JOURNAL OF
MEDICINAL
CHEMISTRY

Volume 32, Number 9 September 1989

Perspective

DNA and Associated Targets for Drug Design†

Laurence H. Hurley

Department of Medicinal Chemistry, Drug Dynamics Institute, College of Pharmacy, The University of Texas Austin, Austin, Texas 78712. *Received January 11, 1989*

It is over 60 years since the first nitrogen mustards were used in cancer chemotherapy. Since then a structurally diverse group of DNA-reactive drugs has been discovered largely through cytotoxic screens. In virtually every case the molecular target was only tentatively identified as DNA after antitumor activity in animal model systems was demonstrated. On the basis of what is known concerning the molecular interactions of these compounds with DNA, three major groups of clinically important DNA reactive agents have been identified.[1] The three groups are the alkylating agents exemplified by Cytoxan, Cisplatin, and mitomycin C, the DNA strand breakage compounds such as bleomycin, and the intercalating agents typified by Adriamycin and actinomycin D. A fourth group typified by chromomycin, anthramycin, and CC-1065 are drugs that modify DNA in the minor groove of DNA. However, these compounds have yet to be proven to have clinical utility. Some of the discoveries that could be considered to be seminal in our present understanding of the molecular basis for antitumor activity of DNA-reactive compounds are the DNA intercalation model for acridine proposed by Lerman,[2] the in vitro inhibition of RNA polymerase produced by complexation of actinomycin D with DNA demonstrated by Goldberg,[3] and the demonstration of interstrand DNA–DNA cross-linking by mitomycin C due to Szybalski.[4] More recent landmark discoveries are the identification by Liu[5] and Kohn[6] of the importance of topoisomerase II in mediating the antitumor activity of compounds such as adriamycin, and the contributions by various groups to understanding the molecular basis for the DNA sequence specificity of DNA-reactive drugs.[7-10]

What is clear as we approach the end of the 20th century is that structural tools such as X-ray crystallography, high-field NMR, and computational chemistry, alongside techniques from molecular biology such as DNA sequencing, DNA construction strategies, and gene cloning, have ushered in a new era in conception and design of new drugs. The availability of these tools together with the recently uncovered structural heterogeneity in DNA and restricted access of domains in eukaryotic genomes makes the human genome an attractive target for drug design. In this Perspective I will address four important issues that relate to strategies for improving the therapeutic selectivity of existing groups of agents and discovery of new groups of important drugs that either interact directly with DNA or an associated target. These are (1) identification and characterization of the biologically important lesions on DNA produced by clinically important DNA-reactive drugs, (2) molecular strategies to improve the sequence selectivity of existing groups of DNA-reactive drugs, (3) the design of mechanism-based pharmacological screens to identify new classes of drugs that react with DNA and

† This article is not meant to be a comprehensive review of the drug–DNA interaction area. Nor does this Perspective consider how therapeutic selectivity can be achieved through pharmacokinetic and pharmacodynamic parameters. The brevity of discussion is imposed by space constraints.

(1) Gale, E. F.; Cundliffe, E.; Reynolds, P. E.; Richmond, M. H.; Waring, M. J. *The Molecular Basis of Antibiotic Action*, 2nd ed.; John Wiley & Sons: New York, 1981; pp 258–401.
(2) Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18.
(3) Goldberg, I. H.; Rabinowitz, M. *Science* **1962**, *136*, 315.
(4) Iyer, V. N.; Szybalski, W. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 355.
(5) Nelson, E. M.; Tewey, K. M.; Liu, L. F. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1361.
(6) Ross, W. E.; Glaubiger, D. L.; Kohn, K. W. *Biochem. Biophys. Acta* **1978**, *519*, 23.
(7) Dervan, P. B. *Science* **1986**, *232*, 464.
(8) Kohn, K. W.; Hartley, J. A.; Mattes, W. B. *Nucleic Acids Res.* **1987**, *15*, 10531.
(9) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1876.
(10) Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315 and references therein.

0022-2623/89/1832-2027$01.50/0

associated targets, and (4) the design of new drug classes based upon the identification of pharmacologically selective receptors on, or associated with, DNA.

If we are successful in the design of new therapeutic classes that have improved selectivity for defined cellular receptors, these agents will have applications beyond treatment of cancer. For example, they may be useful in treating infectious diseases or genetic disorders that result from aberrant expression of cellular proteins.

I. Identification and Characterization of the Biologically Important Lesions Produced on DNA by Important DNA-Reactive Drugs

Identification and characterization of the types of DNA modification produced by DNA-interactive drugs is a prerequisite for understanding the biochemical and biological responses induced by these agents. Since the majority of DNA-reactive drugs produce a variety of lesions on DNA, it is an important objective to determine which lesions are biologically important. In recent years studies aimed toward these goals have used both short (6–12 base pairs), intermediate (100–300 base pairs), and circular DNA fragments obtained through oligodeoxynucleotide synthesis, restriction enzyme digestions, and plasmid constructions, respectively. While this represents progress from studies with individual bases or nucleotides, it still stops short of the desired goal of using human chromatin as the target structure. The formidable analytical problems associated with analyzing whole human genomic DNA because of the large size of the molecular target argue strongly for designing plasmid systems that can be constructed in vitro and that then use the replication and transcriptional machinery of eukaryotic cells. In systems such as SV40 DNA the nucleosomal structure is retained and the circular DNA can be isolated free of nuclear DNA for subsequent analytical manipulation.[11] Using a combination of short oligomers for structural work,[12-14] restriction enzyme fragments for studies on DNA sequence specificity,[15-17] and plasmid systems[18] for in vivo experiments, considerable information relevant to the manner in which drugs may interact and modify DNA structure and function in human genomic DNA can be obtained. Thus it is now possible to gain structural and mechanistic data that allow the investigator to gain considerable insight into which are the biologically important lesions on DNA and how these cause the observed potent biological effects. Some examples of these approaches for DNA-reactive drugs that produce DNA strand breakage, intercalate into DNA, or alkylate DNA are given below.

The oxidative mechanisms leading to DNA strand breakage by such drugs as bleomycin[19,20] and neocarzinostatin[21] are quite complex and some of the finer points for bleomycin remain controversial. However, the chemical products of such reactions are well described. Hecht and co-workers designed a dodecamer with which they characterized the reaction of bleomycin with DNA.[22] The products of bleomycin-induced chemistry under a variety of conditions were determined by comparison with authentic standards. Longer DNA fragments have been used to determine other aspects of bleomycin reaction with DNA, such as the strand specificity of DNA breakage[23] and the effect of DNA methylation on sequence specificity.[24] The sequence specificity of bleomycin cleavage in SV40 DNA has also been compared in purified and intracellular systems.[25] Using a pLTL-1 plasmid containing a herpes simplex virus thymidine kinase gene grown in mouse mammary tumor cells, bleomycin and neocarzinostatin have both been demonstrated to cleave preferentially within regions of DNA that are actively transcribed following glucocorticoid induction.[26] While these experiments do not provide definitive answers to the mechanism of bleomycin cytotoxicity they provide a background for the design of even more exacting studies. Goldberg and co-workers have demonstrated for neocarzinostatin that mutational events can be directly related to the spectrum of damage produced on DNA.[27] This is an important example of how in vitro and in vivo data can be correlated.

The DNA-intercalation model proposed by Lerman[2] and confirmed by X-ray structure analysis for drugs such as Adriamycin[28] and actinomycin D[29] has long been accepted as a biologically important event in the cytotoxicity and antitumor activity of these compounds. However, it has always been disconcerting that the structure–activity relationships for anthracyclines and other intercalators have never strongly supported this claim.[30] For some compounds, such as the anthracyclines, this has been rationalized by involving non-DNA related targets such as membranes and oxidative DNA damage mechanisms.[31] In reality, the truth may be somewhere in between. DNA per se may not be the ultimate target, but intercalation may serve as a mechanism for holding the drug "on location" until a critical event, such as a change in supercoiling catalyzed by a DNA topoisomerase, occurs. The intercalating agent that is intimately associated with the DNA molecule then interferes with this process at a critical step, resulting in protein-associated strand breaks in DNA.[6] Exactly how intercalating agents such as Adriamycin and *m*-AMSA interfere with topological processing is still unknown, but it is quite possible that topoisomerase II or the enzyme–DNA complex are the actual targets.[32] This observation may behoove medicinal chemists to examine

(11) Hirt, B. *J. Mol. Biol.* **1967**, *26*, 365.
(12) Patel, D. J.; Shapiro, L. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 423.
(13) Harris, T. M.; Stone, M. V.; Harris, C. M. *Chem. Res. Toxicol.* **1988**, *1*, 79.
(14) Neidle, S. *Topics in Nucleic Acid Structure*; Neidle, S., Ed.; Halsted Press: New York, 1981; pp 177–198.
(15) Hurley, L. H.; Boyd, F. L. *Annu. Rep. Med. Chem.* **1987**, *22*, 259.
(16) Dabrowiak, J. C. *Life Sci.* **1983**, *32*, 2915.
(17) Wilkins, R. J. *Mol. Cell. Biochem.* **1984**, *64*, 111.
(18) Basu, A. K.; Essigmann, J. M. *Chem. Res. Toxicol.* **1988**, *1*, 1 and references therein.
(19) Stubbe, J.; Kozarich, J. W. *Chem. Rev.* **1987**, *87*, 1107.
(20) Hecht, S. M. *Acc. Chem. Res.* **1986**, *19*, 383.
(21) Chin, D-H.; Zeng, C-h.; Costello, C. F.; Goldberg, I. H. *Biochemistry* **1988**, *27*, 8106.
(22) Sugiyama, M.; Kilknskie, R. F.; Hecht, S. M.; van der Marcel, G. A.; van Boom, J. M. *J. Am. Chem. Soc.* **1985**, *107*, 7765.
(23) Krishramoothy, C. R.; Vanderwall, D. E.; Kozarich, J. W.; Stubbe, J. *J. Am. Chem. Soc.* **1988**, *110*, 2008.
(24) Hertzberg, R. P.; Caranfa, M. J.; Hecht, S. M. *Biochemistry* **1985**, *24*, 5285.
(25) Grimwade, J. E.; Cason, E. B.; Beerman, T. A. *Nucleic Acids Res.* **1987**, *15*, 6315.
(26) Beckmann, R. P.; Agostino, M. J.; McHugh, M. M.; Sigmund, R. D.; Beerman, T. A. *Biochemistry* **1987**, *26*, 5409.
(27) Povirk, L. F.; Goldberg, I. H. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3182.
(28) Quigley, G. L.; Wang, A. M. J.; Ughetto, G.; van der Marel, G.; van Boom, J. H.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 7204.
(29) Takusagawo, F.; Dabrow, M.; Neidle, S.; Beerman, H. M. *Nature* **1982**, *296*, 466.
(30) DiMarco, A. *Cancer Chemother. Rep.* **1975**, *6*, 91.
(31) Schwartz, H. S. In *Molecular Aspects of Anti-cancer Drug Action*; Neidle, S., Waring, M. J., Eds.; Macmillan Press Ltd.: London, 1983; pp 93–126.
(32) Lock, R. B.; Ross, W. E. *Anticancer Drug Des.* **1987**, *2*, 151.

other ternary systems (drug–DNA–protein) in order to understand more completely the biochemical and biological effects of drug–DNA interactions.

In some respects alkylation of DNA is the most challenging area in which to make correlations between DNA damage and its biochemical and biological consequences.[33] The complexity of these systems can at first appear to be overwhelming. Many alkylating agents produce a variety of lesions on DNA, some of which are chemically unstable and thus difficult to characterize.[34] Virtually all covalently bonding drugs show some degree of DNA sequence selectivity that can be attributed to either the noncovalent (binding) interaction, the covalent (bonding) reaction, or a combination of these processes.[10] DNA repair recognition and subsequent repair of lesions can be variable, depending upon the adduct characteristics,[35] tissue type,[36] and transcriptional state of the damaged area.[37] Nevertheless progress is being made in unraveling these complexities by making "site-directed adducts" in plasmids.[18] Different types of adducts are engineered into predetermined sites in the genome and chemically and enzymatically characterized. DNA repair recognition, DNA replication, transcriptional activity, phage survival, or mutagenesis can then each be examined separately. Although this may seem to be a considerable effort, it is probably the only viable way of unraveling the complexity of an otherwise intractable system. In this regard Essigmann and Lippard have provided important data on the construction and characterization of a Cisplatin adduct on DNA.[38,39] This area has been recently reviewed by Essigmann.[18]

There is still a large gap in our knowledge about the structure of the drug–DNA complex or adduct and its relationship to the biological response, e.g., cell death, mutagenicity. Considerably headway has been made at the structural end of the problem, but only in rare cases[27] has this provided insight into rationalizing the biological effects. There is an urgent need to develop biochemical systems that will extend our knowledge beyond DNA as the target for drug action. For example, are DNA-binding proteins such as those involved in regulation of transcription, replication, or repair necessary to express the biological potency of some DNA-reactive drugs? I will return to some of these questions later in this Perspective.

II. Molecular Strategies To Improve the Sequence Selectivity of Existing Groups of DNA-Reactive Drugs

Cellular and molecular selectivity are key aspects of drug action. For cytotoxic drugs such as DNA-reactive compounds selective uptake at the cellular level would appear to be crucial unless unique intracellular targets can be identified in, for example, cancer cells. Strategies for selective cellular retention are outside the scope of this article but include such mechanisms as targeting using monoclonal antibodies to tumor antigens.[40] At the intracellular level the ultimate selectivity of a DNA-reactive drug should be at the sequence level. In mammalian cells the DNA receptor is an elusive target.[41] Besides its large size (2.9×10^9 base pairs), human nuclear DNA is largely covered with proteins[42] and the accessibility of potential drug binding regions may be dependent upon DNA processing events such as replication and transcription. The reactivity of open regions undergoing processing to DNA-binding proteins[43] and DNA-reactive drugs[44] may also be modulated by superhelical stress. These are certainly complicating factors in the design of more selective DNA-reactive drugs. However, this increased complexity could be an advantage since such factors may allow greater sequence selectivity to be achieved. In this section I will provide examples of how processes, such as transcription activity and DNA repair may lead to greater selectivity for drug action. The main focus will, however, be on the design of DNA-reactive drugs with increased sequence selectivity.

While nuclear DNA is generally cited as the target for all three classes of DNA-reactive drugs (intercalating agents, DNA-degradative drugs, or alkylating agents), competing targets such as mitochondrial DNA and cellular RNA may also be significant targets. For drugs that are A-T selective or whose binding to DNA is restricted by nucleosomal structure, mitochondrial DNA (which is A-T rich and relatively "naked") may be the preferred target. Mitochondrial DNA may also be the preferred target for drugs which are selectively traken up through the mitochondria membrane rather than nuclear membrane. Also the absence of DNA repair in mitochondria may predispose mitochondrial DNA to selective toxicity of DNA reactive compounds.[45,46]

Actively transcribed regions of DNA are the preferred target for a number of DNA-reactive molecules. The potent carcinogenic fungal toxin aflatoxin B_1 selectively reacts with such regions[37] and the DNA strand cleavage agents, bleomycin and neocarzinostatin, selectively cut within transcriptionally active regions.[26] Drugs or carcinogens that react selectively with single-stranded DNA would also be expected to have increased reactivity with actively transcribed or replicated regions of DNA.[47] Since actively transcribed regions of DNA are preferred targets for some DNA-reactive drugs, then agents such as steroids that can selectively induce transcription[48] may increase the selectivity of DNA-reactive drugs for these regions. This may provide a rationale for the combined use of steroids and cytotoxic agents such as DNA-reactive drugs in hormone-responsive cancers. Where enzymes such as topoisomerase II are involved in modulating the superhelical density of transcriptionally active regions and are also a requirement for drug action, then this may significantly increase the selectivity of such agents.[49] Finally there is evidence from experiments with certain DNA-reactive drugs that nucleosomal structure[50] and superhelical den-

(33) See for example: Singer, B.; Grunberger, D. In *Molecular Biology of Mutagens and Carcinogens*; Plenum Press: New York, 1983 and references therein.
(34) Reference 33, pp 45–94 and 143–190.
(35) Reference 33, pp 199–216.
(36) Reference 33, pp 221–245.
(37) Irvine, R.; Wogen, G. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *81*, 664.
(38) Pinto, A. L.; Naser, L. J.; Essigmann, J. M.; Lippard, S. J. *J. Am. Chem. Soc.* **1986**, *108*, 7405.
(39) Naser, L. J.; Pinto, A. L.; Lippard, S. J.; Essigmann, J. M. *Biochemistry*, in press.
(40) See for example: In *Immunoconjugates: Antibody Conjugates in Radioimaging and Therapy of Cancer*; Vogel, C.-W., Ed.; Oxford University Press: New York, 1987.
(41) Hurley, L. H.; Boyd, F. L. *Trends Pharmacol. Sci.* **1988**, *9*, 402.
(42) Felsenfeld, G.; McGhee, J. D. *Cell* **1986**, *44*, 375.
(43) Gellert, M.; Monzel, R. In *New Frontiers in the Study of Gene Functions*; Poste, G., Crooke, S. T., Eds.; Plenum Press: New York, 1987; pp 51–59.
(44) Boles, T. C.; Hogan, M. E. *Biochemistry* **1986**, *25*, 3039.
(45) Clayton, D. A.; Doda, J. N.; Friedberg, E. C. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2777.
(46) Segal-Bendirdjian, E.; Couland, D.; Roques, B. P.; LePecq, J-B. *Cancer Res.* **1988**, *48*, 4982.
(47) Muench, K. F.; Misra, R. P.; Humayun, M. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6.
(48) Yamamoto, K. R. *Annu. Rev. Genetics* **1985**, *19*, 209.
(49) Wang, J. C. *Annu. Rev. Biochem.* **1985**, *54*, 73.

sity[44,51] may influence sequence selectivity. Consequently sequence selectivity data obtained on linear DNA molecules may not always be directly applicable to nuclear DNA.

When a defined sequence in DNA is considered as a possible drug receptor (see later), it has some properties that are in sharp contrast to the more conventional proteinaceous receptor molecules. For example, when covalently modified by a drug, DNA may be restored by very efficient repair processes,[52] whereas proteinaceous receptors must be resynthesized. It follows that cells deficient in DNA-repair processes may be particularly sensitive to DNA-reactive drugs. Indeed there is evidence that some cancer cells defective in DNA repair are more susceptible to cross-linking agents such as nitrosoureas than repair-proficient cells.[53] Even within the cell nucleus selective repair may take place in certain regions.[54] In repair-proficient cells it is possible to potentiate the cytotoxic effects of DNA-reactive drugs by combining alkylating agents with inhibitors of DNA repair.[55] In this case, selectivity of the cytotoxic effect for tumor cells may not necessarily be achieved. Poly(ADP)ribosylation of histone molecules is associated with DNA damage by agents that cause DNA strand breaks directly or as a consequence of repair.[56] Poly(ADP)ribosylation is believed to be important for the survival of cells with damaged DNA, although the exact role is not clear. The discovery that benzamide derivatives are selective inhibitors of poly(ADP)ribosylation provided a rationale for laboratory experiments to potentiate the effect of alkylating agents, and at least in vitro significant increases in the cytotoxic potency of nitrosoureas have been achieved.[57] However, the in vivo effectiveness of this strategy has yet to be demonstrated. The multifunctional aspects of poly(ADP)ribosylation in cellular processes and nonselective inhibition by benzamide derivatives such as 3-aminobenzamide *at the high dosage levels used* may complicate the picture for rational combination therapy.[58] Nevertheless the demonstration of potentiation of cytotoxic potency by DNA-repair-associated inhibitors argues for the search for other inhibitors that may be useful in combination chemotherapy with DNA-reactive drugs.

The site-size DNA sequence specificity is a critical feature of drugs that exert their selectivity by reaction with target sequences. Virtually all DNA-reactive drugs exert some degree of sequence selectivity, although in the case of highly reactive alkylating agents such as the nitrogen mustards, this may be very modest.[10] Other drugs such as CC-1065 show a surprising degree of sequence selectivity.[59] Sufficient DNA sequence specificity is achieved by proteins to permit their participation in the precise control of gene expression and other genetic events.[60] Therefore the ability to design and synthesize a nonpeptide small molecular weight molecule that would bind to a desired DNA sequence of reasonable site size (up to about 15–16 base pairs) would be a significant achievement in medicinal chemistry.

Before discussing the design of drugs that have improved chemotherapeutic selectivity by virtue of increased sequence specificity, it is instructive to examine how natural biological molecules such as proteins and oligonucleotides achieve sequence specificity for duplex DNA. The most familiar DNA recognition motif for proteins that bind selectively to defined duplex DNA sequences is the "helix-turn-helix".[61] In this case, one-α-helix of the protein is held in the major groove of DNA by the second α-helix that lies across the back side of the first. Direct readout of major groove information such as hydrogen-bonding patterns and van der Waals contacts can be achieved by a complementary reading frame in the amino acid residues of the protein.[62] In some cases a sequence-dependent conformational change is required before the reading frames of the DNA and protein are brought into sync.[63] This has been termed "indirect readout" and in the case of the *trp* repressor results in specific hydrogen-bonding recognition of the phosphate backbone of DNA. As far as I am aware the design and synthesis of nonprotein mimics of this form of sequence-specific recognition (i.e. taking place in the major groove or on the phosphate backbone) have not yet been achieved. Oligonucleotides can also bind to duplex DNA in a sequence-selective fashion. Pyrimidine oligonucleotides bind to duplex DNA sequence specifically at homopurine sites to form a "triple helix" structure,[64] and RNA oligonucleotides can form a similar structure with a G-rich polypurine sequence.[65] Charge and stability problems of the oligonucleotides remain to be solved before these latter molecules can be used as drugs to attain selective pharmacological action in living cells. However, these are attractive ideas for cases in which high selectivity will be required (e.g. oncogene inactivation; see below).

Nonpeptide or nonoligomeric molecules that bind to DNA with sequence selectivity generally do so either by intercalation or by minor groove binding. While polyintercalators (i.e. molecules possessing two or more planar ring systems spaced by linkers to insert every two or more base pairs into DNA) can achieve a modest degree of sequence selectivity,[66] their inherently poor discrimination between A-T and G-C base pairs and the "site exclusion rule" prohibit their practical application as sequence-specific probes. Nonintercalating minor groove binding agents use planar recognition words for sequence recognition within the minor groove of DNA.[67] In contrast to the direct sequence-specific complementarity of protein

(50) Low, C. W. L.; Drew, H. R.; Waring, M. J. *Nucleic Acids Res.* **1986**, *14*, 6785.
(51) Mirabelli, C. K.; Huang, C. H.; Crooke, S. T. *Biochemistry* **1983**, *22*, 300.
(52) Friedberg, E. C. In *DNA Repair*; W. H. Freeman: San Francisco, 1984.
(53) Erickson, L. C.; Laurent, G.; Sharkey, N. A.; Kohn, K. W. *Nature* **1980**, *288*, 727.
(54) Hanawalt, P. *Cell* **1989**, *40*, 359.
(55) Collins, A. R. S.; Squires, S.; Johnson, R. T. *Nucleic Acids Res.* **1982**, *4*, 1203.
(56) Ueda, K.; Hayaishi, O. *Annu. Rev. Biochem.* **1985**, *54*, 73.
(57) Berger, N. A.; Berger, S. J.; Gerson, S. L. *Anticancer Drug Des.* **1987**, *2*, 203.
(58) Rankin, P. W.; Jacobson, E. L.; Benjamin, R. C.; Moss, J.; Jacobson, M. K. *J. Biol. Chem.*, in press.
(59) Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Mitchell, M.; Warpehoski, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886.
(60) Ptashne, M. *A Genetic Switch*; Cell Press and Blackwell Scientific Publications, Cambridge, MA, 1986.
(61) Ohlendorf, D. H.; Anderson, W. F.; Lewis, M.; Pabo, C. O.; Matthews, B. W. *J. Mol. Biol.* **1983**, *169*, 757.
(62) McClarin, J. A.; Frederick, C. A.; Wang, B.-C.; Geene, P.; Boyes, H. W.; Grable, J.; Rosenberg, J. M. *Science* **1986**, *234*, 1526.
(63) Otwinowski, Z.; Schevitz, R. W.; Zhang, R.-G.; Lawson, C. L.; Joachimiak, A.; Marmorstein, R. Q.; Luisi, B. F.; Sigler, P. B. *Nature* **1988**, *335*, 321.
(64) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645.
(65) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456.
(66) Wakelin, L. P. G. *Med. Res. Rev.* **1986**, *6*, 275.
(67) Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31.

recognition of DNA, nonintercalating minor groove-binding agents generally "read" DNA sequence by indirect mechanisms such as groove geometry, secondary structure, and electrostatic interactions.[67] These processes inherently give rise to a lower sequence specificity, or more aptly termed, sequence selectivity, than direct hydrogen bonding or van der Waals interactions. Nevertheless A-T and G-C selective "words" do exist that discriminate based upon recognition features in the minor groove of DNA.[15,67] A variety of natural and synthetic products including netropsin, Hoechst 33258, SN6999, and CC-1065 exemplify the utilizing of planar, sometimes fused aromatic rings that can be sandwiched within the minor groove of DNA.[67] The compounds listed above are all A-T "words" and heroic efforts have been made to convert A-T words into G-C words by Lown and co-workers.[68] While some degree of success has been achieved, the converted A-T word does not have a high selectivity for G-C base pairs. Some natural minor groove G-C words also exist, i.e., chromomycin and the anthramycins. Chromomycin has a complex recognition motif for G-C base pairs involving a dimer of antibiotic molecules and magnesium ion.[69] Anthramycin covalently bonds to the exocyclic 2-amino group of guanine in a nondistortive but helix-stabilizing manner.[70] Consequently in either case the use of these recognition words is complicated by those features. For this problem to be addressed successfully we will have to await the discovery or design of a simple G-C word that is chemically compatible with the existing A-T words. "Sentences" can then be constructed of suitable phased A-T and G-C words to read any chosen sequence. Even with the availability of suitable A-T and G-C words, it remains to be seen whether the minor groove of DNA inherently has sufficient sequence specificity information to give rise to the level of selectivity required to successfully read DNA. The lesson from Nature (i.e., protein–DNA interactions) would suggest the major groove is a better target.

At this time, any increase in selectivity of existing agents that react with DNA appears to depend upon either modulating the relative efficiency at which a target sequence is hit, or preferential inhibition of DNA repair in select cells, e.g. cancer cells. The relative efficiency of targeting a chosen cognate sequence vs other noncognate sequences is dependent upon the extent of sequence selectivity of the DNA-reactive drug and the relative frequency of occurrence of the chosen sequence in target cells and its accessibility to drug modification. These variables are attractive features for drug development. If the genomic target for drug reaction can be limited to functional domains such as transcriptionally or replicationally active regions, then the overall size of the target can be dramatically reduced. An alternative to the preferential inhibition of DNA repair in target cells as an approach to obtain improved selectivity is the design of less well recognized and consequently excised lesions on DNA. Efforts will surely be made in these directions, but it is difficult to predict the possibilities for improvement at the therapeutic level. In the next two sections I will focus on strategies for the discovery and design of new therapeutic entities.

III. The Design of Mechanism-Based Screens To Identify New Classes of Drugs That React with DNA and Associated Targets

DNA-reactive compounds such as the bleomycins, Adriamycin, and actinomycin D were originally identified through in vitro cytotoxic screens using for example L-1210 leukemia cells. These drugs are representatives of just a handful of clinically useful drugs that were selected by cytotoxic screens and show in vivo antitumor activity. Most cytotoxic agents identified by these screens show little selective activity in in vivo systems. The poor correlation between in vitro and in vivo activity has forced the drug discovery community to explore alternatives to cytotoxic screens. Two major screening strategies have evolved; the *disease-oriented screen* and *mechanism-based screen*. The NCI has recently adopted a disease-oriented approach to attempt to identify more selective antitumor agents.[71] In this screening program compounds are tested against a large panel of human tumor cell lines derived from a broad spectrum of solid tumors. Compounds that only show activity in vitro in *select* tumor lines are then earmarked for in vivo screening. It is still too early to evaluate how effective this new strategy will be in identifying new classes of clinically useful antitumor agents. In principle, mechanism-based strategies can use DNA or a DNA-mediated process as the screening event. I will describe these strategies in more detail.

The mechanism-based screen identifies a specific target or process that when modulated is likely to produce a desired pharmacological response, e.g., antiviral or antitumor activity. A review of mechanism-based screens for the discovery of new DNA-reactive drugs has been published by Johnson et al.[72] A broad screen that is sensitive to virtually all of the known DNA-reactive drugs is the biochemical induction assay (BIA).[73] The BIA is a modification of the lysogenic phage induction assay in which λ phage repressor regulates expression of β-galactosidase from a lambda promotor fused to a lac Z gene in *Escherichia coli*.[74] As a consequence of exposure to DNA-damaging agents the SOS response (DNA-repair pathway) is triggered, which leads to induction of expression of β-galactosidase. Bartus et al.[73] have increased the sensitivity of this screen by construction of a subclone with an increased copy number of the β-galactosidase transcriptional unit. This has been shown to be a very effective prescreen for active cultures that can then be further examined by more specific mechanism-based screens.

Many potent DNA-reactive drugs can produce single or double strand breaks in DNA. These reactions are generally oxygen dependent and may also depend on the presence of metal ions such as magnesium or copper. Other DNA-reactive compounds cause changes in the topological conformation of DNA by, for example, intercalation. By use of agrose gel electrophoresis of covalently closed circular DNA (cccDNA), these DNA strand breaks and topological changes can be sensitively detected,[75,76] e.g., conversion of cccDNA to nicked, circular, or linear forms that can be resolved by electrophoresis. DNA-reactive drugs that do not produce strand breaks or topological

(68) Lee, M.; Krowicki, K.; Hartley, J. A.; Pon, R. T.; Lown, J. W. *J. Am. Chem. Soc.* **1988**, *110*, 3641 and references therein.
(69) Gao, Z.; Patel, D. J. *Biochemistry* **1989**, *28*, 751.
(70) Hurley, L. H.; Needham-VanDevanter, D. R. *Acc. Chem. Res.* **1986**, *19*, 230.
(71) Alley, M. C.; Scudiero, D. A.; Monks, A.; Hursey, M. L.; Czerwinski, M. J.; Fine, D. L.; Abbot, B. J.; Mayo, J. G.; Shoemaker, R. M.; Boyd, M. R. *Cancer Res.* **1988**, *48*, 589.
(72) Johnson, R. K.; Bartus, H. F.; Hofmann, G. T.; O'Leary Bartus, J.; Mong, S.-M.; Faucette, L. F.; McCabe, F. L.; Chan, J. A.; Mirabelli, C. K. In *In Vitro and In Vivo Models for Detection of New Antitumor Drugs*, Proceedings of the 14th International Congress of Chemotherapy, Kyoto, 1985, pp 15–26.
(73) Bartus, H. F.; Mirabelli, C. K.; Auerbach, J. I.; Shatzman, A. R.; Taylor, D. T.; Roseberg, M.; Crooke, S. T. *Antimicrob. Agents Chemother.* **1984**, *25*, 622.
(74) Elespuru, R. K.; White, R. J. *Cancer Res.* **1983**, *43*, 2819.
(75) Mong, S.; Strong, J. E.; Busch, J. A.; Crooke, S. T. *Antimicrob. Agents Chemother.* **1979**, *16*, 398.
(76) Scannell, R. T.; Barr, J. R.; Murty, V. S.; Reddy, K. S.; Hecht, S. M. *J. Am. Chem. Soc.* **1988**, *110*, 3650.

changes in DNA either spontaneously or in the presence of metals may still be detected by using thermal treatment (e.g., CC-1065).[77] By use of a combination of these chemical, physical, or enzymatic treatments, it is possible to detect and categorize most of the known DNA-reactive drugs. Rapid differentiation of potentially new activities from already existing compounds can be made at this stage.

Mechanism-based screens that depend upon yeast mutants that have decreased or increased activity of key proteins involved in DNA processes such as repair are very useful in identifying both new compounds as well as classifying activities.[78] Because yeasts are eukaryotic organisms they are genetically and biochemically nearer to mammalian cells than bacterial cells. However, they retain the advantages of short generation times and genetic manipulability of prokaryotic organisms. The array of DNA-repair mutants available that have been characterized biochemically provide excellent opportunities to design mechanism-based screens to select compounds that interact with DNA to produce defined lesions on DNA such as double strand breaks. Johnston et al.[72] have used a battery of DNA-repair-deficient mutants of *Saccharomyces cerevisia* to discriminate between various classes of DNA-reactive drugs.

More recently yeast mutants have been isolated that are defective in the expression of topoisomerase I.[79,80] These mutants are either resistant to camptothecin[80] or when carrying a plasmid that overproduces topoisomerase I[79] are *hypersensitive* to camptothecin. These mutants have obvious roles in mechanism-based screens to discover other topoisomerase I inhibitors. With this principle in mind it should be possible to design other mechanism-based screens, providing the absence of enzyme does not prevent growth of the mutant. However, where the activity (e.g., topoisomerase I) is necessary to express the potent biological effects, it is overproduction rather than underproduction of the protein that will give rise to increased sensitivity to the agent.[79] While there are a variety of topoisomerase II inhibitors, camptothecin is unique as a topoisomerase I inhibitor. Cytotoxic screens apparently do not select well for topoisomerase I inhibitors; therefore, the mechanism-based approach described here may be a more effective screen.

Compounds that modulate gene expression either in a positive or negative way would seem to have potential therapeutic utility for treatment of diseases where either the absence or overexpression of proteins results in deleterious effects. Molecular biologists have cloned eukaryotic transcriptional regulatory systems into plasmids that can be conveniently manipulated to determine the sequences that are involved in activation of the system.[81] The glucocorticoid-inducible LTL gene[82] has been used as a model target to evaluate preferential drug effects on gene expression.[83] With this system the possibility that bleomycin, neocarzinostatin, and actinomycin D would induce alterations in either transcription or posttranscriptional gene expression was assessed. Although no preferential effects were observed for these particular drugs, this would seem to be a useful model system. In a more sophisticated system, mutants in the transcriptional activation system might be used as a mechanism-based screen to select for DNA-reactive drugs that would affect DNA structure or conformation and annul the effect of the mutation on transcriptional activity. Similarly, compounds that increase recombinational events might be detected in a mechanism-based screen. Clearly the application of molecular biology holds many attractive possibilities for detection of compounds with interesting biological properties.

IV. The Design of New Drug Classes Based upon the Identification of Pharmacologically Selective Receptors on, or Associated with, DNA

The ultimate objective in any drug design program is the synthesis of a specific drug entity that binds with sufficient selectivity to a receptor resulting in a defined pharmacological response without adverse side effects. For the initiation of pharmacological events that originate on DNA, the receptor molecule might be DNA itself, a DNA-binding protein, or a protein–DNA complex.

A receptor in the strict pharmacological sense must have both cognitive and response features.[84] The majority of sequences on DNA are acceptors rather than true receptors, since they lack response characteristics[41] (except in a toxicological sense, e.g., mutation, deletion, etc.). Receptor sequences may, for example, be part of regulatory regions of DNA that are normal DNA-binding regions for proteins or even short oligonucleotides such as RNA. For example, the transcriptional factor SpI binds to GC boxes in the 21 base pair repeat region of DNA in the transcriptional regulatory regions of various genes.[85] Many other proximal and distal signal regions have been identified in cis-acting regulatory and transcriptional control regions.[86] Such sequences are potential receptor sequences for drugs that might modulate gene expression.

While the sequence-dependent microheterogeneity of DNA gives rise to a vast array of recognition features for sequence specific binding of proteins,[87] there are also a variety of unusual DNA structures that can form within AT- or GC-rich regions of DNA. For example, guanine-rich DNA sequences can form "four-stranded" complexes in which the strands run in a parallel fashion.[88] Homopurine–homopyrimidine sequences $[d(C\text{-}T)_n\text{-}d(A\text{-}G)_n]$ are proposed to form a "triple-helical" structure (C:G:T).[89] G-rich strands occur at chromosome ends and form novel intramolecular structures.[90] It is believed that these unusual DNA structures are involved in such processes as meiosis (four-stranded complex)[88] transcriptional control (three-stranded complexes)[89] and chromosomal maintenance (G-rich strands).[90] These structures are also therefore potential receptors for selective drug action.

The chemical modification of DNA sequences that are specific DNA–protein binding sites could, in principle, lead

(77) Reynolds, V. L.; Molineux, I. J.; Kaplan, D.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228.
(78) Mirabelli, C. K.; Bartus, H.; Bartus, J. O. L.; Johnson, R. K.; Mong, S. M.; Sung, C. P.; Crooke, S. T. *J. Antibiot.* **1985**, *38*, 758.
(79) Eng, W. K.; Faucette, L.; Johnson, R. K.; Sternglanz, R. *Mol. Pharm.*, in press.
(80) Nitiss, J.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7501.
(81) Brown, D. *Science* **1981**, *211*, 667.
(82) Zaret, K. S.; Yamamoto, K. R. *Cell* **1984**, *38*, 29.
(83) Beckmann, R. P.; Beerman, T. A. *Mol. Pharm.*, in press.
(84) Green, J. P. *Trends Pharmacol. Sci.* **1987**, *8*, 90.
(85) Kadonaga, J. T.; Briggs, M. R.; Tjian, R. In *New Frontiers in the Study of Gene Function*; Poste, G., Crook, S. T., Eds.; Plenum Press: New York, 1987; pp 87–98.
(86) Freifelder, D. In *Molecular Biology*, 2nd ed.; Jones & Bartlett Publ. Inc.: Boston, 1987; pp 501–550.
(87) von Hippel, P. H.; Berg, O. G. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 1608.
(88) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364.
(89) Johnston, B. H. *Science* **1988**, *241*, 1800.
(90) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J.-R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6622.

to an increase or decrease in protein binding. It has been elegantly demonstrated that distamycin A or netropsin binding to a specific nonalternating AT sequence in the promotor region of the phage I leads to activation of transcription initiation.[91] This is because distamycin binding to DNA leads to an increase in RNA polymerase binding to the promotor region to form the "open" or active complex. This can be rationalized since distamycin produces a bending of the DNA helix in the spacer region (a region not contacted by the enzyme) that is also a requirement for formation of the RNA polymerase–promoter open complex. Consequently other sequence-specific DNA-binding drugs that modulate the structural or conformational forms of DNA binding protein regions might also be expected to increase or decrease protein binding and presumably the processes they control.

Altered forms of cellular proto-oncogenes have been implicated in the development of human cancer.[92] These transforming genes (oncogenes) are often found in solid tumors and leukemias. The *ras* oncogene is frequently associated with myeloid leukemias and various carcinomas.[93] This *ras* gene family encodes for 21kDa proteins that bind GTP (G-proteins).[94] Because the positions of oncogenic mutation are known for a number of *ras* genes, these seem possible selective targets for drug design. However, the single base substitutions that differentiate the normal cellular proto-oncogenes from their oncogenic counterparts make selective targeting a formidable task. Conceivably the triple-strand approach (see above) may be useful if this principle can be broadened to accommodate mixed A-T and G-C sequences. Even so, the previously mentioned problems with using oligonucleotides as therapeutic agents still remain.

As an alternative to targeting isolated unique DNA sequences, it may be possible to gain increased selectivity by targeting DNA structures and conformations *which are the consequence* of protein–DNA interactions. For example, the binding of RNA polymerase to DNA has been proposed to produce a β-kinked structure[95] that may be an optimal target for intercalators. The binding of topoisomerase II to DNA may create special receptor binding sites on DNA for intercalation agents such as *m*-AMSA.[32] Since these binding sites are *only created* as a consequence of protein binding to DNA, increased receptor selectivity may result.

While not directly a DNA target, the binding of drugs to the DNA recognition motifs of DNA-binding proteins such as zinc-fingers[96] or helix-turn-helix[61] motifs would be expected to modulate processes such as transcriptional control. Since a defined sequence duplex DNA molecule is the specific binding molecule for these proteins, it is an interesting, and as far as I am aware, untested idea to design surrogate duplex molecules that would compete for these sequence specific proteins. There are, of course, inherent problems in this approach such as duplex stability and cellular uptake. These however, are being addressed in the anti-sense area.[97,98] In a more complex case it might be possible to synthesize mechanism-based inhibitors of DNA-cleavage enzymes such as topoisomerases by designing modified duplex DNA molecules. This is particularly attractive where both the topoisomerase consensus sequence and the DNA-cleavage site within the region is also known.[99]

Conclusions

In this Perspective I have attempted to summarize some of the opportunities available to medicinal chemists through DNA and its associated drug targets. Modern techniques in structural chemistry (high-field NMR, X-ray crystallography, molecular modeling) and techniques and concepts from molecular biology provide new opportunities to reexamine the mechanism of action of existing compounds that are thought to interact with DNA, as well as select and design new drug classes. Many of the recent discoveries in molecular biology have yet to be applied to drug discovery. Although there are still tremendous voids in our understanding of how existing drugs that are thought to interact with DNA really do work, experiments can now be designed that can potentially narrow this gap. The therapeutic opportunities offered through DNA and associated targets are not limited to anticancer and antiviral diseases, but also include genetic disorders that result in over- or underexpression of gene products, gene therapy, and autoimmune diseases. Because DNA-reactive ligands can be carcinogens as well as chemotherapeutic agents, separation of these activities will be an important objective. Even in cases where a therapeutic product may not ultimately be forthcoming, the receptor-selective ligands designed and developed by medicinal chemists will likely be very useful as biochemical tools to unravel the complexities of processes such as transcriptional control in eukaryotic cells.

Acknowledgment. The research in the author's laboratory is supported by grants from the National Cancer Institute (CA-30349, CA-35318, R35-CA-49751) and NIH (A1-23229), the Welch Foundation, and the Texas Advanced Technology program. I gratefully acknowledge the critical evaluation by Pat McGovren (Upjohn) and Terry Beerman (Roswell Park) of the ideas contained in this paper. I thank Randall Johnson (SKF) and Myron Jacobson (NTS) for access to unpublished material, and, finally, I thank Joann Haddock for careful preparation of this manuscript.

(91) Bruzik, J. P.; Auble, D. T.; deHaseth, P. L. *Biochemistry* **1987**, *26*, 950.
(92) Cairns, J. *Nature* **1981**, *289*, 352.
(93) Barbacid, M. *Carcinogenesis* **1986**, *7*, 1037.
(94) Gibbs, J. B.; Sigal, I. S.; Scolnick, E. M. *Trends Biochemical Sci.* **1985**, *10*, 350.
(95) Sobell, H. M. In *Molecular Actions and Targets for Cancer Chemotherapeutic Agents*; Sartorelli, A. C., Lazo, J. S., Bertino, J. R., Eds.; Academic Press: New York, 1981; pp 119–135.
(96) Rhodes, D.; Klug, A. In *Nucleic Acids and Molecular Biology*, Vol. 2; Eckstein, F., Lilley, D. M. J., Eds.; Springer-Verlag: West Berlin, 1988; pp 149–166.
(97) Sun, J.-s.; Francois, J.-C.; Lavery, R.; Saison-Behmoaras, T.; Montenay-Garestier, T.; Thuong, N. T.; Helene, C. *Biochemistry* **1988**, *27*, 6039.
(98) Van der Krol, A. R.; Mol, J. N. M.; Stuitje, A. R. *Biotechniques* **1988**, *6*, 958.
(99) Bonven, B. J.; Gocke, E.; Westergaard, O. *Cell* **1985**, *41*, 541.