained with pyridopsoralens differs from the map obtained for 8-MOP or other derivatives in that more bars (bands on autoradiograms in Figure 4) appear at each site in the map of pyridopsoralens photoadducts.

The number of well-defined bands at each site may have several origins. It may represent oligonucleotides bearing one or more pyridopsoralen moieties. Although the level of photobinding in these experiments has not been determined, this hypothesis is not very likely. First, the maximum number of MePyPs photoadducts has been estimated to 2 per 100 base pairs (Blais et al., 1984). According to the low solubility of pyridopsoralens (about 3×10^{-6} M in aqueous solution), to their affinity for DNA (Blais et al., 1984), and to characteristics of other psoralen derivatives given in Hearst (1981), one can estimate that the covalent binding of pyridopsoralens would plateau at no more than 4 per 100 bp. Furthermore, if we had several adducts per DNA fragment in our experiments, we would see a strong decrease, if not a disappearance, of the bands corresponding to the 5′ site ATATAA on the autoradiogram, since the 3′–5′ exonuclease is blocked by the first adduct it meets. Photoadditions at cytosine residues located at the 3′ end of most of the reactive sites in HN3 DNA fragment are unlikely to happen. They would not be in cross-linkable sites as we determined for HMT–cytosine photoadducts. Moreover, the two diastereoisomeric thymine photoadducts account for at least 90% of the covalent bound MePyPs (major product in 5′-TpX) (Moysan, 1987). A preferred hypothesis is that these numerous bands at each site on autoradiograms represent pauses of the exonuclease in the vicinity of photoadducts. The polymerase activity of the T_4 DNA polymerase is known to make pauses or stops at secondary structure like hairpin loops in the template (Huang et al., 1981; Hillebrand & Beattie, 1985). Pyridopsoralen residues bound to DNA are buried in the double helix and exhibits a strong stacking interaction with adjacent bases (Blais et al., personal communication). The exonuclease has to unwind the two strands of the double helix and may stutter when approaching a pyridopsoralen photoadduct. In addition, in

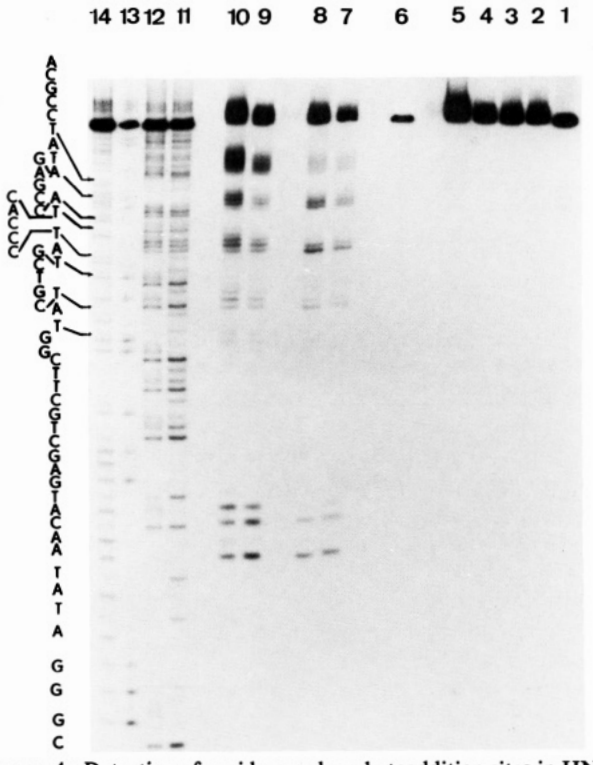

FIGURE 4: Detection of pyridopsoralen photoaddition sites in HN3 DNA fragment. 5′-End-labeled DNA was incubated with 2.5×10^{-4} M MePyPs (lanes 1–3 and 6–8) or PyPs (lanes 4, 5, 9, and 10) and either irradiated at 365 nm at doses of 72 kJ m^{-2} (lanes 2, 4, 7, and 9) and 144 kJ m^{-2} (lanes 3, 5, 8, and 10) or not irradiated (lanes 1 and 6). After 2-h digestion with T_4 DNA polymerase 3′–5′ exonuclease, termination products are resolved on a 16% sequencing gel. In lanes 1–5, enzyme treatment was omitted. Lanes 11–14 correspond to Maxam and Gilbert sequencing products for C + T, C, G, and G + A, respectively.

Figure 4, it is noticed that the digestion of modified samples by the exonuclease is largely incomplete when it is compared to the control sample (34%, 84%, 73%, 84% and 93% of undigested plus partially digested DNA molecules in lanes 6, 7, 8, 9, and 10 of Figure 4, respectively, as determined by trichloroacetic acid precipitation after enzymatic hydrolysis). This is in favor of a stable double-helical structure of the modified DNA fragments. For comparison, HMT monoaddition is able to stabilize the double helix formed by two non-self-complementary oligonucleotides of inverted sequence (Shi & Hearst, 1986). Stops or pauses of the exonuclease at adducts located on the unlabeled strand cannot be completely excluded in the case of pyridopsoralens. Quantitation of termination products shows a plateau in photoaddition at such high-UVA doses (data not shown), and it also indicates that PyPs forms slightly more adducts than MePyPs. A comparison in the photobinding of the three pyridopsoralens on a DNA fragment EPI is given in Table I. The order of reactivity is MePyPs < PyPs < BCH 394.

Table I also indicates that runs of thymines are a good target. The decreased photobinding accompanying increasing irradiation seen in certain cases probably reveals photoreversal of adducts as already observed under certain conditions (Isaacs et al., 1984; Cimino et al., 1986). A new pyridopsoralen, 7-methylpyrido[4,3-*c*]psoralen (2N-MePyPs), differing from MePyPs in the position of the nitrogen atom in the pyridine ring has been used to modify the 76 bp DNA fragment. It exhibits exactly the same specificity in the photoreaction as other pyridopsoralens, i.e., the same usual sites, adjacent thymines as targets, and early stops at the site ATATAA. It is much less reactive, as mentioned in Blais et al. (1987).

Table I: Comparison of Frequency of Pyridopsoralen Photoaddition[a] at Specific Sites on the EP1 DNA Fragment

5′CTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGG-
AAACCCTCGCGTTACCCAACTTAATCGCCTTCC3′

	UVA dose (kJ m^{-2})	TT	TTT	TA	TTA	TTAAT	TT
MePyPs	24	0.4	1.3	0.2	7	3.3	3.5
	72	1.4	2	1.7	13.4	13.4	8.3
PyPs	24	2.1	7.6	3.4	9.9	14.4	6
	72	0.6	2.8	0.9	9	19.5	6.7
BCH 394	24	0.6	2.2	1.2	10.4	10.4	7.1
	72	2.4	6	8.3	20	15	10.3

[a] Corrected percentage of radioactivity in each band to the total radioactivity in the lane after electrophoresis of the digested pyridopsoralen-treated DNA samples.

Curiously, in vivo this compound has much lower genotoxic effects than MePyPs and is comparable to 3-CPs (Blais et al., 1987).

In contrast to 3-CPs, the bulky but hydrophobic pyridine ring of pyridopsoralens assures a better intercalation of the furocoumarin in the double helix and a strong stacking interaction with adjacent bases. These photoadducts do not produce S1 nuclease sensitive sites (Gaboriau, personal communication). The three derivatives have a high affinity for DNA and photoreact even with sites other than 5′-TpA (i.e., TT). Our results show that the blockage of either the pyrone or the furan reactive sites of the furocoumarin by a pyridine group does not change the sequence specificity of the photobinding to DNA. It implies that the ability of a furocoumarin to intercalate in the double helix and to interact with adjacent bases will determine its photoreaction process with DNA. Indeed, the key step in the photoreaction of furocoumarins is the intercalation.

Concept of "Weak Sites" and "Strong Sites". The notion of weak and strong sites derives from our observation that all thymine moieties are not equally reactive to psoralens. We have found basically the same sequence specificity in the photoaddition to thymine of three bifunctional compounds, 8-MOP, 5-MOP, HMT, and six monofunctional compounds, angelicin, 3-CPs, MePyPs, PyPs, BCH 394, and 2N-MePyPs. The sequence selectivity in the formation of a furan-side or a pyrone-side monoadduct is similar.

Weak sites are T or TT in a GC context; they are typically hit by compounds having a high association constant in the intercalation step, i.e., pyridopsoralens. A typical example of strong sites is given in the photobinding of 3-CPs. This derivative has a low affinity for DNA and is found covalently bound only to AT-rich sites, almost exclusively to repeated $(AT)_n$ sequences. We have already mentioned that repeated $(AT)_n$ region are hot spots for the photobinding of 8-MOP (Sage & Moustacchi, 1987). Sobell et al. (1982) have proposed β-kinked DNA, a hyperflexible structure, as a site for drug intercalation. These authors have defined alternating $(AT)_n$ sequences, through their 5′-TpA units, as "kinky sequences", which are specific sites for various nucleases' action. Pearlman et al. (1985) have emphasized a strongly favored psoralen photobinding with sequences able to form a kink, since they have suggested that psoralen cross-links will bend DNA. Nevertheless, analysis of 8-MOP-cross-linked DNA fragments on polyacrylamide gels did not reveal any substantial bend in the adducted DNA (Sinden & Hagerman, 1984). A recent study (McClellan et al., 1986) has demonstrated a nonstandard B-helix for alternating $(AT)_n$ sequences (i.e., alternating conformation). In such a model, a poor 5′-TpA stacking would lead to a better accessibility of the C5–C6 double bond of thymine. $(AT)_n$ sequences are targets for micrococcal nuclease S1 and mung bean nucleases and are subject to easy torsional deformation. This deformability would explain how 3-CPs with its bulky nonpolar carbethoxy group could be accommodated in such sequences. This model would also account for the preference for 5′-TpA sites in the covalent binding of furocoumarins.

More generally, $(AT)_n$ sequences are sites of great instability. They are often located upstream to expressed genes [reviewed in McClellan et al. (1986)]. They are present in promoters (TATA box) but are also involved in recombination, integration sites, and origin of DNA replication, all processes in which unstacking and strand separation are required. Interestingly, for transcription, the TAT −10 site of the *tyr*T promoter of *E. coli* has been shown to be the nucleation center for the opening of the double helix controlled by superhelicity, the RNA polymerase being involved in the unwinding step (Drew et al., 1985). The $(AT)_n$ sequences are of biological significance. If these sequences are hot spots for the in vivo photoreaction of psoralens, the effects of psoralen lesions should be considered in much specific terms. Because of the polymorphic and dynamic properties of these sequences, the reparability of psoralen lesions in such regions of the genome should be reconsidered.

The rules for adduct formation that we have defined allow us to predict the localization of photoadducts along a gene, once the affinity constant of the compound is estimated. This is particularly interesting in view of establishing a correlation between adducts and mutations. Indeed, mutations photoinduced by 8-MOP and angelicin in the *lac*I gene of *E. coli* have been mapped (Yatagai & Glickman, 1986; Yatagai et al., 1987; Miller & Eisenstadt, 1987). Some of the hot spots for mutation occur at sequences that we define as strong sites (E. Sage, unpublished observation). Nevertheless, it now clearly appears that the frequency of mutations at a particular site is not always directly related to the frequency of adduct formation (Brash et al., 1987; Fuchs, 1984). These authors highlight the importance of the sequence and the local DNA structure around a damage on the induction of mutation. The strong sites of psoralen photoaddition are a favorable case for studying such involvement in the processing of a lesion.

Attempt To Correlate Structural and Biological Data. We try to make use of the small differences that appear in the photoreaction of various psoralen derivatives with DNA, to give some clues on the reparability and mutagenicity of their respective lesions induced in DNA in vivo. We will mainly discuss the genotoxicity of 3-CPs and pyridopsoralens (MePyPs) in comparison with 8-MOP. These two monofunctional furocoumarins, while they have the same type of furan-side monoadducts in a cis–syn configuration (Moysan, 1987), show the most different photoaddition spectra (see Figures 2 and 3) and exhibit a rather large difference in their genetic effects. Obviously the genotoxicity of a compound depends in part of the amount of its DNA binding, so we try to establish a comparison at equal numbers of DNA adducts.

Photobinding, repair, and mutagenesis have been studied for 8-MOP, 3-CPs, and MePyPs in yeast (Averbeck, 1985; Magaña-Schwencke & Moustacchi, 1985) and in human cells (Papadopoulo et al., 1986; Nocentini, 1986). Quantitatively, the biological effects of these compounds are not identical according to the system used; nevertheless, they can be summarized as follows. 3-CPs has low cytotoxicity and is poorly mutagenic, and 3-CPs photoadducts are rapidly removed with no accumulation of single-strand breaks. MePyPs forms more adducts than 8-MOP. It seems less cytotoxic and mutagenic than 8-MOP in yeast but more toxic and mutagenic in human

cells. In yeast, the rate of MePyPs and 8-MOP photoadducts removal is comparable, but single-strand breaks accumulate after MePyPs treatment, whereas breaks are readily resealed after 8-MOP treatment. Accumulation of DNA breaks in MePyPs-treated human cells is also observed, as well as persistence of photoadducts. These results imply that a DNA repair process is impaired on MePyPs photoadducts. Furthermore, it seems that MePyPs photoadditions induce less base substitutions than 8-MOP monoadducts in yeast but are as recombinogenic as 8-MOP monoadducts (Averbeck, personal communication). We recall that MePyPs has a high antiproliferative activity, is efficient in the treatment of psoriasis, and is less carcinogenic than 8-MOP [Papadopoulo et al. (1986) and references cited therein].

Structural data correlate with a good repair of 3-CPs photoadducts. These adducts produce S1 nuclease sensitive sites in DNA. We predict that they are located in AT-rich regions or better in $(AT)_n$ repeated sites. Such sequences are unstable and are good targets for endonucleases (Sobell et al., 1982; McClellan et al., 1986). Furthermore, 3-CPs adducts will probably be clustered in such regions. These observations suggest that 3-CPs photoadducts constitute good recognition signals for DNA repair endonucleases. This could explain the low toxicity and mutagenicity of 3-CPs. In agreement with that assumption, Fukuhara et al. (1978) have observed a dramatically amplified deletion of the mitochondrial genes *oxi3*, *oli2*, *pho1* of *Saccharomyces cerevisiae*, leading to *rho*$^-$ mutants, after treatment with 3-CPs + UVA. These regions consist of long AT-rich sequences (A+T content >80%) with many $(AT)_n$ stretches (Macino & Tzagoloff, 1980). We believe that an extensive excision of stretches of adducted DNA would lead to deletion of these genes.

MePyPs-adducted DNA is not sensitive to S1 nuclease, and adducts are buried in the double helix. Because of its high affinity for DNA, MePyPs will not react in preferential regions of the genome, i.e., not exclusively in $(AT)_n$ sequence which are probably easily repaired. We expect adducts distributed all along the genomic DNA. This would not be favorable for repair, since a processive mode of action of repair enzymes is suggested (Bohr et al., 1986). Moreover, we observed that the exonuclease associated with the T_4 DNA polymerase does not function easily on pyridopsoralen photoadducts. Taken together, these results indicate that MePyPs photoadducts are not good candidates for excision repair. Adducts located in random sequences may persist and give rise to mutations at replication.

Our results suggest that the lower ratio of cross-links to monoadducts formed in vivo in the highly repetitive α DNA than in the bulk DNA [Smith (1987) and references cited therein] reflects the low level of cross-linkable 5′-TpA sites, compared to the richness in runs of thymine residues in the α DNA (Waye & Willard, 1986).

A recent review focuses on mechanisms of repair of furocoumarin DNA adducts in bacteria, yeast, and mammalian cells (Smith, 1987). The same types of repair systems operate in the different organisms. Furocoumarin adducts, at least monoadditions, are processed like other bulky lesions. Endonucleases from the excision repair system have first to recognize the damage in DNA. As we mentioned above, these enzymes will work more or less efficiently, depending on the local perturbation created in the DNA double helix at the site of the damage. This hypothesis has also been suggested for another damaging agent, (acetylamino)fluorene (Sage & Leng, 1980; Sage, 1981). In mammalian cells, some regions of the genome are preferentially repaired in comparison to others, probably in relation with the chromatin state (Bohr et al., 1986). The extent of psoralen adduct removal and repair synthesis in the nontranscribed α DNA is lower than in the bulk DNA [reviewed in Smith (1987)]. Differential repair of HMT photoadducts in different genes, in transcribed versus nontranscribed genes, as for UV lesions, is discussed in this review. We would like to suggest a difference in the extent of repair at the sequence level. We speculate that some sequences, i.e., alternating $(AT)_n$ tracks, are more subject to excision repair, because of the local conformation of the DNA.

Excision repair is considered as error free. The unremoved lesions can then either be repaired by an error-free recombinational process or give rise to mutations by an error-prone synthesis over the noninstructional site by a polymerase. Again, this last process would depend on the sequence flanking the damage as proposed by Brash et al. (1987). Ultimately, the knowledge of the sequence specificity in psoralen photobinding will be an important step in understanding better the biological effects of this class of compound.

ACKNOWLEDGMENTS

We thank Drs. D. Averbeck, J. Blais, F. Gaboriau, A. Moysan, and P. Vigny for helpful discussion, Dr. B. Glickman for critical reading of the manuscript, and D. Chardonnieras for excellent assistance.

Registry No. 8-MOP, 298-81-7; 5-MOP, 484-20-8; HMT, 62442-59-5; 3-CPs, 20073-24-9; PyPs, 85878-62-2; MePyPs, 85878-63-3; BCH 394, 85878-80-4; angelicin, 523-50-2; EPI DNA fragment, 113351-73-8; HN3 DNA fragment, 113351-74-9; HN2 DNA fragment, 113351-72-7.

REFERENCES

Averbeck, D. (1985) *Mutat. Res. 151*, 217–233.

Ben-Hur, E., & Song, P. S. (1984) *Adv. Radiat. Biol. 11*, 131–171.

Blais, J., Vigny, P., Moron, J., & Bisagni, E. (1984) *Photochem. Photobiol. 39*, 145–156.

Blais, J., Averbeck, D., Moron, J., Bisagni, E., & Vigny, P. (1987) *Photochem. Photobiol. 45*, 465–472.

Bohr, V. A., Okumoto, D. S., Ho, L., & Hanawalt, P. C. (1986) *J. Biol. Chem. 261*, 16666–16672.

Brash, D. E., Seetharam, S., Kraemer, K. H., Seidman, M. M., & Bredberg, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3782–3786.

Calvin, N. M., & Hanawalt, P. C. (1987) *Photochem. Photobiol. 45*, 323–330.

Cimino, G. D., Shi, J., & Hearst, J. (1986) *Biochemistry 25*, 3013–3020.

Drew, H., Weeks, J., & Travers, A. (1985) *EMBO J. 4*, 1025–1032.

Fukuhara, H., Moustacchi, E., & Wesolowski, M. (1978) *Mol. Gen. Genet. 162*, 191–201.

Gaboriau, F., Vigny, P., Averbeck, D., & Bisagni, E. (1981) *Biochimie 63*, 899–905.

Gaboriau, F., Vigny, P., Cadet, J., Voituriez, L., & Bisagni, E. (1987) *Photochem. Photobiol. 45*, 199–207.

Gasparro, F. P. (1987) in *Psoralen DNA Photobiology* (Gasparro, F. P., Ed.) CRC Press, Boca Raton, FL (in press).

Gordon, L. K., & Haseltine, W. A. (1980) *J. Biol. Chem. 255*, 12047–12050.

Gruenert, D. C., Ashwood-Smith, M., Mitchell, R. H., & Cleaver, J. E. (1985) *Cancer Res. 45*, 5394–5398.

Hearst, J. (1981) *J. Invest. Dermatol. 77*, 39–44.

Hillebrand, G. G., & Beattie, K. L. (1985) *J. Biol. Chem. 260*, 3116–3125.

Huang, C. C., Hearst, J., & Alberts, B. (1981) *J. Biol. Chem. 256*, 4087-4094.
Isaacs, S., Wiesehahn, G., & Hallick, L. (1984) *Natl. Cancer Inst. Monogr. 66*, 21-31.
Kanne, D., Straub, K., Rapoport, H., & Hearst, J. (1982) *Biochemistry 21*, 861-871.
Koffel-Schwartz, N., Verdier, J. M., Bichara, M., Freund, A. M., Daune, M., & Fuchs, R. P. P. (1984) *J. Mol. Biol. 177*, 33-51.
Macino, G., & Tzagoloff, A. (1980) *Cell (Cambridge, Mass.) 20*, 507-517.
Magaña-Schwencke, N., & Moustacchi, E. (1985) *Photochem. Photobiol. 42*, 43-49.
Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499-560.
McClellan, J. A., Palecek, E., & Lilley, D. (1986) *Nucleic Acids Res. 14*, 9291-9309.
Miller, S., & Eisenstadt, E. (1987) *J. Bacteriol. 169*, 2724-2729.
Moysan, A. (1987) Ph.D. Thesis, Université Pierre et Marie Curie, Paris, France.
Nocentini, S. (1986) *Mutat. Res. 161*, 181-192.
Papadopoulo, D., Averbeck, D., & Moustacchi, E. (1986) *Photochem. Photobiol. 44*, 31-39.
Pearlman, D., Holbrook, S., Pirkle, D., & Kim, S. H. (1985) *Science (Washington, D.C.) 227*, 1304-1308.
Piette, J., Decuyper-Debergh, D., & Gamper, H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7355-7359.
Sage, E. (1981) Ph.D. Thesis, Université d'Orléans, Orléans, France.
Sage, E., & Leng, M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4597-4601.
Sage, E., & Moustacchi, E. (1987) *Biochemistry 26*, 3307-3314.
Shi, Y., & Hearst, J. (1986) *Biochemistry 25*, 5895-5902.
Sinden, R. R., & Hagerman, P. J. (1984) *Biochemistry 23*, 6299-6303.
Smith, C. A. (1987) in *Psoralen DNA Photobinding* (Gasparro, F. P., Ed.) CRC Press, Boca Raton, FL (in press).
Sobell, H. M., Sakore, T. D., Jain, S. C., Banerjee, K. K., Bhandary, K. K., Reddy, B. S., & Lozansky, E. D. (1982) *Cold Spring Harbor Symp. Quant. Biol. 47*, 293-314.
Straub, K., Kanne, D., Hearst, J., & Rapoport, H. (1981) *J. Am. Chem. Soc. 103*, 2347-2355.
Vigny, P., Blais, J., Ibanez, V., & Geacintov, N. (1987) *Photochem. Photobiol. 45*, 601-607.
Waye, J. S., & Willard, H. F. (1986) *Mol. Cell. Biol. 6*, 3156-3165.
Wiesehahn, G., & Hearst, J. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2703-2707.
Yatagai, F., & Glickman, B. W. (1986) *Mutat. Res. 163*, 209-224.
Yatagai, F., Horsfall, M. J., & Glickman, B. W. (1987) *J. Mol. Biol. 194*, 601-607.

Interaction of Synthetic Analogues of Distamycin and Netropsin with Nucleic Acids. Does Curvature of Ligand Play a Role in Distamycin–DNA Interactions?†

K. Ekambareswara Rao, Dipak Dasgupta,‡ and V. Sasisekharan*
Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India
Received July 27, 1987; Revised Manuscript Received December 11, 1987

ABSTRACT: Distamycin and netropsin, a class of minor groove binding nonintercalating agents, are characterized by their B-DNA and A-T base-specific interactions. To understand the conformational and chemical basis of the above specificities, the DNA-binding characteristics of a novel synthetic analogue of distamycin have been studied. The analogue, mPD derivative, has the requisite charged end groups and a number of potential hydrogen-bonding loci equal to those of distamycin. The difference in the backbone curvatures of the ligands, distamycin, the mPD derivative, and NSC 101327 (another structurally analogous compound), is a major difference between these ligands. UV and CD spectroscopic studies reported here show the following salient features: The mPD derivative recognizes only B-DNA, to which it binds via the minor groove. On the other hand, unlike distamycin, it binds with comparable affinities to A-T and G-C base pairs in a natural DNA. These DNA-binding properties are compared with those reported earlier for distamycin and NSC 101327 [Zimmer, Ch., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31-112]. The backbone structures of these three ligands were compared to show the progressive decrease in curvatures in the order distamycin, mPD derivative, and NSC 101327. The plausible significance of the backbone curvature vis-à-vis the characteristic B-DNA and AT-specific binding of distamycin is discussed. To our knowledge, this is the first attempt (with a model synthetic analogue) to probe the possible influence of backbone curvature upon the specificity of interactions of the distamycin class of groove-binding ligands with DNA.

The binding of the class of minor groove binding nonintercalating antibiotics, like distamycin (Dst) and netropsin (Nt), to synthetic and natural DNAs has evoked much interest in view of its significance in understanding the elements of specific recognition between nonintercalating ligands and DNA (Zimmer & Wahnert, 1986). X-ray crystallographic studies and the physicochemical investigations in solutions have indicated that the A-T base-specific and B-DNA selective nature of binding of Dst (and Nt) might be ascribed to electrostatic, van der Waals, and hydrogen-bonding interactions (Kopka et al., 1985; Zimmer & Wahnert, 1986). The theoretical cal-

† This work was supported by Grant ICMR/MBU/VS/24 from the Indian Council of Medical Research. K.E.R. is a recipient of a research associateship from the Council of Scientific and Industrial Research, India.
* Author to whom correspondence should be addressed.
‡ Present address: Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, Bidhannagar, Calcutta 700 064, India.

0006-2960/88/0427-3018$01.50/0