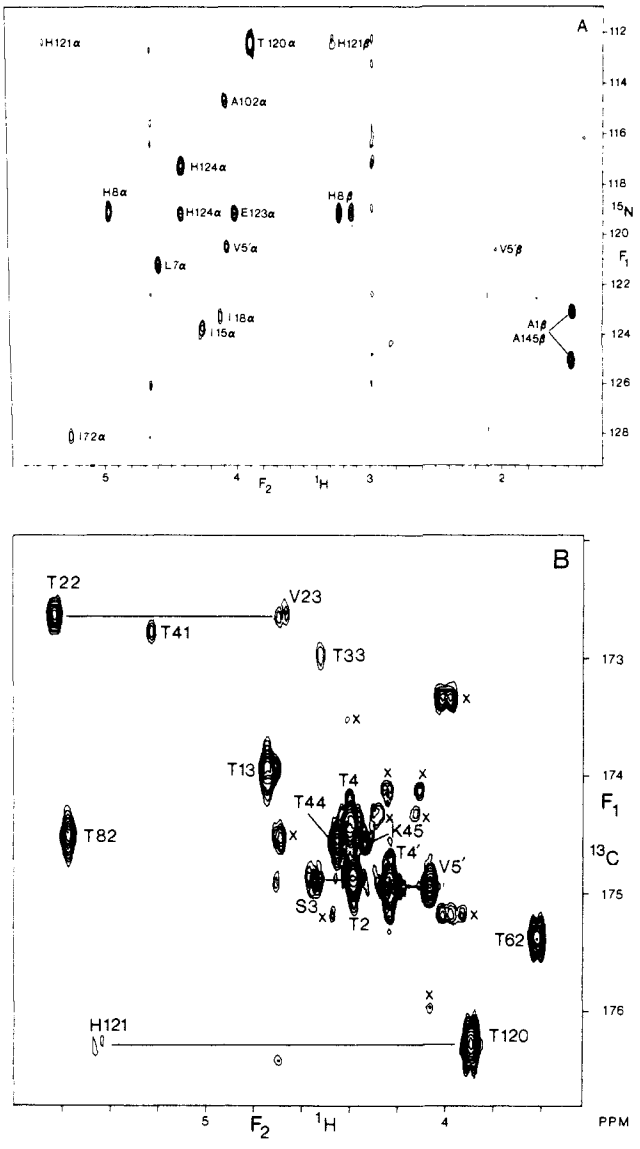


Figure 1. HMBC spectra of a 1.5 mM staphyloccocal nuclease/pdTp/Ca^{2+} complex in D_2O, recorded at 600 MHz on a Bruker AM-600 spectrometer. (A) ^1H–^{15}N correlation of a sample labeled with ^{15}N for all Leu, Ile, and His residues. Low levels of cross labeling are also found for Val (5%) and Ala (1.5%). (B) ^1H–$^{13}C_1$ correlation of a sample labeled with ^{13}C in the carbonyl position of all Thr residues. Low levels of ^{13}C cross labeling to other amino acids are also observed and give rise to correlations marked "x". Residues labeled with primes correspond to the heptapeptide attached to the N-terminus. Both spectra are presented in the mixed mode, absorption in F_1, and absolute value in F_2 (^1H). Acquisition times for both spectra were 110 (t_2) and 70 ms (t_1). Total measuring times were (A) 14 h and (B) 6.5 h. The delay Δ was set to 40 ms for both experiments; unshifted sine bell filtering was used in t_2 and a 60° shifted sine bell in t_1. A difference of $\approx$0.7 ppm in ^{15}N chemical shifts relative to values reported earlier[17] is caused by the ^2H isotope effect.

(Thr-4'/Val-5', Thr-2/Ser-3, Thr-22/Val-23, Thr-44/Lys-45, Thr-120/His-121). For α-helical domains, the $C^\alpha H(i)$–$^{13}C_1$ (i − 1) J coupling (<2 Hz) is too small to yield an observable correlation. For β-strands ($J \approx 3.5$ Hz) weak correlations are expected. For ^{13}C it also is possible to record the correlation spectrum with the ^1H detected heteronuclear correlation scheme, originally proposed by Maudsley and Ernst.[14,15] Because this type of spectrum can be obtained in the pure 2D absorption mode, this scheme offers higher resolution albeit at a cost in sensitivity (Supplementary Material).

As demonstrated here, combined use of ^1H-detected heteronuclear long range correlation techniques and isotopic labeling of selected amino acids can yield significant additional information for the study of medium-size proteins. For smaller proteins, complete ^{15}N labeling can be helpful for resolving ambiguities in the sequential assignment of α-helical domains.[16] Intraresidue couplings between C^β protons and C_1 or $N\alpha$, combined with qualitative knowledge of the couplings between C^α and C^β protons, can provide information about stereospecific assignment and about the χ angle.[7,8] Two-bond ^1H–^{15}N couplings (combined with one-bond N–H correlation) yield intraresidue $C^\alpha H$–NH connectivities. In the case of small J_{HH} couplings, the intraresidue connectivities cannot always be determined from COSY or HOHAHA spectra because the relatively large line widths of the NH resonances in medium-size proteins makes magnetization transfer inefficient.

Note Added in Proof. In a recent paper, Westler et al. (*J. Am. Chem. Soc.* **1988**, *110*, 4093) reported the observation of sequential $^{13}C_1$–$C^\alpha H$ HMBC connectivity in the streptomyces subtilisin inhibitor.

Acknowledgment. We thank Prof. E. Lattman and Dr. P. Loll for providing us with the X-ray coordinates of S. Nase prior to publication. This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health.

Supplementary Material Available: Figure showing ^1H–$^{13}C_1$ correlation recorded with the Maudsley–Ernst sequence (2 pages). Ordering information is given on any current masthead page.

(14) Maudsley, A. A.; Ernst, R. R. *Chem. Phys. Lett.* **1977**, *50*, 368.
(15) Sklenar, V.; Miyashiro, H.; Zon, G.; Miles, H. T.; Bax, A. *FEBS Lett.* **1986**, *208*, 94.
(16) Clore, G. M.; Bax, A.; Wingfield, P.; Gronenborn, A. M. *FEBS Lett.*, in press.
(17) Torchia, D. A.; Sparks, S. W.; Bax, A. *Biochemistry* **1988**, *27*, 5135.

Double-Strand Cleavage of Genomic DNA at a Single Site by Triple-Helix Formation

Scott A. Strobel,† Heinz E. Moser,‡ and Peter B. Dervan*

Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology Pasadena, California 91125

Received July 22, 1988

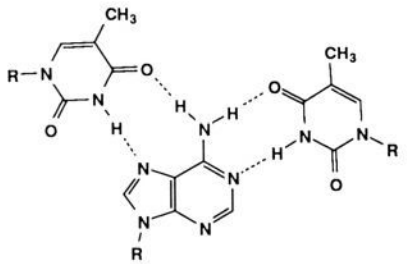
The sequence-specific cleavage of double helical DNA by restriction endonucleases is essential for many techniques in molecular biology, including gene isolation, DNA sequencing, and recombinant DNA manipulations.[1,2] With the advent of pulsed-field gel electrophoresis, the separation of large segments of DNA is now possible.[3,4] However, the recognition sequences of naturally occurring restriction enzymes are in the range of 4–8 base pairs, and hence their sequence specificites may be inadequate for isolating genes from large chromosomes (10^8 base pairs in size) or mapping genomic DNA.[5]

Pyrimidine oligonucleotides bind duplex DNA sequence specifically at homopurine sites to form a triple helix structure.[6,7]

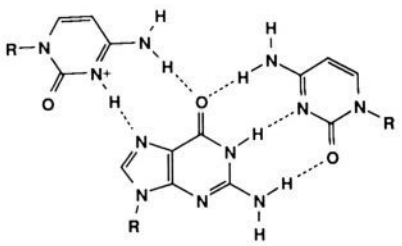
† Howard Hughes Medical Institute Doctoral Fellow.
‡ Swiss National Foundation Postdoctoral Fellow.

(1) Smith, H. O. *Science* (*Washington, D. C.*) **1979**, *205*, 455.
(2) Modrich, P. *Crit. Rev. Biochem.* **1982**, *13*, 287.
(3) Schwartz, D.; Cantor, C. R. *Cell* **1984**, *37*, 67.
(4) Carle, G. F.; Frank, M.; Olson, M. V. *Science* (*Washington, D. C.*) **1986**, *232*, 65.
(5) (a) Dervan, P. B. *Science* (*Washington, D. C.*) **1986**, *232*, 464. (b) Dervan, P. B. In *Nucleic Acids and Molecular Biology* (*Vol. 2*); Eckstein, F., Lilley, D. M. J., Eds.; Springer-Verlag: Heidelberg, 1988; pp 49–64.
(6) Moser, H. E.; Dervan, P. B. *Science* (*Washintgon, D. C.*) **1987**, *238*, 645.

0002-7863/88/1510-7927$01.50/0



TAT-base triplet



C^+GC-base triplet

Figure 1. Isomorphous base triplets of TAT and C+GC. The pyrimidine oligonucleotide is bound by Hoogsteen hydrogen bonds in the major groove to the purine strand in the Watson–Crick duplex.

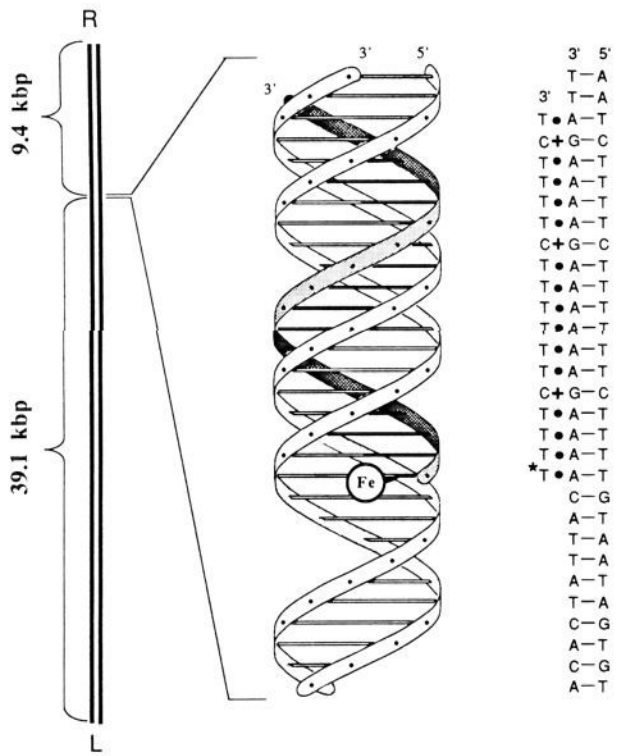


Figure 2. (Left) Double strand cleavage of λ DNA at the 18 base pair purine site affords two DNA fragments, 39.1 and 9.4 kbp in size. (Center) Simplified model of triple helix complex between the Watson–Crick homopurine–homopyridine site on λ DNA and the Hoogsteen bound oligonucleotide–EDTA·Fe.

The pyrimidine oligomer is oriented in the major groove of DNA parallel to the Watson–Crick purine strand by Hoogsteen base pairing (Figure 1).[6-8] We have shown that oligonucleotides 15 bases in length, equipped with a cleaving function EDTA·Fe at the 5′ end, cause sequence specific double strand breaks at one site in plasmid DNA which is 4.0 kilobase pairs (kbp) in size.[6] Due to the length of the recognition site, in a formal sense, *this is 10^6 times more sequence specific than restriction enzymes.*[5,6] It is important to determine whether the full potential of this cleavage specificity can be realized in larger DNA. As an initial step beyond the cleavage of plasmid DNA, we report here the specificity and efficiency of double strand cleavage by an oligonucleotide–EDTA·Fe probe targeted to a naturally occurring 18 base pair sequence, 5′-$A_4GA_6GA_4GA$-3′, that occurs once within the 48 502 base pairs (48.5 kbp) of the bacteriophage λ genome.[9,10] The target sequence, located at base pair positions 39 138-56 from the left (L) end, is an essential purine cluster in λ's origin of replication (ori) (Figure 2).[11,12] Double strand cleavage at the 18 base pair target site would afford two DNA fragments, 39.1 and 9.4 kbp in size.

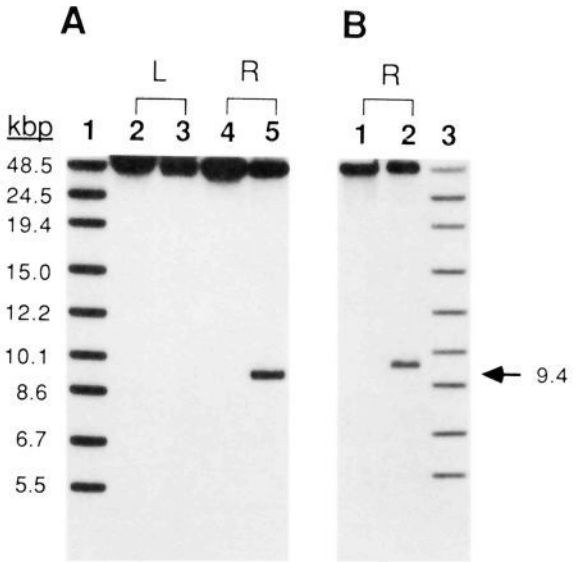


Figure 3. (A) Site-specific double strand cleavage of bacteriophage λ DNA. Autoradiogram of ^{32}P right (R) and left (L) end-labeled DNA, 48.5 kbp in length, on a 0.4% nondenaturing agarose gel. Lane 1, DNA size markers obtained by digestion of L end-labeled λ DNA with BamH I, Sma I, Apa I, SnaB I, and Xba I; digestion of right end-labeled DNA with BamH I, Sma I, and Xho I: 48.5 (undigested DNA): 24.5, 19.4, 15.0, 12.2, 10.1, 8.6, 6.7, and 5.5 kbp. Lane 2, L end-labeled intact λ DNA control. Lane 3, L end-labeled λ DNA with 0.8 μM oligonucleotide–EDTA·Fe. Lane 4, R end-labeled intact λ DNA control. Lane 5, R end-labeled λ DNA with 0.8 μM oligonucleotide–EDTA·Fe. (B) Site specific double strand cleavage of bacteriophage λ DNA in a 1% LMP agarose matrix. Autoradiogram of ^{32}P R end-labeled DNA on a 0.4% nondenaturing agarose gel. Lane 1, R end-labeled intact λ DNA control. Lane 2, R end-labeled λ DNA with 0.8 μM oligonucleotide–EDTA·Fe. Lane 3, DNA size markers as described above. Arrow on the right indicates 9.4 kbp cleavage fragment.

(7) (a) Doan, T. L.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucl. Acids Res.* **1987**, *15*, 7749. (b) Praseuth, D.; Perrouault, L.; Doan, T. L.; Chassignol, M.; Thuong, N.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349.

(8) For background on triple helical polynucleotides see: (a) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023. (b) Michelson, A. M.; Massoulié, J.; Guschlbauer, W. *Prog. Nucleic Acids Res. Mol. Biol.* **1967**, *6*, 83. (c) Felsenfeld, G.; Miles, H. T. *Annu. Rev. Biochem.* **1967**, *36*, 407. (d) Lipsett, M. N. *Biochem. Biophys. Res. Commun.* **1963**, *11*, 224. (e) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256. (f) Howard, F. B.; Frazier, J.; Lipsett, M. N.; Miles, H. T.; *Biochem. Biophys. Res. Commun.* **1964**, *17*, 93. (g) Miller, J. H.; Sobell, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1201. (h) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63. (i) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073.

(9) Sanger, F.; Coulson, A. R.; Hong, G. F.; Hill, D. F.; Petersen, G. B. *J. Mol. Biol.* **1982**, *162*, 729–773.

(10) Strain cI857ind1Sam7 from New England Biolabs.

(11) Thompson, K. D.; Moore, D. D.; Kruger, K. E.; Furth, M. E.; Blattner, F. R. *Science (Washington, D. C.)* **1977**, *198*, 1051–1056.

(12) Furth, M. E., Wickner, S. H. In *Lambda II*; Hendrix, R. W.; Roberts, J. W.; Stahl, F. W.; Weisberg, R. A., Eds.; Cold Spring Harbor, NY 1983; pp 145–173.

Linearized bacteriophage λ DNA was labeled with ^{32}P individually at the left (L) and right (R) ends (3′ fill-in reaction) with AMV reverse transcriptase. An 18 base homopyrimidine oligonucleotide, 5′-T*$T_3CT_6CT_4$CT-3′, with thymidine–EDTA (T*)[13] at the 5′ end, was synthesized by automated methods. The oligonucleotide–EDTA (0.8 μM) was mixed with $Fe(NH_4)_2(SO_4)_2$ (1.0 μM) and spermine (1.0 mM). The oligonucleotide–EDTA·Fe(II)/spermine was added to a solution of L (or R) ^{32}P end-labeled λ DNA (approximately 4 μM in base pairs) in 100 mM NaCl and 25 mM tris–acetate at pH 7.0.[14] After 0.5 h incubation at 24 °C, 4 mM dithiothreitol was added to initiate strand cleavage. The reaction was allowed to continue for 6 h (24 °C). The double strand cleavage products were separated on a 0.4% vertical agarose gel which resolves DNA segments up to 25 kbp in size (Figure 3). Therefore, DNA uniquely labeled at the R and L ends allows cleavage site analysis of the entire 48.5 kbp of the λ genome.

No cleavage was observed for the first 25 kbp of the L end-labeled DNA (Figure 3A, lane 3). However, a single major cleavage site 9.4 kbp from the right (R) end of the 48.5 kbp DNA was visualized on the autoradiogram (Figure 3A, lane 5).[15] This was quantitated by scintillation counting of the individual bands. The oligonucleotide–EDTA·Fe probe afforded double strand cleavage of λ DNA at the target sequence, 5′-$A_4GA_6GA_4$GA-3′, with an efficiency of 25% (Figure 3).[16] Within the limits of our detection, overexposure of the autoradiogram revealed no secondary cleavage sites. We estimate that all secondary sites were cleaved at least 30-fold less efficiently than the primary sequence.[17]

In order to prevent random shearing of the DNA caused by pipetting or vortexing,[3,4] large DNA is routinely manipulated by embedding it in low melting point (LMP) agarose and diffusing reagents into the matrix. To test whether triple helix mediated site-specific cleavage can occur with oligonucleotide–EDTA·Fe *within an agarose matrix*, end-labeled λ DNA was embedded in 1% LMP agarose gel, and the cleavage reactions were performed as described above (24 °C, pH 7.0, 6 h, initial incubation time was extended to 1 h to allow for complete diffusion into the matrix) followed by electrophoresis. Primary site-specific double strand cleavage occurred in the agarose matrix with an efficiency of 25% (Figure 3B, lane 2).

In conclusion, this work has implications for human genetics. The way is now clear for the development of a triple helix strategy to isolate large segments of genomic DNA for mapping and sequencing.

Acknowledgment. This work was supported in part by the National Institutes of Health (GM 35724), the Caltech Consortium in Chemistry and Chemical Engineering (Founding members: E. I. duPont de Nemours and Company, Inc., Eastman Kodak Company, Minnesota Mining and Manufacturing Company, and Shell Oil Company Foundation), and a National Research Service Award (T32GM07616) from the NIH-GM to S.A.S.

(13) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968.

(14) Concentrations given are final dilutions.

(15) The complementary fragment, 39.1 kbp from the L terminus, was not resolved from the intact λ DNA but can be visualized in underexposed autoradiograms as a shoulder with intensity equal to the 9.4 kbp fragment.

(16) 25% efficiency means that one-quarter of all the DNA underwent double strand cleavage at one site.

(17) Although the above conditions are optimal for single-site cleavage, reaction parameters can be modified for the purposes of searching large DNA for sequences of *partial* homology with the primary target site. For example, identical reaction conditions performed at the reduced temperature of 0 °C for 24 h resulted in three minor sites of cleavage (25-fold less efficient) which map to 22.4, 37.9, and 47.7 kbp from the L terminus. Sequences partially homologous with the primary site are found at each of these positions. One site, 5′-AA*CA*AAAAAAG-3′, contains two mismatches in the first 12 nucleotides from the 5′ end of the oligonucleotide–EDTA·Fe probe. Two other sites, 5′-AAAAGAAAAA*T*GAA-3′ and 5′-AAA*T*GAA*T*AAAGAA-3′, contain one and two mismatches in the first 14 base pairs, respectively. All other sequences in λ DNA with a similar degree of sequence homology (two mismatches in 12) were not cleaved within the limits of our detection and have in common the mismatch at or adjacent to the T* which contains the EDTA·Fe(II) cleaving moiety.

Preparation of the 8,9-Epoxide of the Mycotoxin Aflatoxin B_1: The Ultimate Carcinogenic Species

Steven W. Baertschi, Kevin D. Raney, Michael P. Stone, and Thomas M. Harris*

Department of Chemistry and Center in Molecular Toxicology, Vanderbilt University Nashville, Tennessee 37235

Received August 15, 1988

Aflatoxin B_1 (AFB_1, **1**) is formed by ubiquitous fungi and is one of the most potent environmental mutagens and carcinogens, causing worldwide concern about contamination of agricultural products.[1] It requires metabolic activation to a highly labile, as yet unisolated, electrophilic species which forms covalent adducts with DNA.[2] Metabolic activation can be mimicked by in situ chemical activation of AFB_1 by peroxy acids,[3] by photooxidation,[4] or by solvolysis of 8-acyloxy-9-hydroxy derivatives of AFB_1.[5] Indirect evidence derived from structures of adducts with nucleic acids and other nucleophiles points to the fugitive intermediate being the *exo*-8,9-epoxide **2**.[3-6] Numerous attempts to prepare this epoxide have failed,[7-10] which has raised questions as to its existence and role in carcinogenesis.[7] The epoxide could be expected to be highly labile by virtue of the fact that it is fused to a tetrahydrofuran ring and would be prone to acid-catalyzed solvolysis at C8.[11,12] We wish to report the preparation of **2** by a chemical oxidation of AFB_1 and the finding that the compound is stable enough to permit storage and manipulation.[13] Epoxide **2** has the properties associated with in vivo activation of AFB_1, supporting the hypothesis that it is the ultimate carcinogen. The availability of the epoxide will permit direct investigation of its chemistry and metabolism.[14]

Synthesis of epoxide **2** involved oxidation of AFB_1 with dimethyldioxirane (Scheme I), which had been prepared from potassium peroxysulfate as a distilled 0.05 M solution in acetone.[15] Approximately 1.5 equiv of the dioxirane was added to AFB_1 dissolved in either acetone or methylene chloride.[16] The reaction

(1) More than 5000 articles on aflatoxins have been abstracted by *Chemical Abstracts*. For a review, see: Busby, W. F., Jr.; Wogan, G. N. In *Chemical Carcinogens*, 2nd ed.; Searle, C., Ed.; American Chemical Society Series, 1984, Vol. 182, pp 945–1136.

(2) Garner, R. C.; Miller, E. C.; Miller, J. A.; Garner, J. V.; Hansen, R. S. *Biochem. Biophys. Res. Commun.* **1971**, *45*, 774–780.

(3) Martin, C. N.; Garner, R. C. *Nature (London)* **1977**, *267*, 863–865.

(4) Büchi, G.; Fowler, K. W.; Nadzan, A. M. *J. Am. Chem. Soc.* **1982**, *104*, 544–547.

(5) Coles, B. F.; Welch, A. M.; Hertzog, P. J.; Lindsay Smith, J. R.; Garner, R. C. *Carcinogenesis* **1980**, *1*, 79–89.

(6) (a) Essigmann, J. M.; Croy, R. G.; Nadzan, A. M.; Busby, W. F., Jr.; Reinhold, V. N.; Büchi, G.; Wogan, G. N. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 1870–1874. Lin, J. K.; Miller, J. A.; Miller, E. C. *Cancer Res.* **1977**, *37*, 4430–4438. (b) Garner, R. C.; Miller, E. C.; Miller, J. A. *Cancer Res.* **1972**, *32*, 2058–2066.

(7) Gorst-Allman, C. P.; Steyn, P. S.; Wessels, P. L. *J. Chem. Soc., Perkin Trans. 1* **1977**, 1360–1364.

(8) Coles, B. F.; Lindsay Smith, J. R.; Garner, R. C. *J. Chem. Soc., Perkin Trans. 1* **1979**, 2664–2671.

(9) Garner, R. C.; Martin, C. N.; Lindsay Smith, J. R.; Coles, B. F.; Tolson, M. R. *Chem.-Biol. Interact.* **1979**, *26*, 57–73.

(10) Miller, J. A. *Cancer Res.* **1970**, *30*, 559–576.

(11) Stevens, C. L.; Tazuna, J. *J. Am. Chem. Soc.* **1954**, *76*, 715–717.

(12) Havel, J. J.; Chan, K. H. *J. Org. Chem.* **1976**, *41*, 513–516.

(13) **Caution**: The toxicological properties of the epoxide **2** are presently being investigated. On the basis of structure and the chemical properties reported herein it should be assumed that **2** is highly toxic and carcinogenic. Manipulations should be carried out in a well-ventilated hood using disposable latex gloves. Particular care should be taken when handling crystalline material due to electrostatic properties. Aflatoxins can be destroyed by treatment with commercial NaOCl solutions.

(14) The stability of **2** is sufficient that vendors of fine chemicals should be able to market it to investigators not wanting to prepare their own material.

(15) Murray, R. W.; Jeyaraman, R. *J. Org. Chem.* **1985**, *50*, 2847–2853. Adam, W.; Chan, Y.-T.; Cremer, D.; Gauss, J.; Scheutzow, D.; Schindler, M. *J. Org. Chem.* **1987**, *52*, 2800–2803.

(16) Although water has been shown to accelerate epoxidations with dimethyldioxirane,[17] the dimethyldioxirane solution and the reaction solvent were dried (molecular sieves or K_2CO_3) to minimize hydrolysis of the epoxide.

0002-7863/88/1510-7929$01.50/0