STUDIES OF DNA LESION-PROTEIN INTERACTIONS: TOWARD THE DEVELOPMENT OF DNA REPAIR-INHIBITING CHEMOTHERAPEUTIC AGENTS

by

ZHENG-HUAN CHEN

B.S., Fudan University, 1987

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 1995

© Massachusetts Institute of Technology, 1995
All rights reserved

Signature of Author_____
Division of Toxicology
September 13, 1995

Certified by_____
John M. Essigmann
Thesis Advisor

Accepted by_____
Peter C. Dedon
Chairperson, Committee on Graduate Students

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEP 19 1995

Science

LIBRARIES

This doctoral thesis has been examined by a Committee of the Division of Toxicology as follows:

Professor Gerald N. Wogan_____
Chairperson

Professor John M. Essigmann_____
Thesis Supervisor

Professor Robert Langer_____

Professor James R. Williamson_____

STUDIES OF DNA LESION-PROTEIN INTERACTIONS: TOWARD THE DEVELOPMENT OF DNA REPAIR-INHIBITING CHEMOTHERAPEUTIC AGENTS

by

Zheng-Huan Chen

Submitted to the Division of Toxicology on September 13, 1995, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

ABSTRACT

The first subject of this dissertation addresses the substrate specificity of a human ultraviolet (UV) light-damaged DNA recognition protein (UV-DRP). This protein is implicated in the repair of UV-induced DNA lesions by its absence in cells from complementation group E patients with xeroderma pigmentosum, a human DNA repair disorder. The results suggest that the pyrimidine (6-4) pyrimidone photoproduct (6-4 photoproduct), a minor product of UV irradiation of DNA, is a superior substrate for the UV-DRP, while the major photoproduct, the cyclobutane pyrimidine dimer (CPD), is not recognized by the protein. These findings may provide a molecular mechanism to explain the observed biased repair of 6-4 photoproducts over CPDs in mammalian cells.

The second subject of this dissertation focuses on a possible new way of designing anticancer drugs. There is some evidence that certain drug-DNA adducts have their toxicity enhanced because they attract specific cellular proteins. It is suspected that those proteins shield the adducts from DNA repair enzymes. Based on that concept, it is proposed that a drug could be synthesized consisting of a DNA-attacking chemical warhead linked to a ligand attractive to a tumor specific protein (TSP; e.g., a mutant p53 protein). It is predicted that in tumor cells where a TSP is present, the lethal DNA lesions formed by the drug will be protected by the TSP recruited to the lesion sites. In normal cells, by contrast, these lesions will be quickly removed by DNA repair systems due to the absence of the TSP. The result of this selective persistence of drug-DNA adducts in tumor cells is selective killing of tumor cells. We termed this drug design strategy "Fatal Engineering" because we believe it allows the "engineering" or "programming" of tumor-targeting drugs based on knowledge of the genetic changes responsible for particular tumors.

In an attempt to prove the aforementioned hypothesis, I have used a conjugate of psoralen (a DNA crosslinker) and biotin (an avidin- or streptavidin-binding ligand) as a test molecule. I have shown that both DNA-reacting and

protein-binding activities of the unconjugated precursors were retained in the bifunctional conjugate. Streptavidin, which can bind to the biotin moiety attached to DNA through the psoralen domain, not only slowed the removal of a nearby uracil by uracil DNA glycosylase but also protected a region on the DNA molecule flanking the adduct site. The size of the covered region is comparable to the size of the DNA patch released by the nucleotide excision repair pathway.

The uracil protection experiment helped to establish the concept of Fatal Engineering. The possibility of using the commonly-occurring mutant p53 proteins as the tumor specific proteins was explored. As a first step to obtain ligands for mutant p53 proteins, an in vitro selection procedure involving rounds of selection and amplification was carried out to select DNA aptamers from a randomized single-stranded DNA pool for a mutant p53-specific epitope. Groups of DNA aptamers were obtained with selective affinity for columns with immobilized peptides containing the mutant p53-specific epitope, Arg-His-Ser-Val-Val. The selected DNA aptamers had little affinity for columns containing a control peptide in which the sequence of the epitope region was scrambled. The selected aptamers, however, also failed to bind the target peptide free in solution, indicating that both the sequence of the target peptide and the immobilization step probably contributed jointly to the observed binding. Sequence analysis did not reveal any obvious consensus motifs or common secondary structures among the selected aptamer sequences. Steady-state binding analysis by using a cloned aptamer sequence as a model molecule revealed that the apparent dissociation constant ($K_{d(app)}$) between this aptamer and the immobilized peptide is approximately 9 μM. No binding between the epitope-binding aptamers and a mutant p53 protein, however, was demonstrated by either gel mobility shift or immunoprecipitation assays.

Thesis Supervisor: Dr. John M. Essigmann

Title: Professor of Chemistry and Professor of Toxicology

ACKNOWLEDGMENTS

During my six years at MIT as a graduate student, there are many people to whom I owe a great deal of appreciation. I thank my advisor, John Essigmann, for his guidance, support and patience in me. I also thank him particularly for the story he told us during a Genetic Toxicology class about his impressive performance in the isolation and characterization of the major aflatoxin B_1 DNA adduct, because it was then I decided to join his lab. I haven't regretted that decision a bit. I also thank the members of my thesis committee, Gerry Wogan, Bob Langer, and Jamie Williamson, for their help and interest in my project.

Many thanks go to all the past and current members of the Essigmann lab. I have learned that, just like there are many Chens in Chinese names, there are many Bills, Brians, Dans, Daves, Debs, Jills, Jims, Kevins, Lisas, Marjies, Marshalls, and Mikes in American names. However, the Bills, Brians, Dans, Daves, Debs, Jills, Jims, Lisas, Marjies, Marshalls, and Mikes I have met are surely the brightest, the most thoughtful, and the most generous people of all. I owe special thanks to Dan, who was my mentor when I first joined the lab and has helped me a lot, both technically and intellectually, over the years. I thank Karen and Anna for their help in my project. I also acknowledge Kim, Laurie, Denise, and Debra, for their administrative skills and generosities.

I also thank Xiao-Quan, Yuejun (Joe), Zhifeng (the other Z), Zhuang (John), Jin-Hai, and Chi-Hong, for the opportunities to talk about life, society and science in Chinese!

Finally, I would like to thank the people who mean the most to me, my family members. My wife, Hui-Juan, has been there for me during the happy and difficult times of my life. I couldn't thank her enough for her love and support. My lovely daughter, Ying-Ying, has made me the luckiest, the happiest, and the proudest father in the world. I also thank my parents-in-law, who came all the way from Shanghai to help us take care of Ying-Ying. I owe special thanks to my own parents, whose love and unconditional sacrifice have made all this possible. I dedicate this thesis to them. I also thank my other family members, especially my brother, Zheng-Fu, and my sister, Zheng-Lan, for their contributions to my educational costs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AP	apurinic/apyrimidinic
bp	base pair
CPD	cyclobutane pyrimidine dimer
cisplatin	*cis*-diamminedichloroplatinum(II)
DDP	diamminedichloroplatinum(II)
DRP	damage recognition protein
ds	double-stranded
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EP240	NH_2-Phe-Arg-His-Ser-Val-Val-Val-COOH (the mutant p53-unique epitope region is underlined)
EP240-Cys	NH_2-Thr-Phe-Arg-His-Ser-Val-Val-Val-Pro-Cys-COOH (the mutant p53-unique epitope region is underlined)
EP240S	NH_2-Phe-*Val-His-Val-Ser-Arg*-Val-COOH (the residues in italics represent a scrambled sequence of the mutant p53-unique epitope)
EP240S-Cys	NH_2-Thr-Phe-*Val-His-Val-Ser-Arg*-Val-Pro-Cys-COOH (the residues in italics represent a scrambled sequence of the mutant p53-unique epitope)
ER	estrogen receptor
HEPES	4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid
HMG	high mobility group
HPLC	high pressure liquid chromatography
hUBF	human upstream binding factor
IP	immunoprecipitation
IPTG	isopropyl-β-D-thiogalactopyranoside
K_d	dissociation constant
$K_{d(app)}$	apparent dissociation constant
kDa	kilodaltons
nt	nucleotide
NP-40	Nonidet P-40 (a detergent)
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonylfluoride
RF	replicative form
SDS	sodium dodecyl sulfate
SRY	sex-determining region Y
ss	single-stranded
TBE	89 mM Tris, 89 mM boric acid, 0.2 mM EDTA

TE	10 mM Tris-HCl pH 8.0, 1 mM EDTA
TMP	trimethylpsoralen
Tris	Tris(hydroxymethyl)aminoethane
TSP	tumor specific protein
U-17	oligonucleotide 5'-d(CGGCCGTACGUGCGCCG)-3'
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
XP	xeroderma pigmentosum

I. GENERAL INTRODUCTION AND RATIONALE FOR THE WORK

Recent studies have revealed the existence of proteins in eukaryotic cells that recognize and bind to DNA damaged by a variety of chemical or physical agents. Although the cellular functions of many of these damaged DNA recognition proteins (DRPs) remain unclear, it is reasonable to speculate that they may play roles in DNA repair. For example, the UV light-damaged DNA recognition protein (UV-DRP) is linked to DNA repair because of its apparent absence in cells from patients who belong to the genetic complementation group E of the human DNA repair disorder xeroderma pigmentosum (XP) (Chu and Chang, 1988). One goal of my thesis work was to determine the precise role of this protein in DNA repair. Work with the widely used anticancer drug *cis*-diamminedichloroplatinum(II) (cisplatin or *cis*-DDP), however, has suggested that DRPs do not always participate in the DNA repair process (Toney et al., 1989; Donahue et al., 1990; Pil and Lippard, 1992; Clugston et al., 1992; Brown et al., 1993; Treiber et al., 1994). Indeed, it has been speculated that the *cis*-DDP-DRPs may actually hinder the DNA repair process by shielding the DNA lesions from repair enzymes (Donahue et al., 1990, 1991). This intriguing suggestion has inspired a novel approach to drug design that will be the major focus of this dissertation.

The first goal of my thesis was to elucidate the possible role of the UV-DRP in the repair of UV-induced DNA lesions. UV irradiation can form a number of different DNA adducts, among which cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4 photoproducts) are the two major ones (Wang, 1976; Cadet and Vigny, 1990). Each of these UV DNA lesions can cause distortions to the DNA structure and is believed to be both cytotoxic and mutagenic. Studies with mammalian cells and cellular extracts have indicated that the relative repair rates of these lesions are biased by a factor of 4–12 in favor of repair of the 6-4 photoproducts (Mitchell et al., 1985, 1990; Thomas et al., 1989; Wood, 1989). I have demonstrated, in collaboration with

Daniel Treiber, that the difference in the repair of the two major lesions could be attributed to the differential recognition of them by the DNA repair systems. By molecular and biochemical analyses, I have shown that the UV-DRP has little affinity for the cyclobutane dimers, but binds selectively to the 6-4 photoproducts.

The major focus of my thesis work, however, has been on a project we call Fatal Engineering. The goal of this project is to design a novel class of anticancer drugs that will attack the genomic DNA and then have the ability to recruit tumor specific proteins (TSPs) onto the resultant DNA damages, thus enhancing the toxic effect of the drug selectively in tumor cells. The intellectual foundation of this project comes from recent studies on the possible mechanism(s) for the action of cisplatin. By using the electrophoretic gel mobility shift assay, a group of proteins were identified in human cell extracts that bound selectively to DNA fragments modified with cisplatin (Donahue et al., 1990, 1991). Although the initial speculation on the possible functions of these cisplatin-damaged DNA recognition proteins (*cis*-DDP-DRPs) was that they play roles in the repair of the cisplatin-induced DNA adducts, further studies revealed little evidence supporting this speculation (Pil and Lippard, 1992; Brown et al., 1993; Treiber et al., 1994). On the contrary, the binding of these proteins to the cisplatin-induced DNA adducts was proposed to be a determinant in the cytotoxicity of the drug through observations that these proteins bound specifically to DNA fragments modified by *cis*-DDP but did not bind to DNA modified by the clinically ineffective platinum compounds, *trans*-DDP and [Pt(dien)Cl]Cl (dien, diethylenetriamine). More specifically, the proteins bound to the 1,2-d(GpG) and -d(ApG) intrastrand crosslinks that are unique to *cis*-DDP modification of DNA but did not bind to 1,3-d(GpNpG) intrastrand crosslinks that can also be formed by *trans*-DDP modification of DNA (Donahue et al., 1990, 1991).

Two essentially nonexclusive models were proposed to explain the possible

roles of these *cis*-DDP-DRPs in cisplatin cytotoxicity (Donahue et al., 1990, 1991). One model proposes that these proteins are not DNA repair proteins; on the contrary, they may actually block the repair of the DNA adducts formed by cisplatin and therefore enhance the genotoxicity of the drug. A second model hypothesizes that these proteins have their own critical cellular functions, for example, regulating the transcription of genes essential for cell survival. Binding of these proteins to platinated DNA prevents them from binding to their natural substrates thereby altering the regulation of critical genes and resulting in cell death.

Further studies have been carried out in our, as well as other, laboratories to test these two models. The results of these studies have been detailed in the Literature Survey chapter of this dissertation. Although the precise mechanism(s) of the anticancer activity of cisplatin has yet to be firmly established, studies designed to elucidate the possible mechanism(s) did, however, inspire us with new ways of thinking about drug design. We postulate that selective killing of tumor cells can be achieved by the selective persistence of drug-DNA adducts in tumor cells. We propose the drug to be a bifunctional molecule containing a DNA-attacking domain (a chemical warhead) linked to a ligand-binding domain that is designed to bind with high affinity to a tumor specific protein (TSP). As illustrated in Figure 1, in tumor cells where a TSP is present, the DNA lesions formed by the drug will form complexes with the tumor specific protein through the TSP-binding domain of the drug. Binding of the TSP to the lesion sites is expected to hinder their efficient removal by the DNA repair systems. A practical result of this DNA repair blocking is the persistence of the drug in the cellular DNA of tumor cells which should enhance the killing effect of the drug. In normal cells, although DNA lesions are also formed by the drug, they will be quickly repaired by the versatile DNA repair systems due to the absence of the tumor specific protein in these cells.

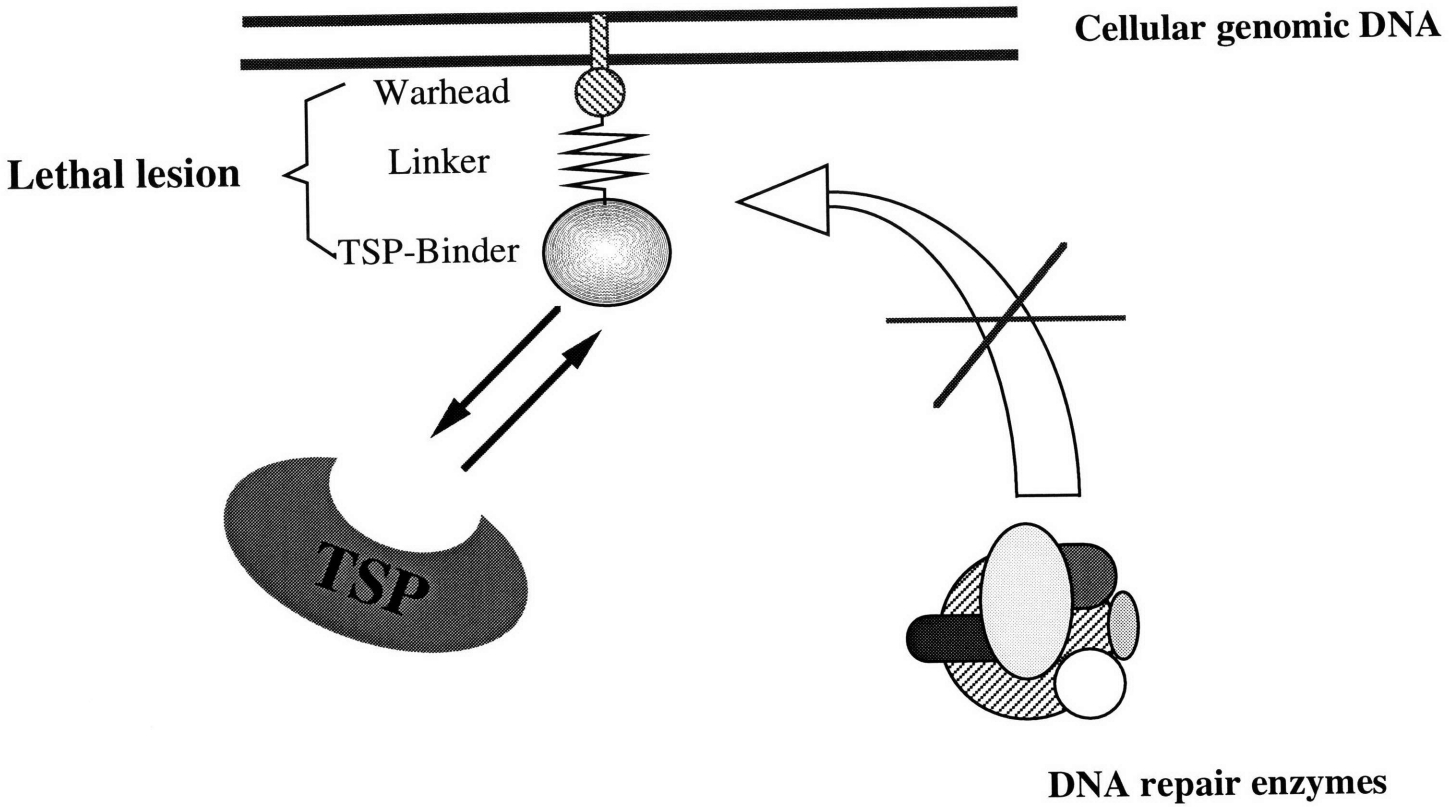
A tumor specific protein here is defined as a protein that is expressed either exclusively or at a higher level in tumor cells as compared to normal cells. Many overly-expressed oncogene or other cell-growth-promoting gene products, and mutated oncogene or tumor-suppressor gene products can be considered as tumor specific proteins. For example, estrogen receptor (ER) has been found to be uniquely expressed in some breast tumors (McGuire, 1976). Ligands such as estrogen derivatives, therefore, can be used as the protein-binding domain to design drug that can treat these breast tumors according to the scheme just described. Experiments leading toward this objective are being carried out in our laboratory.

The long term goal of the Fatal Engineering project is, however, to design drugs that can be used to treat many different types of tumors. For this purpose, a tumor specific protein that is common to a variety of different tumor types is needed. Mutant forms of tumor-suppressor p53 protein can serve as excellent candidates, because p53 mutations are the most common genetic changes during tumorigenesis and most of the mutant p53 proteins are only observed in tumor cells (see the Literature Survey chapter of this dissertation). Unlike estrogen receptor, however, mutant p53 proteins do not bind to molecules that are suitable for the drug design of Fatal Engineering. We need, therefore, to obtain ligands that can recognize mutant p53 proteins with high affinity and specificity through other means. In vitro selections involving combinatorial molecular libraries have been successfully used to obtain novel binding ligands against a variety of target molecules including proteins (Szostak, 1992). Thus, one of the goals of my thesis work was to explore the possibility of using combinatorial library selection to obtain molecules that bind tightly and selectively to mutant p53 proteins. I chose to use a randomized single-stranded DNA library, because the construction of this type of library was technically feasible in our laboratory and also because there had been reports of its successful applications.

Before I attempted to select novel mutant p53-binders, however, we needed to know whether the concept of Fatal Engineering project was valid. A bifunctional molecule containing a psoralen linked to a biotin derivative was used as a test molecule to prove the concept of Fatal Engineering. Psoralen and its derivatives can form covalent DNA adducts under near UV light irradiation; and biotin can bind tightly to avidin or streptavidin. After carrying out a series of biochemical experiments, I was able to show that this bifunctional molecule behaved as predicted by the scheme of Fatal Engineering.

In this dissertation, I have carried out a literature survey, which describes the background information for my thesis work (the chapter of Literature Survey). I then have reported my thesis work in the following three, relatively independent, chapters, each having its own sections of Introduction, Materials and Methods, Results, Discussion, and Conclusions. The central theme connecting these three chapters is the implication of DNA lesion-protein interactions for DNA repair and drug design. The specific interaction between the UV-DRP and the 6-4 photoproduct provides a possible molecular basis for the biased repair of this photoproduct in mammalian cells. Results of studying the interactions between streptavidin and the DNA adducts of a psoralen-biotin conjugate support the concept, which is the central premise of the Fatal Engineering approach, that a DNA lesion complexed with a protein could become resistant to repair and thus have an enhanced genotoxicity. The in vitro selection of DNA aptamers for a mutant p53-specific epitope represents a first step toward our long term goal: to design ligands with high and selective affinity for mutant p53 proteins. These ligands will be linked to DNA-attacking warheads to form bifunctional compounds that could be used to treat tumors where a mutant p53 protein is present.

Figure 1. A schematic illustration of the anticancer action proposed by the Fatal Engineering approach



II. LITERATURE SURVEY

A. OVERVIEW OF DNA REPAIR

The genome is constantly subject to alterations in its chemistry and sequence. Many of the changes arise as a consequence of errors introduced during normal biological processes such as DNA replication, recombination and even repair. Other alterations in DNA chemistry and sequence arise from the inherent instability of specific chemical bonds in DNA under physiological conditions of temperature and pH (e.g., depurination). Finally, a large variety of chemical compounds, either existing in an extracellular environment or produced endogenously, and a smaller number of physical agents can react with DNA, resulting in various forms of DNA damage.

DNA repair is a cellular response associated with the restoration of the normal chemistry of DNA following damage. Various DNA repair pathways have been identified and some have been characterized (Friedberg, 1985). Depending on the nature of the damage, a specific pathway is designed to deal with a specific damage; alternatively, several pathways may work in concert to repair a specific type of damage. On the other hand, some repair pathways are versatile and are capable of dealing with a variety of different types of DNA damage (e.g., the nucleotide excision repair pathway). In general, the various DNA repair pathways can be divided into two groups. In the first, the damaged sites are directly reversed without incision of the DNA strands. Examples of enzymes that are involved in this type of repair include: the bacterial and yeast photolyases, which photoreactivate the ultraviolet (UV) light-induced pyrimidine dimers; the alkyltransferases, which remove the alkyl groups on bases subjected to alkylation; and the DNA ligases, which directly ligate a single-strand break. In the second group of repair pathways, incision of the DNA strand is required. Following the incision step, this type of DNA repair also requires the synthesis of new DNA involving DNA polymerases and DNA ligases. Belonging to this group are the

base excision repair pathways involving various glycosylases (removing the damaged bases) and apurinic/apyrimidinic (AP) endonucleases, some of the mismatch repair pathways, the recombinational repair pathways, and the nucleotide excision repair pathways.

Since one goal of my thesis work was to study how the UV-induced DNA lesions are repaired by the nucleotide excision repair pathway, I have gone into depth in discussing this pathway in the following section of this dissertation.

B. ULTRAVIOLET (UV) LIGHT IRRADIATION OF DNA

1. DNA lesions formed by UV irradiation

The carcinogenic effect of ultraviolet light (200–300 nm) has long been attributed to its ability to damage DNA (Ananthaswamy and Pierceall, 1990). Of the many DNA lesions caused by UV irradiation, the *cis-syn* cyclobutane pyrimidine dimer (CPD) is the most common, followed by the pyrimidine (6-4) pyrimidone photoproduct (6-4 photoproduct) (Figure 2) (Wang, 1976; Cadet and Vigny, 1990). CPDs form at all dipyrimidine sequences, but the efficiency of dimer formation at a particular dipyrimidine sequence is strongly influenced by the flanking sequence context. In general, dimer formation is favored in runs of pyrimidines with the favored site for CPD formation at adjacent thymines (Brash and Haseltine, 1982). The primary site for 6-4 photoproduct formation, which occurs at roughly one third the frequency of CPDs (Friedberg, 1985), is at Py-cytosine (PyC) and CC sequences positioned 3' to runs of pyrimidines. The 6-4 photoproducts form rarely at other dipyrimidine sequences, including TT sites (Lippke et al., 1981).

Both CPDs and 6-4 photoproducts are believed to cause distortions to the

DNA architecture (Taylor et al., 1990; Smith and Taylor, 1993). As illustrated in Figure 2, the 6-4 photoproduct has an unusual structure in which the 6 carbon of T is covalently linked to what was the 4 carbon of the 3' neighboring C. The 5,6 double bond of T, but not C, is saturated, and the exocyclic amino group from the 4 position of C has been transferred to the 5 position of T (Franklin et al., 1985).

2. Repair of UV-induced DNA lesions

In *E. coli* cells, both types of pyrimidine dimers induced by UV light are repaired by the *E. coli* UvrABC excision nuclease (Sancar and Rupp, 1983). The UvrABC nuclease repairs a DNA lesion in double-stranded DNA by first binding to the lesion and then incising the lesion-containing strand flanking the lesion. More specifically, the eighth phosphodiester bond 5' to the lesion and the fourth or fifth phosphodiester bond 3' to the lesion are incised (Sancar and Rupp, 1983, Sancar and Tang, 1993). A current model (Sancar and Tang, 1993) for the mechanism of UvrABC nuclease action proposes that UvrA is the damage recognition unit. It acts as a dimer to deliver a monomer of UvrB to the site of damage. In doing so, UvrA binds transiently to DNA in a complex with UvrB. After UvrA dissociates from the complex, a stable UvrB-DNA complex is formed, which will then attract a UvrC protein. The incisions are performed by the bound UvrC (Orren and Sancar, 1989). Other gene products are involved in displacing the repair patch, synthesizing new DNA to fill in the gap, and ligating the 3' end to complete the repair process (Sancar and Tang, 1993).

In *E. coli*, CPDs are also repaired by photolyase, a photoreactivating enzyme, that does not act on 6-4 photoproducts (Brash et al., 1985), and T4 endonuclease, which is both a CPD-specific glycosylase and an apurinic site (AP site) endonuclease (Nickell et al., 1992). In photoreactivation, photolyase binds to CPD and, upon excitation by blue light (400 nm), splits the cyclobutane ring and

restores the intact bases (Kim and Sancar, 1993; Sancar, 1994). The action of T4 endonuclease on a CPD will result in a strand break, which can be then repaired by other bacterial enzymes (Nickell et al., 1992).

In mammalian cells, both types of UV photoproducts are repaired by the excision pathway (Aboussekhra and Wood, 1994). The mammalian DNA excision pathway, however, is more complex than the *E. coli* UvrABC system. Using a plasmid containing synthetic thymine dimers at unique sites in a human cell-free system capable of nucleotide excision repair, it has been established that the repair complex incises the 22-24th and the 5th phosphodiester bonds 5' and 3', respectively, to the photodimer, resulting in the release of a 27- to 29-nucleotide oligomer (Huang et al., 1992; Svoboda et al., 1993).

The best-characterized components of the mammalian excision repair pathway are the xeroderma pigmentosum (XP) factors that are defective in the seven complementation groups of XP cells, XP-A to XP-G (Tanaka and Wood, 1994). XP is a human genetic disorder that is defective in the first steps of the excision repair process. Patients with XP are characterized by high UV sensitivity and a strong predisposition to skin cancer (Cleaver and Kraemer, 1989). During the past several years, great progress has been made in cloning and characterization of the XP proteins. XPA is a DNA damage-binding protein (Robins et al., 1991; Jones and Wood, 1993; Aboussekhra and Wood, 1994). XPC is a single-stranded DNA-binding protein and could be involved in binding to an opened DNA structure during incision (Masutani et al., 1994; Aboussekhra et al., 1995). XPG is an endonuclease (O'Donovan et al., 1994a,b), while XPB and XPD are DNA helicases. Interestingly, XPB AND XPD are components of the RNA polymerase transcription factor TFIIH (Schaeffer et al., 1993, 1994; Drapkin et al., 1994a), suggesting that DNA repair and transcription are mechanistically coupled (Drapkin et al., 1994b). This finding may help to establish the molecular basis for the

preferential repair of DNA lesions in the coding stand of actively transcribed genes (Mellon et al., 1987). This finding may also provide a molecular mechanism for another human disorder, Cockaynes syndrome, which is believed to result from a defect in transcription-coupled repair, because XP-B patients also suffer from this syndrome (Schaeffer et al., 1993, 1994; Drapkin et al., 1994a).

The XPE protein has also been extensively studied. It has been shown that cells from two consanguineous XP-E patients are deficient in a protein that recognizes UV-damaged DNA (Chu and Chang, 1988; Hirschfeld et al., 1990; Kataoka and Fujiwara, 1991; Keeney et al., 1992; Treiber et al., 1992). More specifically, as discussed in this dissertation, the XPE protein binds to the 6-4 photoproducts induced by UV irradiation. The exact role for the XPE protein in DNA repair is unknown. It is suggested, however, from the mild phenotype of XP-E patients that this protein probably plays a supporting role in DNA repair. This suggestion is supported by a recent study in which in vitro excision repair assays with purified protein components showed that the XPE protein stimulated repair but was not essential (Aboussekhra et al., 1995).

Although the genes encoding many of the proteins involved in mammalian excision repair have been isolated, experiments studying the biochemistry of these proteins have just been started (Aboussekhra et al., 1995). The information on the repair of UV-induced DNA lesions has come mainly from in vitro repair assays using crude cell extracts (Wood, 1989) and in vivo studies with the repair of CPDs and 6-4 photoproducts measured by using antibodies against UV-damaged DNA (Mitchell et al., 1985). Both types of studies revealed that 6-4 photoproducts are repaired more efficiently than CPDs. What is responsible for the biased repair of 6-4 photoproducts remains unknown but, as shown in this dissertation, a damaged DNA binding protein that is absent in XP-E cells may play a role. The protein encoded by the XP-A gene may also be involved in the selective repair of 6-4

photoproducts because it also has affinity for UV-damaged DNA (Robins et al., 1991).

3. UV-damaged DNA recognition proteins

The existence of a protein in mammalian cells that can bind to UV-damaged DNA was first demonstrated by Feldberg and Grossman (1976) almost twenty years ago. Feldberg et al. (1982) then isolated a UV-damaged DNA recognition protein (UV-DRP) but failed to uncover the significance of this protein. More recently, Chu and Chang apparently rediscovered the UV-DRP activity and, importantly, suggested a role for this protein in DNA repair by showing that this activity was absent in two cell lines from XP group E patients (Chu and Chang, 1988). The generality of this finding was undermined by studies in which the UV-DRP activity was detected in several XP-E cell lines derived from Japanese XP patients (Kataoka and Fujiwara, 1991; Keeney et al., 1992). It is possible, however, that these cell lines contain a form of the UV-DRP that is defective in a function other than binding to UV-damaged DNA.

Since each XP group is deficient in the incision step of DNA excision repair (Cleaver and Kraemer, 1989), the UV-DRP may be part of the preincision complex, possibly acting at the level of damage recognition, in a role analogous to that of the *E. coli* UvrA protein (Orren and Sancar, 1989). The UV-DRP is induced 2-4 fold above constitutive levels in primate cells pretreated with UV or other agents that block DNA polymerase or damage DNA (Protić et al., 1989). These cells also reactivate UV-treated plasmids with increased efficiency. The UV-DRP is also overly expressed in human cells that display resistance to the anticancer drug cisplatin (Chu and Chang, 1990). However, it was also shown by Chu and Chang (1990), and more recently by Treiber (1993), that the UV-DRP activity did not bind to cisplatin-modified DNA with high affinity. Moreover, the

UV-DRP does not seem to bind to DNA modified by other agents (Abramic et al., 1991; Reardon et al., 1993; van Assendelft et al., 1993). These observations certainly do not support a role for the UV-DRP as a general DNA damage recognition protein. In contrast, the XPA protein can also recognize DNA modified by UV (Robins et al., 1991) and other agents (Jones and Wood, 1993; Aboussekhra and Wood, 1994) with moderate affinities, making it a candidate for a part of the general DNA damage recognition system (Jones and Wood, 1993). The UV-DRP/XPE factor, as discussed in Chapter III of this dissertation, probably has an auxiliary role in recognition, and/or in turnover of the incision/excision complex, during repair of UV-irradiated DNA.

Since the isolation of a human UV-DRP protein by Feldberg et al. in 1982, two other groups have reported the purification of UV-DRPs from human and monkey sources (Hwang and Chu, 1993; Abramic et al., 1991). These proteins appear to be identical based on their molecular weights (~127 kDa) and their ability to bind to UV-damaged DNA. Moreover, they are most likely the protein that is defective in XP-E cells (Takao et al., 1993), because microinjection of the p127 protein into XP-E fibroblasts corrected the mild repair defect of these cells (Keeney et al., 1994). A cDNA clone encoding the monkey UV-DRP has been recently isolated (Takao et al., 1993). The predicted amino acid sequence shows homology to a slime mold protein of unknown function, but no homology was found to any proteins, including DNA repair proteins, of known function. The UV-DRP that was studied in this dissertation appears to be the XPE protein because it was demonstrated that the activity was absent in the extracts of some XP-E cell lines.

C. ANTICANCER DRUG CISPLATIN

Cisplatin (*cis*-diamminedichloroplatinum(II) or *cis*-DDP) is an important

anticancer drug used mainly to treat ovarian and testicular cancers (Loehrer and Einhorn, 1984). It has been proposed that the therapeutic efficacy of cisplatin is due to its ability to react with DNA to form adducts that block replication, transcription, or both processes (Roberts and Thomson, 1979; Sorenson and Eastman, 1988a,b).

1. DNA adducts formed by cisplatin

cis-DDP binds to DNA at the N7 position of purines and primarily forms intrastrand crosslinks between adjacent nucleotides. The predominant adducts formed by *cis*-DDP are 1,2-intrastrand crosslinks at d(GpG) and d(ApG) sequences (Fichtinger-Schepman et al., 1985; Eastman, 1986), which make up 65% and 25%, respectively, of cisplatin DNA adducts (Table 1). Crosslinks at d(GpA) sequences are not observed. Minor adducts are 1,3 d(GpNpG) intrastrand crosslinks, monofunctional dG species, and interstrand crosslinks linking guanine residues on opposing strands.

trans-Diamminedichloroplatinum(II) (*trans*-DDP), a therapeutically ineffective isomer of *cis*-DDP, also binds to DNA at the N7 position of guanine to form both intrastrand and interstrand crosslinks (Eastman and Barry, 1987). *trans*-DDP cannot form 1,2-intrastrand crosslinks for steric reasons, but can form intrastrand crosslinks in which the guanines are separated by at least one nucleotide. These 1,3-intrastrand crosslinks appear to be the most abundant DNA adducts formed by *trans*-DDP (Table 1). The observation that *trans*-DDP cannot form 1,2-intrastrand crosslinks at adjacent nucleotides, which are the major DNA adducts formed by *cis*-DDP, has led to speculation that it is these adducts that are responsible for the potent therapeutic efficacy of *cis*-DDP (*vide infra*).

2. Cisplatin-damaged DNA recognition proteins

Although it has been suggested that cisplatin-DNA adducts exert their effects by inhibiting DNA and RNA synthesis (Bruhn et al., 1990) and by inducing programmed cell death (Eastman, 1993), the precise molecular mechanism for the significant chemotherapeutic efficacy of the drug remains elusive. Of possible importance to the cytotoxic mechanism of cisplatin is a family of eukaryotic proteins that bind selectively to cisplatin-modified DNA fragments (Chu and Chang, 1988; Donahue et al., 1990). Further studies have revealed several additional members of this protein family, including structure-specific recognition protein 1 (SSRP1) (Toney et al., 1989; Bruhn et al., 1992), high mobility group 1 (HMG1) (Pil and Lippard, 1992), human single-stranded DNA binding protein (hSSBP) (Clugston et al., 1992), the yeast protein IXR1 (Brown et al., 1993), the testis determining factor (SRY) (E. Trimmer, personal communication), and the human upstream binding factor (hUBF) (Treiber et al., 1994). With the exception of hSSBP, all of these cisplatin-damaged DNA recognition proteins (*cis*-DDP-DRPs) contain protein domains called high mobility group (HMG) boxes. The possible roles of these proteins in the cytotoxicity of cisplatin are discussed below.

3. Possible roles of the cisplatin-damaged DNA recognition proteins in the cytotoxicity of cisplatin

Since all the aforementioned proteins can bind to cisplatin-DNA adducts, it is reasonable to speculate that these proteins may be involved in the repair of cisplatin-induced DNA lesions (Donahue et al., 1990, 1991). However, of all the *cis*-DDP-DRPs that have been cloned, or otherwise identified, only hSSBP (Coverley et al., 1991) is implicated in DNA repair. Moreover, there is little evidence to suggest that this protein is involved the repair of cisplatin-DNA adducts.

As mentioned earlier, all of the remaining *cis*-DDP-DRPs contain protein motifs called HMG boxes. The HMG box is an 80-amino acid region that has conserved basic and aromatic residues and is the structural motif of a novel class of DNA-binding proteins (Jantzen et al., 1990; Lilly, 1992; Bianchi et al., 1992). An unusual feature of the HMG domain is its affinity for noncanonical DNA structures with sharp angles, such as four-way junctions (Ferrari et al., 1992). It is, therefore, believed that DNA bending and unwinding induced by the cisplatin adducts provide the recognition cues for these HMG-containing proteins (Bruhn et al., 1992, Pil and Lippard, 1992). However, HMG1 binds selectively to the cisplatin d(GpG) and d(ApG) 1,2 intrastrand crosslinks but lacks specificity for d(GpNpG) 1,3 crosslinks, indicating that the HMG box does not bind to all DNA structures bent by platinum coordination (Pil and Lippard, 1992). As already mentioned, the clinically ineffective isomer of *cis*-DDP, *trans*-DDP, forms 1,3- but not 1,2-intrastrand crosslinks; consequently, DNA modified by this compound is not recognized by the HMG-box proteins (Toney et al., 1989; Pil and Lippard, 1992). All these observations suggest a possible role(s) for these HMG-box proteins in cisplatin antitumor activity.

Since the HMG-box proteins bind specifically to cisplatin-induced DNA adducts but apparently are not involved in the repair of these adducts, a second model, in contrast to the first model described earlier, proposes that the bound HMG-box proteins actually prevent DNA repair enzymes from removing these adducts, thus sensitizing the cells to the toxic effect of the persisting cisplatin DNA lesions (Donahue et al., 1991). This hypothesis gains support from both in vitro and in vivo studies. Results from in vitro DNA repair assays (Wood et al., 1988) showed that the cisplatin d(GpG) 1,2 intrastrand crosslinks, to which the HMG-box proteins can bind, are poorly repaired by mammalian cell extracts (Szymkowski et al., 1992) while d(GpNpG) 1,3 intrastrand crosslinks, to which the proteins cannot bind, are repaired (R. Wood, personal communication). It is

noteworthy, however, that a conflicting point of view has emerged recently (Huang et al., 1994). Genetic data in support of the repair shielding hypothesis come from studies of a yeast strain lacking a HMG-box protein, IXR1. The IXR1 defective strain is 2- to 3-fold more resistant to the cisplatin than the wild-type strain, suggesting that IXR1 confers sensitivity to cisplatin presumably by shielding the cisplatin DNA adducts from being efficiently removed (Brown et al., 1993).

Studies outside the cisplatin field have also revealed evidence supporting the concept of the second, repair blocking, model. Chambers et al. (1985) reported that O^6-methylguanine was more mutagenic in cells proficient in DNA repair than in cells deficient in DNA excision ($UvrA^-$) and recombination ($RecA^-$) repair. This observation was quite unexpected because O^6-methylguanine was believed to be repaired mainly by the alkyltransferase systems (Samson and Cairns, 1977; Lindahl et al., 1988; Potter et al., 1987), and to a far less extent by excision repair (Samson et al., 1988). One possible explanation of this phenomenon is that both excision and recombinational repair are involved in repairing the O^6-methylguanine adducts, but with much less efficiency as compared to the alkyltransferases. So in cells proficient in all DNA repair systems, excision and recombinational repair enzymes actually impede the removal of the lesions from DNA by blocking the true repair enzyme (the methyltransferase) from gaining access to the lesion sites. Persistence of the O^6-methylguanine lesions on the DNA leads to misincorporation of bases by DNA polymerase, and hence higher mutational frequencies are observed. This explanation is supported by the work of Rossi et al. (1989). Studies on the role of poly(ADP-ribose) formation in DNA repair also lend support to the DNA repair blocking model. It has been known that poly(ADP-ribose) polymerase, the enzyme catalyzing the synthesis of poly(ADP-ribose) from nicotinamide adenine dinucleotide (NAD^+) (Althaus and Richter, 1987), binds strongly to DNA strand breaks. The bound enzyme subsequently undergoes auto-poly(ADP-ribosyl)ation and departs from the DNA breaks, allowing the repair

enzymes to correct the lesions. However, in the absence of NAD^+ or in the presence of a poly(ADP-ribosyl)ation inhibitor, 3-aminobenzamide, the enzyme does not undergo automodification but rather persists on the DNA, resulting in little repair of the DNA breaks presumably because the damages sites are shielded by the bound enzyme (Satoh and Lindahl, 1992). A study from the Sancar laboratory provided even more direct evidence that the removal of an adduct can be blocked if the accessibility of the adduct to repair enzymes is limited (Sibghat-Ullah and Sancar, 1990). In this study, the authors showed that *E. coli* photolyase, which can bind to and photoreactivate the cyclobutane dimers, inhibited excision repair activities in the absence of photoreactivating light, presumably by binding to the pyrimidine dimers and interfering with the recognition of the lesions by the excision repair system.

A third model, not exclusive of the second model, for the possible role(s) of the cisplatin adduct-binding HMG-box proteins suggests that the normal functions of these HMG-box proteins are essential for tumor cell growth. Platinum adducts, by affording structures that mimic the natural substrates of these proteins, would titrate these proteins away from their normal functional pathways, and thus compromising the welfare of tumor cells. Normal cells, in which the functions of these genes might be less critical, would be less vulnerable to the drug (Donahue et al., 1991; Treiber et al., 1994). Because some of the HMG-box proteins, such as hUBF, the lymphoid enhancing factor (LEF-1), and SRY are transcription factors (Jantzen et al., 1990; Giese et al., 1991; Harley et al., 1992), this model is also referred to as the transcription factor hijacking model (Treiber et al., 1994). A recent study showed that hUBF, a critical positive regulator of ribosomal RNA transcription (Jantzen et al., 1990), bound to DNA fragments containing the cisplatin d(GpG) 1,2 intrastrand crosslinks and its cognate rRNA promoter sequence with comparable affinities (Treiber et al., 1994). More recently, it has been shown that the product of the retinoblastoma susceptibility gene (Rb), a tumor

suppressor gene, likely inhibits cell growth by down-regulating the activity of this transcription factor and thus the synthesis of rRNA (Cavanaugh et al., 1995). This observation indicates that high levels of rRNA expression are needed for rapid cell growth. Titration of hUBF away from its normal function by cisplatin-DNA adducts would result in a decrease in the expression of rRNA and thus compromise the viability of fast-growing tumor cells.

D. EVIDENCE THAT COMPROMISED DNA REPAIR WOULD INCREASE THE SENSITIVITY OF CELLS TO DNA DAMAGING AGENTS

The Fatal Engineering approach proposed in this dissertation is based on the assumption that a DNA damaging agent can serve as an effective anticancer drug if the DNA lesions formed by this agent are resistant to repair in tumor cells. In other words, when the genome is damaged DNA repair is required to restore the integrity of the genome and thus the cell growth process; if the DNA repair machinery fails to restore genomic integrity quickly, persistence of the DNA lesions will inhibit the normal cellular processes (e.g., replication and transcription) and eventually lead to cell death. There has been accumulating evidence to support the vital role of adequate DNA repair in normal cell growth.

1. Cells genetically defective in DNA repair are hypersensitive to DNA damaging agents

The most direct and profound evidence for the important role of DNA repair in normal cell growth is from genetic observations. As discussed earlier, the characteristic features of patients of the DNA repair disorder xeroderma pigmentosum are their extreme sensitivity to sunlight and an overwhelming predisposition to skin cancers (Cleaver and Kraemer, 1989). Numerous studies

also document that cells derived from these patients are sensitive to the killing effect of UV light and other DNA-damaging agents. Moreover, the hypersensitivities of XP cells to DNA-damaging agents can be corrected by, for example, injection of the normal gene products into the cells (Keeney et al., 1994).

Additional genetic evidence comes from more recent studies on mismatch repair pathways. Defects in DNA mismatch correction have been implicated in a significant portion of hereditary nonpolyposis colorectal cancer (HNPCC) incidence (Fishel et al., 1993; Papadopoulos et al., 1994; Bronner et al., 1994). Cells derived from such tumors display microsatellite instability (Parsons et al., 1993; Modrich, 1994), thus providing a direct link between mismatch repair and genomic integrity and stability.

2. Cells can increase their viability either by enhancing DNA repair or by decreasing lethal DNA lesion production

There are numerous examples of cells acquiring resistance to DNA damaging agents by boosting their DNA repair capacities (e.g., Eastman and Schulte, 1988; Protić et al., 1989; Chu and Chang, 1990). Another way for cells to resist the genotoxicity of DNA-damaging agents is by avoiding the lethal DNA lesions. For example, the cytotoxicity of some DNA alkylating agents is attributable to the DNA strand breaks possibly generated by the involvement of DNA mismatch repair systems acting on alkyl adducts (Goldmacher et al., 1986; Karran and Marinus, 1992). One can imagine that cells can acquire resistance to this type of agent by limiting the production of the lethal DNA strand breaks, even at the risk of allowing the alkyl DNA adducts, which are less genotoxic, to persist in their genomes. This is apparently the case with a TK6 lymphoblast-derived cell line, MT1, which is resistant to the cytotoxicity of alkylating agent *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (Goldmacher et al., 1986; Kat et al., 1993). The cell line

has been recently shown to carry a defect in the G/T mismatch binding protein (GTBP), a subunit of a mismatch repair complex that is believed to be involved in repair of the DNA alkylation adducts (Drummond et al., 1995; Palombo et al., 1995; Papadopoulos et al., 1995).

3. Compromised DNA repair may result in cell death: genomic integrity and apoptosis (programmed cell death) are linked

For the past several years, great progress has been made in the genetics and biochemistry of programmed cell death (apoptosis). Interestingly, the link between apoptosis and genomic integrity becomes more and more evident. Upon DNA damage, cells often stop at G1 phase of the cell cycle, presumably allowing their DNA repair systems to restore the genomic integrity (Kastan et al., 1991). If, however, the DNA repair systems fail to do the job in time or properly, the cell death pathways will be activated and, quite often, these death pathways are dependent on the function of p53 protein (for representative studies in this field, refer to Lowe et al., 1993a,b). As to be discussed in the following section of this dissertation, the multifunctional p53 protein is probably the center of the network connecting DNA repair, normal cell growth progression, and apoptosis, providing the molecular basis for designing effective anticancer drugs that act by persisting in the DNA of tumor cells.

E. MUTANT p53 PROTEIN AS A TUMOR SPECIFIC PROTEIN

1. Overview

The *p53* tumor suppressor gene has come to the forefront of cancer research because it is the most commonly mutated gene in human cancers and the spectra of its mutations in these cancers are providing clues to the etiology and

molecular pathogenesis of neoplasia (Hollstein et al.,, 1991; Levine et al., 1991, Harris et al., 1993; Greenblatt et al., 1994). Over two thousand *p53* mutations have been observed in human cancers (Hollstein et al., 1994; Levine et al., 1994) and over half of these are missense mutations. Furthermore, except for a few germ line mutations that have been observed in patients with Li-Fraumeni syndrome, an inherited human cancer syndrome (Malkin et al., 1990; Srivastava et al., 1990), most of the mutations in the *p53* gene are somatic and have mostly been observed only in tumor cells. Two consequences of *p53* mutations have made the mutant form of p53 protein a good target for designing new cancer diagnostics and treatments. Because many mutant p53 proteins are often more stable than the wild-type protein (Finlay et al., 1988; Hinds et al., 1990), higher concentrations of the mutant proteins are often observed in tumor cells (Iggo et al., 1990; Moll et al., 1992). In addition, up to 40% of all the point *p53* mutations result in a common conformational change to the protein that can be detected by the monoclonal antibody PAb240 (Gannon et al., 1990).

2. Biology of the *p53* gene

The nuclear phosphoprotein p53 was originally discovered in rodent cells transformed by simian virus 40 (SV40) in a complex with the SV40 large T antigen (Lane and Crawford, 1979; Linzer and Levine, 1979). Subsequently, p53 was shown to be complexed with adenovirus and oncogenic papillomavirus oncoproteins (Sarnow et al., 1982; Werness et al., 1990). Because these viral proteins are needed to maintain the transformed phenotype, it was suggested that their interactions with the p53 protein are important for transformation. Since a variety of genomic and complementary DNA clones of *p53* were isolated that could immortalize cells in culture (Jenkins et al., 1984) and cooperate with the *ras* oncogene to transform primary rat embryo fibroblasts in cell culture (Eliyahu et al., 1984; Parada et al., 1984), *p53* was classified as an oncogene.

However, all of these transforming *p53* cDNA clones turned out to be mutant forms of the *p53* gene (Hinds et al., 1989). Furthermore, the *p53* gene is now classified as a tumor suppressor gene, negatively regulating the cell cycle and requiring loss-of-function mutations for tumor formation. Expression of cDNA or genomic clones of wild-type *p53* suppresses the transformation of cells in culture by other oncogenes (Finlay et al., 1989; Eliyahu et al., 1989), the growth of transformed cells in culture (Baker et al., 1990; Diller et al., 1990; Mercer et al., 1990; Michalovitz et al., 1990; Martinez et al., 1991), and the tumorigenic potential of these cells in animals (Chen et al., 1990). Deletions or mutations of wild-type *p53* alleles occur in several animal (Mowat et al., 1985) and human (Baker et al., 1989) tumors, and mice lacking *p53* are prone to the spontaneous development of a variety of tumors at a very early age (Donehower et al., 1992).

The human *p53* gene, localized to the short arm of chromosome 17 (17p13), spans about 20 kb of genomic DNA and contains 11 exons (Lamb and Crawford, 1986; Soussi et al., 1990). The human p53 protein is composed of 393 amino acids (Matlashewski et al., 1984) and can be structurally and functionally divided into three domains (Pavletich et al., 1993; Bargonetti et al., 1993; Wang et al., 1993) (Figure 4). The amino terminal (N-terminal) 42 amino acids can act as a transcriptional transactivation domain (Fields and Jang, 1990; Raycroft et al., 1990). The transactivation function can be inhibited by the bindings of the mdm-2 protein and the adenovirus 5 E1B 55 kDa protein to this region (Lin et al., 1994). The middle region of the protein, amino acid residues 120 to 290, is the sequence-specific DNA binding domain (Pavletich et al., 1993; Bargonetti et al., 1993; Wang et al., 1993). The carboxyl-terminal (C-terminal) domain, composed of residues 311 to 393, contains the nuclear localization signals for the protein (Shaulsky et al., 1991), the p53 protein oligomerization (tetrameric) sequences (Pavletich et al., 1993; Stürzbecher et al., 1992), several sites for phosphorylation by cellular protein kinases (Bischoff et al., 1990; Stürzbecher et al., 1990), a region

responsible for the non-sequence-specific DNA binding activity of the protein (Wang et al., 1993), and sequences that enable the protein to promote the annealing of complementary single-stranded nucleic acids (Oberosler et al., 1993; Bakalkin et al., 1994; Wu et al., 1995). Interestingly, Lee et al. (1995) have recently shown that the C-terminal domain of the p53 protein recognizes primary DNA damage in the form of insertion/deletion mismatches, suggesting a role for the protein as a sensor in monitoring DNA damage.

The cloning and sequencing of *p53* cDNAs from a variety of species has allowed researchers to analyze structural and evolutionary features of the p53 protein. It has been noted that some regions of the protein are highly conserved at the amino acid level throughout vertebrate evolution (Soussi et al., 1990). These conserved regions fall into five widely spaced amino-acid clusters termed domains I to V (Figure 4). In some of these domains the sequence conservation among all the species is perfect over stretches of 14 or more amino acids, indicating a high degree of functional importance. Interestingly, as will be discussed in greater detail later, the mutations in *p53* associated with human tumors tend to cluster in the conserved domains II through V (Levine et al., 1991; Hollstein et al., 1991)

p53 protein has been implicated in the control of the cell cycle, DNA synthesis and repair, cell differentiation, genomic plasticity, and programmed cell death (recently reviewed by Donehower and Bradley, 1993; Harris and Hollstein, 1993; Zambetti and Levine, 1993). This functional pleiotropy of p53 protein may be due to its ability to bind specific DNA sequences (Bargonetti et al., 1991; Kern et al., 1991; El-Deiry et al., 1992; Funk et al., 1992) and act as a transcriptional activator (Fields and Jang, 1990; Raycroft et al., 1990; Farmer et al., 1992; Kern et al., 1992; Unger et al., 1992), to function as a transcriptional repressor suppressing the transcription of genes with the TATA promoter sequence (Seto et al., 1992; Mack et al., 1993), and to form complexes with cellular and viral proteins (see

reviews by Donehower and Bradley, 1993; Harris and Hollstein, 1993).

With a PCR-based survey of genomic DNA sequences, the Vogelstein group revealed a consensus sequence for p53 binding. They showed that wild-type p53 binds to two copies of a 10-bp motif 5'-PuPuPuC(A/T)(T/A)GPyPyPy-3' separated by 0-13 base pairs (El-Deiry et al., 1992). Structural studies indicate that p53 protein binds as a tetramer to its target sequence (Cho et al., 1994). This type of binding is supported by the discovery of an oligomerization domain in the C-terminal region of the protein (Pavletich et al., 1993) and results from structural studies on the domain (Jeffrey et al., 1995). Searches for genes that contain such p53-binding motifs in their regulatory regions and are therefore potentially regulated by the p53 protein have revealed some interesting results. The ability of the p53 protein to arrest cells at G1 phase upon exposure of the cells to DNA-damaging agents is well known (Kastan et al., 1991). A more recent study has shown that wild-type p53 is an upstream factor that can lead to upregulation of a human growth arrest and damage-inducible gene, *GADD45* (Kastan et al., 1992). Although the function of *GADD45* remains unclear, it was postulated that expression of this and other effector genes allows cells to arrest in late G1 phase until the DNA damage is repaired. Also of interest is the fact that p53 protein can, as mentioned above, bind to single-stranded DNA ends and mismatches induced by insertion/deletion, indicating that p53 could function as a DNA damage sensor (Bakalkin et al., 1994; Lee et al., 1995). p53 can also induce apoptosis when it is overexpressed (Yonish et al., 1991; Shaw et al., 1992), or when cells are exposed to certain DNA damaging agents (Lowe et al., 1993a,b). A recent work showed that p53 is a direct transcriptional activator of the human *bax* gene (Toshiyuki and Reed, 1995) which encodes a protein that forms heterodimers with the human *bcl-2* gene product (Oltvai et al., 1993). Bcl-2 and Bax are homologous proteins that have opposing effects on cell life and death, with Bcl-2 serving to prolong cell survival and Bax acting as an accelerator of apoptosis (reviewed by Reed, 1994).

Since the inhibitory effect of Bcl-2 on apoptosis can be attributed to its ability to form heterodimers with Bax protein (Yin et al., 1994; Sato et al., 1994), upregulation of the *bax* gene would allow the Bax protein to exert its apoptosis-promoting activity. p53 has also been found to activate the transcription of the murine double minute 2 gene (*mdm-2*) (Wu et al., 1993), the product of which can, interestingly, bind to the p53 protein and inhibit its transcriptional activity (Momand et al., 1992; Oliner et al., 1993). The p53 and MDM-2 proteins form an autoregulatory feedback loop that controls the activity of the *p53* gene. Other genes that contain the p53-binding sites in their regulatory regions include the muscle creatine kinase gene (Weintraub et al., 1991; Zambetti et al., 1992), *WAF1* gene (El-Deiry et al., 1993), and a *GLN* retroviral element (Zauberman et al., 1993). The relationship of these genes to suppression of cell growth by p53 remains unclear. However, it was reported that introduction of *WAF1* cDNA suppressed the growth of human brain, lung, and colon tumor cells in culture (El-Deiry et al., 1993).

There have been numerous reports on the ability of p53 protein to repress transcription in a wide variety of cellular and viral promoters, including those from PCNA (proliferating-cell nuclear antigen), *c-fos*, beta actin, *c-jun*, IL-6, hsc70, Rb (retinoblastoma susceptibility gene), MDR (multidrug resistance gene), Rous sarcoma virus, HIV, HSV-1 (herpes simplex virus), SV40, and CMV (cytomegalovirus) (Subler et al., 1992; Deb et al., 1992; Ginsberg et al., 1991; Santhanam et al., 1991; Kley et al., 1992; Shiio et al., 1992; Chin et al., 1992). Whether p53 represses genes in a manner similar to the way it activates them, by binding to a specific DNA sequence, is unclear. However, a repeated cis-acting sequence, GGAACTGG, was identified which confers the susceptibility to the repression by wild-type p53 (Shiio et al., 1992). It is also possible that p53 mediates its repression effect through another DNA-binding protein. It has been shown that wild-type p53 binds to TATA-binding protein (TBP) which is an

essential member of the transcription-initiation complex for the TATA sequence-containing promoters (Seto et al., 1992; Mack et al., 1993). As mentioned earlier, p53 inhibits transcription from minimal promoters containing only the TATA sequence (Seto et al., 1992).

The effect of the p53 protein on gene transcription is influenced by several factors. Firstly, loss-of-function mutations in the *p53* gene have been shown to abolish for the most part the ability of its protein to bind sequence-specifically to DNA. The sequence-specific DNA binding activity of the p53 protein is believed to be essential for its regulatory function in gene transcription (Cho et al., 1994). Secondly, DNA-binding activity of p53 can be regulated by post-translational modification, such as phosphorylation of the protein. Hupp et al. (1992) showed that the sequence-specific DNA-binding activity of bacterially produced p53 was inhibited by the C-terminal region of the protein, and that this inhibition could be overcome by phosphorylation of the C-terminal region with casein kinase II. The authors hypothesized that alteration of the C-terminal region by phosphorylation changes the conformation from an inactive form to an active form for DNA binding. As discussed earlier, the C-terminal region of the p53 protein possesses a non-sequence-specific nucleic acid binding activity (Oberosler et al., 1993; Bakalkin et al., 1994; Wu et al., 1995). Bayle et al. (1995) have recently shown that regulation of the sequence-specific DNA binding activity of the p53 protein by its C-terminal region is through the non-sequence-specific nucleic acid binding activity of the region. The p53 protein can also be phosphorylated on other serine residues, located either in the C-terminal region or other parts of the protein, by other cellular kinases including $p34^{cdc2}$ kinase (Addison et al., 1990; Bischoff et al., 1990), casein kinase I (Milne et al., 1992), double-stranded DNA activated kinase (Lees-Miller et al., 1990; Wang and Eckhart, 1992), and MAP kinase (Milne et al., 1994, 1995). The roles of these phosphorylation events remain to be investigated. Thirdly, the activity of p53 protein can be modulated by its ability to bind to a

variety of cellular and viral proteins. The aforementioned complexing of the MDM-2 protein to p53 inhibits the tumor suppression function of p53 protein. Indeed, amplification of *mdm-2* gene was observed in a good portion of tumors studied (Oliner et al., 1992), indicating that *mdm-2* plays a role in cancer development probably by inhibiting the function of the *p53* gene. Mutant but not wild-type forms of p53 have been shown to bind to a heat shock or chaperon protein called hsc70 (Pinhasi-Kimhi et al., 1986; Hinds et al., 1987). The role of this particular binding is not clear, but it has been speculated that it may chaperon or transport p53 protein into the nucleus (Levine, 1990). There are three classes of DNA tumor virus oncoproteins that can complex with p53 (see a recent review by Donehower and Bradley, 1993). The large T antigen of SV40 and the E1B 55 kDa protein of adenovirus inhibit the activity of p53 apparently by directly binding to the protein. The third class of oncoproteins that have been found to complex with p53 are the E6 proteins of the oncogenic human papillomaviruses HPV-16 and HPV-18. These viral oncoproteins abrogate p53 function by binding to p53 and stimulating degradation of the bound p53 through ubiquitin-dependent proteolysis. Interestingly, other than containing oncoproteins that complex with p53, all three classes of DNA viruses also have proteins that form complexes with the retinoblastoma susceptibility (Rb) protein, another prototypical tumor suppressor (also see the review by Donehower and Bradley, 1993). This observation suggests that the function of both tumor suppressor proteins must be abrogated simultaneously in order to create a cellular environment, for example S phase, more amenable to the growth of these DNA tumor viruses. In the case of SV40, T antigen forms a complex with RB and p53 while adenoviruses and oncogenic papillomaviruses use different proteins (adenovirus E1A and HPV E7, respectively) to complex RB.

3. *p53* mutations in human cancers

Inactivation of the p53 tumor suppressor protein is a common event in the development of many human cancers. About half of all cancer cases involve missense mutations of one *p53* allele coupled to the deletion of the second allele (Hollstein et al., 1991; Levine et al., 1991; Donehower and Bradley, 1993; Harris and Hollstein, 1993; Greenblatt et al., 1994), and many of the remaining cases involve cellular or viral oncogenies that inactivate p53 (Scheffner et al., 1990; Oliner et al., 1992). Furthermore, transgenic mice lacking *p53* are prone to the spontaneous development of tumors at a very early age (Donehower et al., 1992), demonstrating the indispensable role *p53* plays in preventing cancer.

The nature or type of mutation observed depends on the tissue origin of the cancer. In carcinomas, 75–80% of the mutations are missense mutations producing a faulty protein product. The second allele in these cells is then lost, reducing the mutant allele to homozygosity (Baker et al., 1989; Nigro et al., 1989). In sarcomas, deletions, insertions, and rearrangements are more common while point mutations are rare (Levine, 1993). In some sarcomas there is an amplification of the oncogene, *mdm2*, whose product inactivates wild-type p53 function (Momand et al., 1992; Oliner et al., 1992). This phenomenon, however, has not been observed in carcinomas. The apparent dependence of mutations observed in the *p53* gene may reflect the mechanisms of mutagenesis, which could differ among cell or tissue types (see below).

Missense mutations in carcinomas have been detected in tumors of the anus (Crook et al., 1991), brain (Nigro et al., 1989), breast (Prosser et al., 1990), esophagus (Hollstein et al., 1990; Hollstein et al., 1991), stomach (Tamura et al., 1991), liver (Bressac et al., 1991; Hsu et al., 1991), lung [small-cell and nonsmall-cell carcinomas (Takahashi et al., 1989; Chiba et al., 1990; Iggo et al., 1990; Hensel et al., 1991)], lymphoid system (Ahuija et al., 1989; Gaidano et al., 1991), ovary (Marks et al., 1991), and prostate (Isaacs et al., 1991). The distribution of

these mutations in the gene is not random. All of the mutations observed in these studies are localized between residues 120 and 290, with most mutations found in the four conserved regions (Figure 4). Recent biochemical studies have demonstrated that the core portion of p53 (residues 102–292) folds into a compact structural domain that contain the sequence-specific DNA binding activity of the protein (Pavletich et al., 1993; Bargonetti et al., 1993; Wang et al., 1993). Mutations in this region would alter the folding process and result in the loss of the DNA binding activity that is believed to be critical for the normal functions of the protein (Cho et al., 1994).

In addition to the uneven distribution of the observed mutations, some specific codons are mutated very frequently (45–50% of the time), and these map to codons 175, 248, 249, 273 and 282 (Hollstein et al., 1991; Levine 1993; Greenblatt et al., 1994). Furthermore, different hot-spot mutations tend to be found in different cell or tissue types. For example, in liver cancers from southern China and southern Africa where there is a high incidence of Hepatitis B virus (HBV) infection and dietary exposure to aflatoxin B_1, 10 out of 13 mutations observed were G to T transversions that occurred at the third base pair of codon 249, and it was hypothesized that this unusual distribution of hot-spot mutations is due to a synergistic effect of HBV infection and dietary aflatoxin B_1 consumption (Bressac et al., 1991; Hsu et al., 1991). This hypothesis is supported by the prominence of G to T transversions in mutagenesis studies of aflatoxin B_1 (Foster et al., 1983; Aguilar et al., 1993; E. Bailey, personal communication).

Results from studies with skin, lung and colon cancers also support the hypothesis that spectra of *p53* mutations reflect the exposure of a tissue to a selective class of mutagens. In the case of skin squamous-cell carcinoma, the main risk factor is exposure to sunlight, which contains ultraviolet (UV) light that is highly mutagenic owing to the pyrimidine dimer premutagenic lesions it generates

(Drobetsky et al., 1987). Of all the experimentally examined mutagens, UV radiation leaves the most distinctive fingerprint in DNA: unrepaired cytosine dimers induce tandem mutations, in which two adjacent cytosine residues (cytosine-cytosine) are replaced by two thymine bases (thymine-thymine), an event that occurs very rarely unless there is exposure to UV radiation. Three of the first 15 mutations observed in the *p53* gene of skin squamous-cell carcinomas were such tandem substitutions (Brash et al., 1991).

In the case of lung carcinomas, it has been established that cigarette smoking is responsible for 90% of lung carcinomas in men and 78% in women (Shopland et al., 1991), and results from the mutational spectrum analysis of the *p53* gene are consistent with this relationship. *p53* mutations are common in lung cancer, with the highest prevalence (70%) in small cell lung carcinomas (SCLC) and the lowest in adenocarcinomas (Greenblatt et al., 1994). The prevalent mutation is the G:C→T:A transversion with a predominance of guanine residues on the nontranscribed strand, and the frequency of transition at CpG sites (9%) is lower than in almost all other cancers (Greenblatt et al., 1994). This spectrum is consistent with data for several different types of chemical carcinogens found in tobacco smoke. The highly mutagenic metabolites of polycyclic aromatic hydrocarbon (PAH) compounds such as benzo(*a*)pyrene preferentially attack deoxyguanosines and lead to mutations (Mazur and Glickman, 1988; Yang et al., 1991; Ruggeri et al., 1993), and one of the quantitatively minor adducts from the tobacco-specific *N*-nitrosamine 4-(methylnitrosamine-1-(3-pyridyl)-1-butanone also leads to G:C→T:A transversions (Ronai et al., 1993). It should be pointed out, however, that there is also an array of chemicals, DNA adducts, and endogenous processes known to elicit G to T transversions. For example, 8-hydroxy-2'-deoxyguanosine, which is found in considerable amounts in human DNA (Shigenata et al., 1989; Ames and Gold, 1990), also generates G to T transversions (Wood et al., 1990; Shibutani et al., 1991).

Finally, mutational analysis of the *p53* gene also suggests that endogenous mutagens could contribute to the development of cancers. Of all the mutational events found in the *p53* gene in human tumors, the most commonly observed are transitions (C:G→T:A and G:C→A:T) at CpG dinucleotides (Hollstein et al., 1991). The unusual mutability of CpG dinucleotides is well documented and is attributed to the disproportionally high mutation rate of 5-methylcytosine residues found at these dinucleotides (Jones et al., 1992; Spruck et al., 1993; Laird and Jaenisch, 1994). Indeed, 24% of all *p53* mutations described are transitions at CpG dinucleotides. In the case of colorectal tumors, 47% of the characterized point mutations occur at CpG sites (Greenblatt et al., 1994). This frequency contrasts with the fact that CpG dinucleotides occur within the coding region of the wild-type of *p53* at a frequency of 3.3% (39 out of 1179) (Greenblatt et al., 1994; Laird et al., 1995). The target size for mutagenesis at either nucleotide within CpG dinucleotides is therefore 6.6%. Mutations at CpG dinucleotides are thus overrepresented approximately 7-fold in human colorectal tumors (Greenblatt et al., 1994). Methylation of mammalian DNA at the C-5 position of cytosine in CpG dinucleotides is a biological process that is essential for normal mammalian development (Li et al., 1992). This biological process has also been implicated in oncogenesis through numerous observations (Spruck et al., 1993; Laird and Jaenisch, 1994). A recent study that provides by far the most compelling evidence implicating a role for DNA methylation in cancer development was done by Laird et al. (1995) in which they demonstrated that animals with reduced activities of DNA methyltransferase, the enzyme that methylates C-5 position of cytosine in CpG dinucleotides, showed reduced levels of intestinal neoplasia formation.

While these somatic mutations in the *p53* gene are found in 50–60% of human cancers, *p53* mutations are also detected in the germ line of some families. The Li-Fraumeni syndrome is characterized by a family with a proband having a sarcoma at a young age, and two first-degree relatives in the same family with

cancer (Li, 1988a; Li et al., 1988b). Several such families have now been shown to have missense and nonsense mutations in one *p53* allele in the germ line (Malkin et al., 1990; Srivastava et al., 1990; Iavarone et al., 1992; Malkin et al., 1992; Toguchida et al., 1992). Individuals in these families most often have a proband with an osteosarcoma and other relatives with adrenocortical carcinomas, breast tumors, or brain cancers. Colon cancers, with very high levels of somatic *p53* mutations, are not prevalent in these families (Levine, 1993)

F. IN VITRO SELECTIONS INVOLVING MOLECULAR LIBRARIES

1. Overview

Drug discovery typically involves the identification of a lead compound and the subsequent optimization of the lead compound to meet the criteria of being clinically effective. Traditionally, identification of a lead compound has been achieved either by random screening of natural products or synthetic chemicals, or by modifications of chemicals with known physiological activities, usually the natural agonist for the targeted molecule. Although advances in medicinal chemistry, structural and molecular biology, and molecular pharmacology have made the drug discovery process progressively more rational, i.e., rational drug design, random screening of natural products is still one of the technological bases of the modern pharmaceutical industry (Ecker and Crooke, 1995). However, the ratio of novel to previously discovered compounds through this approach has diminished with time (Ecker and Crooke, 1995).

The recent development of novel molecular libraries and novel ways of using these libraries has regenerated the excitement about random screening as a tool for drug discovery. There are, however, significant differences between the new development and the traditional way of random screening. First, the size and

complexity of the new molecular libraries are much greater than any of the previously used pools of natural compounds due to the use of combinatorial ways of constructing the new molecular libraries. Second, the strategy for screening often mimics the natural selection processes involving competition, selection and amplification that ultimately result in somewhat homogeneous populations with desired properties. More specifically, the new synthesis and screening methods, which are currently grouped under the term "combinatorial," often involve reiterative selection cycles. These cycles include the construction of libraries through chemical, biochemical or biological (organism-based) synthesis, the selection for molecules with desired characteristics, and the amplification of these selected molecules.

As illustrated in Figure 5, there are a variety of molecules that can be generated through various means. These diverse populations can grouped into monomeric and oligomeric libraries. Either chemical or biochemical synthesis can be used for the construction of both groups of libraries. For example, peptides, oligonucleotides, and sugar derivatives can be synthesized either through direct organic chemical reactions, or through enzymatic reactions (e.g., RNA and DNA polymerases for the synthesis of RNA and DNA, respectively). It is interesting to point out that biological organisms can be also used to produce a diverse population of biopolymers. For example, by using genetic engineering techniques, degenerate DNA fragments can be inserted into a phage surface gene, and the resulting recombinant phage library will display a diverse population of short peptide segments on their surface with each individual phage expressing a unique peptide segment. This technology is often called phage display (Scott and Smith, 1990) and has been widely and successfully used to screen for peptide fragments attractive to a variety of target molecules (Clackson and Wells, 1994).

After a molecular library is generated, a proper selection, or screening,

procedure needs to be designed to obtain those molecules with desired characteristics. The exact method of selection often depends on the overall goal of the project. In general, the various screening methods can be divided into two groups. The first group of screening methods is based on functional separation. For example, enzyme inhibition assays, receptor binding assays, whole cell assays or animal assays are often used to determine the value of a compound(s) under testing (Figure 6). In the past, natural product and chemical files were simply screened, usually individually, with these functional assays. The search rate of this screening method is obviously limited. Currently, screening methods based on functional separation have incorporated the power of combinatorial strategies; the compounds are screened in mixtures and the step of synthesis is an integral of the screening strategy. Selection strategies in which the steps of synthesis and screening are linked include iterative subset deconvolution (Geysen et al., 1986; Geysen and Mason, 1993; Owens et al., 1991; Blake and Litzi-Davis, 1992; Houghten et al., 1991, 1992; Ecker et al., 1993; Carell et al., 1994a,b; Carell et al., 1995; Erb et al., 1994; Wyatt et al., 1994; Zuckermann et al., 1994), positional scanning (Pinilla et al., 1992; Dooley et al., 1993), and partial release from a solid support (Lebl et al., 1993; Salmon et al., 1993). All of these methods can be used in functional assays because the compound mixtures synthesized in these methods can be tested in solution. The second group of selection methods is based on physical separation. This strategy employs binding assays as a means for isolating desired compounds from combinatorial molecular libraries (Figure 6). In binding assays, the targets can either be fixed on a solid phase while the compounds are in solution, or vice versa (Figure 6). The strategy with the target fixed on a solid support operates similarly to affinity chromatography; an affinity column containing the immobilized target molecules is used to attract ligands for the target from a molecular pool. This strategy is commonly used in the in vitro selection field (e.g., Ellington et al., 1990, 1992; Brenner and Lerner, 1992; Bock et al., 1992). The work reported in this dissertation also used this strategy. The strategy

in which the compounds are fixed while the targets are in solution often needs a way to identify the most active compound. Quite often, methods using identifiable codes, such as addressable matrices or pins, are used to facilitate the identification of the active compounds on the solid support (recently reviewed by Janda, 1994; Desai et al., 1994). Also based on physical separation, but different from the above two groups, a new way of isolating the desired species from a binding assay is to carry out the binding assay with compounds and targets both in solution, and subsequently to identify the complexed parties by using the dramatically improved micro-separation techniques and mass spectrometry methods (Chu et al., 1993).

Since only a very small portion of a molecular library will possess the desired properties, amplification after each selection is often needed to generate enough material for further selections. Amplification can be achieved simply by new syntheses after each selection step. Another popular way of amplifying the chosen molecules is to use enzymes or biological systems to make more copies of the selected molecules. Polymerase chain reaction (PCR) procedures are often used to amplify selected populations consisting of nucleic acids (Famulok and Szostak, 1992a) or molecules tagged with nucleic acids (Brenner and Lerner, 1992). In those involving RNA molecules, the combination of reverse transcription which converts RNA to DNA, PCR, and in vitro transcription can accomplish the amplification. In the approach of phage display, amplification can be achieved by simply picking up the plaques containing positive phage and growing these phages in bacterial cells.

Since part of my thesis work involved the use of single-stranded DNA libraries, in this portion of my dissertation I shall focus on the selection approaches involving oligonucleotide libraries. I shall then briefly describe some published work using combinatorial small chemical library, because, as discussed elsewhere, approaches involving these small chemical libraries probably represent the future

direction of this thesis.

2. In vitro selection of oligonucleotide aptamers

The word "aptamer", derived from the Latin "aptus" (to fit), was coined by Ellington and Szostak (1990) in a pioneering study in which RNA ligands with novel binding activities were obtained. By chemically synthesizing DNA molecules consisting of totally randomized and fixed regions, and subsequently carrying out in vitro transcription, they reported the construction of a randomized RNA library with the complexity of 10^{15} molecules, each having a distinct sequence. From this RNA pool, the researchers carried out rounds of selection and amplification (Figure 7) against a group of small chemical ligands and were able to pull out RNA sequences that apparently folded into complex ligand-binding structures. The frequency of obtaining specific RNA ligands was reported to be roughly one in 10^{10} random RNA sequences (Ellington and Szostak, 1990).

Simultaneously, Tuerk and Gold (1990) reported the use of an approach called SELEX (Systematic Evolution of Ligand by EXponential enrichment) which essentially shares the same procedures as the one illustrated in Figure 7. In this study, they used a much smaller RNA pool, with a calculated size of 65,536 species, constructed by randomizing an eight-base region of an RNA that interacts with the bacteriophage T4 DNA polymerase. From this pool, two different sequences were obtained. One was the wild-type sequence found in the bacteriophage mRNA to which the T4 DNA polymerase can bind, and the other differed from the wild-type at four positions. The binding constants of these two RNAs to the enzyme were equivalent (Tuerk and Gold, 1990). The authors concluded that their approach could be applied with other target proteins, not necessarily having nucleic acid-binding activities.

Since these two studies, RNA aptamers have been raised against a variety of other target molecules, including HIV Rev protein (Bartel, 1991), D-tryptophan (Famulok and Szostak, 1992b), bacteriophage R17 coat protein (Schneider et al., 1992), antibodies (Tsai et al., 1992; Doudna et al., 1995), HIV reverse transcriptase (Tuerk et al., 1992), arginine (Connell et al., 1993), fibroblast growth factor (Jellinek et al., 1993), ATP (Sassanfar and Szostak, 1993), theophylline (Jenison et al., 1994), protein kinase C (Conrad et al., 1994), cyanocobalamin (Lorsch and Szostak, 1994), and substance P (a peptide) (Nieuwlandt et al., 1995). In another study, RNA sequences were even obtained with dual specificity for both arginine and guanosine (Connell and Yarus, 1994). In addition to the demonstration of the biophysical diversity of RNA interactions, these studies have also provided clues to the role of the so-called "RNA world" in the evolution of life (Szostak, 1992). Moreover, the therapeutic potentials of RNA aptamers [e.g., as antiviral agents (HIV reverse transcriptase- or Rev-binding RNA aptamers), or as anticancer agents (growth factor-binding RNA aptamers)] have been discussed in many of these studies.

During this overwhelming "RNA rush", selection of DNA aptamers from randomized single-stranded DNA libraries has also shown some promising signs. For example, using selection approaches similar to those applied for RNA aptamer selections (Figure 7), DNA aptamers were obtained against chemical dyes (Ellington and Szostak, 1992) and ATP (Huizenga and Szostak, 1995). It is interesting to point out that both types of molecules had been previously used as the targets in RNA selections (Ellington and Szostak, 1990; Sassanfar and Szostak, 1993). Sequence analysis revealed little similarity between the RNA and DNA aptamers (Ellington and Szostak, 1992). Furthermore, RNA molecules with the same sequences as the selected DNA aptamers did not bind to the targets. The binding difference between identical RNA and DNA aptamer sequences is attributable to the presence of a ribose 2' hydroxyl group in RNA (or the absence

of the hydroxyl group in DNA), but not to the presence of methyl groups at the 5-position of thymidine in DNA (or the absence of the methyl groups in RNA) (Ellington and Szostak, 1992). The potential of DNA as a binding ligand is also demonstrated by a study in which DNA aptamers as short as 15 nucleotides in length with fairly tight affinities (the K_d values were around 200 nM) for human thrombin were obtained (Bock et al., 1992). These DNA aptamers can form compact structures (Macaya et al., 1993; Padmanabhan et al., 1993; Wang et al., 1993) and inhibit the function of human thrombin (Bock et al., 1992).

These results suggest that both single-stranded RNA and DNA molecules can fold, apparently equally well, into complex structures that have novel ligand-binding activities. There are, however, several advantages for DNA selection over RNA selection. Most importantly, DNAs are more stable and easier to handle than RNAs, and selection for DNA aptamers is, therefore, experimentally simpler (Figure 7).

3. Oligonucleotides as potential therapeutic agents

Oligonucleotides that have pharmacological and therapeutic potentials can be divided into three classes: DNA and RNA aptamers, antisense oligonucleotides, and triple-helix forming oligonucleotides. Antisense oligonucleotides bind to an mRNA ("sense") sequence by Watson-Crick hybridization. They serve to block RNA processing or translation, and thus inhibit the synthesis of a disease-mediating protein (for reviews see Uhlmann and Peyman, 1990; Eguchi et al., 1991; Agrawal, 1992). Triple-helix forming oligonucleotide sequences consist of mostly thymidines and bind to double-stranded purine-rich target sequences through Hoogsteen bonding (recently reviewed by Chubb and Hogan, 1992). Without discussing in length the development of these oligonucleotides and the mechanisms by which they act, in this section of my dissertation I shall briefly

describe the progress that has been made in increasing the feasibility of using these oligonucleotides as therapeutic agents.

There are two basic problems preventing oligonucleotide-based agents from becoming therapeutically effective. One is stability, and the other is delivery. Stability of oligonucleotides has always been a concern because of the presence of cellular exonucleases and endonucleases that can rapidly hydrolyze unmodified oligonucleotides. Since the degradation of oligonucleotides mainly comes from the action of 3'-exonucleases (Tidd and Warenius, 1989), 3'-end modifications with various capping groups have been made to overcome the degradation problem (Uhlmann and Peyman, 1990). Other ways of slowing down the degradation process that have been shown to be successful include modifying the phosphodiester backbone with methylphosphonate, phosphorothioate and phosphoramidate analogs, and changing the sugar moiety of oligonucleotides from a β-anomeric configuration to an α-anomeric configuration (Uhlmann and Peyman, 1990; Chubb and Hogan, 1992). The cellular uptake of oligonucleotides was also a cause for concern early in the development of antisense and triple helix technologies. Intuitively, large polyanionic, hydrophilic compounds with molecular weights between 5,000 and 15,000 would not be expected to cross lipid membranes and enter cells (Chubb and Hogan, 1992). A number of studies has demonstrated, however, that oligonucleotides are quickly and actively transported into cells probably by endocytosis, and rapidly diffused into the nucleus once inside cells (Uhlmann and Peyman, 1990; Chubb and Hogan, 1992). Methods designed to improve the cellular uptake of oligonucleotides include the use of liposomes, the coupling oligonucleotides to positively charged molecules such as poly-L-lysines, the application of specific carriers such as a protein that has its receptor in the cell membrane or an antibody that has specificity for a surface antigen of tumor cells (Uhlmann and Peyman, 1990), and the use of a viral vector such as an adeno-associated virus (AAV) (Chatterjee et al., 1992).

Of course, for oligonucleotides to be considered as potential therapeutic agents, they must also possess favorable pharmacokinetics and toxicological properties.

4. In vitro screening of combinatorial small-molecule libraries

Other than using libraries of biopolymer molecules (e.g., oligonucleotides and peptides) as the sources for in vitro selections, great progress has also been made in the development of potential therapeutic agents using libraries of small chemical compounds (recently reviewed by Ecker and Crooke, 1995). In this section of my dissertation, I will use the approach utilized by Carell et al. (1994a,b; 1995) as an example to demonstrate the effectiveness and potential applications of screening small chemical libraries, partially because similar experiments are being carried out in our own laboratory.

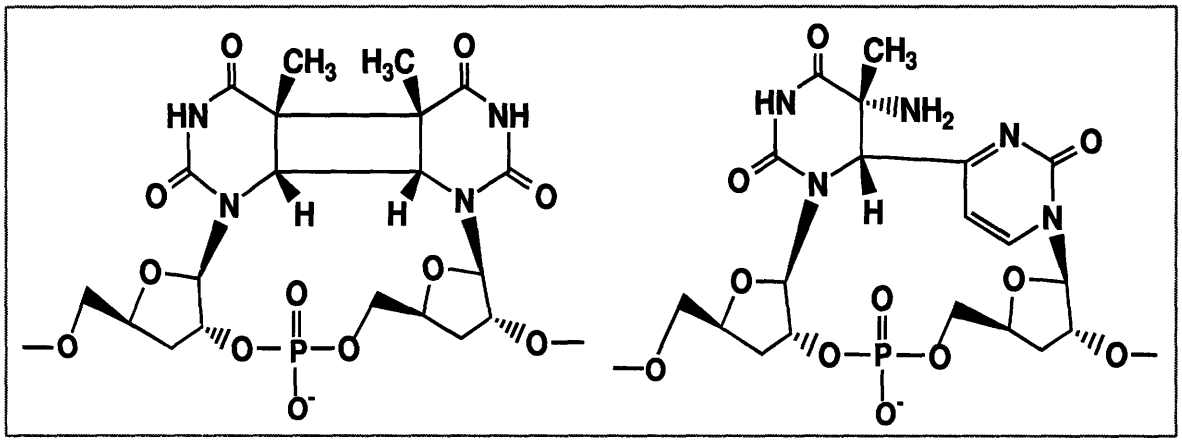
Carell and co-workers have used molecular scaffolds based on cubane, xanthene and benzene containing multiple acid chloride functionalities (Figure 8A) to create libraries of conformationally restricted structures (Carell et al., 1995). More specifically, the highly symmetric and compact cubane derivative (1 in Figure 8A) displays its four acid chloride groups in a tetrahedral array, while the larger xanthene core molecule (2 in Figure 8A) presents a less symmetric arrangement and is planar in overall shape. The benzene triacid chloride core molecule (3 in Figure 8A) offers a different orientation of its functional groups from the xanthene or cubane scaffolds, being less rigid than either. To prepare molecular libraries from these core molecules, the authors mixed, as schematically illustrated in Figure 8B, one equivalent of each acid-chloride core compound with three or equivalents, depending on the number of acid-chloride groups on the core, of an amine mixture. The amine mixtures were prepared by mixing equimolarly four to twenty L-amino acid-derived building blocks.

By applying a screening scheme first employed by Houghten and coworkers (Houghten et al., 1991; Pinilla et al., 1992), Carell et al. (1994a,b; 1995) used the protease trypsin as a model molecule to select for its inhibitors from the molecular libraries. The screening strategy involved rounds of synthesizing compound libraries and assaying these libraries for their ability to inhibit trypsin activity. If, for example, the number of building blocks for the library construction is twenty, the first step of round one will be to synthesize six sublibraries with each resulted from condensing 15 of the 18 building blocks with the core molecule (i.e., in each sublibrary three building blocks have been omitted). Activity assays are then performed with each sublibrary and the relative potential importance of the omitted building blocks to the inhibition activity is evaluated. An exhibition of a good inhibition activity by a sublibrary means that the omitted building blocks could be of less significance; an exhibition of a weak inhibition, on the other hand, means that these omitted building blocks may be important and should be subjected to further investigations. Subsequent rounds of selection are performed in the same way as the first round until a minimal number of building blocks is achieved. The final compound that has potent inhibition activity can be obtained by synthesizing all the possible compounds individually with the remaining building blocks and testing these compounds for their inhibition activities. Using this approach, Carell et al. (1995) generated a single xanthene-derived compound giving micromolar inhibition of trypsin. A broad range of applications for this approach was suggested by the same researchers.

Small chemical-based compounds displaying desired activities could be advantageous over molecules such as oligonucleotide-based aptamers as potential therapeutic agents. The lack of requirements for the use of enzymes (e.g., RNA and DNA polymerases) during the syntheses of these small chemicals allows the aforementioned problems of stability and delivery to be overcome by inducing chemical groups of favorable nature in the synthesis procedures. In addition, even

though the cost of synthesizing oligonucleotides has be reducing significantly (Chubb and Hogan, 1992), the small chemicals are much cheaper to produce. Although a selection for DNA aptamers has been carried out in this dissertation, the possibility of using small chemical libraries to achieve the same goal has been discussed (Chapter VI).

Figure 2. Structure and properties of the major UV-induced photoproducts *



	TT *cis-syn* Cyclobutane Pyrimidine Dimer (TT Dimer)	TC Pyrimidine (6-4) Pyrimidone Photoproduct (6-4 Photoproduct)
Cytotoxicity and Mutagenicity	+	+
Abundance	+++	+
Repair in Human Cells	+	+++
Repair by *E. coli* Photolyase	+	–
Alkali Lability	–	+
T4 Endonuclease V Sensitivity	+	–

* *Adapted from Treiber (1993).*

Figure 3. The DNA repair blocking model: a possible role for the *cis*-DDP-DRPs in *cis*-DDP anticancer activity

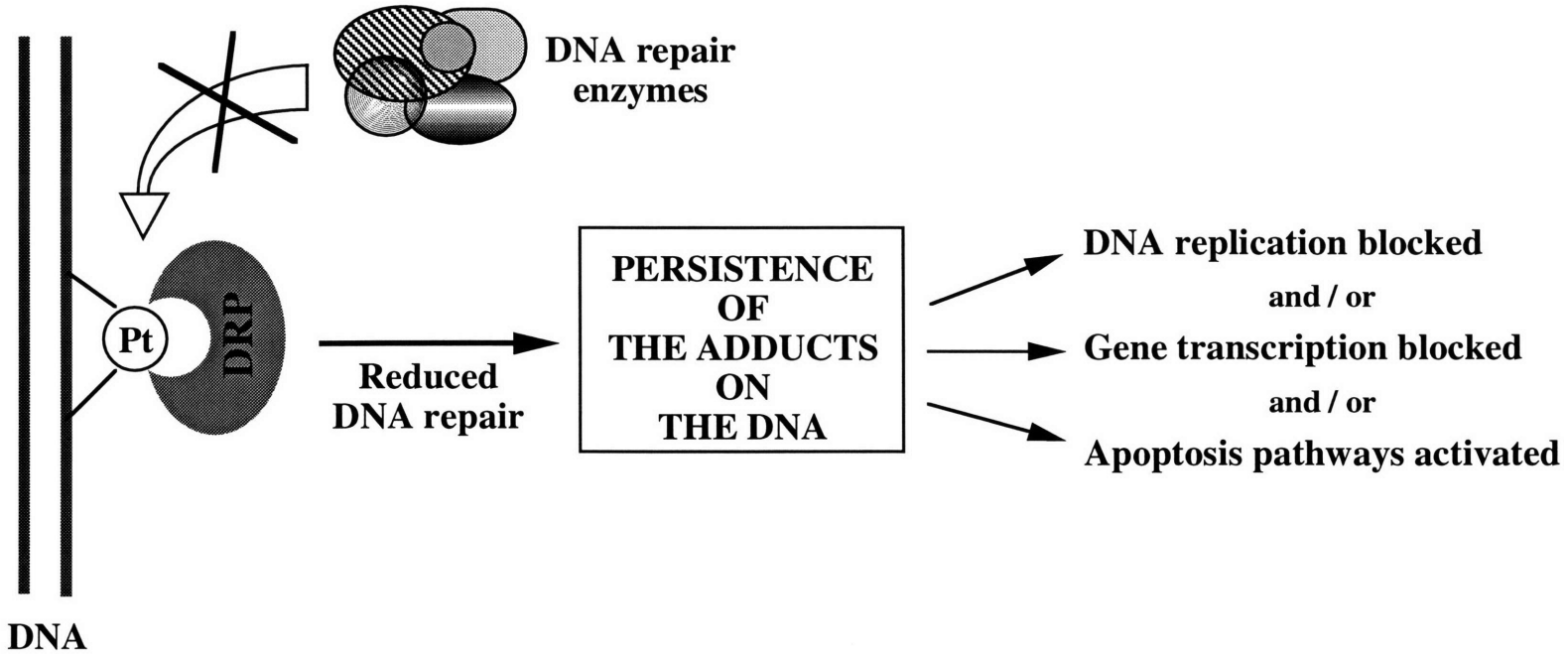
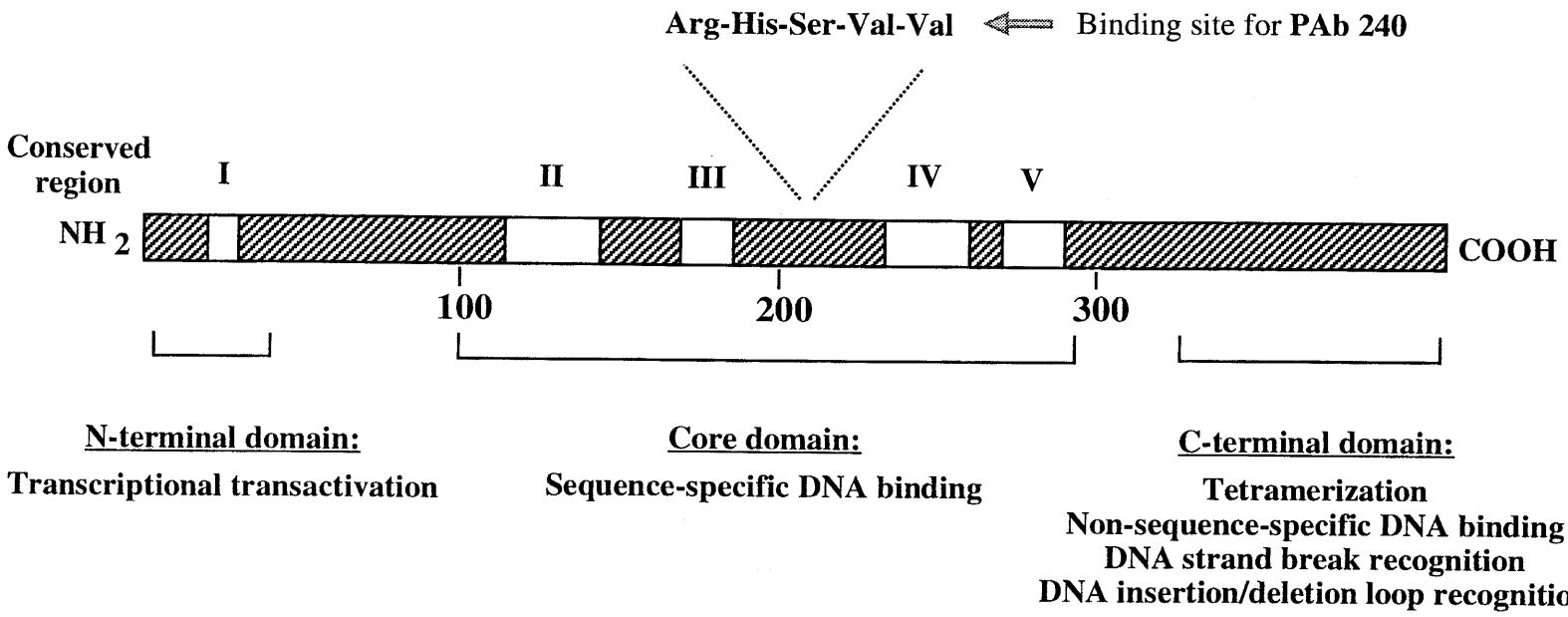
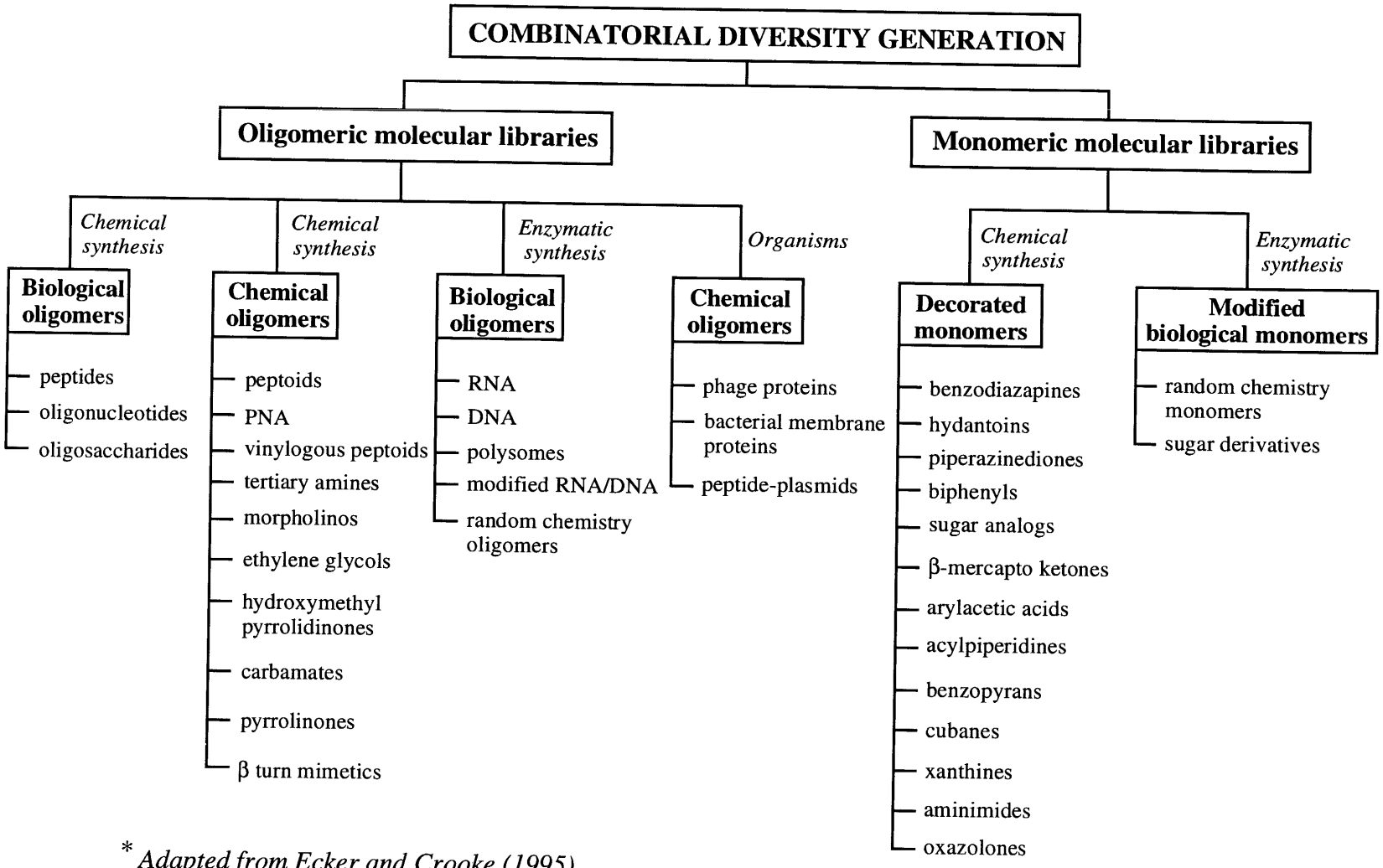


Figure 4. The human p53 protein



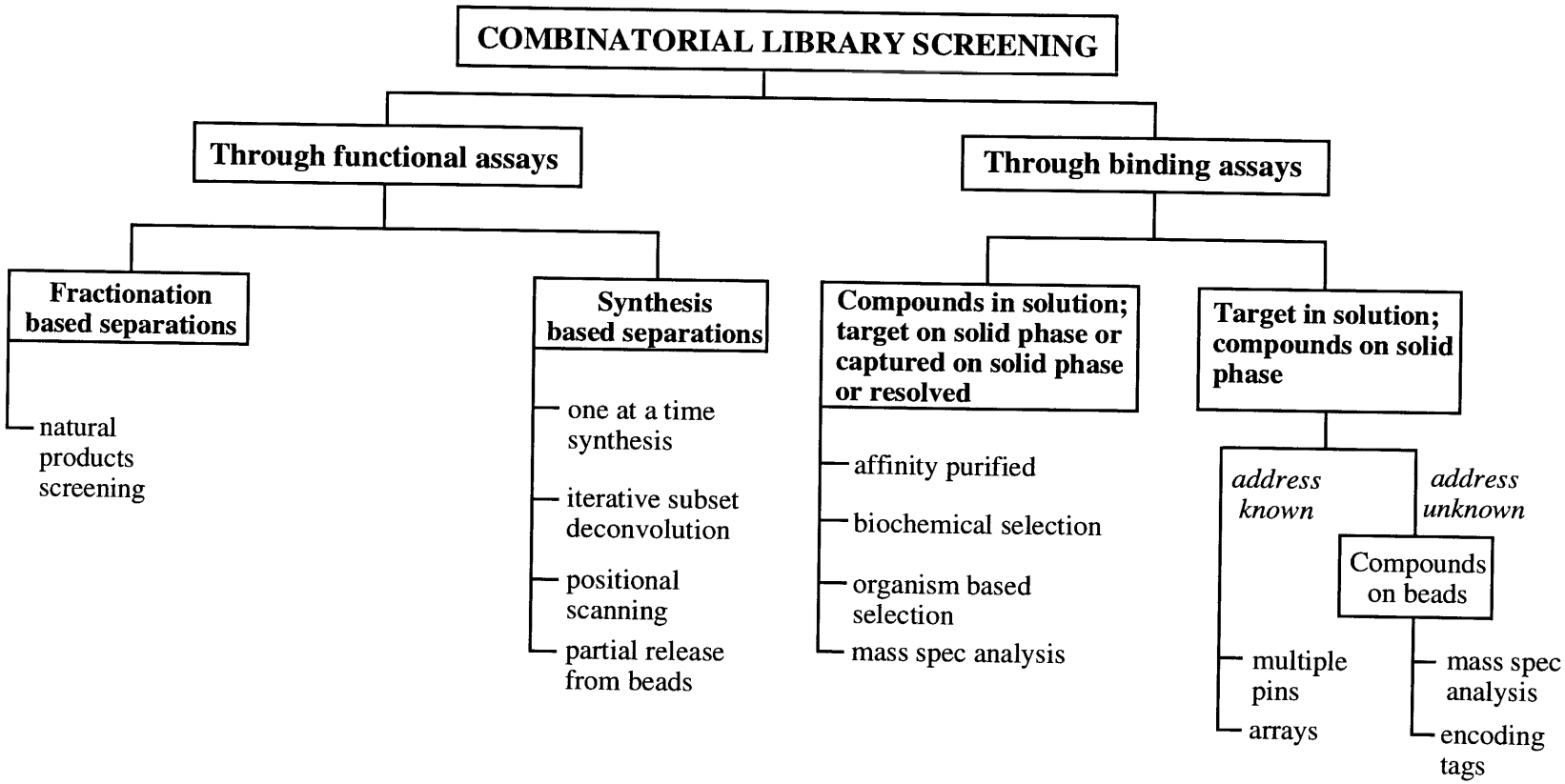
Mutant p53 proteins: Mostly found only in tumor cells
Localized in the nucleus
Elevated concentrations

Figure 5. Combinatorial diversity generation *



* *Adapted from Ecker and Crooke (1995).*

Figure 6. Methods for screening combinatorial libraries *



* *Adapted from Ecker and Crooke (1995).*

Figure 7. In vitro selection of DNA/RNA aptamers

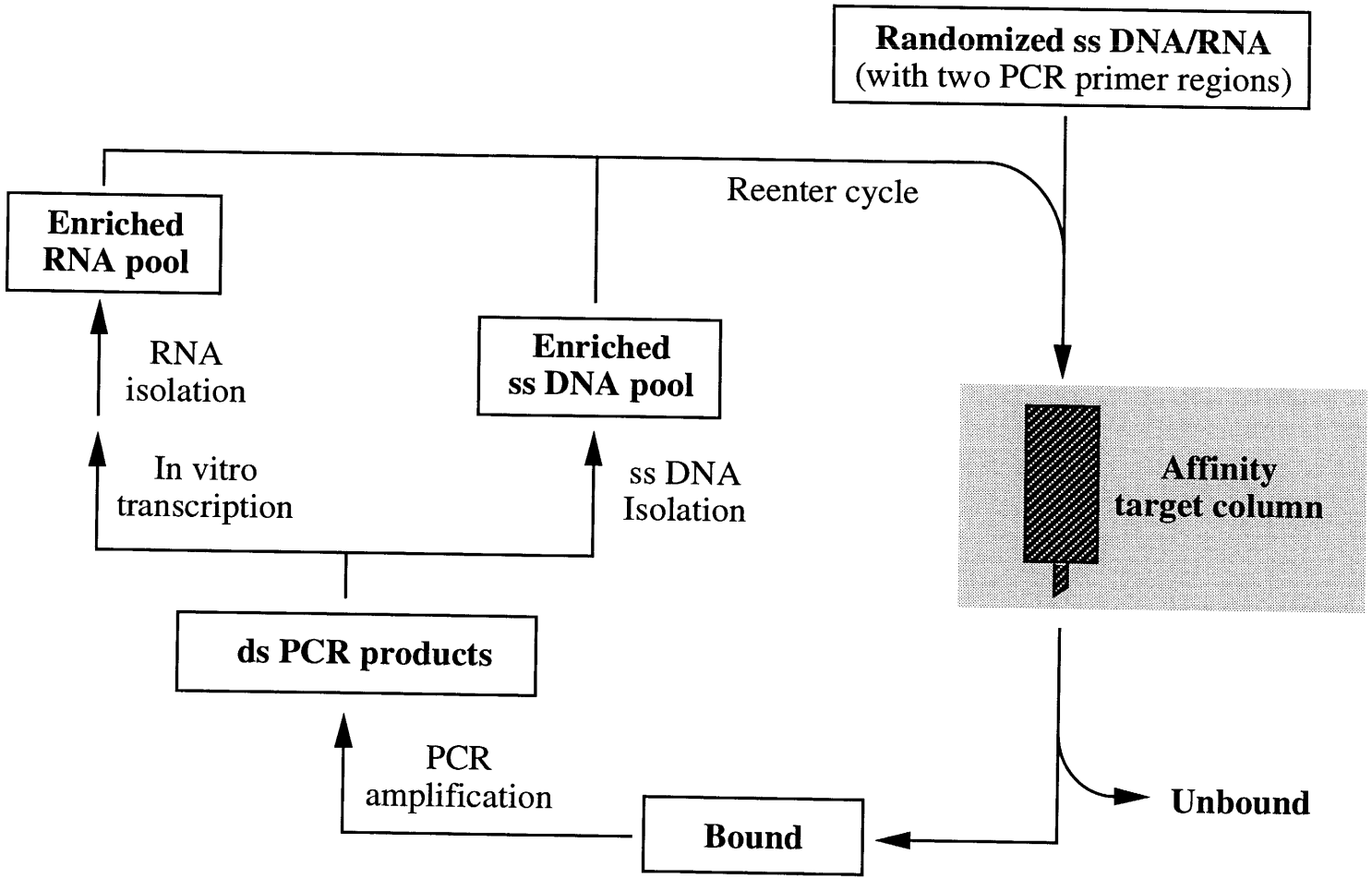
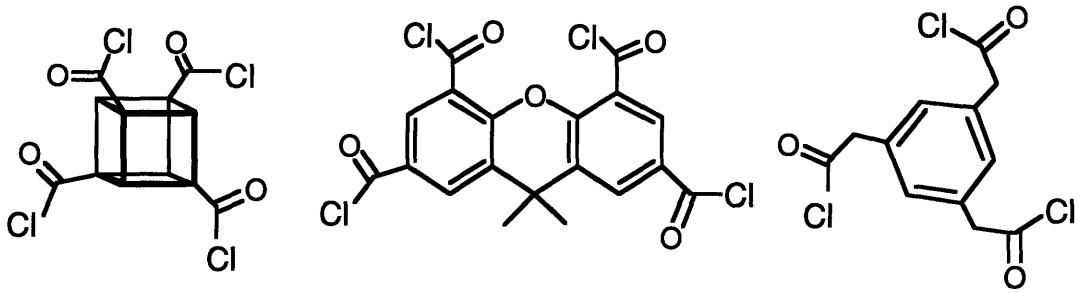
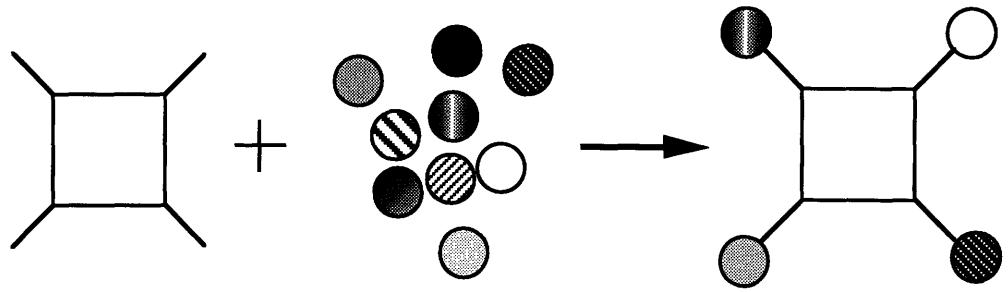


Figure 8A. Core molecules that can be used to generate diverse molecular libraries *



1 2 3

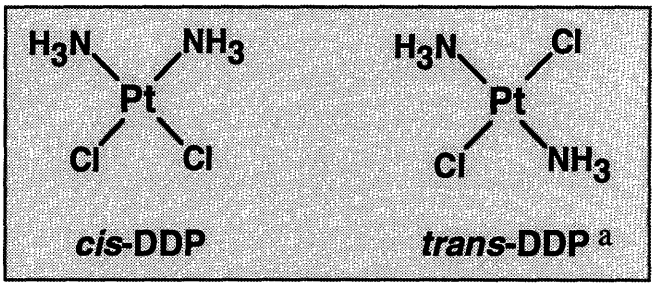
Figure 8B. A schematic demonstration of the generation of a molecular library *



Core-molecule Building blocks Molecular library

* *Adapted from Carell et al. (1995).*

Table 1. DNA adducts formed by *trans*- and *cis*-DDP



DNA monofunctional adducts		
dG	3 [b]	yes
DNA intrastrand crosslinks		
d(GpG)	65	no
d(ApG)	25	no
d(GpXpG)	∠8	yes
DNA interstrand crosslinks		
	∠1	yes
DNA-protein crosslinks		
	∠1	yes

---

[a] *The DNA adducts formed by trans-DDP have not been quantitated.*

[b] *Numbers indicate percentage of each adduct as compared to total platinum bound. Cases in which reliable quantitative data are not available are indicated as "yes" and "no."*

III. AN ULTRAVIOLET LIGHT-INDUCED DNA DAMAGE RECOGNITION PROTEIN ABSENT IN XERODERMA PIGMENTOSUM GROUP E CELLS BINDS SELECTIVELY TO PYRIMIDINE (6-4) PYRIMIDONE PHOTOPRODUCTS

A. INTRODUCTION

As indicated in the previous section, a protein exists in mammalian cells that binds specifically to ultraviolet (UV) light-irradiated DNA. The absence of this UV-damaged DNA recognition protein (UV-DRP) in cells from genetic complementary group E patients of xeroderma pigmentosum (XP) (Chu and Chang, 1988) suggests a link between the protein and DNA repair. As also discussed in the previous section, UV irradiation of DNA can give rise to several forms of DNA lesions, including CPDs and 6-4 photoproducts. Each of these lesions causes distortions in DNA architecture (Taylor et al., 1990; Smith and Taylor, 1993) and is believed to be both cytotoxic and mutagenic (Brash and Haseltine, 1982; Glickman et al., 1986; Protić-Sabljić et al., 1991). Interestingly, the relative repair rates of these DNA damages in mammalian cells (Mitchell et a., 1985, 1990; Thomas et al., 1989) and cellular extracts (Wood, 1989) have been shown to be biased by a factor of 4–12 in favor of repair of the 6-4 photoproducts. One plausible mechanism for this differential repair is that the UV-DRP protein recognizes the 6-4 photoproducts with the highest affinity as compared to the other photolesions allowing the fastest removal of the 6-4 photoproducts from DNA. In this portion of the thesis, I have reported results of experiments I carried out, in collaboration with Daniel Treiber, in an attempt to test the above hypothesis.

I had primary responsibility for the experiments described below involving the preparation of the HeLa cell nuclear extracts, the preparation of UV-treated synthetic oligonucleotide probes, the characterization of DNA lesions after UV-irradiation of the synthetic oligonucleotides, analytical and preparative electrophoretic mobility shift assays, and the 6-4 photoproduct assay. Dr. Treiber's contribution to the work involved the preparation of the XP-E cell nuclear extracts, the synthesis of the oligonucleotide probes, and the cyclobutane pyrimidine dimer assay. I have gone into particular depth in describing the experiments conducted

by myself.

B. MATERIALS AND METHODS

1. Materials

Polynucleotide kinase was obtained from New England Biolabs, poly(dI-dC)·poly(dI-dC) was from Pharmacia and dithiothreitol (DTT) was from GIBCO-BRL. T4 endonuclease V (T4 endo V) was kindly provided by R.S. Lloyd (Vanderbilt University) and HeLa cells were obtained from M. Chow and P. Sharp (Massachusetts Institute of Technology). A human lymphoid line [XP-E (GM02450D)] and a fibroblast cell line [XP-E (GM2415B)] were obtained from Human Genetic Mutant Cell Repository. Media and fetal calf serum were purchased from GIBCO-BRL.

2. Preparation of nuclear extracts

HeLa cells were grown in suspension in S-MEM supplemented with 5% fetal calf serum. The lymphoid XP-E cell line (GM02450D) was grown in suspension in RPMI 1640 medium containing 15% heat-inactivated fetal calf serum and 2 mM L-glutamine. The fibroblast XP-E cell line (GM2415B) was grown in minimal essential medium (MEM) supplemented with 20% fetal calf serum and 2 mM L-glutamine. Nuclear extracts were prepared according to a published procedure (Dignam et al., 1983). Protein concentrations were determined by the Bradford assay (Bio-Rad).

3. Preparation of UV-treated synthetic oligonucleotide probes

Sixty nine-mer oligonucleotides T_4C, T_5 and Δ (*vide infra*) were synthesized

on an Applied Biosystems 381-A DNA synthesizer. The full-length oligomers were isolated on 10% denaturing (7 M urea) polyacrylamide gels (acrylamide/bis, 19:1) and electroeluted by using an Amicon centrilutor. Urea was removed from the DNA samples by two distilled water washes in Amicon Centricon 10 microconcentrators. The purified oligonucleotides were end labeled with [γ - ^{32}P]ATP (6000 Ci/mmol, New England Nuclear) by using T4 polynucleotide kinase. Unincorporated label was removed by centrifugation through pre-packed G-25 spun columns (Boeringer Mannheim). The labeled strands were annealed to their unlabeled complementary strands, and the resulting duplexes were purified by native polynucleotide gel electrophoresis (PAGE) to remove any contaminating single-stranded DNA. Labeled probes were irradiated on a chilled surface with a 15-W germicidal lamp (maximum output at 254 nm) at a fluence of 5 J/m^{2}s until the desired dose was achieved (fluence measured with a UVX-25 sensor).

4. Analytical and preparative electrophoretic mobility shift assays with the UV-treated DNA probes

Binding between the UV-DRP and the UV-treated DNA probes was detected by using gel mobility shift assays (Carthew et al., 1985). For analytical gel mobility shift assays, end-labeled probes (7×10^4 cpm per incubation) either unmodified or UV-treated were incubated in the presence of crude nuclear extracts (10 μg) and poly (dI-dC) (6 μg) for 15 min at 37°C in binding buffer [2 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) pH 7.5, 10 mM NaCl, 50 mM ethlenediaminetetraacetic acid (EDTA) pH 8.0, 1% glycerol and 0.2 mM DTT] in a final volume of 15–25 μl. Protein-DNA complexes were then resolved on 6% native polyacrylamide gels (acrylamide/bis, 29:1). Gels were electrophoresed at room temperature in Tris-glycine buffer (50 mM Tris-HCl pH 8.5, 380 mM glycine and 2 mM EDTA pH 8.0) for 4 hr at 30 mA. Gels were dried and autoradiographed. In some cases, gels were also imaged and quantified

by using a Molecular Dynamics PhosphorImager (Sunnyvale, CA).

By scaling up the protein-DNA binding reaction by one order of magnitude, preparative electrophoretic mobility shift assays were performed to purify UV-DRP-bound T_4C probes. Electrophoresis was carried out for 14 hr to sufficiently separate the UV-DRP-specific band from a nonspecific band. The UV-DRP-specific band and the non-protein bound band were excised from the gel and electroeluted. Gel-purified bound and free probes were deproteinized by a phenol:chloroform:isoamyl alcohol (29:29:1) extraction followed by a chloroform extraction. The DNA samples were then ethanol precipitated and resuspended in TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA).

5. 6-4 Photoproduct and cyclobutane pyrimidine dimer assays

6-4 Photoproducts were detected as alkali-labile sites in irradiated DNA (Franklin et al., 1982; Mitchell et al., 1990). Freshly prepared 1 M piperidine (Fisher) was added to end-labeled DNA samples, and the mixtures were then heated at 90°C for 30 min and, in some cases the incubation time was extended to several hr. Piperidine was removed by vacuum centrifugation and samples were washed by resuspension in distilled water followed by vacuum centrifugation. Washed samples were resuspended in denaturing loading buffer [80% (v/v) recrystallized formamide, 0.1% (w/v) xylene cyanol and 0.1% (w/v) bromophenol blue] and subjected to denaturing PAGE to quantify and to map 6-4 photoproducts. Chemical sequencing reactions (Banaszuk et al., 1983) were used as calibration standards.

CPDs were detected as T4 endo V-sensitive sites in DNA. End-labeled DNA fragments were treated with T4 endo V (1 μl of a 0.02 mg/ml solution of enzyme in TE and 100 μg/ml BSA) in 10–20 μl of 10 mM Tris-HCl pH 8.0, 50

mM EDTA, 50 mM NaCl and 1 mg/ml BSA at 37°C for 1 hr. Digestion products were analyzed by 10% denaturing polyacrylamide gels. The T4 endo V reaction was shown to proceed to completion by 1 hr.

C. RESULTS

1. The UV-damaged DNA recognition protein (UV-DRP) under study was the XPE factor

It has been reported that a protein was present in human nuclear extracts that retarded the electrophoretic mobility of UV-damaged DNA fragments (Chu and Chang, 1988; Hirschfeld et al., 1990), and this protein was missing in cells from two XP patients belonging to complementary group E (Chu and Chang, 1988). Using the gel mobility shift assay, both Dr. Daniel Treiber, then a graduate student in the Essigmann group, and myself have also detected a protein activity in human nuclear extracts that was absent in extracts from the two XP-E cell lines, indicating that the protein we were studying was identical to the one originally identified by Chu and Chang (1988) [Figure 9; results with extracts from HeLa and one of the XP-E cell lines, XP-E (GM02450D), are shown].

2. Characterization of DNA lesions after UV-irradiation of the synthetic oligonucleotides

From the observation that UV-treated DNA depleted of CPDs by enzymatic photoreactivation was recognized with undiminished affinity by the UV-DRP (Hirschfeld et al., 1990; D. Treiber, personal communication), it has become clear that CPDs, the most abundant photoproducts, are <u>not</u> high-affinity binding sites for the UV-DRP. Although it has been hypothesized that 6-4 photoproducts might be the true physiological target for the protein, the importance of the 6-4 lesions in

UV-DRP binding had not been directly and definitively addressed before our work.

In order to evaluate the relative importance of CPDs, 6-4 photoproducts, and non-dimer damage in UV-DRP binding, three nearly identical 69-mer oligonucleotides T_4C, T_5 and Δ were synthesized (Figure 10). Sequences were designed such that adjacent pyrimidines are found exclusively in the 5-bp 'UV boxes' that reside near the center of the oligomers. GC clamps on both ends of the oligonucleotides increased the annealing efficiency but did not alter the outcome of experiments performed with the sequences (our unpublished observations). Since 6-4 photoproducts form rarely at TT sequences (Mitchell et al., 1990; Lippke et al., 1981), irradiation of the T_5 probe, which has a run of Ts, was expected to result the formation of mainly CPDs. The T_4C sequence differs from T_5 by a single T to C transition, which provides a TC dinucleotide following a run of pyrimidines. Such sequences can serve as hotspots for the formation of 6-4 photoproducts (Brash and Haseltine, 1982; Lippke et al., 1981). Therefore, T_4C should yield a much higher frequency of 6-4 photoproducts than T_5. The oligomer Δ contains no adjacent pyrimidines and therefore can not form pyrimidine dimers.

The predicted distribution of photoproducts in the three probes after UV irradiation was confirmed by assaying the UV-treated probes for 6-4 photoproducts and cyclobutane dimers (see this chapter's Materials and methods). The results are shown in Figure 10. T4 endo V analysis of UV-treated T_5 showed that it formed a high frequency of CPDs mapping primarily to the 3' TT sites in the UV box (lane 7 in Figure 10). Conversely, 6-4 photoproducts were not detected by piperidine cleavage in the T_5 probe (compare lanes 9 and 10). The T_4C sequence formed both CPDs and 6-4 photoproducts (lanes 13 and 15, respectively). CPDs were most abundant in the central TT sites, but were also evident in the flanking TT and TC dipyrimidines. The exclusive hotspot for the formation of 6-4 photoproducts was the TC site as evidenced by a UV specific piperidine cleavage product (lane

15) migrating one nucleotide above the corresponding C in the chemical sequencing lanes (lane 17). This banding pattern was expected because piperidine cleavage of TC 6-4 photoproducts present on a 5' end-labeled DNA fragment has been shown to yield a product that has a reduced mobility relative to sequencing standards (Brash et al., 1985). The irradiated oligomer contained, as expected, neither CPDs nor 6-4 photoproducts (lanes 1–4).

3. UV-DRP preferentially binds to UV-treated T_4C

To compare binding of the UV-DRP to T_4C, T_5 and Δ, a UV dose-response gel mobility shift assay was carried out, and the results are shown in Figure 11. UV-treated T_5 and Δ were equally poor substrates for the UV-DRP, indicating that the protein's high-affinity binding sites are non-cyclobutane pyrimidine dimers. The conclusion that cyclobutane dimers are insignificant in UV-DRP binding is in agreement with the aforementioned observation that deletion of CPDs from the UV-treated DNA fragment did not diminish the affinity of the probe for the UV-DRP (Hirschfeld et al., 1990; D. Treiber, personal communication). On the other hand, T_4C was shown to be a superior target for the UV-DRP by a five to seven fold margin at all doses tested. On the basis of these findings we concluded that the major UV-DRP binding site is probably the 6-4 photoproduct while a unidentified non-pyrimidine dimer photoproduct(s) represents a rare or low affinity class of binding sites.

4. UV-DRP-bound T_4C is enriched for 6-4 photoproducts

To provide further evidence for that 6-4 photoproducts represent the major class of UV-DRP binding sites, experiments were carried out to determine whether 6-4 photoproducts were disproportionally represented in the material retarded in gel shift assays (UV-DRP-bound T_4C). Since only a small fraction (2–5%) of the total

T_4C used in an analytical gel shift assay became bound to the UV-DRP, it was necessary to scale up the assay by an order of magnitude to generate sufficient bound DNA for analysis. Both the UV-DRP-bound T_4C and free T_4C from a preparative gel shift assay were purified and subjected to T4 endo V or piperidine treatment. As illustrated in Figure 12, a prolonged exposure of the these DNAs to hot piperidine revealed that UV-DRP-bound material enriched substantially for 6-4 photoproducts mapped to the UV box sequence of T_4C. As much as 90% of the bound DNA contained a 6-4 photoproduct (detected as an alkali-labile site) as compared to 20% for the unbound material. The major site of 6-4 photoproduct formation was at the TC dipyrimidine site as evidenced by the electrophoretic mobility of the piperidine cleavage product relative to standards (not shown, refer to Figure 10). As also shown in Figure 12, a minor cleavage product, ~1 nucleotide smaller than the more abundant fragment, was also enriched in the bound population. This enrichment was possibly due to the binding of the UV-DRP to a rare 6-4 photoproduct occurring between the third and fourth Ts. These findings clearly demonstrate that the UV-DRP can distinguish between DNA possessing or lacking 6-4 photoproducts. Photoproduct frequencies in bound T_5 and oligonucleotides were not determined because these sequences were recognized by the UV-DRP so weakly even at a very high UV dose that it was not feasible to obtain sufficient material for analysis.

As demonstrated earlier in Figure 10, the internal TT dipyrimidines in the UV box of T_4C sequence were hotspots for CPD formation, and CPDs formed to a lesser extent at the flanking TT and TC sites. To monitor the behavior of the cyclobutane dimers in the gel shift assays, T4 endo V was used to analyze the UV-DRP-bound and free T_4C DNA obtained as aforementioned with a preparative gel shift assay. Interestingly, bound T_4C was impoverished by two fold for the CPD photoproducts relative to free T_4C as evidenced by the intensities of bands generated from T4 endo V treatments (Figure 13). In light of the concomitant

enrichment for 6-4 photoproducts, this result may be attributable to two non-mutually exclusive phenomena. First, the UV-DRP may not bind with high affinity to a UV box containing both a CPD and a 6-4 photoproduct. Second, the presence of a 6-4 photoproduct in the UV box may significantly lower the propensity of upstream Ts to form CPDs. As also shown in Figure 13, the UV-DRP-bound DNA migrated as a single band with the same mobility as the full-length control DNA in denaturing gels, indicating that the UV-DRP did not appear to promote the formation of strand breaks or apurinic sites (AP sites) in the bound T_4C. The lack of AP sites in the bound DNA is evidenced by its resistance to the AP endonuclease activity of the T4 endo V.

D. DISCUSSION

The existence of a UV-DRP was first demonstrated by Feldberg and co-workers (Feldberg and Grossman, 1976; Feldberg et al., 1982). This protein was found to be absent in some XP-E cells, suggesting its role in DNA repair (Chu and Chang, 1988). Since each XP group is deficient in the incision step of DNA excision repair (Cleaver and Kraemer, 1989), the UV-DRP may be a component of the preincision complex recognizing damaged DNA sites, a role analogous to that of the *E. coli* UvrA protein (Orren and Sancar, 1989). The involvement of the UV-DRP in DNA repair is also suggested by observations that the protein is inducible by UV or other DNA-damaging agents (Protić et al., 1989), and that it is overly expressed in cells that display resistance to the anticancer drug cisplatin (Chu and Chang, 1990). However, it has been shown that the UV-DRP activity did not bind to DNA modified cisplatin (Chu and Chang, 1990; Treiber, 1993), or by other agents (Abramic et al., 1991; Reardon et al., 1993; van Assendelft et al., 1993). Even more convincingly, results presented in this dissertation suggest that the UV-DRP does not bind to all the UV-induced DNA lesions. It binds selectively to the 6-4 photoproduct, a minor UV DNA adduct, but has little affinity

for the major adduct, the cyclobutane dimer. These observations certainly do not support a role for the UV-DRP as a general DNA damage recognition protein. In contrast, the XPA protein can also recognize DNA modified by UV (Robins et al.,, 1991) and other agents (Jones and Wood, 1993; Aboussekhra and Wood, 1994) with moderate affinities, making it a candidate for a part of the general DNA damage recognition system (Jones and Wood, 1993).

Nevertheless, the results presented here still suggest a role for the UV-DRP/XPE factor in DNA repair. More specifically, the binding specificity of this protein may be of importance to the mechanism by which 6-4 photoproducts are repaired more efficiently than CPDs in mammalian cells. It is postulated that the UV-DRP recognizes DNA regions containing 6-4 photoproducts and acts as an "attractant" to recruit repair complexes. The proposed attractant function of the protein could be mediated by protein-protein interactions, or by structural changes (e.g., bending and/or unwinding) in the DNA that would increase accessibility to repair proteins. This model is consistent with the XP-E phenotype, which is very mild as compared to the other XP groups, because it predicts a repair-enhancing role for the UV-DRP that may increase the overall rate of repair but is non-vital to the repair systems. The suggestion that the UV-DRP/XPE factor acts as an auxiliary component of the repair complexes is directly confirmed by a most recent study by Aboussekhra et al. (1995). In that study, the researchers performed in vitro mammalian DNA nucleotide excision repair with purified components and showed that while XPE protein stimulated repair, it was not essential.

E. CONCLUSIONS

In vitro experiments were carried out to determine the substrate specificity of a human ultraviolet (UV) light-damaged DNA recognition protein (UV-DRP), the activity of which is absent in the XP complementation group E cells. The

major DNA lesions induced by UV irradiation are cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4 photoproducts). A major role for 6-4 photoproducts in UV-DRP binding was suggested in studies showing that a UV-treated oligonucleotide containing a T4C UV box sequence, which efficiently forms a TC 6-4 photoproduct, was a superior substrate for the UV-DRP when compared to a similar irradiated oligomer containing a T5 sequence. The latter sequence forms CPDs much more frequently than 6-4 photoproducts. Biochemical analysis of the T4C-containing oligomer probe complexed with the UV-DRP revealed that 6-4 photoproducts were highly enriched in the protein-bound probe species as compared to the frequency of 6-4 photoproducts in the unbound oligomer species. These results support the conclusion that 6-4 photoproducts are the principal binding targets for the UV-DRP. The substrate specificity of the UV-DRP may provide a clue to the biochemical and molecular mechanism for the observed biased repair of 6-4 photoproducts over CPDs in mammalian cells.

Figure 9. Absence of the UV-DRP in XP-E cell extracts. [^{32}P] end-labeled T_4C probe, either untreated (lanes 1 and 3) or UV-irradiated (1 kJ/m^2, lanes 2 and 4), was incubated with HeLa (lanes 1 and 2) or XP-E (GMO2450D) (lanes 3 and 4) nuclear extracts and the mixtures were then subjected to native PAGE. A damage specific low mobility band (B1) is present in the HeLa sample (lane 2) but not in the XP-E sample (lane 4). A non-damage specific low mobility band (NS) is evident in both HeLa and XP-E samples (lanes 1 and 3). Also evident in all four lanes are some other minor bands that are not related to UV treatment of the probe and thus are not further pursued.

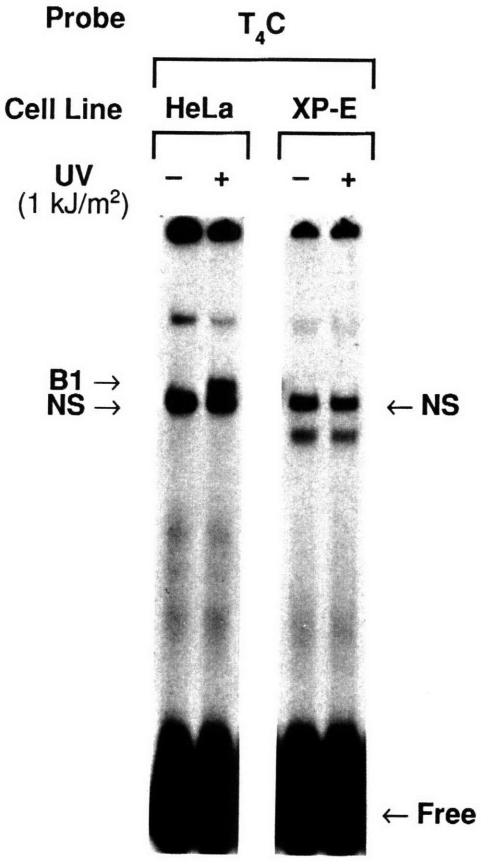

Probe
T4C
Cell Line
HeLa
XP-E
UV
(1 kJ/m2)
–
+
–
+
B1 →
NS →
← NS
← Free

Figure 10. Characterization of DNA lesions after UV-irradiation of the UV box oligonucleotides. The relevant portions of the oligomer sequences are shown to the right. Sixty nine-mer oligonucleotides T_4C, T_5 and Δ were designed such that adjacent pyrimidines are absent in both strands except in the 'UV boxes' shown in large type near the center of each sequence. Δ contains no adjacent pyrimidines as the UV box was replaced with a purine/pyrimidine stretch. Ten bp GC clamps are present on the ends of each oligomer to facilitate strand annealing. The intervening DNA sequences between the GC clamps and UV box were ATATGCGTACATGTGCG for the region upstream of (5' to) the UV box and ACGCACGCACGTACATGTGCACGTGTAT for the region downstream. The autoradiogram to the left shows the photoproduct distribution in these irradiated UV box oligomers. Control and irradiated (10^4 J/m^2) oligomer probes were treated with either T4 endo V (Endo V) for 1 hr at 37°C or with 1 M piperidine (PIP) for 0.5 hr at 90°C to determine the distribution of CPDs and 6-4 photoproducts, respectively. The resulting cleavage products were resolved on a DNA sequencing gel adjacent to the appropriate chemical sequencing markers. Lanes 1–6, Δ, lanes 7–12, T_5 and lanes 13–18, T_4C. T4 endo V analysis of irradiated probes is shown in lanes 1 (Δ), 7 (T_5) and 13 (T_4C) and a similar analysis of unirradiated probes is shown in lanes 2 (Δ), 8 (T_5) and 14 (T_4C). Analysis of piperidine-treated irradiated oligonucleotides is shown in lanes 3 (Δ), 9 (T_5) and 15 (T_4C) and in lanes 4 (Δ), 10 (T_5) and 16 (T_4C) a similar analysis of unirradiated oligonucleotides is shown. C + T chemical sequencing reactions are shown in lanes 5 (Δ), 11 (T_5) and 17 (T_4C), and A + G reactions are shown in lanes 6 (Δ), 12 (T_5) and 18 (T_4C).

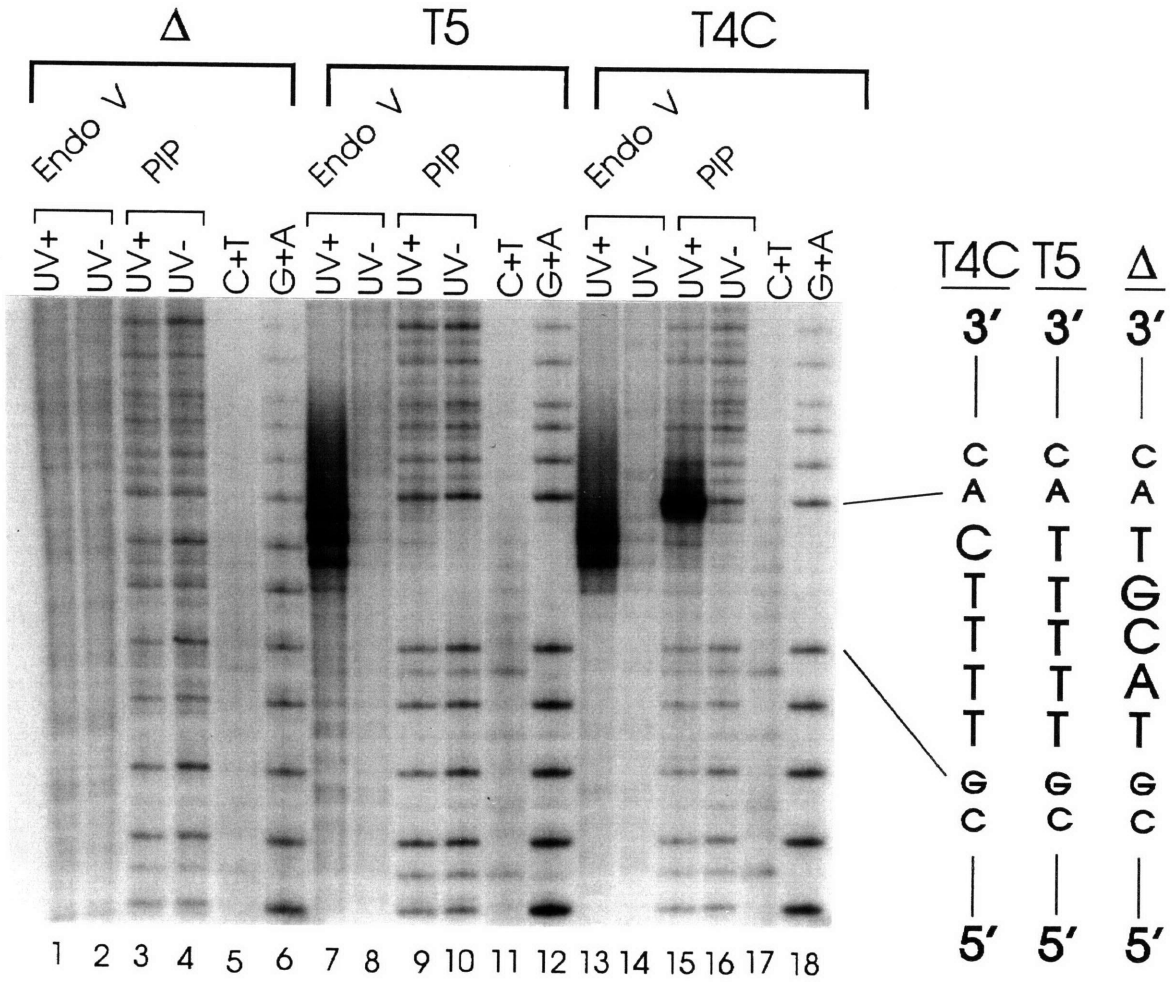

Δ
T5
T4C
Endo V
PIP
UV+
UV-
C+T
G+A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
T4C
T5
Δ
3′
5′
C A C T T T T G C
C A T T T T T G C
C A T G C A T G C

Figure 11. UV dose-response analysis of UV-DRP binding to the UV box oligonucleotides. UV box oligonucleotides T_4C, T_5 and Δ were irradiated to doses of 1–20 kJ/m^2 and incubated with 10 μg of HeLa nuclear extract. The mixtures were then analyzed by native PAGE. The amounts of probe incorporated into the damage-specific protein/DNA complexes were quantified and are reported in arbitrary PhosphorImager units (PIU).

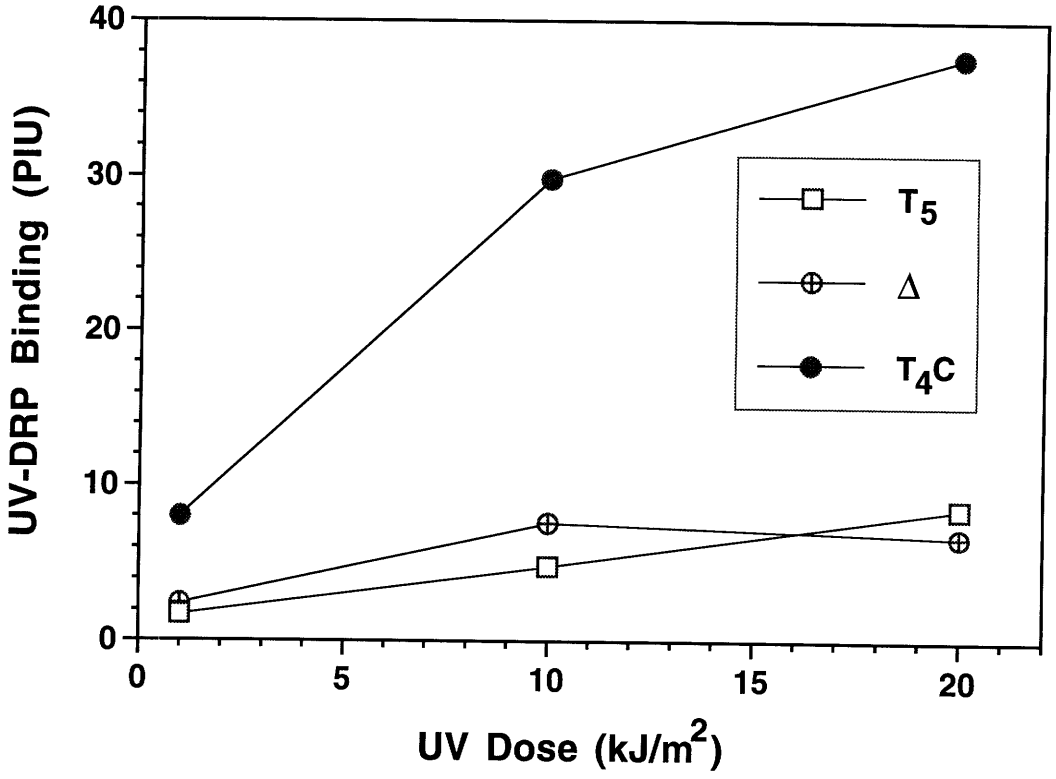

UV-DRP Binding (PIU)
40
30
20
10
0
0
5
10
15
20
UV Dose (kJ/m2)
T5
Δ
T4C

Figure 12. 6-4 photoproduct analysis of UV-DRP-bound T_4C. T_4C, either non-protein bound (Free, lanes 1–5) or complexed with the UV-DRP (Bound, lanes 6–10), was purified from a preparative electrophoretic mobility shift assay. After being deproteinized, the probes were subjected to 1 M piperidine cleavage at 90°C for 0.5–6 hr. Equal amounts of each reaction mixture were analyzed by denaturing PAGE. The two major degradation products mapped to the UV box as determined by comparison to standards (not shown, refer to Figure 10). The percentage of T_4C containing a 6-4 photoproduct in the UV box was determined by quantifying the combined radioactivity of the two major cleavage products after 6 hr and dividing by the amount of full length DNA (69 nucleotides; 69N) in the 0 hr lane.

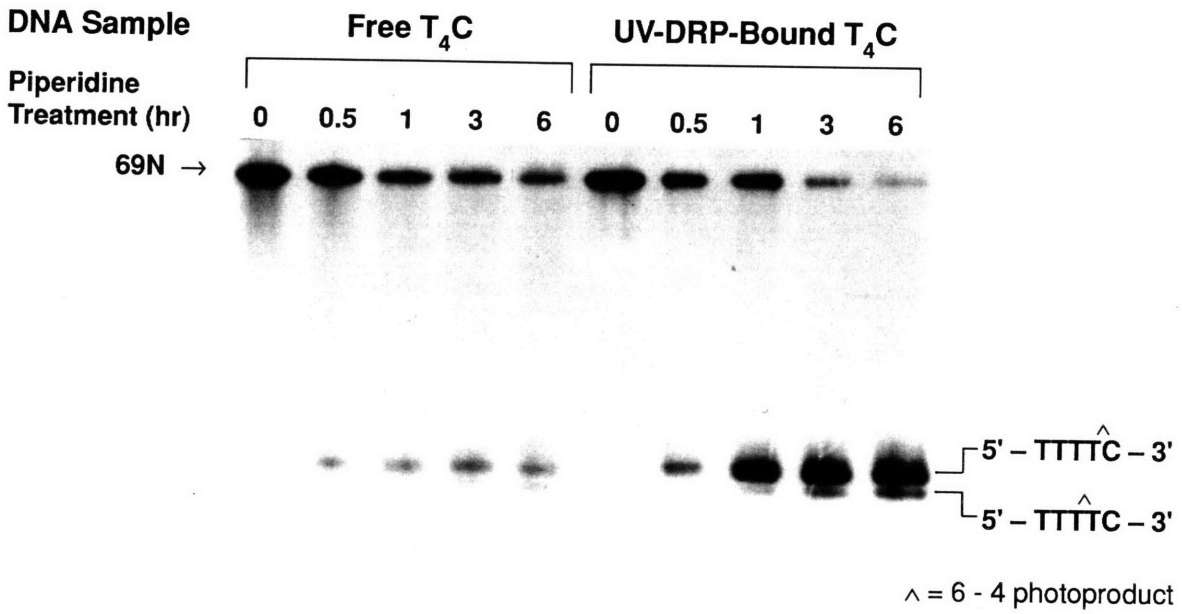
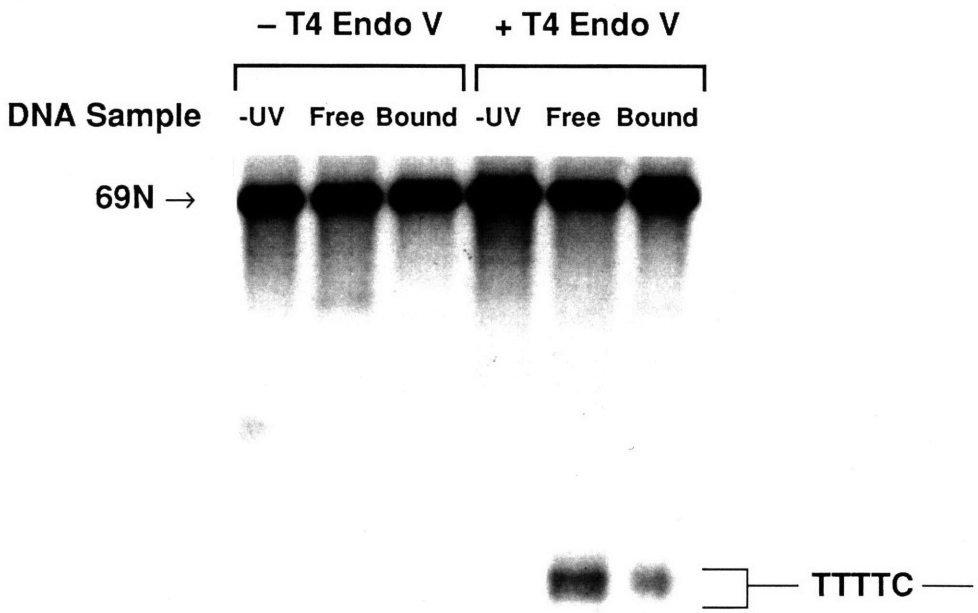

DNA Sample
Free T_4C
UV-DRP-Bound T_4C
Piperidine Treatment (hr)
0 0.5 1 3 6 0 0.5 1 3 6
69N →
5' – TTTT^C – 3'
5' – TTTT^C – 3'
^ = 6 - 4 photoproduct

Figure 13. CPD analysis of UV-DRP-bound T_4C. Free and bound T_4C were purified as described in the Materials and Methods section of this chapter. The samples were digested by T4 endo V and subjected to denaturing PAGE (lanes 5 and 6). Equal amounts of DNA were loaded in each lane. The four major cleavage products mapped to the UV box as determined by comparison to standards (not shown, refer to Figure 10). The percentage of T_4C containing a CPD in the UV box was determined by quantifying the amount of full length DNA (69N) remaining after digestion and dividing by the amount of full length DNA in the T4 endo V-treated, unirradiated control (lane 4). In lanes 1–3, the DNAs were not digested with T4 endo V to show that the cleavage products in lanes 5 and 6 were not caused by either UV light or UV-DRP binding.


– T4 Endo V
+ T4 Endo V
DNA Sample
-UV
Free
Bound
-UV
Free
Bound
69N →
TTTTC

IV. FATAL ENGINEERING I: A BIFUNCTIONAL PSORALEN-BIOTIN CONJUGATE AND ITS IMPLICATIONS FOR THE DESIGN OF ANTICANCER DRUGS

A. INTRODUCTION

Damage to the genome of cells can destroy the genetic information essential for their growth and survival. Tumor cells are particularly susceptible to this kind of attack because their continuous and unregulated growth requires vigorous DNA synthesis and gene transcription. Indeed, drugs designed to damage cellular DNA, such as derivatives of nitrogen or sulfur mustards and heavy metal compounds such as cisplatin, have been important therapeutic tools used in cancer treatments. A major drawback of these anticancer drugs, however, is that they also cause similar damage in normal non-cancerous cells, causing severe dose-limiting toxicity to the patient. A solution to this problem would be to minimize the lethal effects of DNA damage caused by the drug in normal cells while allowing the lethal effects to be fully expressed in tumor cells. To achieve this goal, we have proposed a novel bifunctional form of anticancer drugs consisting of DNA-attacking and protein-binding domains. The DNA-attacking domain (a chemical warhead) allows the drug to form lethal DNA lesions, while the protein-binding domain is designed to attract a protein that will bind to the lesions and prevent their efficient removal by cellular DNA repair systems. As a consequence, the cytotoxicity of the drug will be greatly enhanced by its persistence on DNA. If the protein-binding domain of the bifunctional drug is designed to attract a tumor specific protein (TSP) that is differentially or even exclusively expressed in tumor cells, the drug will target tumor cells with great selectivity because the aforementioned DNA repair blocking effect will only occur in tumor cells where the TSP is present. We term this type of drug design Fatal Engineering.

In this part of my thesis I have used a bifunctional conjugate, in which a trimethylpsoralen (TMP) is covalently linked to a biotin derivative, as a model molecule to help prove the concept of Fatal Engineering. Psoralen and its derivatives can intercalate into double-stranded (ds) DNA and form mono- and

diadducts (interstrand crosslinks) under near UV irradiation (Cimino et al., 1985). Biotin can bind tightly to avidin or streptavidin (Green, 1975). With this bifunctional molecule, I have established that TMP (a DNA reacting agent), when conjugated to biotin (a protein-binding ligand), retains its DNA ability to form DNA adducts, and that streptavidin can bind to DNA modified by the conjugate through the biotin domain. I have also obtained evidence indicating that binding of streptavidin to DNA-adducted biotin would hinder the removal of a nearby uracil deoxynucleotide (a DNA damage) by uracil glycosylase (a DNA repair enzyme). The results support the concept, which is used in the Fatal Engineering approach, that a DNA lesion complexed with a protein could become resistant to DNA repair and thus have its genotoxicity enhanced.

B. MATERIALS AND METHODS

1. Materials

Restriction endonucleases and T4 polynucleotide kinase were purchased from New England Biolabs. Phosphatases and DNases were purchased from Sigma. Dithiothreitol (DTT) was obtained from GIBCO-BRL. Poly(dI-dC)·poly(dI-dC) was obtained from Pharmacia.

2. Preparation of a double-stranded oligonucleotide containing a single monoadduct of TMP-biotin conjugate

A 17-mer oligonucleotide, U-17, with a sequence of 5'-d(CGGCCGTACGUGCGCCG)-3', and its complementary strand were synthesized by the MIT Biopolymers Laboratory. In addition to a single TA dinucleotide site, U17 also contains a uracil deoxynucleotide (underlined) that is located three bases away from the TA site on the 3' side. The oligomer was purified on a 20%

denaturing (7 M urea) polyacrylamide gel (acrylamide/bis, 19:1) and electroeluted by using an Amicon centrilutor. Urea was removed from the oligomer by several distilled water washes in Amicon Centricon 3 microconcentrators. Purified U-17 was end labeled with [γ-^{32}P]ATP (6000 Ci/mmol, New England Nuclear) by using T4 polynucleotide kinase (New England Biolabs). Unincorporated label was removed by centrifugation through a pre-packed G-25 spun column (Boehringer Mannheim). Labeled U-17 was then annealed to its unlabeled complementary strand in TSE buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl and 1 mM EDTA). TMP-biotin conjugate (Figure 14A), which was synthesized by M. Morningstar (1994) according to a published procedure (Saffran et al., 1988), was dissolved in 50% (v/v) acetonitrile/water, and then added to the duplex oligomer solution with the molar ratio of TMP-biotin to base pair (both targeted and nontargeted) at about 1000 : 1. After 10 min of incubation at room temperature, the mixture was placed on a chilled surface and subjected to near UV irradiation with a 15-W General Electric lamp (maximum output at 365 nm). The final irradiation dose was about 85 kJ/m^2. The irradiated mixture was separated on a 20% denaturing polyacrylamide gel. A gel slice containing the U-17 species that were monoadducted by the TMP-biotin conjugate (B3 in Figure 14C) was excised, and the modified DNA was purified by electroeluting as described above. Finally, the monoadducted U-17 was annealed to its intact and unlabeled complementary strand to form the double-stranded modified probe.

3. Gel mobility shift assay

The binding of streptavidin to U-17 monoadducted with the TMP-biotin conjugate was measured by incubating the probe with streptavidin (Pierce) in 10 µl of binding buffer (25 mM Tris-HCl pH 7.4, 100 mM NaCl and 1.5 mM $MgCl_2$) at room temperature for 10 min, and electrophoresing the mixtures on a 5% nondenaturing polyacrylamide gel (acrylamide/bis, 29:1) at 4°C. A fixed amount

of the probe (3200 cpm, ~0.1 nM) was used in each incubation with the concentrations of streptavidin varied from 0 to 50 nM. Free d-biotin (0.4 mM) was added into the incubation(s) where indicated. After electrophoresis, the gel was dried and exposed to X-ray film with an intensifying screen at −80°C. The dried gel was also analyzed with a Molecular Dynamics PhosphorImager (Sunnyvale, CA).

4. Uracil glycosylase protection assay

Double-stranded modified U-17 obtained as described earlier was used as the probe in this assay. The probe (4000 cpm, ~0.15 nM for each reaction) was incubated first with 12 μl of glycosylase buffer (30 mM Tris-HCl pH 7.4, 50 mM KCl and 5 mM $MgCl_2$) at room temperature for 10 min in the presence of streptavidin (36 ng, ~50 nM) where indicated. In some incubations, 0.4 mM free d-biotin was added. After the incubation, 3 μl (0.15 units) of diluted uracil glycosylase (Boehringer Mannheim) was added, and the mixtures were then incubated at 37°C for 5–40 min. Upon the completion of each incubation, 85 μl freshly prepared 1.25 M piperidine (Fisher) was added and the samples were subsequently heated at 90°C for 1 hr. Since an apurinic site in a DNA molecule is labile to alkali cleavage (Lindahl and Andersson, 1972), piperidine treatment as stated above would result in DNA breaks if any apurinic sites were generated from the uracil glycosylase treatment. The samples were vacuum centrifuged to remove the piperidine and washed by resuspension in distilled water and followed by vacuum centrifugation again. The washed samples were resuspended in denaturing loading buffer [80% (v/v) recrystallized formamide, 0.1% (w/v) xylene and 0.1% (w/v) bromophenol blue], and analyzed on a 20% denaturing gel. The gel (without being dried) was exposed to an X-ray film with an intensifying screen at −80°C. The frozen gel was also analyzed with the PhosphorImager instrument.

5. DNase I protection assay

A DNase I protection assay was used to see whether streptavidin would physically bind to the U-17 that was monoadducted with the TMP-biotin conjugate. Again, [^{32}P] end-labeled double-stranded U-17 modified with the TMP-biotin was used in this assay. Briefly, the probe (5000 cpm, ~0.15 nM) was first incubated with various amounts of streptavidin (0–50 nM) at room temperature for 10 min in 10 µl of binding buffer (25 mM Tris-HCl pH 7.4, 100 mM NaCl and 1.5 mM $MgCl_2$). Where indicated, 0.4 mM free d-biotin was included in one of the incubations. At the end of each incubation, 2 µl of freshly diluted DNase I (Worthington Biochemical, Freehold, NJ) was added to a final concentration of 0.4 mg/ml, and the digestions were carried out at room temperature for 2 min before being quenched by the addition of 50 µl stop solution (20 mM EDTA pH 8.0, 1% SDS and 50 µg/ml yeast total RNA). The samples were then ethanol precipitated. After being washed once with 80% chilled ethanol, the DNA pellets were air dried and resuspended in denaturing loading buffer. The resuspensions were analyzed on a 20% DNA sequencing gel. The gel was dried and exposed to an X-ray film with an intensifying screen at −80°C.

C. RESULTS

1. Reactivity of the TMP-biotin conjugate

A trimethylpsoralen-biotin (TMP-biotin) conjugate (Figure 14A) was synthesized by M. Morningstar (1994) according to a published procedure (Saffran et al., 1988). Since earlier studies with a pentanediol linking arm between the two functional groups resulted in negligible reactivity of the conjugate with DNA, presumably due to the poor water solubility of the compound, a dialkylamine-containing spacer arm was used in the above TMP-biotin molecule to increase the

polarity of the spacer and the subsequent water solubility of the compound. The DNA reactivity of the conjugate was tested by first mixing the compound with double-stranded U-17 oligonucleotide which was ^{32}P-radiolabeled at the 5'-end of the top strand (Figure 14B), then irradiating the mixture with near UV light, and finally analyzing the irradiation products on 20% denaturing polyacrylamide gels. An autoradiograph of a representative gel is shown in Figure 14C. Band B1 represents the unmodified probe. B4 represents the crosslinked DNA species as evidenced by its migration position on the denaturing polyacrylamide gel and the results of its photoreactivation with 253 nm wavelength light (Cimino et al., 1985; see below). Bands B2 and B3 indicate the migration positions of the monoadducted species. Gel mobility shift assays with streptavidin showed that almost all of the B3 DNA contained the biotin functionality, while only a small portion of B4 showed binding activity for streptavidin and B2 showed no binding activity (data not shown). Photoreactivation of B4 DNA with 253 nm wavelength light resulted in the appearance of B1, B2, but much less B3 (data not shown), indicating that the majority of B4 was the result of the formation of diadducts (crosslinks) from the monoadducted species of B2. Further analyses with denaturing polyacrylamide gels showed that B2 DNA co-migrated with a marker DNA that had been monoadducted with an unconjugated psoralen derivative, while B3 DNA migrated slightly slower than the marker. Since the TMP-biotin was purified twice by HPLC and the compound appeared to be stable over a long period of time in the 50% (v/v) acetonitrile/water used, it was speculated that B2 was result of the interaction of a trace of amount unconjugated psoralen derivative present in the compound preparation with DNA and B4 was simply the result of further photoreaction of B2 DNA molecules. This speculation was supported by the observation that the conjugated psoralen showed much less reactivity with DNA (by a factor of at least 100) as compared to either trimethylpsoralen or a derivative of this psoralen molecule (data not shown). It was also observed that the conjugated psoralen was less active in forming crosslinks than its unconjugated

derivative or trimethylpsoralen (data not shown). Nevertheless, it appeared that enough material could be obtained for further studies by excising the gel slice (B3 in Figure 14C) containing DNA molecules that had been monoadducted by the TMP-biotin.

2. Binding of Streptavidin DNA modified with TMP-biotin

To investigate whether biotin covalently linked to a ds DNA molecule through a psoralen derivative can still bind to streptavidin, gel mobility shift assays were performed. Streptavidin retarded the electrophoretic mobility of U-17 fragments monoadducted with the TMP-biotin conjugate (Figure 15A). The retardation was caused by the binding of streptavidin to the biotin inserted into the DNA because free biotin reverted the retardation, presumably by competing with the immobilized biotin for streptavidin (lane 8 in Figure 15A). Quantitation of the data by using a Molecular Dynamics PhosphorImager gave rise to a binding curve (Figure 15B). The streptavidin concentration for the half-maximum binding was about 1.5 nM suggesting that the apparent dissociation constant ($K_{d(app)}$) between streptavidin and the immobilized biotin was also about 1.5 nM. Streptavidin showed little binding activity to either intact U-17 oligomer or U-17 monoadducted with just a psoralen derivative (data not shown). It should be pointed out that the K_d value differs quite significantly from that of free biotin and streptavidin or avidin [the K_d for free biotin with streptavidin or avidin is about 10^{-6} nM (Green, 1975)]. If the biotin domain in the TMP-biotin conjugate behaves similar to free biotin, the great increase of the K_d value, or the great decrease in the binding affinity between biotin and streptavidin, was probably caused by the adduction of the TMP-biotin conjugate to DNA. To minimize the impact of DNA adduction on the affinity of a ligand to its binding target, the optimal distance between the ligand and the DNA helix is among the factors to be carefully investigated. An optimal distance between a ligand and the DNA helix should allow tight binding

between the ligand and the protein as well as the adequate protection of the adjacent DNA region by the bound protein.

3. Uracil glycosylase is inhibited by a nearby adduct-protein complex

A uracil deoxynucleotide in DNA is considered as a genetic alteration, and uracil glycosylase is the enzyme responsible for its removal from DNA (Friedberg, 1985). To determine whether a protein bound to a DNA adduct would interfere with the activity of a DNA repair enzyme on the DNA region close to the adduct, a uracil deoxynucleotide was inserted into U-17 sequence three bases away from the TA dinucleotide site where TMP-biotin can form adducts. As indicated in Figure 16, streptavidin, when complexed with the TMP-biotin DNA adduct, slowed the removal of a nearby uracil base by uracil glycosylase; the inhibition was reverted to a certain extent when free biotin was added (compare among lanes 3–6, 7–10, and 11–14; and also refer to the corresponding plots). The TMP-biotin DNA adducts were stable after being heated at 90°C for 1 hr (lane 1), but a small amount of TMP-biotin adducts were degraded when the probe was subjected to piperidine treatment as evidenced by the appearances of band (b) in lane 2 and bands (b) and (d) in lanes 3–14.

To investigate whether the inhibition of uracil removal was due to physical shielding of the base by the binding of streptavidin to the modified probe, a DNase I protection assay was applied. As shown in Figure 17, when streptavidin was added, the modified U-17 became more resistant to DNase I cleavage. Enhanced G^{17} and C^{16} bands indicated that streptavidin protected these fragments (full-length and one base less, respectively) from being further cleaved by DNase I. Since the probe was 5' end-labeled, and fragments shorter than five bases could not be recovered by the ethanol precipitation procedure that was used in the experiment, only two bases (C^5 and G^6) on the 5' side of modified thymidine were observed to

be protected by streptavidin. On the 3' side, however, the protected region was more extended. As discussed above, the protected region extended to the full-length of the probe which was ten nucleotides away from the modified base. It is plausible that a similar length on the 5' side of the probe was protected by streptavidin. These observations suggested that streptavidin, which is a protein of about 60 kDa of size (Green, 1975), protected possibly a region of twenty nucleotides flanking the TMP-biotin modified thymidine where a TMP-biotin was monoadducted. It is noteworthy in this regard that the mammalian nucleotide excision repair pathway repairs a variety of DNA lesions by first recognizing them and then excising a patch of 27- to 29-nucleotide oligomer flanking the lesion site on the damaged strand (Huang et al., 1992; Svoboda et al., 1993). It was also observed that, in the absence of streptavidin (lane 2 in Figure 17), a small region on the 3' side of the modified thymidine was resistant to DNase I cleavage probably due to steric barrier caused by the adducted TMP-biotin conjugate to the enzyme.

D. DISCUSSION

1. Intellectual foundation of Fatal Engineering

The intellectual foundation of the Fatal Engineering project was based upon one of the several still unproven models for the mechanism(s) of action of the anticancer drug, *cis*-diamminedichloroplatinum(II) (cisplatin or *cis*-DDP). Although both *cis*-DDP and its clinically ineffective geometric isomer, *trans*-diamminedichloroplatinum(II) (*trans*-DDP), can react with DNA, a group of HMG-box containing proteins were identified in mammalian cell extracts that bound to DNA modified with the cis, but not the trans, isomer (Donahue et al., 1990, 1991). Three major types of DNA adducts can be observed with *cis*-DDP modification: d(ApG), or d(GpG) 1,2-intrastrand crosslinks, and d(GpNpG) 1,3-intrastrand

crosslinks where N is dA, dC or dT (Fichtinger-Schepman et al., 1985; Eastman, 1986). The 1,2-intrastrand crosslinks are unique to *cis*-DDP modification, while the 1,3-intrastrand crosslinks are also observed with *trans*-DDP modification (Eastman and Barry, 1987). Interestingly, studies showed that DNA-HMG-box protein binding occurred in mammalian cell extracts with DNA fragments that contained the first two types of adducts, but not with those that contained the third type of adduct (Donahue et al., 1990). Moreover, early results obtained from in vitro DNA repair assays (Wood et al., 1988) showed that cisplatin d(GpG) 1,2-intrastrand crosslinks, to which the HMG-box proteins can bind, were slowly if at all repaired by mammalian cell extracts (Szymkowski et al., 1992). By contrast, d(GpNpG) 1,3-intrastrand crosslinks, to which the HMG-box proteins cannot bind, were repaired at a faster speed (R. Wood, personal communication). More recent data (Huang et al., 1994) do not, however, lead to the same conclusion, so the case for hindered repair is not as strong as it once was. That point notwithstanding, a sizeable fraction of literature indicates a possible connection between binding of these cellular proteins to the *cis*-DDP-unique DNA adducts and the chemotherapeutic effectiveness of the drug. In one hypothesis that was proposed to address this possible connection (Donahue et al., 1990, 1991), it was postulated that DNA adducts formed by cisplatin are lethal, and this lethal effect can be enhanced by those cellular proteins that can bind to the adducts and impede the removal of the these adducts by DNA repair systems. This repair blocking hypothesis gains support from a yeast strain in which the IXR1 protein, the yeast homolog of a mammalian cisplatin damage recognition protein (*cis*-DDP-DRP), was deleted (Brown et al., 1993). This mutant yeast strain was two-fold less sensitive to cisplatin than its wild-type parent and accumulated three-fold fewer platinum-DNA adducts, possibly as a consequence of more facile repair in the mutant strain that lacked the adduct-shielding protection of the IXR1 protein (Brown et al., 1993).

Although it remains uncertain that the repair blocking model is indeed the mechanism underlying the action of cisplatin anticancer activity, this intriguing hypothesis provided the intellectual rationale behind the Fatal Engineering project. The central premise of Fatal Engineering is that a protein, when recruited to bind to a DNA lesion site, would enhance the toxicity of the lesion by preventing its efficient removal by DNA repair systems. One way to enhance the binding of a protein to a site of DNA damage is to tether a protein-attractive ligand to a DNA-damaging chemical warhead to make a bifunctional toxin. The tumor selectivity of this drug can be achieved by using a ligand domain that is only attractive to a tumor specific protein that is expressed exclusively, or at higher level, in a tumor cell as compared to a normal cell.

2. Overall goal of Fatal Engineering

The goal of Fatal Engineering is to design chemotherapeutic agents that are highly tumor selective. The tumor selectivity of the proposed drug is achieved by its ability to recruit tumor specific proteins onto its own DNA adducts. Identifying a tumor specific protein and subsequent identifying or designing its molecular attractants are, therefore, essential to this type of drug design.

Estrogen receptor is shown to be overly expressed in many breast tumors (McGuire, 1976). The natural ligand of this protein, estrogen, can be, therefore, tethered to a DNA-damaging chemical to form a complete bifunctional drug. In these breast tumors where estrogen receptor is at high level, the DNA lesions formed by the drug are protected from efficient removal through the binding of the receptor protein to the lesion sites. The practical result is the high and selective toxicity of the drug in these breast tumor cells.

It is desirable, however, to identify a protein that exists uniquely in various

types of human tumors in order to design drugs that possess "universal" antitumor activity. Mutant p53 proteins are a logical choice for this long range drug design because p53 mutations have been observed in a significant proportion of human tumors (Hollstein et al., 1994; Levine et al., 1994), and also because mutant p53 proteins, like the wild-type protein, are localized in the cell nuclei, but their concentrations are often higher than that of the wild-type (Iggo et al., 1990; Moll et al., 1992). Since there are no natural p53-binding molecules that are suitable for the purpose of Fatal Engineering drug design, it is necessary to find a way to obtain molecules that bind to mutant p53 proteins with high affinity. A recent development called "in vitro genetics" could present a way of obtaining such molecules (Szostak, 1992). In vitro genetics involves first constructing a degenerate molecule library (e.g., ss DNA, RNA, peptide, or small chemical libraries) and subsequently selecting from the library the molecules that can bind to a chosen target. The long term goal of Fatal Engineering was to use an in vitro selection strategy to obtain ligands that bind to a common TSP such as a mutant p53 protein with high affinity, and then to link the ligand to a DNA-damaging chemical warhead to form a complete drug that will have a wide range of anticancer activities.

3. Relevance of this work to the overall goal of Fatal Engineering

This portion of my thesis describes the use of a model-testing molecule, a psoralen-biotin conjugate, in an attempt to prove the concept of Fatal Engineering. In this work, it has been established that a molecule designed according to the scheme of Fatal Engineering behaved as predicted. The conjugate retained the DNA reactivity of its chemical warhead domain, and its protein-attractive activity was also retained after being covalently linked to a double-stranded DNA fragment. Streptavidin, when bound to the modified DNA, protected a region from a DNA repair enzyme. The size of the protected region was comparable to that of

DNA patch released by the nucleotide excision repair pathway.

It should be noted that decreases in both DNA reactivity and protein affinity were observed with the bifunctional molecules when compared to the unconjugated precursors. This could be attributed to various factors such as the solubility of the compound and the length of the linker between the two functional domains. These factors can be systematically analyzed by using the TMP-biotin test molecule; the results of these analyses could be of great importance to the future design of real anticancer drugs.

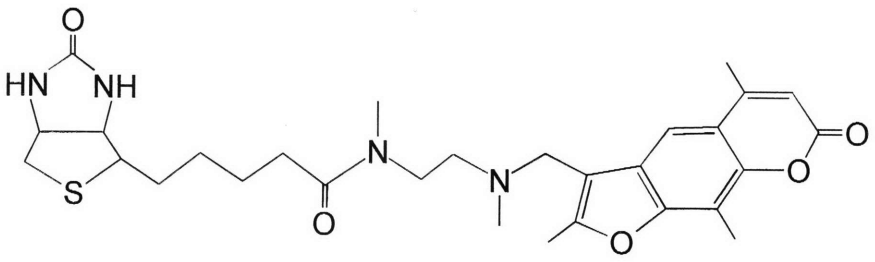
E. CONCLUSIONS

A synthetic molecule containing a biotin domain linked to a psoralen derivative was used to help prove the concept of a novel approach of designing effective anticancer drugs that act through their selective persistence on the DNA of tumor cells. The drugs are proposed to be bifunctional molecules that have a DNA-attacking chemical warhead tethered to a ligand domain that is designed to bind with high affinity to a TSP. It is predicted that in tumor cells where a TSP is present, the lethal DNA lesions formed by the drug will be protected by the binding of the TSP to the lesion sites, while in normal cells these lesions will be quickly removed by DNA repair systems due to the absence of the TSP. The result of this selective persistence of drug-DNA adducts in tumor cells is the preferential killing of tumor cells. In an attempt to prove the feasibility of above hypothesis, I have used a conjugate of psoralen (a DNA crosslinker) and biotin (an avidin or streptavidin-binding ligand) as a test molecule. I have shown that both DNA-reacting and protein-binding activities of the unconjugated precursors were retained in the bifunctional conjugate. Streptavidin, which can bind to the biotin that was adducted to a DNA molecule through the psoralen domain, not only slowed the removal of a nearby uracil by DNA uracil glycosylase, but also

protected a region on the DNA molecule flanking the adduct site. The size of the covered region is comparable to the size of the DNA patch released by the DNA excision repair pathway.

Figure 14. The TMP-biotin conjugate and its reactivity with DNA. **A.** The structure of the TMP-biotin molecule. The synthesis of this molecule was carried out by Morningstar (1994). A dialkylamine-containing spacer arm was used to increase polarity, and thus the water-solubility, of the compound. **B.** The sequence of double-stranded U-17. The top strand, containing a deoxyuracil nucleotide (underlined), was labeled with [^{32}P] at the 5'-end. The sequence has a single TA dinucleotide site at which psoralen and its derivatives are known to form DNA adducts (Cimino et al., 1985). **C.** Modification of DNA by the TMP-biotin compound. TMP-biotin in 50% (v/v) acetonitrile/water was mixed with ^{32}P-labeled U-17 at approximately a 1000 : 1 molar ratio of TMP-biotin to base pair. After incubation at room temperature for 10 min, the mixture was irradiated with 365 nm wavelength light on a chilled surface to a final dose of about 85 kJ/m^2, followed by an analysis on a 20% denaturing PAGE. Bands B1 and B2 represent the migration positions of unmodified U-17 and U-17 crosslinked to its complementary strand, respectively. B2 represents the U-17 species modified by an unconjugated psoralen derivative present in a trace amount in the TMP-biotin preparation (refer to text for details). B3 represents the desired U-17 species that contain a monoadduct of TMP-biotin conjugate on the thymine of the TA dinucleotides.

A



B

^{32}P-CGGCCGTACGUGCGCCG
GCCGGCATGCACGCGGC

C

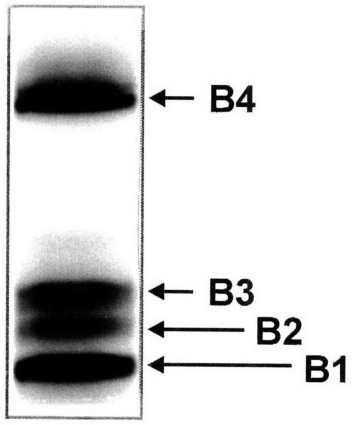


Figure 15. Binding of streptavidin to U-17 monoadducted with the TMP-biotin conjugate. **A.** Gel mobility shift assay. 3200 cpm (~0.1 nM) of the probe was used in each lane. 0, 0.4 nM, 1 nM, 2 nM, 5 nM, 10 nM, 50 nM and 50 nM of streptavidin were used in lanes 1–8, respectively. In lane 8, free d-biotin was also included in the incubation to the final concentration of 0.4 mM. **B.** Binding curve created by plotting the percentages of bound probe against streptavidin concentrations.

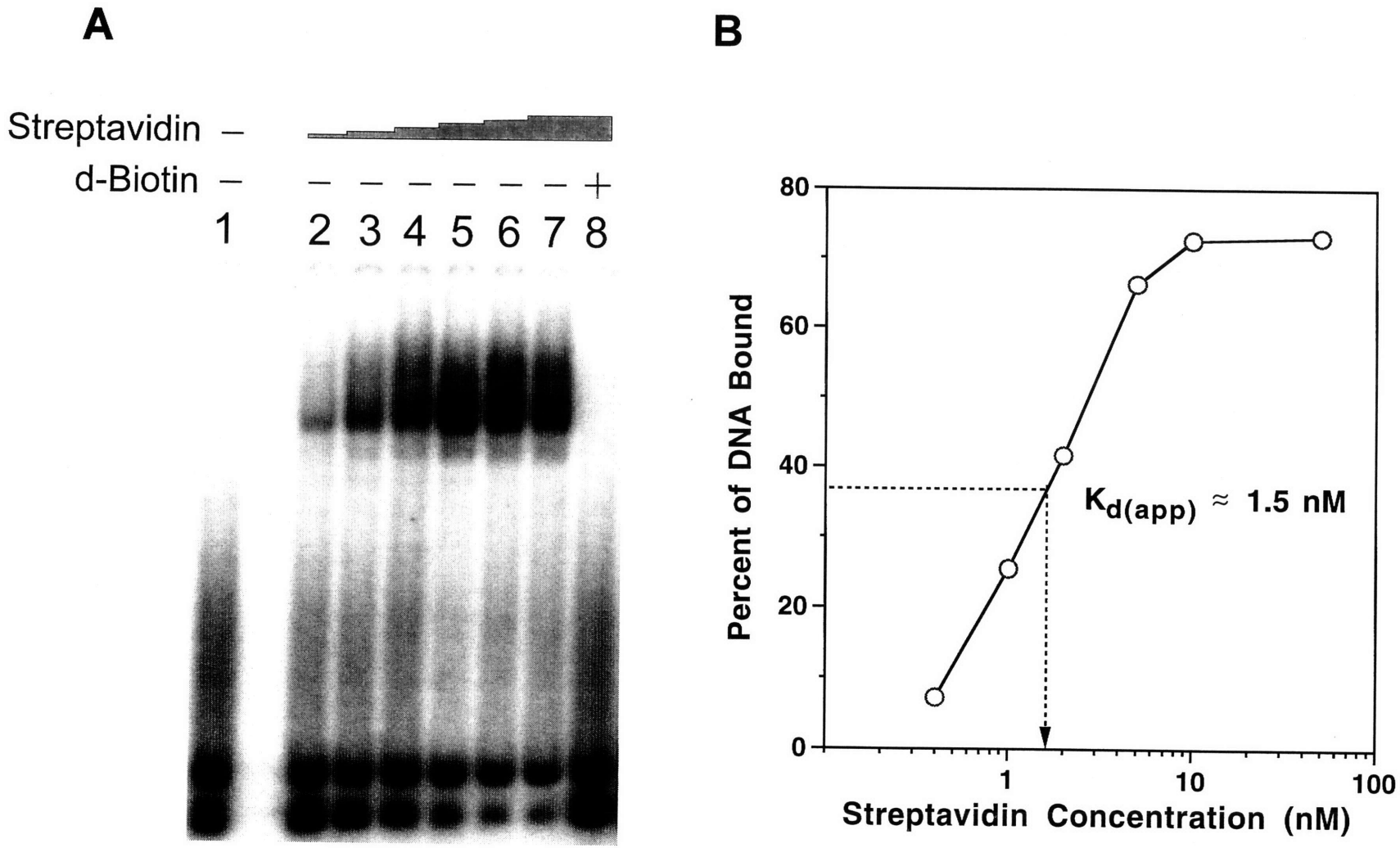
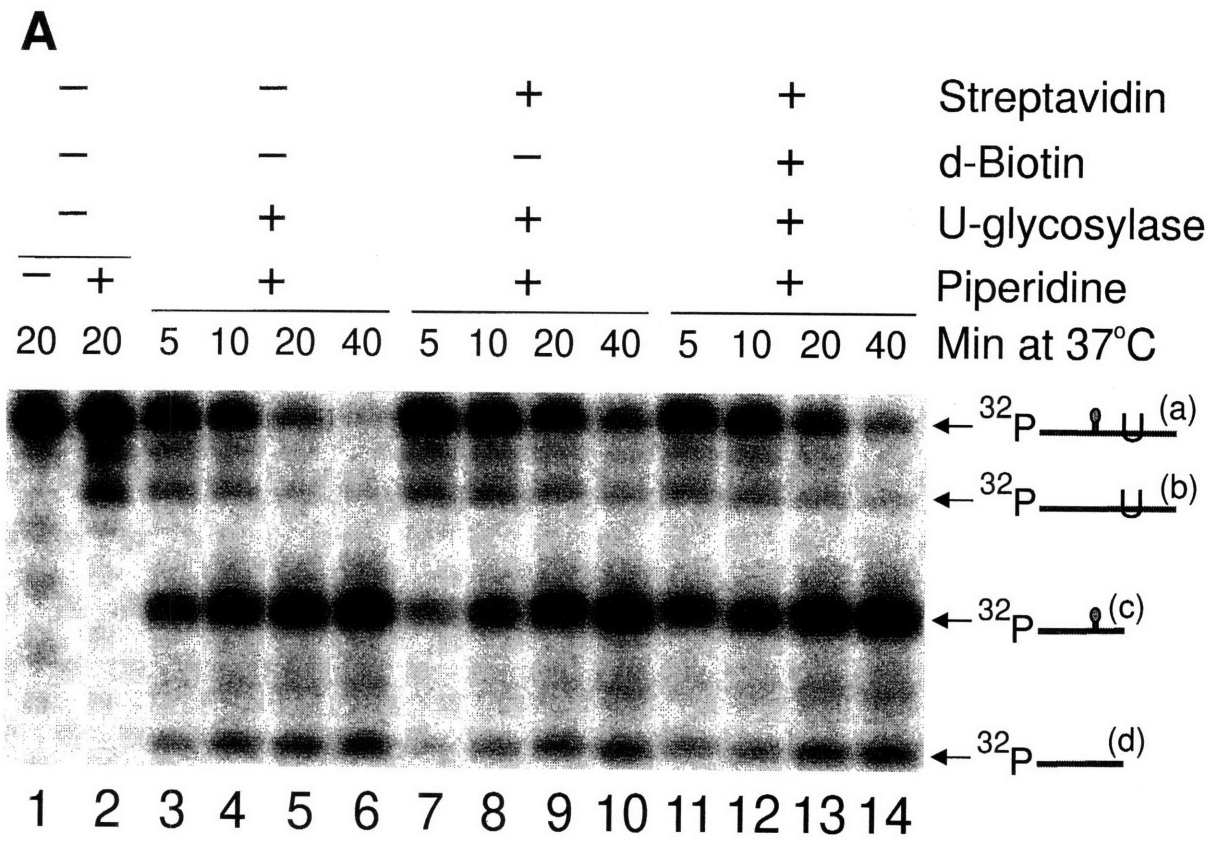

A
Streptavidin –
d-Biotin – – – – – – – +
1 2 3 4 5 6 7 8
B
Percent of DNA Bound
80
60
40
20
0
$K_{d(app)} \approx 1.5$ nM
1
10
100
Streptavidin Concentration (nM)

Figure 16. Inhibition of uracil glycosylase by streptavidin. 4000 cpm (~0.15 nM) was used in each lane. The probe was incubated with (lanes 7–14) or without (lanes 1–6) streptavidin at room temperature for 5 min (lanes 3, 7 and 11), 10 min (lanes 4, 8 and 12), 20 min (lanes 1, 2, 5, 9 and 13), or 40 min (lanes 6, 10 and 14). Except for the sample in lane 1, which serves as marker for the intact full-length probe, the rest of the samples were subjected to piperidine cleavages. In lanes 11–14, free d-biotin was added during the incubations with streptavidin to the final concentration of 0.4 mM. The samples were analyzed on a 20% denaturing polyacrylamide gel. Bands (a) represent the full-length and intact probe used in each reaction. Bands (c) represent the products of the uracil glycosylase treatments and the subsequent piperidine cleavages. Bands (b) and bands (d) are the breakdown products of bands (a) and bands (c), respectively, due to the alkali lability of the adducts.


A
− − + + Streptavidin
− − − + d-Biotin
− + + + U-glycosylase
− + + + + Piperidine
20 20 5 10 20 40 5 10 20 40 5 10 20 40 Min at 37°C
32P (a)
32P (b)
32P (c)
32P (d)
1 2 3 4 5 6 7 8 9 10 11 12 13 14

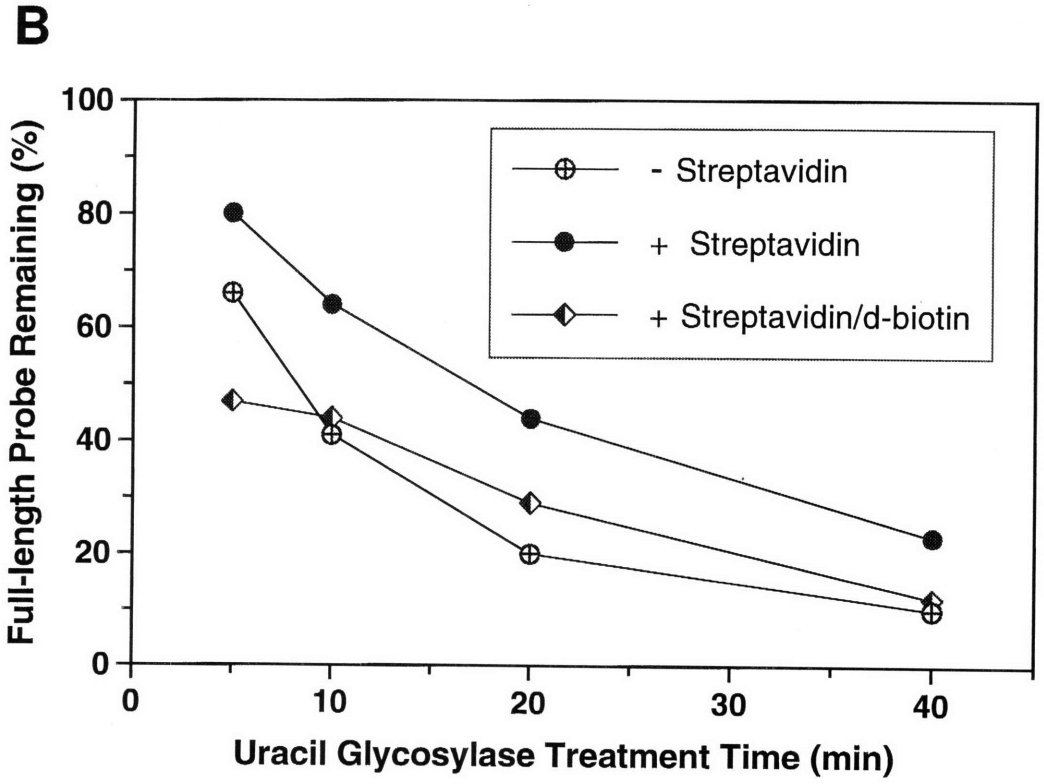
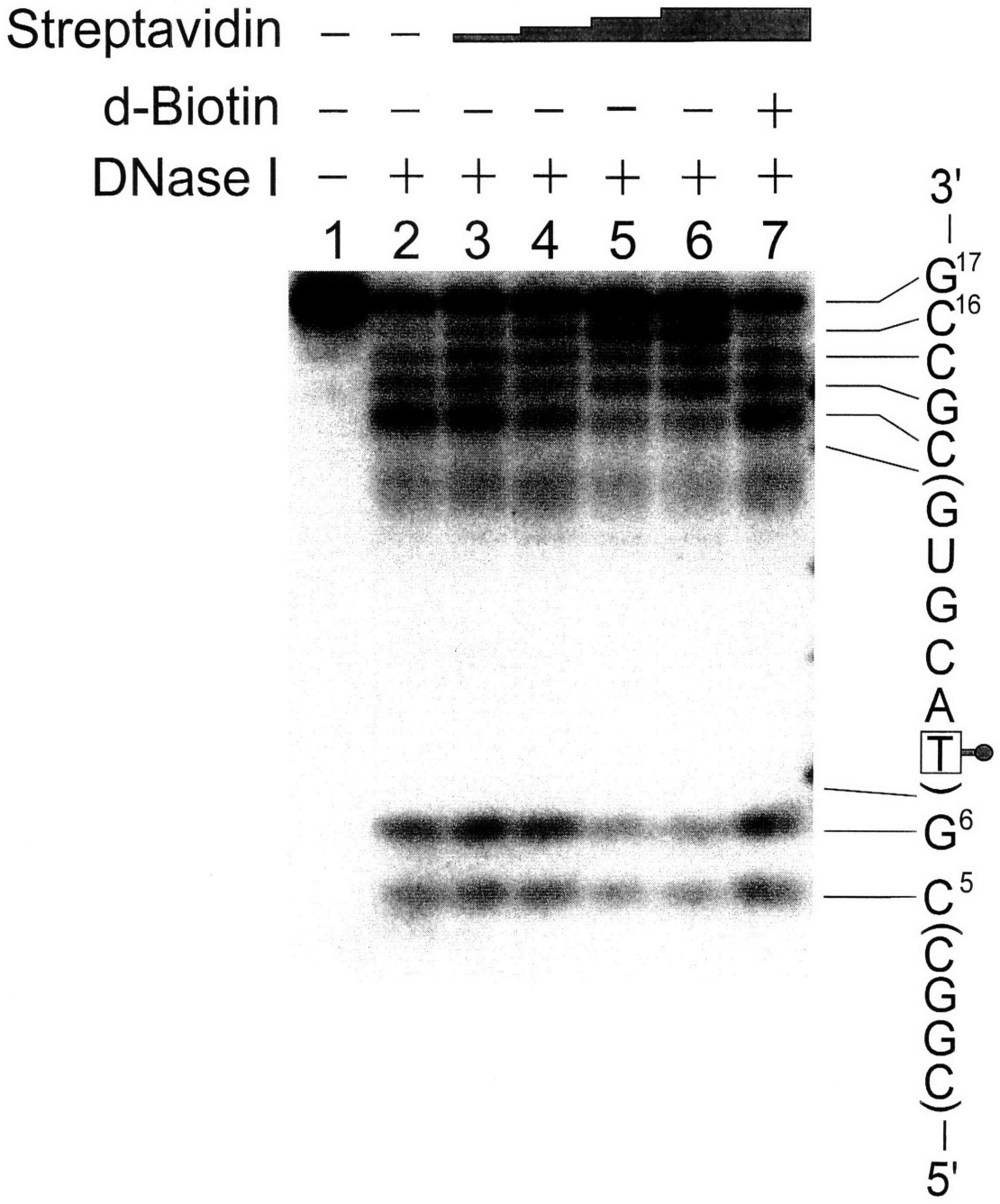

B
Full-length Probe Remaining (%)
100
80
60
40
20
0
0
10
20
30
40
Uracil Glycosylase Treatment Time (min)
- Streptavidin
+ Streptavidin
+ Streptavidin/d-biotin

Figure 17. DNase I protection assay. 5000 cpm (~1.5 fmoles, ~1.5 nM) of the [^{32}P] end-labeled probe was used in each lane. 0, 0.4 nM, 2 nM, 10 nM, 50 nM and 50 nM of streptavidin were used in lanes 2–7, respectively. In lane 7, 0.4 mM of free d-biotin was also included in the incubation. For the lanes where indicated, 5 μg of DNase I (final concentration, 0.4 mg/ml) was added into each incubation and the digestions were carried out at room temperature for 2 min. Lane 1 represents the intact full-length probe. The samples were analyzed on a 20% denaturing polyacrylamide gel. The marked thymidine base demonstrates the position where a TMP-biotin was adducted.


Streptavidin – – +
d-Biotin – – – – – – +
DNase I – + + + + + +
1 2 3 4 5 6 7
3'
G17
C16
C
G
C
(G
U
G
C
A
T
)
G6
C5
C
(C
G
G
C)
5'

V. FATAL ENGINEERING II: IN VITRO SELECTION OF DNA APTAMERS FOR A MUTANT p53-SPECIFIC EPITOPE

A. INTRODUCTION

Under the scheme of Fatal Engineering for drug design, the drug is a DNA-attacking chemical that also contains a functional domain attractive to a tumor specific protein (TSP). A TSP here is defined as a protein that is expressed either exclusively or at a higher level in tumor cells as compared to normal cells. It is predicted that binding of TSPs to DNA lesions formed by the drug will prevent the efficient removal of these lesions and, hence, enhance the toxicity of the drug selectively in tumor cells. Since most tumor specific proteins are not known to bind molecules that are suitable for this type of drug design, it is desirable to find a way to obtain ligands that have affinity for a chosen tumor specific protein. Recent progress in the field of in vitro genetics presents an interesting approach by which such ligands can be obtained (Szostak, 1992).

In vitro genetics involves the construction of a large combinatorial library of molecules and the subsequent selection from the library for those molecules with desired characteristics. Degenerate oligonucleotide libraries have been successfully used to select for nucleic acid sequences (called RNA or DNA aptamers) that bind to a variety of target molecules including organic dyes (Ellington and Szostak, 1990, 1992), cofactors (Lorsch and Szostak, 1994), bases and nucleotides (Sassanfar and Szostak, 1993; Connell and Yarus, 1994; Jenison et al., 1994; Huizenga and Szostak, 1995), amino acids (Famulok and Szostak, 1992b; Connell et al., 1993; Famulok, 1994), peptides (Nieuwlandt et al., 1995), and proteins (Tuerk and Gold, 1990; Bock et al., 1992; Tsai, et al., 1992; Tuerk et al., 1992; Jellinek, 1993; Conrad et al., 1994; Doudna et al., 1995). In one study that is relevant to the goal of this project, DNA aptamers were selected from a randomized single-stranded (ss) DNA pool for the protease thrombin of the blood coagulation cascade with binding affinities in the range 25–200 nM (Bock et al., 1992). It should be pointed out that thrombin has no physiological DNA binding

properties.

Many oncogene-encoded proteins, when mutated or overly expressed, can be considered as tumor specific proteins. For example, p53 mutations are the most common cancer-related genetic changes known (Hollstein et al., 1994; Levine et al., 1994) and, except for some germinal mutations (Malkin et al., 1990; Srivastava et al., 1990), most of the p53 mutations are only observed in tumor cells. Furthermore, while both the wild-type and mutant p53 proteins are localized in the nucleus, the concentration of mutant p53 is often higher than that of the wild-type protein (Iggo et al., 1990; Moll et al., 1990). Mutant p53 proteins can therefore be used as the target molecule for aptamer selection. Additional measures, however, have to be taken to ensure that the selected aptamers possess a minimal affinity for the wild-type protein. Alternatively, a mutant p53-specific epitope can be used as the selection target if the selected epitope-binding aptamers also show selective affinity for the mutant p53 protein as a whole. There is indeed such an epitope that is believed to be buried within the wild-type p53 protein. That epitope apparently becomes exposed on the protein surface in many of the observed mutant proteins (Gannon et al., 1990; Cho et al., 1994). The sequence of the epitope itself (Arg-His-Ser-Val-Val) is wild-type. It is believed that mutations in other regions of the protein either induce conformational changes to the protein or block its folding process (Cho et al., 1994).

The epitope mentioned above was therefore used to raise the aptamers described in this thesis. A synthetic peptide containing the epitope was immobilized to agarose beads via a disulfide bond, and the beads were then packed into columns. A randomized single-stranded DNA pool was generated by chemically synthesizing deoxynucleotides consisting of a randomized central region flanked by two fixed PCR primer termini. Through an in vitro selection procedure involving rounds of selection and amplification, groups of ss DNA molecules

(DNA aptamers) were obtained with specific affinity for columns containing the target peptide. The aptamers had little affinity for columns containing a control peptide in which the sequence of the epitope was scrambled. Attempts were made to see whether these selected aptamers would also have affinity for a whole mutant p53 protein.

B. MATERIALS AND METHODS

1. Materials

Restriction endonucleases, T4 polynucleotide kinase and T4 DNA ligase were purchased from New England Biolabs. Dithiothreitol (DTT) was obtained from GIBCO-BRL and DNA sequencing reagents (Sequenase 2.0) were purchased from United States Biochemical Corporation. Salmon sperm DNA, Nonidet P-40 (NP-40) and phenylmethylsulfonylfluoride (PMSF) were obtained from Sigma. Proteinase inhibitors (aprotinin, leupeptin, and α_2-macroplobulin) were purchased from Boehringer Mannheim Biochemicals.

Recombinant baculoviruses expressing either human wild-type or human mutant (Arg175 to His) p53 protein were gifts from C. Prives (Columbia University). Insect cell line High Five was kindly provided by R. Weinberg (MIT), and Grace's insect media was purchased from Invitrogen (San Diego, CA). Monoclonal antibodies PAb421 (recognizes the C-termini of both wild-type and mutant p53 proteins), PAb1801 (recognizes the N-termini of both wild-type and mutant p53 proteins), and PAb240 (recognizes an epitope only existing on the surface of certain mutant p53 proteins) were obtained from Oncogene Sciences (Cambridge, MA).

2. Generation of a randomized deoxynucleotide pool

A randomized ss DNA pool was generated by synthesizing a 100-mer deoxynucleotide of sequence 5'-ATACCAGAATTCACTATT-N_{64}-AGATAGGAATTCGAATCT-3', where N_{64} represents 64 nucleotides with equimolar balance of dG, dA, dT and dC at each position. The synthesis was carried out by the MIT Biopolymers Laboratory on a 0.2 μmol synthesis scale, and the 100-mer was purified on an 8% denaturing polyacrylamide gel (acrylamide/bis, 19:1). The gel was UV shadowed and the band of the full-length DNA was excised, crushed and soaked in distilled water overnight. The gel pieces were removed by filtration, and the recovered DNA solution was subjected first to sec-butanol extraction and then to ethanol precipitation. After being washed twice with cold 80% ethanol, the DNA pellet was resuspended in TE buffer (10 mM Tris-HCl pH 7.5 and 1 mM EDTA). The DNA concentration was calculated according to Maniatis et al. (1982) using an extinction coefficient of 1100 for the oligomer.

Since the total randomization of a 64-nucleotide region would generate 4^{64} ($\sim 3 \times 10^{38}$) possible distinct sequences, and a 0.2 μmol synthesis would give rise to a maximum of about 10^{17} molecules, the chance for two sequences within the 10^{17} molecules to have the same sequence would be one out of 3×10^{21}. Therefore, each synthesized oligomer should have a distinct sequence. The initial amount of DNA used for the first round of selection was 170 pmol. Taking into account that only ~10% of these sequences are amplifiable due to chemical lesions introduced during the synthesis (Ellington and Szostak, 1992; and our unpublished observations), this amount of DNA should contain approximately 10^{13} amplifiable molecules with distinct sequences.

3. Preparation of EP240-Cys peptide affinity columns

Design and synthesis of peptides EP240-Cys and EP240S-Cys. A peptide

(EP240-Cys: NH_2-Thr-Phe-Arg-His-Ser-Val-Val-Val-Pro-**Cys**-COOH) which contains the mutant p53-specific epitope region (underlined) and a control peptide (EP240S-Cys: NH_2-Thr-Phe-Val-His-Val-Ser-Arg-Val-Pro-**Cys**-COOH) in which the sequence of the epitope region was scrambled as indicated by the underlined S were synthesized by the MIT Biopolymers Laboratory. The sequence of EP240-Cys was derived from the human wild-type p53 protein sequence (Lamb and Crawford, 1986). A cysteine residue (highlighted) was added to the C-termini of both peptides to facilitate their coupling to the thiol-containing agarose bead matrix.

Immobilization of peptides EP240-Cys and EP240S-Cys. A buffer of 100 mM Tris-HCl pH 7.5, 50 mM NaCl and 1 mM EDTA was used to prepare peptide solutions at 45 mg/ml. It was necessary, however, to adjust the pH to between 2.5 and 3.5 in order to completely dissolve the peptides. Thiopropyl-activated Sepharose 6B (Pharmacia) gel suspension was prepared according to the manufacture's instructions. Fifteen grams of the thiol-Sepharose powder gave rise to about 40 ml gel bed volume, to which 120 ml of 10 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 mM EDTA was added to bring up the volume of the gel suspension to 160 ml. Coupling reactions were carried out by mixing 45 ml of peptide solution with 80 ml of the gel suspension, and allowing the mixture to be incubated for 45 min at room temperature with end-to-end rotation. The coupling efficiencies were measured by monitoring the displacement of 2-thiopyridone, which has a peak absorption at 343 nm, by the thiol-containing peptides. The concentrations of coupled peptides were determined to be 6 mM (drained gel volume) for EP240-Cys and 5 mM for EP240S-Cys.

Preparation of EP240-Cys and EP240S-Cys columns. Two types of columns were used during the course of this study. One was used in the selection procedures, and the other was used as an analytical column (see 4 and 7,

respectively, in this chapter's Materials and Methods). For the selection column, Sepharose beads with coupled peptides were packed to a final volume of ~1.6 ml with a height of ~3.2 cm; for the analytical column, the bed volume was ~0.25 ml and the height was ~2.2 cm.

4. In vitro selection of DNA aptamers from the randomized ss DNA pool

Approximately 170 pmol of the synthetic 100-mer was PCR amplified. A 3000 µl cocktail consisting of 1.0 µM primer P-1: 5'-ATACCAGAATTCACTATT-3'; 1.0 µM primer B-P-2: 5'-biotin-AGATTCGAATTCCTATCT-3'; 250 µM each of dGTP, dATP, dTTP and dCTP; 20 mM Tris-HCl pH 8.5; 50 mM KCl; 2.5 mM $MgCl_2$; 75 units Taq polymerase (Perkin-Elmer Cetus) and 100 µCi of [α-^{32}P]dCTP (3000 Ci/mmol, New England Nuclear) was distributed in 100-µl aliquots to 30 tubes and cycled through 30 rounds of 94°C, 1 min; 54°C, 1 min; and 72°C, 1 min. The radiolabeled PCR products were isolated on an 8% denaturing polyacrylamide gel. The purified DNA was dissolved in a buffer of 25 mM Tris-HCl pH 7.5, 100 mM NaCl and 1.5 mM $MgCl_2$ and loaded onto a streptavidin affinity column (Pierce, Rockland, IL). The unbiotinylated strand, which was used for the aptamer selection, was eluted off the column by 150 mM NaOH (Hultman et al., 1988; Bock et al., 1992) followed by ethanol precipitation. To allow the resultant single-stranded DNA sequences to fold into proper structures, ~500 pmol of the purified ss DNA was dissolved in 500 µl aptamer selection buffer (ASB; 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 1 mM EDTA), heated at 70°C for 5 min and allowed to cool down to room temperature. The sample was loaded onto an EP240-Cys selection column which had been pre-equilibrated with ASB. After the column was washed with ~10 ml of ASB, the bound DNAs were recovered by cleaving the column-bound peptides with 10–14 ml of ASB containing 40 mM DTT. The flow-through was collected at ~0.5 ml per fraction, and the fractions of DTT elution (with a total

volume of 8–9 ml) were pooled and concentrated with Microcon 10 (Amicon) microconcentrators to approximately 0.5 ml. One fifth of the recovered DNA was used for PCR amplification and, as just described, single-stranded DNA sequences were isolated from the PCR products, folded and subsequently used for the next selection cycle.

A total of 9 cycles of amplification and selection was carried out. For cycles 7, 8 and 9, the volume of aptamer selection buffer used to wash the columns was extended from 10 ml, as used in the previous cycles, to 20 ml. Also in these cycles, a pre-column (an EP240S-Cys column containing the scrambled peptides) was used prior to the selection column. In these cycles, DNA from a previous cycle was first loaded onto the EP240S-Cys pre-column; the flow-through fractions representing the unbound population were combined, concentrated to 800–900 µl and then loaded directly onto the EP240-Cys selection column.

5. Cloning and sequencing of individual DNA aptamer sequences

DNA from the eighth round of selection (pool 8) was amplified by PCR, and the double-stranded PCR products were digested with EcoRI. Individual DNA sequences were cloned by ligating the EcoRI-digested pool 8 DNA to EcoRI-digested RF M13mp18 vectors, transferring the resultant recombinant vectors into *E. coli* DL7 cells by electroporation, and plating the transformed cells in the presence of *E. coli* GW5100 cells, IPTG and X-Gal to obtain independent plaques. A total of 100 plaques was picked and used to produce single-stranded DNA that gave 63 legible sequences using the dideoxy chain-termination method (Sanger et al., 1977).

6. Sequence analysis of the selected DNA aptamers

The homology search was carried out with the Pustell sequence analysis software from IBI (New Haven, CT). The sequences of all the sixty-three isolated clones were entered into the computer program to create a sequence data bank. Each individual sequence was then compared to the rest of the sequences to identify clones sharing vast homologies. Consensus sequence domains were searched by identifying regions 10 or 20 nucleotides long with respective 75% or 60% minimal homology. The search for secondary structure among the cloned aptamer sequences was carried out by folding the sequences using the GCG program (Madison, Wisconsin).

DNA base composition was obtained by HPLC analysis of nucleosides. In each analysis, ~5 μg of DNA was digested at 37°C for 4 hr with snake venom phosphodiesterase and bovine alkaline phosphatase (Sigma; 24 μg and 17 units, respectively) in 80 μl buffer containing 100 mM Tris-HCl pH 8.5 and 1.5 mM $MgCl_2$. The deoxynucleosides were resolved by reverse-phase HPLC using a Beckman C-18 analytical column (0.46 × 25 cm). A linear 40 min gradient from 0–20% B was performed at 1.0 ml/min, with pump A in 0.1 NH_4OAc (aqueous) and pump B in 0.1 M NH_4OAc (1 : 1, H_2O : CH_3CN).

7. Column binding assays

Sepharose columns (analytical; also see section 3 in this chapter's Materials and Methods) containing immobilized EP240-Cys peptides were used to analyze the ability of enriched DNA pools and individual aptamer sequences to bind the immobilized peptides. In this experiment, the column was pre-equilibrated with buffer containing 25 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM EDTA before a DNA sample in 100 μl of the same buffer was loaded. It should be noted that the NaCl concentration used in these column binding assays was 100 mM instead of 200 mM, as was previously used in the selection

procedures. For each loading, between 3–5 μg of DNA was used which consisted of either just the ^{32}P-labeled DNA sequences under study or the labeled DNA mixed with unlabeled denatured salmon sperm DNA (about 300-nucleotides in length). After a DNA sample was loaded, the column was washed with either 2, 3 or 4 ml of the above buffer, followed by elution with 2 ml of the same buffer containing 40 mM DTT.

8. Attempts to determine whether the peptide-binding aptamers can also bind selectively to a mutant p53 protein

a. Preparation of whole cell extracts from insect cells infected with p53-expressing recombinant baculoviruses

High Five insect cells (2.5 x 10^7 cells per 150-mm dish), grown in Grace's insect medium (Invitrogen, San Diego, CA) at 27°C, were infected with recombinant baculoviruses expressing either human wild-type or human mutant (Arg175 to His) p53 protein at a multiplicity of infection (M.O.I.) of 3 plaque-forming units per cell. Infected cells were harvested 48 hr post-infection, and cell pellets were washed twice with cold phosphate-buffered saline (PBS; 1 mM Na_2HPO_4, 10.5 mM KH_2PO_4, 140 mM NaCl, 14 mM KCl, pH 6.2), followed by extraction for 30 min on ice with 1.0 ml of lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 100 μg/ml aprotinin, 30 μg/ml leupeptin, 10 μg/ml α_2-macroplobulin and 0.5 mM PMSF). Lysates were centrifuged at 15,000 g for 45 min, and the clear supernatant was frozen in aliquots at −80°C. The protein concentrations of the cell extracts were determined by the Bradford assay (Bio-Rad).

To metabolically label proteins with [^{35}S], cells were collected 42 hr post-infection, washed twice with methionine-free media (JRH Biosciences, Lenexa,

KS), and resuspended in the methionine-free medium supplemented with 10% heat inactivated fetal calf serum. [^{35}S]methionine (cell-labeling grade, 0.5 mCi per 100-mm dish; Amersham) was then added into the media, and the labeling process was carried out for 6 hr at 27°C before the cells were lysed as previously described.

Immunoprecipitation (IP) assays were carried out to confirm the expression of p53 proteins using the ^{35}S-labeled cell lysates. In each tube, 10 µl of lysate and 4 µl of monoclonal antibodies PAb421, PAb1801 or PAb240 (100 µg/ml in solution; Oncogene Sciences, Cambridge, MA) were mixed in 186 µl of immunoprecipitation buffer (IP buffer; 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol, 0.2% NP-40 and 2 mM EDTA). The mixtures were incubated for 1 hr at 4°C with end-to-end rotation before a 50 µl mixture consisting of 10 µl protein A-Sepharose 4B slurry (Sigma) and 40 µl IP buffer was added to each incubation. The samples were incubated for another hour at 4°C with rotation followed by three washings with cold IP buffer. The resultant immuno-complexes were analyzed on a 12% SDS-PAGE gel. The gel was dried and exposed to X-ray film with an intensifying screen at −80°C. Similar immunoprecipitation experiments were also carried out with the unlabeled extracts, but upscaled 5-fold. The results were also analyzed on a 12% SDS-PAGE gel, but were visualized by staining the gel with silver-staining reagents (Bio-Rad).

b. Gel mobility shift assays

Gel mobility shift assays were performed to determine whether the epitope-binding aptamers could also bind to the His175 p53 mutant protein. Single-stranded DNA aptamer probes were end-labeled with [γ-^{32}P]ATP (6000 Ci/mmol, New England Nuclear) using T4 polynucleotide kinase. After being purified on 8% polyacrylamide gels, the labeled probes were dissolved in 25 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM EDTA, heated at 70°C for 5 min and then cooled

down to room temperature. A gel shift assay was carried out by incubating the DNA aptamer probe (~10^4 cpm per incubation) with 20 μg cell extract at room temperature for 10 min in 20 μl of buffer containing 25 mM Tris-HCl pH 7.5, 1 mM EDTA and either 100 mM, 75 mM or 50 mM NaCl. The mixtures were then subjected to electrophoresis on a 5% nondenaturing polyacrylamide gel (acrylamide/bis, 29:1) in either 0.5 × or 1 × TBE (1 × TBE: 89 mM Tris base, 89 mM boric acid and 1 mM EDTA) at 4°C. After electrophoresis, the gel was dried and analyzed using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

c. Immunoprecipitations

Immunoprecipitation assays were also carried out to determine whether the epitope-binding aptamers have affinity for the His175 mutant p53 protein. In these experiments, end-labeled DNA aptamers (10^4–10^5 cpm per incubation) were incubated with 40 μg of His175 mutant p53 cell extract in 200 μl of 25 mM Tris-HCl pH 7.5, 10% glycerol, 0.2% NP-40, 1 mM EDTA, 5.5 μg/ml sonicated and denatured salmon sperm DNA and either 100 mM, 75 mM or 50 mM NaCl at 4°C with end-to-end rotation for 1 hr. Then, 2 μl of PAb1801 or PAb421 monoclonal antibody was added into each tube, and the mixtures were incubated for another half-hour before 50 μl IP buffer containing 5 μl of the protein A-Sepharose 4B slurry was added. After a 1 hr incubation at 4°C with end-to-end rotation, the putative immuno-complexes were recovered by centrifugation and counted with a scintillation machine.

9. Attempts to measure the steady-state binding affinity of clone C1 by solid-liquid phase partitioning

Steady-state binding analysis was performed in 1.5 ml Eppendorf tubes containing mixtures of EP240-Cys agarose beads and ^{32}P-labeled clone C1 aptamer

DNA at various concentrations in 100 μl of Tris-HCl pH 7.5, 125 mM NaCl and 1 mM EDTA. Different DNA concentrations (0.25–25 μM) were achieved by mixing ^{32}P-labeled clone C1 probe (~2 × 10^5 cpm per incubation) with various amounts of unlabeled clone C1 DNA. In each tube, agarose beads containing a total amount of 50 pmol coupled EP240-Cys peptides were used which gave rise to a 0.5 μM concentration of immobilized peptide. The mixtures were incubated at room temperature with end-to-end rotation for 3 hr before the samples were transferred into 0.45 μm microfilters (Millipore) and subjected to a 2-min spin in a microcentrifuge at room temperature. The filter units containing the retained beads were counted with a scintillation machine.

C. RESULTS

1. Selection of peptide-binding aptamers

An in vitro selection was carried out to obtain DNA aptamers that can bind to a peptide containing a mutant p53-unique epitope, Arg-His-Ser-Val-Val. The peptide was immobilized on a solid support through a disulfide bond. In each cycle, the bound DNA molecules were recovered by DTT cleavage of the disulfide linkage. ^{32}P-labeled DNAs were used to monitor the selection. As demonstrated in Figure 18A, the majority of the DNA loaded in cycle 1 showed little binding to the column. By cycle 6, however, the shape of the peak representing unbound DNA became broader, and its maxima shifted to higher fraction numbers (Figure 18A). However, as also indicated in Figure 18A, fractions collected after DTT elution apparently did not contain any significant amount of DNA. It remained to be the case even in cycle 8 (Figure 18B). These results indicated that after 5 rounds of selection, the DNA pools showed some reactivities with the immobilized target peptide. Furthermore, this interaction was specific for the target peptide because the DNA molecules did not interact with control columns which contained

peptides with a scrambled epitope region (Figure 18B). However, this observed interaction was probably not strong enough to survive the intensive washings with the binding buffer containing 200 mM NaCl.

When a buffer containing a lower salt concentration (100 mM instead of 200 mM NaCl) was used in the column binding assays to analyze the DNA pools, stable interactions were observed between the selected DNA populations and the EP240-Cys columns (Figure 19). The portion of a DNA population interacting with the target column appeared to be increasing along the progression of the selection cycles. By pool 8 (the DNA population after 8 selection cycles), a significant amount DNA appeared in the DTT-eluted fractions with the EP240-Cys column. Moreover, the observed binding was specific for the target column because, as exampled by the elution profile from an EP240S-Cys control column with pool 8 DNA, the selected DNA pools did not bind to the immbolized peptide with a scrambled epitope sequence. These results indicated that aptamer sequences possessing an acquired ability to bind to the immobilized target peptide were enriched, through selection, from the original ss DNA library.

Changes in DNA population during selection were also confirmed by analyzing the DNA pools with restriction endonucleases recognizing four base sequences. As shown in Figure 20, the DNA sequences in the early pools are highly heterogeneous as evidenced by the evenly distributed bands in lanes 1–4. Starting from pool 4 (lanes 5 in Figure 20) some distinct bands started to emerge, suggesting that the heterogeneity of the DNA pools was reducing. The banding patterns, however, did not change much from pool 6 to pool 9 indicating that cycles 6, 7, 8 and 9 did not introduce significant changes to the DNA population. This observation is consistent with the results obtained from the analytical column binding experiments, which showed little difference among pools 6, 7, 8 and 9 in their ability to interact with the columns (data not shown).

PCR products from the eighth cycle (pool 8) were cloned and sequenced. Of the 63 clones sequenced, 52 sequences were unique. Seventeen individual aptamer sequences representing 26 independent sequenced clones were analyzed for their ability to bind to the selection peptide column. The results showed that all of these tested sequences bound the peptide column (data not shown). These 17 sequences are shown in Figure 21. A sequence homology search did not reveal any obvious consensus region among these unique sequences, nor did simulated foldings of these sequences using the GCG program give rise to any conserved secondary structure.

2. Analysis of the selected aptamers

Since in each selection cycle, single-stranded DNA aptamers were isolated from double-stranded PCR products, there was a possibility that a residual amount of double-stranded DNA molecules might be responsible for the observed binding activities. To eliminate this possibility, clone C1 sequence was chemically synthesized and binding of this synthetic single-stranded DNA to the EP240-Cys target column was tested. The results indicated that single-stranded aptamers were responsible for the observed bindings (Figure 22B). This is consistent with the observation that double-stranded PCR products of the aptamers did not bind to the column (data not shown).

The lack of a consensus domain among the cloned sequences makes it difficult to speculate which regions of these sequences are essential for binding. Nevertheless, two sequences (designated as C1m1 and C1m2, respectively) in which several bases are different from clone C1, a designated model aptamer molecule (Figure 22A), were also chemically synthesized and tested for their ability to bind to the EP240-Cys target column. Both C1m1 and C1m2 showed moderate decreases in their affinity for the column as compared to the wild-type

sequence C1 (Figure 22B).

The central region (64-nucleotide) of the 100-mer oligonucleotide should have been totally randomized in the original pool; however, the cloned sequences from pool 8 appear to be depleted of cytosines (Figure 21). The overall base composition of the 64-nucleotide region of the 63 clones was calculated from the sequencing data. As indicated in Table 2, there is an obvious deficiency in cytosines in these sequences. This result could be either due to the selection procedure or because less cytosines were incorporated in the original pool. To determine whether the second possibility is reasonable, base compositions of pool 0 (the original pool) and pool 8 (the pool from which the 63 sequences were cloned) were obtained by HPLC analysis, and the results have been also listed in Table 2. Although cytosines were indeed depleted to a certain extent in the original pool, an enrichment for purines or a decrease in pyrimidines in the selected pool was obvious (Table 2). The HPLC procedures used here were shown to give rise to comparable results when they were applied to analyze a DNA sample with a known sequence (data not shown).

3. The selected aptamers did not bind to a mutant p53 protein

The pool 8 and eleven individual aptamer sequences from pool 8 that had been tested positive in their ability to bind to the EP240-Cys target column were tested for their affinity for a mutant p53 protein using gel-shift and immunoprecipitation (IP) assays. In the IP experiment, monoclonal antibodies PAb421 and PAb1801 which recognize the C-terminal and N-terminal regions, respectively, of both wild-type and mutant p53 proteins were used. It was determined that the presence of aptamer DNAs did not interfere with the binding of these antibodies to the protein (data not shown). In both assays, crude cellular extracts containing baculovirus-expressed p53 proteins were used. The control was

the wild-type p53 protein, while the mutant p53 was an Arg (175) to His mutation which has been shown by others (Gannon et al., 1990) and confirmed by myself to have the exposed 5-mer epitope (lane 6 in Figure 23). No binding to the whole mutant p53 protein by the tested sequences was detected through either assay (data not shown).

4. The selected aptamers did not bind to the target peptide in solution

In order to determine whether the column-selected aptamers have affinity for the target peptide in solution, the loading and washing steps of the column binding assay were carried out in buffer containing either peptide EP240 (sequence: NH_2-Phe-Arg-His-Ser-Val-Val-Val-COOH; the epitope core is underlined) or peptide EP240S (a control peptide with a scrambled epitope region; sequence: NH_2-Phe-Val-His-Val-Ser-Arg-Val-COOH). EP240 and EP240S are truncated versions of EP240-Cys and EP240S-Cys, respectively. At 0.25 mM, neither peptide inhibited the binding of pool 8 DNA to the immobilized target peptide (Figure 24A). At a higher concentration (1 mM), an inhibition was seen with the EP240 peptide but a similar inhibition was also observed with the control peptide (Figure 24B). Since it was determined that the enriched pools, including pool 8, or individual aptamers did not bind to the immobilized control peptide, the inhibition seen at the higher concentration of either peptide was probably due to a nonspecific effect. These observations suggest that pool 8 DNA did not have detectable affinity for the target peptide in solution.

5. Measurement of the dissociation constant (K_d) between clone C1 aptamer and the immobilized EP240-Cys peptide

Steady-state binding analysis was used to measure the affinity between clone C1 aptamer DNA and immobilized EP240-Cys peptides. Different

concentrations of DNA (^{32}P-labeled) were incubated with a fixed but relatively smaller amount of peptide-containing Sepharose beads. The amounts of DNA bound to the beads were obtained by measuring the radioactivity of the beads after these beads were separated from the solution by filtration. Figure 25 demonstrates a representative binding curve generated from the interaction between clone C1 aptamer DNA at concentrations of 0.25–25 μM and immobilized EP240-Cys peptides. From the biding curve, an apparent dissociation constant ($K_{d(app)}$) was estimated to be 9 μM. The binding isotherm appears to be sigmoidal in character.

D. DISCUSSION

1. Why use cleavable disulfide coupling?

At the time when the strategy of selection was being designed, most studies reported in the literature used free target molecules to elute the bound aptamers (affinity elution). The strategy of affinity elution showed some advantages when accessible small chemical compounds such as chemical dyes or cofactors were the selection targets, or when it was of interest to obtain aptamers that can bind to the target free in solution. The goal of this selection was to obtain DNA aptamers with affinity for mutant p53 proteins, and I proposed to achieve that goal through obtaining first aptamers that recognize a mutant p53-unique epitope. Since the configuration of the EP240 epitope in the form of free peptide may differ from the one that is in the context of a mutant p53 protein, it is not a necessity for the aptamers to have affinity for the free EP240-Cys peptide in order to be able to recognize the protein as a whole. On the other hand, an immobilized epitope may share some resemblance to the epitope region in a mutant p53 protein. Thus, it is reasonable to speculate that a DNA population recognizing the immobilized epitope may contain aptamers that have affinity for the protein. In addition, studies in which affinity elution was used to recover bound aptamer molecules

reported affinities of the selected ligands for the targets no tighter than micromolar level, supporting the notion that the so-called tight-binders cannot be easily recovered by an affinity elution procedure. Therefore, it was postulated that if the immobilized target molecules, e.g., the EP240-Cys peptide coupled to the Sepharose matrix in this study, are cleaved off from the solid support during elution, everything, including the putative tight-binders, will be recovered from the matrix. Based upon this rationale, a cysteine residue was added to the p53-derived EP240 peptide to facilitate the coupling of the peptide to thiol-activated Sepharose through a disulfide bond, which is stable under a variety of experimental conditions but susceptible to cleavage by strong thiol-reducing agents such as DTT. Also, there was a technical reason why a covalent elution procedure was chosen. Prior to the covalent elution strategy, I had attempted to use a variety of conditions including high salt concentrations, H_2O, EDTA, and free EP240 peptide (affinity elution) to elute the DNA molecules that bound to columns containing covalently immobilized EP240 peptides (data not reported). For unexplained reasons, high nonspecific binding of the ss DNA libraries to the immobilized target peptide, but not to the control peptide, was observed. Moreover, the nonspecific binding seemed to be increasing along with the progression of the selection process, suggesting that DNA sequences that had nonspecific affinity for the columns were being enriched by the selection procedures. Covalent elution was expected to avoid this problem because it should only recover those sequences that have bound to the target peptide while leaving behind those that are nonspecifically bound to the column matrix.

2. Why no consensus motifs?

In contrast to the results from most published studies, no consensus domain of 10 nucleotides or longer was identified among the 63 cloned aptamer sequences using a computer DNA homology search program. This relatively high

heterogeneity of the selected aptamer sequences could be due to the selection strategy applied in this study. As previously mentioned, the EP240-Cys target peptide was immobilized through a disulfide bond to the Sepharose column support and, after each binding reaction, DTT was used to cleave the peptide off the column support. This procedure would elute everything that had bound to the peptide molecules, including those desired tight-binders, off the column. Since there are probably many of possible binding sites on even a small peptide such as EP240-Cys, the eluted populations were expected to include all the sequences bound to these sites. Furthermore, since there was little competition among DNA molecules for their corresponding target sites on the peptide due to the relatively excessive amounts of peptides used in each selection cycle, it was unlikely any group of DNA species to emerge as a dominant population in the enriched pools. This argument is in agreement with the results from a recent study in which a similar strategy of covalent elution was used to select RNA aptamers against a peptide (Nieuwlandt et al., 1995).

The lack of homogeneity among the selected aptamer sequences could also be due to the possibility that the EP240-Cys peptide (NH_2-Thr-Phe-Arg-His-Ser-Val-Val-Val-Pro-Cys-COOH) is a good ss DNA-binding molecule. It is speculated that the selection procedures using this particular peptide as the target did not favor those DNA molecules that can be recognized sequence-specifically by the peptide. Rather however, the population has enriched for those DNA molecules that simply possess more single-strandedness. This speculation gains support from several experimental observations. First, as reported earlier, changing as many as six bases in the clone C1 sequence just reduced the binding ability to limited degrees, undermining the importance of the sequence itself in binding. Second, a folding analysis of the isolated sequences did not reveal any major stable secondary structure. Third, the complementary sequence of clone C1 demonstrated an affinity for the peptide column comparable to that of the clone C1 sequence itself

(data not shown). Since the folding analysis of the clone C1 sequence did not reveal any stable secondary structure, it is unlikely that the complementary strand of clone C1 bound to the column relatively well because it formed an appreciable secondary structure. Therefore, there might be no sequence-specific binding of clone C1 aptamer, or of all the other selected aptamers, by the immobilized peptide. What may be recognized by the immobilized peptide is the single-strandedness of the selected aptamers.

3. Why no binding to a whole mutant p53 protein?

As mentioned earlier, the ultimate goal of this study was to obtain DNA aptamers that can bind selectively to mutant p53 proteins. A mutant p53-unique epitope was used as the selection target, hoping that, as in the case antibody production, molecules that bind to this epitope would also have selective affinity for mutant p53 proteins that have the epitope exposed on the protein surface, but would not have affinity for the wild-type protein containing a buried epitope. However, attempts to determine whether the selected peptide-binding aptamers possess affinity for a mutant p53 protein did not reveal any positive signal using gel mobility shift and immunoprecipitation assays.

Since an affinity with a K_d value of lower than the micromolar level is usually needed for binding to be detected by these two assays, the negative result could indicate that the possible binding between the selected aptamers and the mutant p53 protein is not tight enough; i.e., the K_d is higher than 1 μM. Results from experiments measuring the affinity between clone C1 DNA and the immobilized EP240-Cys peptide showed a K_d value of ~9 μM, supporting this speculation. Methods of increasing the binding affinity of the selected aptamers for the target molecule include re-randomizing the selected sequences using mutagenic PCR (Cadwell and Joyce, 1992) and incorporating modified nucleotides

into the sequences. Both of these methods are designed to generate more diverse binding capacities. Possible applications of these methods to this study have been discussed in the Suggestions for Future Research chapter of this dissertation. Alternatively, it is possible that the conditions used in the two assays were not optimal for the binding under study. Finally, it is also possible that the configuration of the immobilized peptide was different from what is presented when the peptide is in the context of a mutant p53 protein. If this is the case, different selection strategies should be considered. For example, one could use a whole mutant p53 protein as the selection target to select for molecules that can disrupt the binding between the protein and antibody PAb240. Since PAb240 only recognizes the mutant conformation of p53 protein, those selected molecules that specifically inhibit the antibody's binding would probably contain molecules that compete with the antibody for its binding site on the mutant p53 protein.

4. What else could have been done?

As previously mentioned, the goal of this in vitro selection study was to obtain molecules that have acquired affinity for mutant p53 proteins. That goal was not achieved despite some interesting results. Although there are always some risks of using in vitro selection, the following steps could have been taken which, in this author's opinion, would probably increase the chance of achieving the goal:

a. Using an additional column of a different solid support in the selection procedure

As mentioned before, the selected DNA aptamers did not bind to the target peptide in solution, suggesting that the Sepharose matrix and/or the linker connecting the peptide to the matrix also contributed the observed binding. This effect would most likely diminish the possibility of these aptamers binding to

mutant p53 proteins because, obviously, these elements do not exist when the peptide is in the context of a protein. To eliminate any influence from the matrix or the linker region, one could couple the peptide onto another matrix of a different chemical nature using a different linker and introduce an additional selection step using the resultant column. This way, molecules that can bind to both types of columns will likely recognize the peptide portion and will be selected and amplified.

b. Introducing mutations into the DNA population in the course of selection by PCR mutagenesis

In addition to the step suggested above, mutations could have been introduced into the DNA pools by mutagenic PCR (Cadwell and Joyce, 1992) during the course of selection or, even in the early cycles to increase the complexities of the DNA pools. In this way, new sequences derived from those that have already shown some reactivities with the target peptides can be developed that could have an increased binding ability. This approach has been used in both aptamer selection (Lorsch et al., 1994; Huizenga and Szostak, 1995; Nieuwlandt et al., 1995) and ribozyme selection (Bartel and Szostak, 1993).

c. Using a whole mutant p53 protein as the selection target and taking advantage of the specific interaction between the PAb240 monoclonal antibody and the mutant p53 protein

The monoclonal antibody PAb240 recognizes only mutant p53 proteins by binding to a mutant p53-unique epitope (Gannon et al., 1990). This epitope, Arg-His-Ser-Val-Val, is in a buried position in the wild-type protein but becomes exposed apparently on the surface of the protein when mutations in other regions of the protein either induce conformational changes to the protein or block its

folding (Cho et al., 1994). An approach using a whole mutant p53 protein as the selection target and taking advantage of the mutant p53-specificity of PAb240 could be tried to obtain aptamers with selective affinity for mutant p53 proteins. Cellular extracts containing His175 mutant p53 are incubated with a ss DNA library, and sequences bound to the p53 protein can be recovered by immunoprecipitation with either monoclonal antibody PAb421 (recognizes the C-terminus of the protein) or PAb1801 (recognizes the N-terminus). The recovered DNA population should include the mutant p53-unique epitope binding sequences as well as those DNA molecules that bind to other regions of the mutant p53 protein. Those undesired sequences can be depleted by immunoprecipitation with PAb240 because the epitope-binding sequences cannot bind to the mutant p53 protein in the presence of this antibody. The isolated potential DNA aptamers are then amplified by PCR and subjected to further rounds of selection.

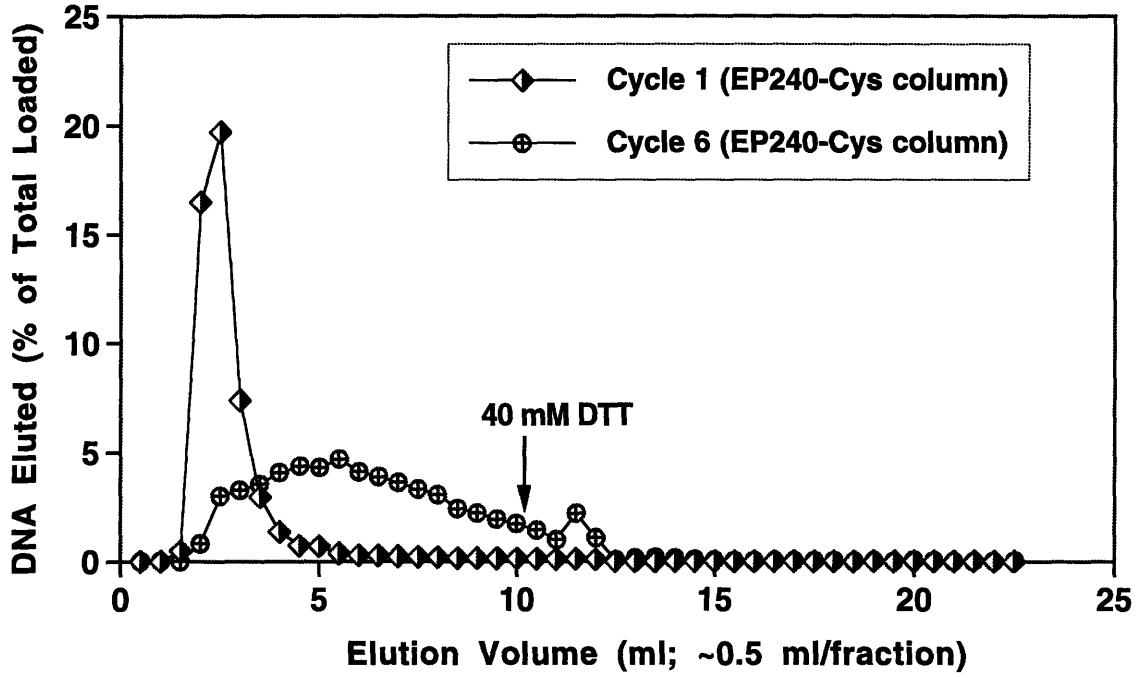
E. CONCLUSIONS

By using an in vitro selection procedure involving rounds of selection and amplification, groups of DNA aptamers were obtained with selective affinity for columns with an immobilized peptide containing a mutant p53-specific epitope. The selected DNA aptamers had little affinity for columns containing a control peptide in which the sequence of the epitope region was scrambled, or for the target peptide free in solution, indicating that both the sequence of the target peptide and the immobilization step probably contributed jointly to the observed binding. Sequence analysis did not reveal any consensus motifs or common secondary structures among the selected aptamer sequences. Steady-state binding analysis using a cloned aptamer sequence (clone C1) as a model molecule revealed that the apparent disassociation constant ($K_{d(app)}$) between this aptamer and the immobilized peptide is approximately 9 μM. However, no binding between the epitope-binding aptamers and a mutant p53 protein was demonstrated by either gel

mobility shift or immunoprecipitation assays. The results suggest that it is possible to obtain molecules through in vitro selection that can recognize a target as small as a peptide. This conclusion provides a rationale for designing therapeutic agents that target a particular region of a disease-related protein (e.g., the ligand binding site of a receptor or the catalytic site of an enzyme).

Figure 18. Selection of DNA aptamers from a randomized ss DNA library. In each cycle, ~500 pmol of ^{32}P-radiolabeled ss DNA population in ~500 µl of aptamer selection buffer (ASB; 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 1 mM EDTA) was loaded onto a selection column with a 1.6-ml bed volume containing EP240-Cys epitope peptides. The columns were washed with either ~10 ml (cycles 1–6) or ~20 ml (cycles 7–9) of ASB before being eluted with 10–14 ml of ASB containing 40 mM DTT. The washing and elution steps were carried out at ~0.5 ml/fraction. Starting from the third fraction after DTT was included in the elution buffer, 16–18 fractions of DTT elution were collected, combined and concentrated. Approximately one-fifth of the recovered DNA population was subjected to further selection cycles. In cycles 7–9, a control column (also referred to as a pre-column) was used prior to the selection column. The pre-column also had a 1.6-ml bed volume but contained the EP240S-Cys control peptide in which the sequence of the epitope region was scrambled. In these cycles, the DNA samples were first loaded onto the pre-columns. Fractions containing the unbound DNAs were combined, concentrated down to ~500 µl and loaded directly onto the selection column. **A.** Column elution profiles from the selection cycles 1 and 6. **B.** Column elution profiles from the control (EP240S-Cys) and selection (EP240-Cys) columns in cycle 8.

A


Cycle 1 (EP240-Cys column)
Cycle 6 (EP240-Cys column)
40 mM DTT
DNA Eluted (% of Total Loaded)
Elution Volume (ml; ~0.5 ml/fraction)
0
5
10
15
20
25

B

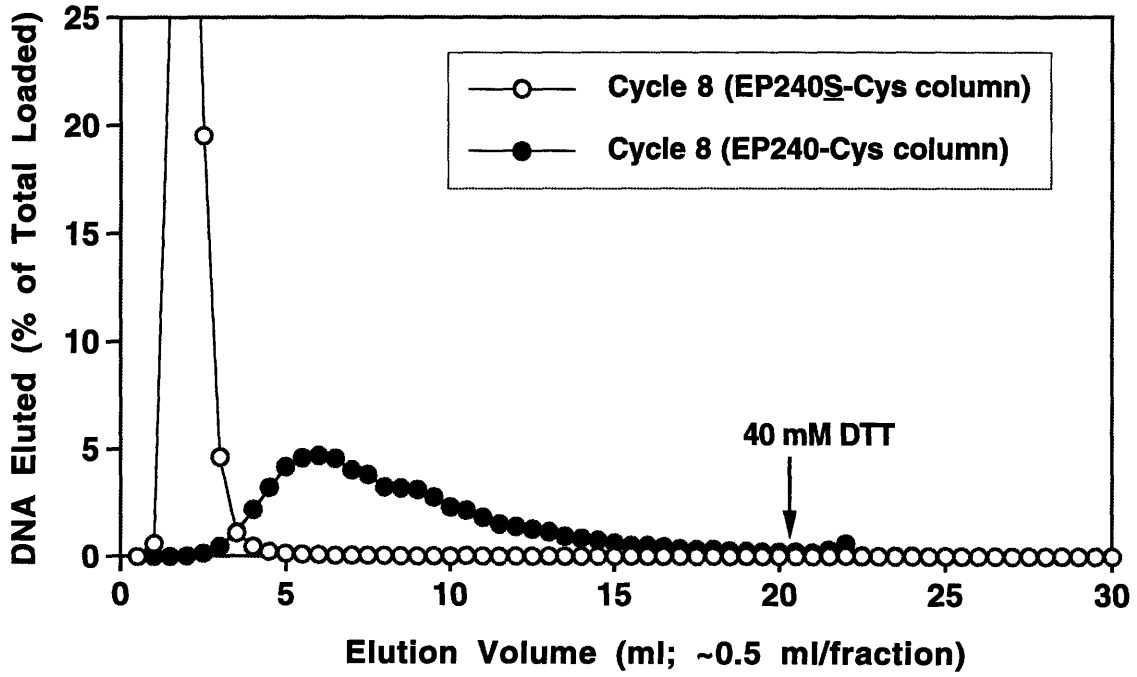

Cycle 8 (EP240S-Cys column)
Cycle 8 (EP240-Cys column)
40 mM DTT
DNA Eluted (% of Total Loaded)
Elution Volume (ml; ~0.5 ml/fraction)
0
5
10
15
20
25
30

Figure 19. Column binding analysis of the DNA populations (pools) derived from selection cycles. EP240-Cys columns of 250-µl bed volume were used to analyze the binding ability of various DNA pools derived from corresponding selection cycles (e.g., pool 3 was the DNA population after 3 cycles of selection) to bind the immobilized target peptide. A buffer of a lower salt concentration (25 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM EDTA) as compared to the aptamer selection buffer (ASB; 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 1 mM EDTA) was used in these experiments. In each analysis, 100 pmol of ^{32}P-radiolabeled DNA in 100 µl of buffer was loaded onto the column, followed by washing and DTT elution. The washing and elution steps were carried out at ~84 µl/fraction.

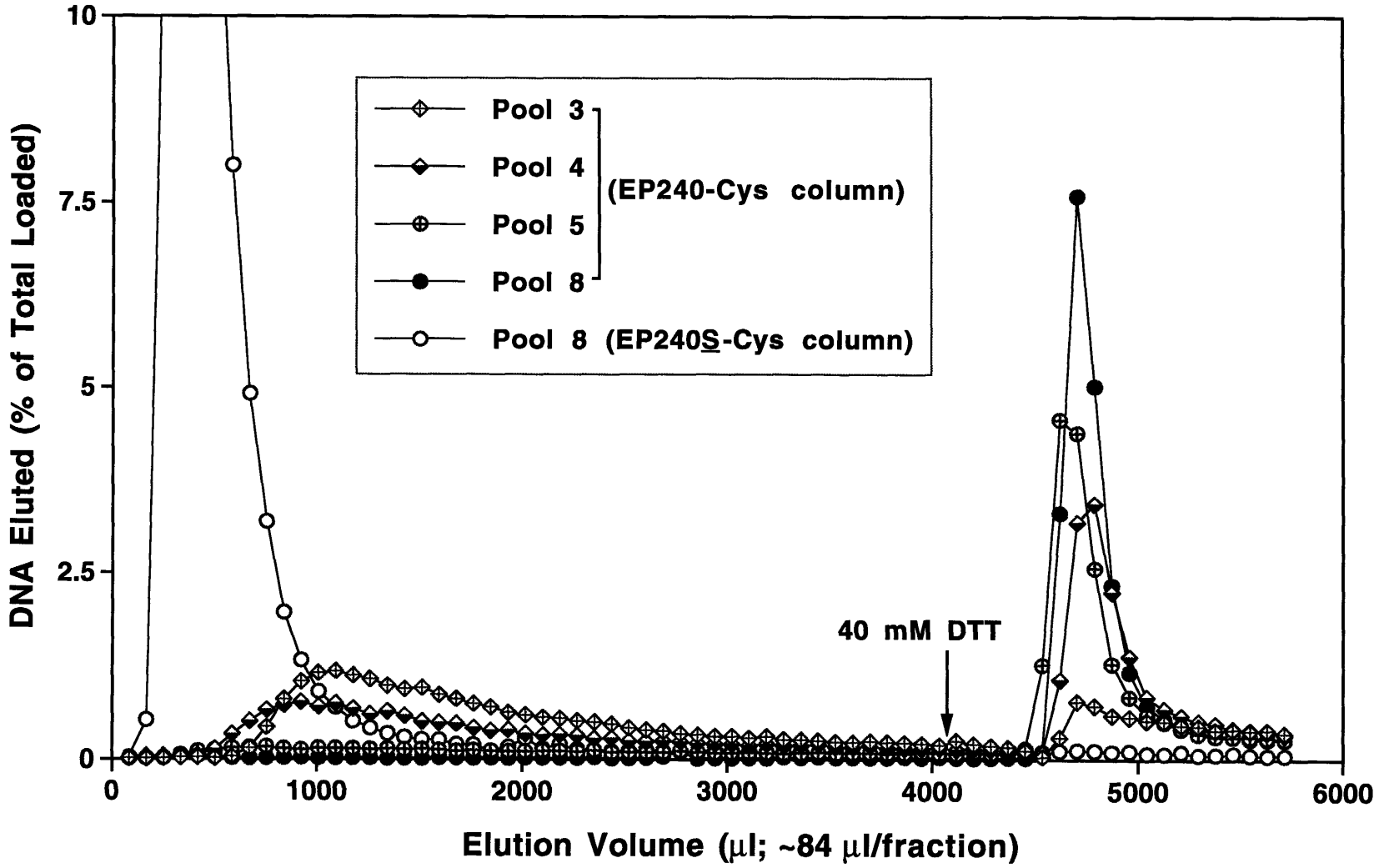

10
7.5
5
2.5
0
DNA Eluted (% of Total Loaded)
Pool 3
Pool 4
Pool 5
Pool 8
(EP240-Cys column)
Pool 8 (EP240S-Cys column)
40 mM DTT
0
1000
2000
3000
4000
5000
6000
Elution Volume (µl; ~84 µl/fraction)

Figure 20. Analysis of the DNA pools with restriction endonucleases. DNA from each pool (0.5–2 pmol) was amplified with 25 PCR cycles as described in the Materials and Methods of this chapter. The primer P-1 was ^{32}P-radiolabeled (~10^6 cpm per PCR reaction). Fractions of ^{32}P-labeled double-stranded PCR products were subjected to digestion with various 4-base restriction endonucleases, followed by electrophoreses on 8% denaturing polyacrylamide gels. Results of the digestion with HaeIII (recognizes 5'-GGCC-3' sites) and MseI (recognizes 5'-TTAA-3' sites) are shown.

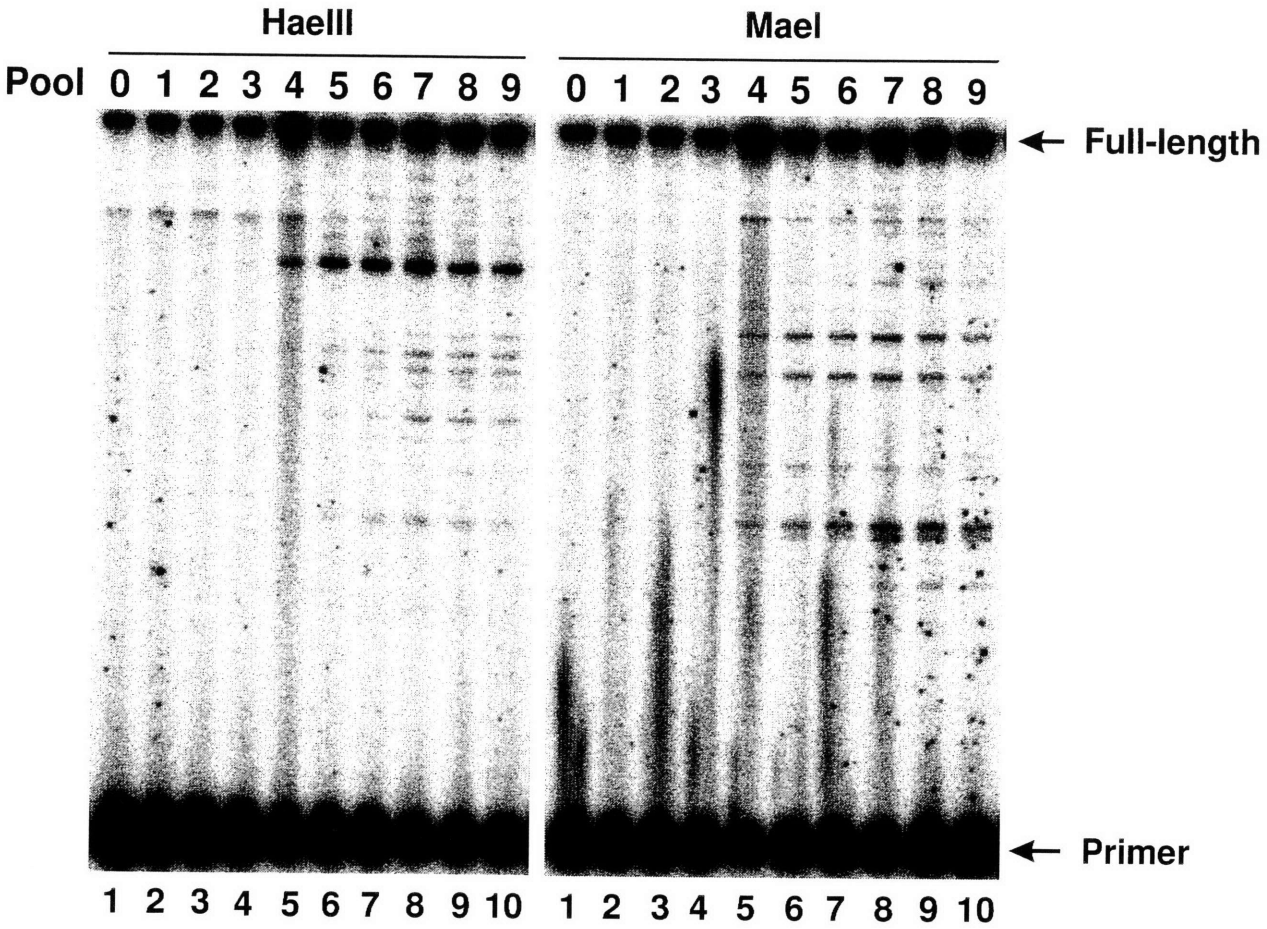

HaeIII
MaeI
Pool
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
Full-length
Primer
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10

Figure 21. Sequences of the selected aptamers. Seventeen of the 52 unique sequences cloned were tested for their ability to bind to the EP240-Cys column, and the results showed that they all bound to the column. The sequences shown are the 64-nucleotide central regions. The number in parentheses indicates the numbers of clones sharing the same sequence.

```
A1        CCGAAACGCGGAGCCCAGGGACACGTTGTTCTAGCCTTTAAGTAAGTAGATAAGCTTACGCGCT
A2  (2)                  -

B1  (2)   GGGGAGAATCGTAACGTAAATTGGAGAATTGGGAGTGAGCAGGAGGATGCTGCGTCGATTCG

C1  (2)   CAGCGAGGCGAAGGATGAGCCGTCGTTTGGATCAAAGTTTTGAAGTTTTTGCAATCGTAGCTAG

D1        GCAACGGAGGGCATAGTGCGCGGGATTTTAGCGGATGAAGACTGAAGTCTGTTGTCCTCTGTAT
D2                                             T

E1  (2)   GACCAGGAAGTCAGGAGCGCGTTTCGCGTTGGTGGACGTGGCCATGAACTAATTCATGGTCTTG

F1  (2)   GGACGATAGCATGAAAGTACTTGCTTTGTCTTGCGAGAGAACGCGATCCGCGATCTCGTTTGTC

G1  (2)   GAGTTAGTGACAGGGGCAATTTATGGGTACGGGCGGTATTATAGCTTCCTAGGGAGGATACTGA

H1        GAGAGTAACGGAGTGAGGTTTAAAGTAGAGTAAAGATAAGTAGGCAATACTGTAAACCGTCGTG
H2                                                                      A

I1        GCCCCTCTTGCATACTATGCGGGAGTGTTTAGATCGGAAAATTTATTGGGTGGACCAAACAGG
I2        CGCAGTGAGATCCTCTTGGGGGTCGTTTGTCGGCACACCTAGAGTGGCGTGGTGACAATAACGA
I3        CGCAGTCGGAAGTGAAGTTGTTGGAAGGAACTGTTGGTTGGTGTCAGGCAGCCATAGCTTCGCT
I4        GGTTAATGCCGAGTGTCTACGTTTGAGGCGACTGTAAGAATTATACGACGAAATAGGTATTGGA
I5        CAGGCCGAGATGGTCCCGACCATCGGTTGGGAAGCATCGTCGGTAGGTTACCGTAGCATGCGT
I6        GGAGTGTAGGGAGCACCGGACACGTTCTAAGTAGGCTGGGTGTGGACGAACGAAGTCACGAT
I7        GAAGGAGTTGTCTTGCAACGGATGTTATGCTGGTTCGCAATTCATTGGGGACTGTTGAGCATAT
I8        GGGTGGATGCAACGGTAGCGAATGACGAGAAGGTTTTAACTTCGTTCGGGTCTTTTATAGATCA
I9        CATGAGAGAACGAAGAGGTTAGGGATGTCAACGAGCTTTGTATCGTTGTAGTGTTATCATGTA
```

Figure 22. Binding of chemically synthesized clone C1 sequence and two sequences mutated from clone C1 to the EP240-Cys column. Clone C1 and two sequences that differ from C1 at 6 positions (**A**) were chemically synthesized, and their relative binding ability to the EP240-Cys target column was tested (**B**).

A

```
C1:  CAGCGAGGCGAAGGATGAGCCGTCGTTTGGATCAAAGTTTTGAAGTTTTTGCAATCGTAGCTAG
C1m1:                              CC            CC          CC
C1m2:                             C  C          C   C  C      C
```

B

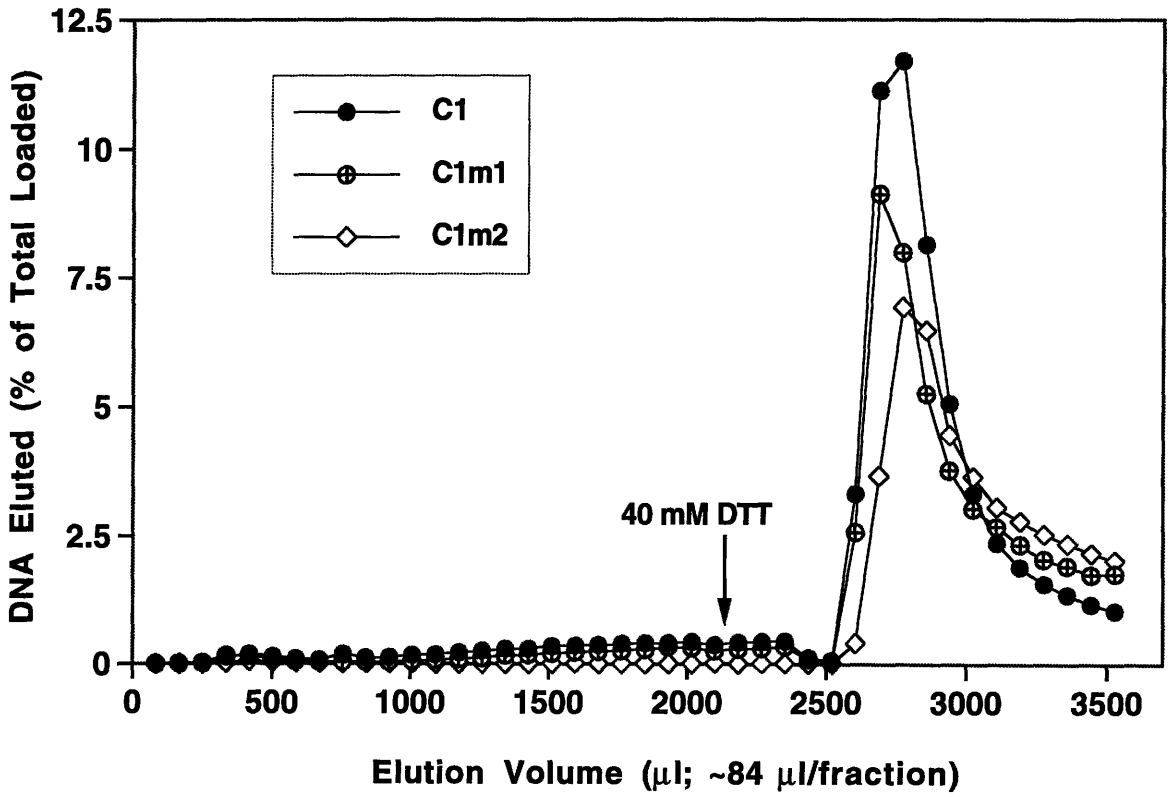


Figure 23. Recognition of the baculovirus-expressed His175 mutant p53 by monoclonal antibody PAb240. Monoclonal antibodies PAb421 (recognizes the C-termini of both wild-type and mutant p53 proteins), PAb1801 (recognizes the N-termini of both wild-type and mutant p53 proteins) and PAb240 (recognizes the Arg-His-Ser-Val-Val epitope that is only available on the surface of some mutant p53 proteins) were used in immunoprecipitation assays to confirm: (1) the expression of the wild-type and His175 mutant p53 proteins in cells infected by the p53-expressing baculoviruses; and (2) the recognition of the His175 mutant p53 protein by the PAb240 monoclonal antibody. Equal amounts of crude ^{35}S-labeled cell extracts were used in this experiment, and the immuno-complexes were analyzed on a 12% SDS-PAGE gel. Since it was not determined that the two types of cell extracts contained an equal level of p53 proteins, no quantitative comparison should be made between the wild-type and the mutant p53 lanes. The relative ability of each antibody to bind either type of the p53 proteins, however, can be estimated by comparing the intensities of the p53 bands among lanes of either wild-type or mutant p53.

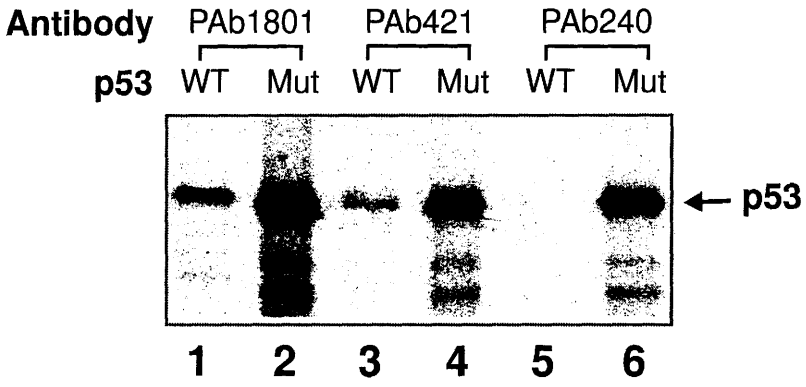
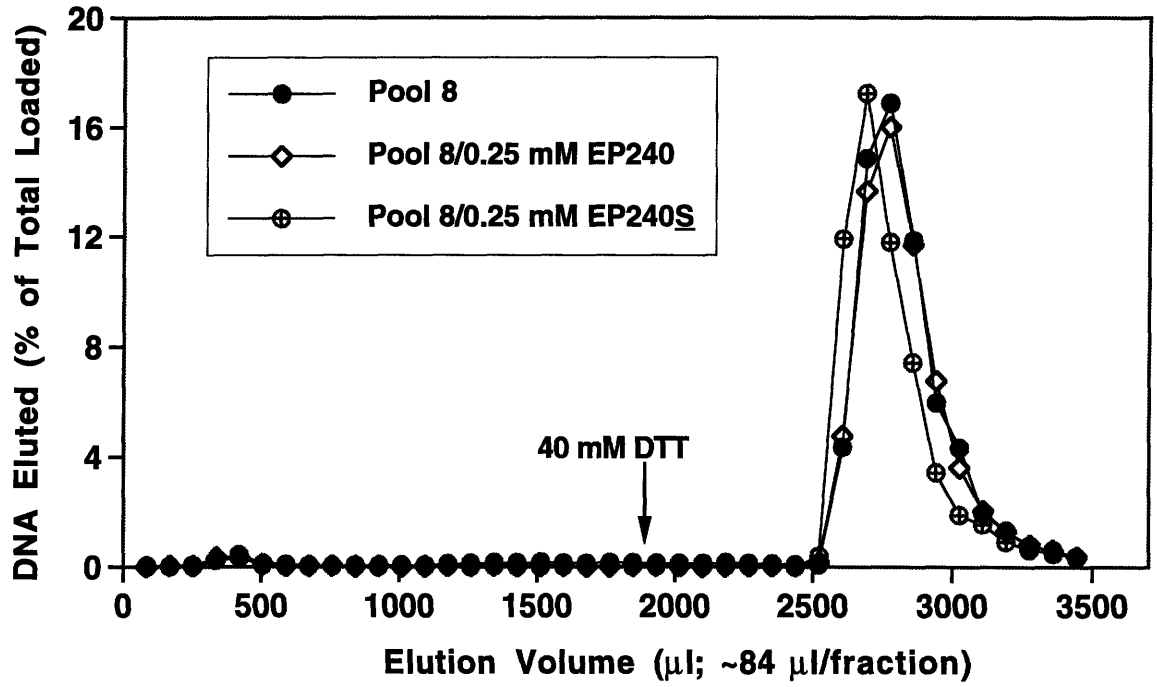

Antibody
PAb1801
PAb421
PAb240
p53
WT
Mut
WT
Mut
WT
Mut
p53
1
2
3
4
5
6

Figure 24. Pool 8 DNA did not bind to the EP240 peptide in solution. The EP240-Cys analytical columns were used in this experiment. The loading and washing steps of column binding assays were carried out with binding buffer containing either 0.25 mM (**A**) or 1 mM (**B**) of EP240 or EP240S peptides. As stated in the text, EP240 peptide (Phe-Arg-His-Ser-Val-Val-Val) is shorter than the target peptide (EP240-Cys; Thr-Phe-Arg-His-Ser-Val-Val-Val-Pro-Cys), but both peptides contain the epitope region (underlined). EP240S shares the same amino acid composition as EP240 but contains a scrambled epitope region.

A


Pool 8
Pool 8/0.25 mM EP240
Pool 8/0.25 mM EP240S
40 mM DTT
DNA Eluted (% of Total Loaded)
Elution Volume (µl; ~84 µl/fraction)
0
4
8
12
16
20
0
500
1000
1500
2000
2500
3000
3500

B

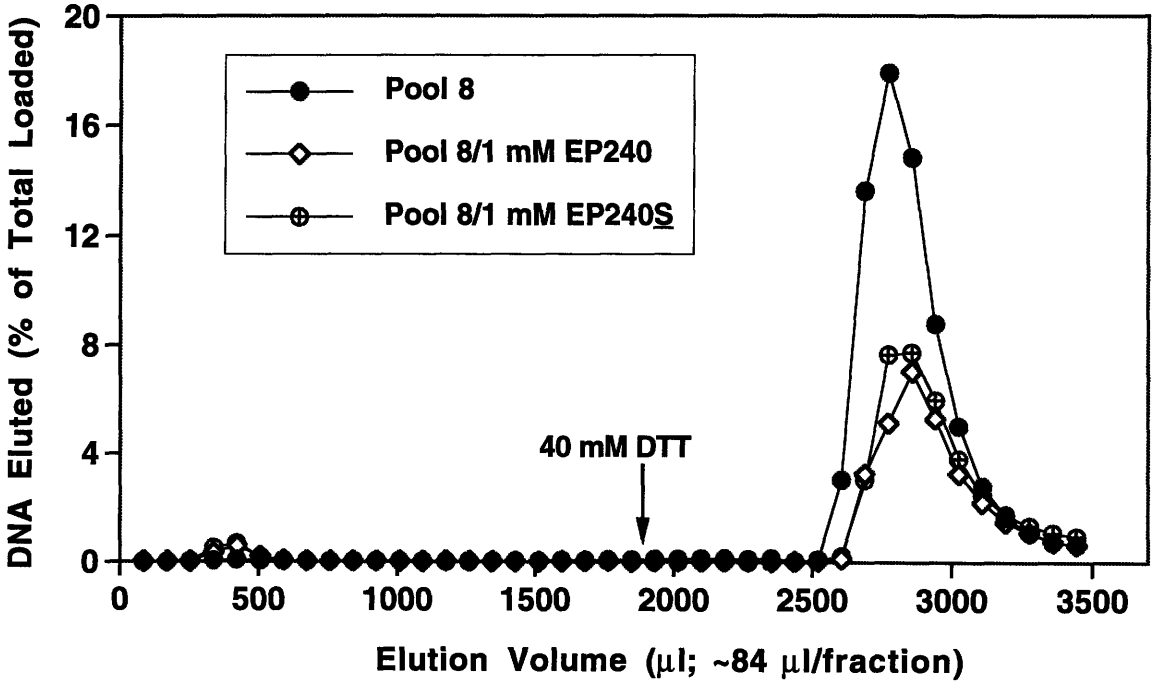

Pool 8
Pool 8/1 mM EP240
Pool 8/1 mM EP240S
40 mM DTT
DNA Eluted (% of Total Loaded)
Elution Volume (µl; ~84 µl/fraction)
0
4
8
12
16
20
0
500
1000
1500
2000
2500
3000
3500

Figure 25. Steady-state binding analysis of clone C1 to the immobilized EP240-Cys peptide. In each binding reaction, Sepharose beads containing EP240-Cys peptides (with a final apparent concentration of ~0.5 μM) were incubated with either 0.25 μM, 1 μM, 2.5 μM, 5 μM, 7.5 μM, 10 μM, 15 μM or 25 μM of ^{32}P-labeled clone C1 DNA in 100 μl of Tris-HCl pH 7.5, 125 mM NaCl and 1 mM EDTA. Different clone C1 concentrations were achieved by mixing a fixed amount of ^{32}P-labeled C1 (~2 × 10^5 cpm per incubation) with various amounts of unlabeled C1 DNA. The incubations were carried out in 1.5 ml eppendorf tubes at room temperature for 3 hr with end-to-end rotation. After incubation the samples were transferred into microfilters, which were then spun for 2 min in a microcentrifuge to separate the liquid and the beads. The retained beads were counted in a scintillation machine and the resultant data were converted to the amounts of bound clone C1 DNA, which were subsequently plotted against the initial C1 DNA concentration used in each incubation.

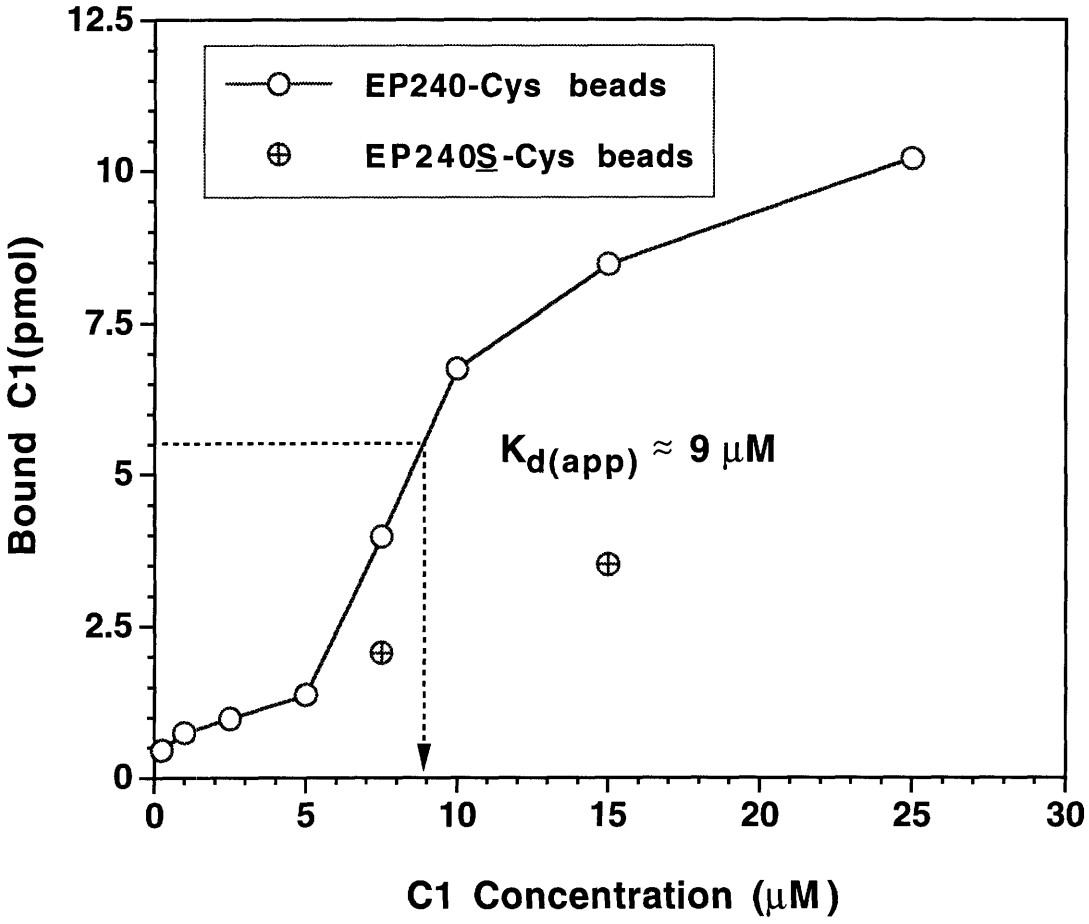

12.5
10
7.5
5
2.5
0
0
5
10
15
20
25
30
Bound C1(pmol)
C1 Concentration (μM)
EP240-Cys beads
EP240S-Cys beads
$K_{d(app)} \approx 9\ \mu M$

Table 2. DNA base compositions of pools 0 and 8. The DNA base composition of pool 8 was obtained either from the sequences of the 63 clones, or by HPLC analysis of a nucleoside digestion. A similar HPLC analysis was also performed to obtain the base composition of pool 0 (the original pool) DNA. The numbers in the shaded row are the relative incorporation efficiencies of an equimolar mixture of dG, dA, dT and dC phosphoramidites (Evaluating and Isolating Synthetic Oligonucleotides, Applied Biosystems).

	dG	dA	dT	dC
Pool 0 (theory)	26.0%	20.0%	30.0%	24.0%
Pool 0 (HPLC)	28.4%	19.2%	30.4%	22.1%
Pool 8 (HPLC)	30.8%	23.4%	26.2%	19.5%
Pool 8 (sequencing)	32.1%	25.1%	24.8%	18.0%

VI. SUGGESTIONS FOR FUTURE RESEARCH

A. THE UV-DRP

Studies can be carried out to investigate further (1) the structural basis of the high affinity recognition of the 6-4 photoproduct by the UV-DRP; (2) the mechanism(s) for this protein's role in facilitating the removal of the 6-4 photoproduct; and (3) the evolutionary role for this protein in repair of DNA lesions formed by UV irradiation and, more broadly, by other agents.

B. THE TMP-BIOTIN BIFUCTIONAL CONJUGATE

The approach of using psoralen-biotin conjugate as a test molecule has the advantage of being straightforward because the biochemical features of the interactions involved (e.g., biotin-streptavidin interaction, uracil glycosylase enzymatic reaction, and psoralen-DNA interaction) have been well characterized. This simplicity allows detailed and systematic studies on factors that will ultimately contribute to the effectiveness of a drug. These factors include the solubility of the drug, the nature of the chemical warhead, and the length and chemical nature of the linker between the warhead and a protein-binding domain.

1. Experimenting with different warheads

An effective warhead should lead the drug to attack the cellular genome, i.e., double-stranded DNA, with high selectivity. Also, its reactivity with DNA should remain high even when tethered to a protein-binding ligand. The requirement of near UV light activation for psoralen to interact with DNA, however, severely limits its potential as a warhead for a real anticancer drug. Currently, the idea of using molecules that alkylate DNA upon intercalation is under consideration.

2. Experimenting with linkers that have different but more favorable chemical properties

Linkers with different chemical properties should be studied for their effects on the solubility of the compound, the reactivity of the compound with both the DNA target and its interacting protein. The possible contributions of the linker localizing the compound to, for example, the cell nucleus should also be studied.

3. Optimizing the length of the linker

The length of the linker is critical for a good balance between the binding of a protein to its ligand and shielding of the DNA adducts by the bound protein. If the linker is too short, the DNA helix would probably impose a sterically inhibitory effect on the binding of the ligand domain to its target protein once the bifunctional compound is conjugated with its DNA target. On the other hand, if the linker is too long, the ability of the protein, even though it can bind readily to the DNA adduct sites, to interfere with the removal of the adducts by DNA repair systems would be compromised because there still could be enough space between the bound protein and the actual DNA lesion site for the DNA repair enzymes to operate.

C. THE SELECTION OF MUTANT p53-BINDING APTAMERS

1. Studying the biophysical factors of the observed binding

Although the original goal of this study has not yet been fully achieved, it is still of biophysical interest to understand why certain single-stranded DNA sequences can bind to a short peptide sequence-specifically.

2. Increasing the binding affinity of a selected DNA aptamer

The possibility of applying PCR mutagenesis (Cadwell and Joyce, 1992) to generate new sequences from the selected sequences, and modified (unnatural) bases to expand the binding capacities of the aptamer molecules should be explored to increase the binding affinities of the selected aptamers toward their targets.

3. Obtaining mutant p53 protein-binding aptamers

The possibility of using ways (e.g., the filter-binding assay), other than gel-shift or immunoprecipitation assays, should also be explored to seek the chances of obtaining potential aptamers that can bind mutant p53 proteins from the selected aptamer populations.

4. Using other molecular libraries to find molecular attractants for TSPs

Given the limited but encouraging results from this study, it could be fruitful to do the in vitro selection using other libraries such as small chemical molecular libraries, such as those recently developed by the Rebel laboratory here at MIT (Carell et al., 1994a,b; 1995), and those of others (Achier and Crook, 1995).

VII. REFERENCES

Abousskehra, A., and Wood, R.D. (1994). Repair of ultraviolet light-damaged DNA by mammalian cells and *Saccharomyces cerevisiae*. Curr. Opin. Genet. Dev. *4*, 212-220.

Aboussekhra, A., Biggerstaff, M., Shivji, M.K.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protić, M., Hübscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859-868.

Abramic, M., Levine, A.S., and Protić, M. (1991). Purification of an ultraviolet-inducible, damage-specific DNA-binding protein from primate cells. J. Biol. Chem. *266*, 22493-22500.

Addison, C., Jenkins, J.R., and Stürzbecher, H.W. (1990). The p53 nuclear localization signal is structurally lined to a $p34^{cdc2}$ kinase motif. Oncogene *5*, 423-426.

Agrawal, S. (1992). Antisense oligonucleotides as antiviral agents. Trends Biotechnol. *10*, 152-158.

Aguilar, F., Hussain, S.P., and Cerutti, P. (1993). Aflatoxin B_1 induces the transversion of G→T in codon 249 of the *p53* tumor suppressor gene in human hepatocytes. Proc. Natl. Acad. Sci. USA *90*, 8586-8590.

Ahuja, H., Bar-Eli, M., Advani, S.H., Benchimol, S., Cline, M.J. (1989). Alterations in the *p53* gene and the clonal evolution of the blast crisis of chronic myelocytic leukemia. Proc. Natl. Acad. Sci. USA *86*, 6783-6787.

Althaus, F.R., and Richter, C. (1987). ADP-ribosylation of proteins, enzymology and biological significance. Mol. Biol. Biochem. Biophys. *37*, 1-237.

Ames, B.N., and Gold, L.S. (1990). Chemical carcinogenesis: too many rodent carcinogens. Proc. Natl. Acad. Sci. USA *87*, 7772-7776.

Ananthaswamy, H.N., and Pierceal, W.E. (1990). Molecular mechanisms of ultraviolet radiation carcinogenesis. Photochem. Photobiol. *52*, 1119-1136.

Bakalkin, G., Yakovleva, T., Selivanova, G., Magnusson, K.P., Szekzly, L., kiseleva, E., Klein, G., Terenius, L., and Wiman, K.G. (1994). p53 binds single-stranded DNA ends and catalyzes DNA renaturation and strand transfer. Proc. Natl. Acad. Sci. USA *91*, 413-417.

Baker, S.J., Fearon, E.R., Nigro, J.M., Hamilton, S.R., Preisinger, A.C., Jessup,

J.M., van Tuinen, P., Ledbetter, D.H., Barker, D.F., Nakamura, Y., White, R., and Vogelstein, B. (1989). Chromosome 17 deletions and *p53* gene mutations in colorectal carcinomas. Science *244*, 217-221.

Baker, S.J., Markowitz, S., Fearon, E.R., Wilson, J.K.U., and Vogelstein, B. (1990). Suppression of human colorectal carcinoma cell growth by wild-type p53. Science *249*, 912-915.

Banaszuk, A.M., Deugau, K.V., Sherwood, J., Michalak, M., and Glick, B.R. (1983). An efficient method for the sequence analysis of oligonucleotides. Anal. Biochem. *128*, 281-286.

Bargonetti, J., Reynisdottir, J., Friedman, P.N., Kern, S.E., Vogelstein, and Prives, C. (1991). Wild-type but not mutant p53 immunopurified proteins bind to sequence adjacent to the SV40 origin of replication. Cell *65*, 1083-1091.

Bargonetti, J., Manfredi, J.J., Chen, X., Marshak, D.R., and Prives, C. (1993). A proteolytic fragment from the central region of p53 has marked sequence-specific DNA binding activity when generated from wild-type but not from oncogenic mutant p53 protein. Genes Dev. *7*, 2565-2574.

Bartel, D.P., Zapp, M.L., Green, M.R., and Szostak, J.W. (1991). HIV-1 REV regulation involves recognition of non-Watson-Crick base pairs in viral RNA. Cell *67*, 529-536.

Bartel, D.P., and Szostak, J.W. (1992). Isolation of new ribozymes from a large pool of random sequences. Science *261*, 1411-1418.

Bayle, J.H., Elenbaas, B., and Levine, A.J. (1995). The carboxyl-terminal domain of the p53 protein regulates sequence-specific DNA binding through its nonspecific nucleic acid-binding activity. Proc. Natl. Acad. Sci. USA *92*, 5729-5733.

Bianchi, M.E., Falciola, L., Ferrari, S., and Lilley, D.M.J. (1992). The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins. EMBO J. *11*, 1055-1063.

Bischoff, J.R., Friedman, P.N., Marshak, D.R., Prives, C., Beach, D. (1990). Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. Proc. Natl. Acad. Sci. USA *87*, 4766-4770.

Blake, J., and Litzi-Davis, L., (1992). Evaluation of peptide libraries: An iterative strategy to analyze the reactivity of peptide mixtures with antibodies. Bioconjugate

Chem. *3*, 510-513.

Bock, L.C., Griffin, L.C., Latham, J.A., Vermass, E.H., and Toole, J.J. (1992). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. Nature *355*, 564-566.

Brash, D.E., and Haseltine, W.A. (1982). UV-induced mutation hotspots occur at DNA damage hotspots. Nature *298*, 189-192.

Brash, D.E., Franklin, W.A., Sancar, G.B., Sancar, A., and Haseltine, W.A. (1985). *Escherichia coli* DNA photolyase reverses cyclobutane pyrimidine dimers but not pyrimidine-pyrimidone (6-4) photoproducts. J. Biol. Chem. *260*, 11438-11441.

Brash, D.E., Rudolph, J.A., Simon, J.A., Lin, A., McKenna G.J., Baden, H.P., Halperin, A.J., and Pontén, J. (1991). A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc. Natl. Acad. Sci. USA *88*, 10124-10128.

Brenner, S., Lerner, R.A. (1992). Encoded combinatorial chemistry. Proc. Natl. Acad. Sci. USA *89*, 5381-5383.

Bressac, B., Kew, M., Wands, J., and Ozturk, M. (1991). Selective G to T mutations of *p53* gene in hepatocellular carcinoma from southern Africa. Nature *350*, 429-431.

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergård, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Nordenskjold, M., Fishel, R., Kolodmer, R., and Liskay, R.M. (1994). Mutation in the DNA mismatch repair gene homolog *hMLH1* is associated with hereditary non-polyposis colon cancer. Nature *368*, 258-261.

Brown, S.J., Kellett, P.J., and Lippard, S.J. (1993). Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin. Science *261*, 603-605.

Bruhn, S.L., Pil, P.M., Essigmann, J.M., Housman, D.E., and Lippard, S.J. (1992). Isolation and characterization of human cDNA clones encoding a high mobility group box protein recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. Proc. Natl. Acad. Sci. USA *89*, 2307-2311.

Cadel, J., and Vigny, P. (1990). The photochemistry of nucleic acids. In Bioorganic Photochemistry, H. Morrison, eds. (Now York: John Wiley & Sons), Vol. 1, pp. 1-272.

Cadwell, R.C., and Joyce, G.F. (1992). Randomization of genes by PCR mutagenesis. PCR Methods Appl. *2*, 28-33.

Carell, T., Wintner, E.A., Bashir-Hashemi, A., and Rebek, J., Jr. (1994). A novel procedure for the synthesis of libraries containing small organic molecules. Angew. Chem. Int. Ed. Engl. *33*, 2059-2061.

Carell, T., Wintner, E.A., and Rebek, J., Jr. (1994). A solution phase screening procedure for the isolation of active compounds from a library of molecules. Angew. Chem. Int. Ed. Engl. *33*, 2061-2164.

Carell, T., Wintner, E.A., Sutherland, A.J., Rebek, J., Jr., Dunayevskiy, Y.M., and Vouros, P. (1995). New promise in combinatorial chemistry: synthesis, characterization, and screening of small-molecule libraries in solution. Chem. Biol. *2*, 171-183.

Carthew, R.W., Chodosh, L.A., and Sharp, P.A. (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell *43*, 439-448.

Cavanaugh, A.H., Hempel, W.M., Taylor, L.J., Rogalsky, V., Todorov, G., and Rothblum. L.I. (1995). Activity of RNA polymerase I transcrption factor UBF blocked by *Rb* gene product. Nature *374*, 177-180.

Chambers, R.W., Sledziewska-Gojska, E., Hirani-Hojatti, S., and Borowy-Borowski, H. (1985). *uvrA* and *recA* mutations inhibit a site-specific transition produced by a single O^6-methylguanine in gene *G* of bacteriophage ϕX174. Proc. Natl. Acad. Sci. USA *82*, 7173-7177.

Chatterjee, S., Johnson, P., Wong, K.K. (1992). Dual-target inhibition of HIV-1 in vitro by means of adeno-associated virus antisense vector. Science *258*, 1485-1488.

Chen, P.-L., Chen, Y., Bookstein, R., and Lee, W.-H. (1990). Genetic mechanisms of tumor suppressor by the human *p53* gene. Science *250*, 1576-1580.

Chiba, I., Takahashi, T., Nau, M.M., D'Amico. D., Curiel, D.T., Mitsudomi, T., Buchhagen, D.L., Carbone, D., Piantadosi, S., Koga, H., Reissman, P.T., Slamon, D.J., Holmes, E.C., and Minna, J.D. (For the Lung Cancer Study Group) (1990). Mutations in the *p53* gene are frequent in primary, resected non-small cell lung cancer. Oncogene *5*, 1603-1610.

Chin, K.-V., Ueda, K., Pastan, I., and Gottesman, M.M. (1992). Modulation of

activity of the promoter of the human MDR1 gene by Ras and p53. Science *255*, 459-462.

Cho, Y., Gorina, S., Jeffrey, P.D., and Pavletich, N.P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science *265*, 346-355.

Chu, G., and Chang, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. Science *242*, 564-567.

Chu, G., and Chang, E. (1990). Cisplatin-resistant cells express increased levels of a factor that recognizes damaged DNA. Proc. Natl. Acad. Sci. USA *87*, 3324-3328.

Chu, Y.-H., Avilla, L.Z., Biebuyck, H.A., and Whitesides, G.M. (1993). Using affinity capillary electrophoresis to identify the peptide in a peptide library that binds most tightly to vancomycin. J. Org. Chem. *58*, 648-652.

Chubb, J.M., and Hogan, M.E. (1992). Human therapeutic based on triple helix technology. Trends Biotechnol. *10*, 132-136.

Cimino, G.D., Gamper, H.B., Isaacs, S.T., and Hearst, J.E. (1985). Psoralen as photoreactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. Ann. Rev. Biochem. *54*, 1151-1193.

Clackson, T., and Wells, J.A. (1994). In vitro selection from protein and peptide libraries. Trends Biotechnol. *12*, 173-184.

Cleaver, J.E., and Kraemer, K.H. (1989). Xeroderma pigmentosum. In The Metabolic Basis of Inherited Diseases, Sixth Edition, C.R. Scriver, A.L. Beaudet, W.S. Sly, and D. Valle, eds. (New York: McGraw-Hill), pp. 2949-2971.

Clugston, C.K., McLaughlin, K., Kenny, M.K., and Brown, R. (1992). Binding of human single-stranded DNA binding protein to DNA damaged by the anticancer drug *cis*-diamminedichloroplatinum(II). Cancer Res. *52*, 6375-6379.

Connell, G.J., Illangesekare, M., and Yarus, M. (1993). Three small ribonucleotides with specific arginine sites. Biochemistry *32*, 5495-5502.

Connell, G.J., and Yarus, M. (1994). RNAs with dual specificity and dual RNAs with similar specificity. Science *264*, 1137-1141.

Conrad, R., Keranen, L.M., Ellington, A.D., and Newton, A.C. (1994). Isozyme-

specific inhibition of protein kinase C by RNA aptamers. J. Biol. Chem. *269*, 32051-32054.

Coverley, D., Kenny, M.K., Munn, M., Rupp, W.D., Lane, D.P., and Wood, R.D. (1991). Requirement for the replication protein SSB in human excision repair. Nature *349*, 538-541.

Crook, T., Fisher, C., Vousden, K.H. (1991). Modulation of immortalizing properties of human papillomavirus type 16 E7 by p53 expression. J. Virol. *65*, 505-510.

Deb, S., Jackson, C.T., Subler, M.A., and Martin, D.W. (1992). Modulation of Cellular and viral promoters by mutant human p53 proteins found in tumor cells. J. Virol. *66*, 6164-6170.

Desai, M.C., Zuckermann, R.N., and Moos, W.H. (1994). Recent advances in the generation of chemical diversity libraries. Drug Dev. Res. *33*, 174-188.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl. Acids Res. *11*, 1475-1489.

Diller, L., Kassel, J., Nelson, C.E., Gryka, M.A., Litwak, G., Gebhardt, M., Bressac, B., Ozturk, M., Baker, S.J., Vogelstein, B., and Friend, S.H. (1990). p53 functions as a cell cycle control protein in osteosarcomas. Mol. Cell. Biol. *10*, 5772-5781.

Donehower, L.A., and Bradley, A. (1993). The tumor suppressor *p53*. Biochimica et Biophysica Acta *1155*, 181-205.

Donhower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.J., Butel, J.S., and Bradley, A. (1992). Mice deficient for *p53* are developmentally normal but susceptible to spontaneous tumor. Nature, *356*, 215-221.

Donahue, B.A., Augot, M., Bellon, S.F., Treiber, D.K., Toney, J.H., Lippard, S.J., and Essigmann, J.M. (1990). Characterization of a DNA damage-recognition protein from mammalian cells that binds specifically to intrastrand d(GpG) and d(ApG) DNA adducts of the anticancer drug cisplatin. Biochemistry *29*, 5672-5880.

Donahue, B.A., Bellon, S.F., Lippard, S.J., and Essigmann, J.M. (1991). A protein from mammalian cells that recognizes platinated DNA. In Platinum and Other Metal Compounds in Cancer Chemotherapy, S.B. Howell, eds. (New York:

Plenum Press).

Dooley, C.T., and Houghten, R.A. (1993). The use of positional scanning synthetic peptide combinatorial libraries for the rapid determination of opioid receptor ligands. Life Sci. *52*, 1509-1517.

Doudna, J.A., Cech, T.R., and Sullenger, B.A. (1995). Selection of an RNA molecule that mimics a major autoantigenic epitope of human insulin receptor. Proc. Natl. Acad. Sci. USA *92*, 2355-2359.

Drapkin, R., Reardon, J.T., Ansari, A., Huang, J.C., Zawel, L., Ahn, K.J., Sancar, A., and Reinberg, D. (1994a). Dual role of TFIIH in DNA excision-repair and in transcription by RNA-polymerase-II. Nature *368*, 769-772.

Drapkin, R., Sancar, A., and Reinberg, D. (1994b). Where transcription meets repair. Cell *77*, 9-12.

Drobetsky, E.A., Grosovsky, A.J., and Glickman, B.W. (1987). The specificity of UV-induced mutation at an endogenous locus in mammalian cells. Proc. Natl. Acad. Sci. USA *84*, 9103-9107.

Drummond, J.T., Li, G.-M., Longley, M.J., and Modrich, P. (1995). Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science 268, 1909-1912.

Eastman, A. (1986). Reevaluation of interaction of *cis*-dichloro(ethylenediamine)platinum(II) with DNA. Biochemistry *25*, 3912-3915.

Eastman, A. (1993). Activation of programmed cell death by anticancer agents: cisplatin as a model system. Cancer Cells *5*, 275-280.

Eastman, A., and Barry, M.A. (1987). Interaction of *trans*-diamminedichloroplatinum(II) with DNA: formation of monofunctional adducts and their reaction with glutathione. Biochemistry *26*, 3303-3307.

Eastman, A., and Schulte, N. (1988). Enhanced DNA repair as a mechanism of resistance to *cis*-diamminedichloroplatinum(II). Biochemistry *27*, 4730-4734.

Ecker, D.J., and Crooke, S.T. (1995). Combinatorial drug discovery: Which methods will produce the greatest value? Biotechnology *13*, 351-360.

Ecker, D.J., Vickers, T.A., Hanecak, R., Driver, V., and Anderson, K. (1993). Rational screening of oligonucleotide combinatorial libraries for drug discovery.

Nucl. Acids Res. *21*, 1853-1856.

Eguchi, Y., Itoh, T., and Tomizawa, J. (1991). Antisense RNA. Annu. Rev. Biochem. *60*, 631-652.

El-Deiry, W.S., Tokino, T., Velculescu, V.E., Levy, D.B., Parsons, R., Trent, J.M., Lin, D., Mercer, W.E., Kinzler, K.W., and Vogelstein, B. (1993). *WAF1*, a potential mediator of p53 tumor suppression. Cell *75*, 817-825.

Eliyahu, D., Raz, A., Gruss, P., Givol, D., and Oren, M. (1984). Participation of p53 cellular tumor antigen in transformation of normal embryonic cells. Nature *312*, 646-649.

Eliyahu, D., Michalovitz, D., Eliyahu, S., Pinhasi-Kimhi, O., and Oren, M. (1989). Wild-type p53 can inhibit oncogene-mediated focus formation. Proc. Natl. Acad. Sci. USA *86*, 8763-8767.

Ellington, A.D., and Szostak, J.W. (1990). In vitro selection of RNA molecules that bind specific ligands. Nature *346*, 818-822.

Ellington, A.D., and Szostak, J.W. (1992). Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. Nature *355*, 850-852.

Erb, E., Janda, K.D., and Brenner, S. (1994). Recursive deconvolution of combinatorial chemical libraries. Proc. Natl. Acad. Sci. USA *91*, 11411-11426.

Famulok, M. (1994). Molecular recognition of amino acids by RNA-aptamers: an L-cirulline binding RNA motif and its evolution into an L-arginine binder. J. Am. Chem. Soc. *116*, 1698-1706.

Famulok, M., and Szostak, J.W. (1992a). In vitro selection of specific ligand-binding nucleic acids. Angew. Chem. Int. Ed. Engl. *31*, 979-988.

Famulok, M., and Szostak, J.W. (1992b). Stereospecific recognition of tryptophan agarose by in vitro selected RNA. J. Am. Chem. Soc. *114*, 3990-3991.

Farmer, G., Bargonetti, J., Zhu, H., Friedman, P.N., Prywes, R., and Prives, C. (1992). Wild-type p53 activates transcription in vivo. Nature, *358*, 83-86.

Feldberg, R.S., and Grossman, L. (1976). A DNA binding protein from human placenta specific for ultraviolet damaged DNA. Biochemistry *15*, 2401-2408.

Feldberg, R.S., Lucas, J.L., and Dannenberg, A. (1982). A damaged-specific DNA binding protein: Large scale purification from human placenta and characterization. J. Biol. Chem. *257*, 6394-6401.

Ferrari, S., Harley, V.R., Pontiggia, A., Goodgellow, P.N., Lovell-Badge, R., Bianchi, M.E. (1992). SRY, like HMG1, recognizes sharp angles in DNA. EMBO J. *11*, 4497-4506.

Fichtinger-Schepman, A.M.J., van der Veer, J.L., den Hartog, J.H., Lohman, P.H., and Reedijk, J. (1985). Adducts of the antitumor drug *cis*-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. Biochemistry *24*, 707-713.

Fields, S., and Jang, S.K. (1990). Presence of a potent transcription activating sequence in the p53 protein. Science *249*, 1046-1049.

Finlay, C.A., Hinds, P.W., Tan, T.-H., Eliyahu, D., Oren, M., and Levine, A.J. (1988). Activating mutations for transformation by *p53* produce a gene product that forms an hsc70-p53 complex with an altered half-life. Mol. Cell. Biol. *8*, 531-539.

Finlay, C.A., Hinds, P.W., and Levine, A.J. (1989). The *p53* protooncogene can act as a suppressor of transformation. Cell *57*, 1083-1093.

Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jemkins, N.A., Garber, J., Kane, M., and Kolodner, R. (1993). The human mutator gene homolog *MSH2* and its association with hereditary nonpoluposis colon cancer. Cell *75*, 1027-1038.

Foster, P.L., Eisenstadt, E., and Miller J.H. (1983). Base substitution mutations induced by metabolically activated aflatoxin B_1. Proc. Natl. Acad. Sci. USA *80*, 2695-2698.

Franklin, W.A., Lo, K.M., and Haseltine, W.A. (1982). Alkaline lability of fluorescent photoproducts produced in ultraviolet light-irradiated DNA. J. Biol. Chem. *257*, 13535-13543.

Franklin, W.A., Doetsch, P.W., and Haseltine, W.A. (1985). Structural determination of the ultraviolet light-induced thymine-cytosine pyrimidine-pyrimidone (6-4) photoproduct. Nucl. Acids Res. *13*, 5317-5325.

Friedberg, E.C. (1985). DNA Repair. (New York: Freeman).

Funk, W.D., Pak, D.T., Karas, R.H., Wright, W.E., and Shay, J.W. (1992). A

transcriptionally active DNA-binding site for human p53 complexes. Mol. Cell. Biol. *12*, 2866-2871.

Gaidano, G., Ballerni, P., Gong, J.Z., Inghirani, G., Neri, A., Newcomb, E.W., Magrath, I.T., Knowes, D.M., and Dalla-Favera, R. (1991). p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA *88*, 5413-5417.

Gannon, J.V., Greaves, R., Iggo, R., and Lane, D.P. (1990). Activating mutations in p53 produce a common conformational effect–a monoclonal antibody specific for the mutant form. EMBO J. *9*, 1595-1602.

Geysen, H.M., and Mason, T.J. (1993). Screening chemically synthesized peptide libraries for biologically-relevant molecules. Bioorg. Med. Chem. Lett. *3*, 397-404.

Geysen, H.M., Rodda, S.J., and Mason, T.J. (1986). *A priori* delineation of a peptide which mimics a discontinuous antigenic determinant. Mol. Immunol. *23*, 709-715.

Giese, K., Cox, J., and Grosschedl, R. (1990). The HMG domain of lymphoid enhancing factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. Cell *69*, 185-195.

Ginsberg, D., Mechta, F., Yaniv, F., and Oren, M. (1991). Wild-type p53 can down-modulate the activity of various promoters. Proc. Natl. Acad. Sci. USA *88*, 9979-9983.

Glickman, B.W., Schaaper, R.M., Haseltine, W.A., Dunn, R.L., and Brash, D.E. (1986). The C-C (6-4) UV photoproduct is mutagenic in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *83*, 6945-6949.

Goldmacher, V.S., Cuzick, R.A., Jr., and Thilly, W.G. (1986). Isolation and partial characterization of human cell mutants differing in sensitivity to killing and mutation by methylnitrosourea and *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine. J. Biol. Chem. *261*, 12462-12471.

Green, N.M. (1975). Avidin. Adv. Protein Chem. *29*, 85-133.

Greenblatt, M.S., Bennett, W.P., Hollstein, M., and Harris, C.C. (1994). Mutations in the *p53* tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res. *54*, 4855-4878.

Harley, V.R., Jackson, D.I., Hextall, P.J., Hawkins, J.R., Berkovitz, G.D.,

Sockanathan, S., Lovell-Badge, R., and Goodfellow, P.N. (1992). DNA binding activity of recombinant SRY from normal males and XY females. Science *255*, 453-456.

Harris, C.C., and Hollstein, M. (1993). Clinical implications of the *p53* tumor-suppressor gene. New Engl. J. Med. *329*, 1318-1327.

Hensel, C.H., Xiang, R.H., Sakaguchi, A.Y., Naylor, S.L. (1990). Use of the single-stranded conformation polymorphism technique and PCR to detect *p53* gene mutations in small lung cancer. Oncogene *6*, 1067-1071.

Herrmann, C.P., Kraiss, S., and Montenarh, M. (1991). Association of casein kinase II with immunopurified p53. Oncogene *6*, 877-884.

Hinds, P.W., Finlay, C.A., Frey, A.B., and Levine, A.J. (1987). Immunological evidence for the association of p53 with a heat shock protein, hsc70, in *p53*-plus-*ras*-transformed cell lines. Mol. Cell. Biol. *7*, 2863-2869.

Hinds, P.W., Finlay, C.A., and Levine, A.J. (1989). Mutation is required to activate the *p53* gene for cooperation with the ras oncogene and transformation. J. Virol. *63*, 739-746.

Hinds, P.W., Finlay, C.A., Quartin, R.S., Baker, S.J., Fearon, E.R., Vogelstein, B., and Levine, A.J. (1990). Mutant *p53* DNAs from human colorectal carcinomas can cooperate with *ras* in transfroming primary rat cells: a comparison of the hot spot mutant phenotypes. Cell Growth Differ. *1*, 571-580.

Hirschfeld, S., Levine, A.S., Ozato, K., and Protić, M. (1990). A constitutive damage-specific DNA-binding protein is synthesized at higher levels in UV-irradiated primate cells. Mol. Cell. Biol. *10*, 2041-2048.

Hollstein, M., Sidransky, D., Vogelstein, B., Harris, C.C. (1991). *p53* mutations in human cancers. Science *253*, 49-53.

Hollstein, M., Metcalf, R.A., Welsh, J.A., Montesano, R., Harris, C.C. (1990). Frequent mutation of the *p53* gene in human esophageal cancer. Proc. Natl. Acad. Sci. USA *87*, 9958-9961.

Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., and Cuervo, J.H. (1991). Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. Nature *354*, 84-86.

Houghten, R.A., Appel, J.R., Blondelle, S.E., Cuervo, J.H., Dooley, C.T., and

Pinilla, C. (1992). The use of synthetic peptide combinatorial libraries for the identification of bioactive peptides. Pept. Res. *5*, 351-358.

Hsu, I.C., Metcalf, R.A., Sun, T., Welsh, J.A., Wang, N.J., Harris, C.C. (1991). Mutational hotspot in the *p53* gene in human hepatocellular carcinomas. Nature *350*, 427-428.

Huang, J.-C., Svoboda, D.L., Reardon, J.T., Sancar, A. (1992). Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. Proc. Natl. Acad. Sci. USA *89*, 3664-3668.

Huang, J.-C., Zamble, D.B., Reardon, J.T., Lippard, S.J., and Sancar, A. (1994). HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease. Proc. Natl. Acad. Sci. USA *91*, 10394-10398.

Huizenga, D.E., and Szostak, J.W. (1995). A DNA aptamer that binds adenosine and ATP. Biochemistry *34*, 656-665.

Hultman, T., Stahl, S., Moks, T., and Uhlén, M. (1988). Approaches to solid phase sequencing. Nucleosides and Nucleotides *7*, 629-638.

Hupp, T.R., Meek, D.W., Midgley, C.A., and Lane, D.P. (1992). Regulation of the specific DNA binding function of p53. Cell *71*, 875-886.

Hwang, B.J., and Chu, G. (1993). Purification and characterization of a human protein that binds to damaged DNA. Biochemistry *32*, 1657-1666.

Iavarone, A., Matthay, K.K., Steinkirchner, T.M., and Israel, M.A. (1992). Germiline and somatic *p53* gene mutations in multifocal osteogenic sarcoma. Proc. Natl. Acad. Sci. USA *89*, 4207-4209.

Iggo, R., Gatter, K., Bartek, J., Lane, D.P., and Harris, A.L. (1990). Increased expression of mutant forms of *p53* oncogene in primary lung cancer. Lancet *335*, 675-679.

Isaacs, W.B., Carter, B.S., Ewing, C.M. (1991). Wild-type p53 suppresses growth of human prostate cancer cells containing mutant p53 alleles. Cancer Res. *51*, 4716-4720.

Janda, K.D. (1994). Tagged versus untagged libraries: methods for the generation and screening of combinatorial chemical libraries. Proc. Natl. Acad. Sci. USA *91*,

10779-10785.

Jantzen, H.-M., Admon, A., Bell, S.P., and Tjian, R. (1990). Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature *344*, 830-836.

Jeffrey, P.D., Gorina, S., and Pavletich, N.P. (1995). Crystal structure of the teramerization domain of the p53 tumor suppressor at 1.7 angstroms. Science *267*, 1498-1502.

Jellinek, D., Lynott, C.K., Rifkin, D.B., and Janjić, N. (1993). High-affinity RNA ligands to basic fibroblast growth factor inhibit receptor binding. Proc. Natl. Acad. Sci. USA *90*, 11227-11231.

Jenison, R.D., Gill, S.C., Pardi, A., and Polisky, B. (1994). High-resolution molecular discrimination by RNA. Science *263*, 1425-1429.

Jenkins, J.R., Rudge, K., and Currie, G.A. (1984). Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. Nature *312*, 651-654.

Jones, C.J., and Wood, R.D. (1993). Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. Biochemistry *32*, 12096-12104.

Jones, P.A., Rideout, W.M., Shen, J.C., Spruck, C.H., and Tsai, Y.C. (1992). Methylation, mutation and cancer. Bioessay *14*, 33-36.

Karran, P., and Marinus, M.G. (1992). Self-destruction and tolerance in resistance of mammalian cells to alkylation damage. Nucl. Acids Res. *20*, 2933-2940.

Kastan, M.B., Onyekwere, O., Sodransky, D., Vogelstein, B., and Craig, R.W. (1991). Participation of p53 protein in the cellular response to DNA damage. Cancer Res. *51*, 6304-6311.

Kastan, M.B., Zhan, Q., El-Deriy, W.S., Carrier, F., Jacks, T., Walsh, W.V., Plunkett, B.S., Vogelstein, B., and Fornace, A.J. (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and *GADD45* is defective in ataxia-telangiectasia. Cell *71*, 587-597.

Kat, A., Thilly, W.G., Fang, W.-H., Longley, M.J., Li, G.-M., and Modrich, P. (1993). An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. Proc. Natl. Acad. Sci. USA *90*, 6424-6428.

Kataoka, H., and Fujiwara, Y. (1991). UV damage-specific DNA-binding protein in xeroderma-pigmentosum complementation group E. Biochem. Biophys. Res. Commun. *175*, 1139-1143.

Keeney, S., Wein, H., and Linn, S. (1992). Biochemical heterogeneity in xeroderma-pigmentosum complementation group-E. Mutat. Res. *273*, 49-56.

Keeney, S., Eker, A.P.M., Brody,k T., Vermeulen, W., Bootsma, D., Hoeijmakers, J.H.J., and Linn, S. (1994). Correction of the DNA repair defect in xeroderma pigmentosum group E by injection of a DNA damage binding protein. Proc. Natl. Acad. Sci. USA *91*, 4053-4056.

Kern, S.E., Kinzler, K.W., Bruskin, A., Jarosz, D., Friedman, P.N., Prives, C., and Vogelstein, B. (1991). Identification of p53 as a sequence specific DNA-binding protein. Science *252*, 1708-1711.

Kern, S.E., Pietenpol, J.A., Thiagalingam, S., Seymour, A., Kinzler, K.W., and Vogelstein, B. (1992). Oncogenic forms of p53 inhibit p53-regulated gene expression. Science *256*, 827-830.

Kim, S.-T., and Sancar, A. (1993). Photochemistry, photophysics, and mechanism of pyrimidine dimer repair DNA photolyase. Photochem. Photobiol. *57*, 895-904.

Kley, N., Chung, R.Y., Fay, S., Loeffler, J.P., and Seizinger, B.R. (1992). Repression of the basal c-*fos* promoter by wild-type p53. Nucl. Acids Res. *20*, 4083-4087.

Laird, P.W., and Jaenisch, R. (1994). DNA methylation and cancer. Hum. Mol. Genet. *3*, 1487-1495.

Laird, P.W., Jackson-Grusby, L., Fazeli, A., Dickinson, S.L., Jung, W.E., Li, E., Weinberg, R.A., and Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. Cell *81*, 197-205.

Lamb, P., and Crawford, L. (1986). Characterization of human *p53* gene. Mol. Cell. Biol. *6*, 1379-1385.

Lane, D.P., and Crawford, L.W. (1979). T antigen is bound to a host protein in SV40-transformed cells. Nature *278*, 261-263.

Lebl, M., Pátek, M., Kocis, P., Krchnák, V., Hruby, V.J., Salmon, S.E., and Lam, K.S. (1993). Multiple release of equimolar amounts of peptides from a polymeric carrier using orthogonal linkage-cleavage chemistry. Int. J. Pept. Protein Res. *41*,

201-203.

LeCerc, J.E., Borden, A., and Lawrence C.W. (1991). The thymine-thymine pyrimidine-pyrimidine (6-4) ultraviolet light photoproduct is highly mutagenic and specifically induces 3' thymine-to-cytosine transitions in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *88*, 9685-9689.

Lee, S., Elenbaas, B., Levine, A., Griffith, J. (1995). p53 and its 14 kDa C-terminal domain recognize primary DNA damage in the form of insertion/deletion mismatches. Cell *81*, 1013-1020.

Lees-Miller, S.P., Chen, Y.-R., and Anderson, C.W. (1990). Human cells contain a DNA-activated protein kinase that phosphorylates Simian Virus 40, T-antigen, mouse p53, and the human Ku autoantigen. Mol. Cell. Biol. *10*, 6472-6481.

Levine, A.J. (1993). Tumor suppressor genes. Annu. Rev. Biochem. *62*, 623-651.

Levine, A.J., Perry, M.E., Chang, A., Silver, A., Dittmer, D., Wu, M., and Welsh, D. (1994). The 1993 Walter Hubert Lecture: the role of the *p53* tumor-suppressor gene in tumorigenesis. Br. J. Cancer *69*, 409-416.

Levine, A.J., Momand, J., and Finlay, C.A. (1991). The *p53* tumor suppressor gene. Nature *351*, 453-456.

Levine, A.J. (1990). The p53 protein and its interaction with the oncogene products of the small DNA tumor viruses. Virology *177*, 419-426.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, F.P. (1988a). Cancer families: human models of susceptibility to neoplasia—the Richard and Hinda Rosenthal Foundation Award lecture. Cancer Res. *48*, 5381-5386.

Li, F.P., Fraumeni, J.F., Mulvihill, J.J., Blattner, W.A., Dreyfus, M.G., Tucker, M.A., and Miller, R.W. (1988b). A cancer family syndrome in twenty-four kindreds. Cancer Res. *48*, 5358-5362.

Lin, J., Chen, J., Elenbaas, B., and Levine, A.J. (1994). Several hydrophobic amino acids in the p53 N-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55kDa protein. Genes Dev. *8*, 1235-1246.

Lindahl, T., and Andersson, A. (1972). Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. Biochemistry *11*, 3618-3623.

Lindahl, T., Sedgwick, B., and Nakabeppu, Y. (1988). Regulation and expression of the adaptive responsive to alkylating agents. Annu. Rev. Biochem. *57*, 133-157.

Linzer, D.I.H., and Levine, A.L. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40 transformed cells and uninfected embryonal carcinoma cells. Cell *17*, 43-52.

Lippke, J.A., Gordon, L.K., Brash, D.E., and Haseltine, W.A. (1981). Distribution of UV light-induced damage in a defined sequence of human DNA: detection of alkaline-sensitive lesions at pyrimidine nucleotide-cytidine sequences. Proc. Natl. Acad. Sci. USA *78*, 3388-3392.

Loehrer, P.J., and Einhorn, L.H. (1984). Cisplatin. Ann. Int. Med. *100*, 704-713.

Lorsch, J.R., and Szostak, J.W. (1994). In vitro selection of RNA aptamers specific cyanocobalamin. Biochemistry *33*, 973-982.

Lowe, S.W., Ruley, H.E., Jacks, T., and Housman, D.E. (1993a). p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. Cell *74*, 957-967.

Lowe, S.W., Schmitt, E.M., Smith, S.W., Osborne, B.A., and Jacks, T. (1993b). p53 is required for radiation-induced apoptosis in mouse thymocytes. Nature *362*, 847-849.

Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A., and Feigon, J. (1993). Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. Proc. Natl. Acad. Sci. USA *90*, 3745-3749.

Mack. D.H., Vartikar, J., Pipas, J.M., and Laimins, L.A. (1993). Specific repression of TATA-mediated but not initiator-mediated transcription by wild-type-p53. Nature *363*, 281-283.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.F., Tainsky, M.A., and Friend, S.H. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science *250*, 1233-1238.

Malkin, D., Jolly, K.W., Barbier, N., Look, A.T., Friend, SH., Gebhardt, M.C., Anderson, T.I., Børrensen, A.-L., Li, F.P., Garber, J., and Strong, L.C. (1992). Germline mutations of the *p53* tumor-suppressor gene in children and young adults

with second malignant neoplasms. New Engl. J. Med. *326*, 1309-1315.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). In Molecular Cloning: A Laboratory Manual, (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Marks, J.R., Davidoff, A.M., Kerns, B.J., Humphrey, P.A., Pence, J.C., Dodge, R.K., Clarke-Pearson. D.L., Iglehart, J.D., Bast, R.C., and Berchuck, A. (1991). Overexpression and mutation of p53 in epithelial ovarian cancer. Cancer Res. *51*, 2979-2984.

Martinez, J., Georgoff, I., Martinez, J., and Levine, A.J. (1991). Cellular localization and cell cycle regulation by a temperature sensitive p53 protein. Genes Dev. *5*, 151-159.

Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P.J., Bootsma, D., Hoeijmakers, J.H.J., and Hanaoka, F. (1994). Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homolog of yeast RAD23. EMBO J. *13*, 1831-1843.

Matlashevski, G., Lamb, P., Pim, D., Peacock, J., Crawford, I., and Benchimol, S. (1984). Isolation and characterization of a human *p53* cDNA clone: expression of human *p53* gene. EMBO J. *3*, 3257-3263.

Mazur, M., and Glickman, B.W. (1988). Sequence specificity of mutations induced by benzo(*a*)pyrene-7,8-diol-9,10-epoxide at endogenous *aprt* gene in CHO cells. Somat. Cell Mol. Genet. *14*, 393-400.

McGurie, W.L. (1978). Hormone receptors: their role in predicting prognosis and response to endocrine therapy. Semin. Oncol. *5*, 428-433.

Meek, D.W., Simon, S., Kikkawa, U., and Eckhart, W. (1990). The p53 tumor suppressor protein is phosphorylated at serine 389 by casein kinase. EMBO J. *9*, 3253-3260.

Mellon, I., Spivak, G., and Hanawalt, P.C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. Cell *51*, 241-249.

Mercer, W.E., Amin, M., Sauve, G.J., Appella, E., Ullrich, S.J., and Romano, J.W. (1990). Wild-type human p53 is antiproliferative in SV40-trasnsformed hamster cells. Oncogene *5*, 973-980.

Michalovitz, D., Halevy, O., and Oren, M. (1990). Conditional inhibition of transformation and of cell proliferation by a temperature-sensitive mutant of p53. Cell *62*, 671-680.

Milne, D.M., Palmer, R.H., Campbell, D.G., and Meek, D.W. (1992). Phosphorylation of the p53 tumor-suppressor protein at three N terminal sites by a novel casein kinase 1-like enzyme. Oncogene *7*, 1361-1369.

Milne, D.M., Campell, D.G., Caudwell, F.B., and Meek, D.W. (1994). Phosphorylation of the tumor suppressor protein p53 by mitogen-activated protein kinases. J. Biol. Chem. *269*, 9253-9260.

Milne, D.M., Campell, L.E., Campell, D.G., and Meeks, D.W. (1995). p53 is phosphorylated in vitro and in vivo by an ultraviolet radiation-induced protein kinase characteristic of the c-Jun kinase, JNK1. J. Biol. Chem. *270*, 5511-5518.

Mitchell, D.L., Haipek, C.A., and Clarkson, J.M. (1985). (6-4) photoproducts are removed from the DNA of UV-irradiated mammlian cells more efficiently than cyclobutane pyrimidine dimers. Mutat. Res. *143*, 109-112.

Mitchell, D.L., Brash, D.E., and Nair, R.S. (1990). Rapid repair kinetics of pyrimidine (6-4) pyrimidone photoproducts in human cells are due to excision rather than conformational change. Nucl. Acids Res. *18*, 963-971.

Miyashita, T., and Reed, J.C. (1995). Tumor suppressor p53 is a direct transcriptional activator of the human *bax* gene. Cell *80*, 293-299.

Modrich, P. (1994). Mismatch repair, genetic stability, and cancer. Science *266*, 1959-1960.

Moll, U.M., Riou, G., and Levine, A.J. (1992). Two distinct mechanisms alter p53 in breast cancer: mutation and nuclear exclusion. Proc. Natl. Acad. Sci. USA *89*, 7262-7266.

Momand, J., Zambetti, G.P., Olson, D.C., George, D., and Levine, A.J. (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53 mediated transactivation. Cell *69*, 1237-1245.

Morningstar, M.K. (1994). Design and synthesis of modified deoxynucleosides, site specifically damaged oligonucleotides, and repair inhibiting chemotherapeutic agents. Ph. D. Dissertation, Massachusetts Institute of Technology.

Mowat, M.A., Cheng, A., Kimura, N., Bernstein, A., and Benchimol, S. (1985).

Rearrangements of the cellular *p53* gene in erythroleukemic cells transformed by Friend virus. Nature *314*, 633-636.

Nickell, C., Prince, M.A., and Lloyd, R.S. (1992). Consequences of molecular engineering enhanced DNA binding in a DNA repair enzyme. Biochemistry *31*, 4189-4198.

Nieuwlandt, D., Wecker, M., and Gold, L. (1995). In vitro selection of RNA ligands to substance P. Biochemistry *34*, 5651-5659.

Nigro, J.M., Baker, S.J., Preisinger, A.C., Jessup, J.M., Hostetter, R., Cleary, K., Bigner, S.H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F.S., Weston, A., Modali, R., Harris, C.C., and Vogelstein, B. (1989). Mutations in the *p53* gene occur in diverse human tumor types. Nature *342*, 705-708.

Oberosler, P., Hloch, P., Ramsperger, U., and Stahl, H. (1993). p53 catalyzed annealing of complementary single-stranded nucleic acids. EMBO J. *12*, 2389-2396.

O'Donovan, A., Davies, A.A., Moggs, J.G., West, S.C., and Wood, R.D. (1994a). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. Nature *371*, 432-435.

O'Donovan, A., Scherly, D., Clarkson, S.G., and Wood, R.D. (1994b). Isolation of active recombinant XPG protein, a human DNA repair endonuclease. J. Biol. Chem. *269*, 15965-15968.

Oliner J.D., Kinzler, K.W., Meltzer, P.S., George, D.L., and Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. Nature *358*, 80-83.

Oliner, J.D., Pietenpol, J.A., thiagalingam, S., Gyuris, J., Kinzler, K.W., and Vogelstein, B. (1993). Oncoprotein MDM2 conceals the activation domain of tumor suppressor p53. Nature *362*, 857-862.

Oltvai, Z., Milliman, C., and Korsmeyer, S.J. (1993). Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. Cell *74*, 609-619.

Orren, D.K., and Sancar, A. (1989). The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. Proc. Natl. Acad. Sci. USA *86*, 5237-5241.

Owens, R.A., Gesellchen, P.D., Houchins, B.J., and DiMarchi, R.D. (1991). The rapid identification of HIV protease inhibitors through the synthesis and screening of defined peptide mixtures. Biochem. Biophys. Res. Commun. *181*, 402-408.

Padmanabhan, K., Padmanabhan, K.P., Ferrara. J.D., Sadler, J.E., and Tulinsky, A. (1993). The structure of α-thrombin inhibited by a 15-mer single-stranded DNA aptamer. J. Biol. Chem. *268*, 17651-17654.

Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., Truong, O., Hsuan, J.J., and Jiricny, J. (1995). GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science *268*, 1912-1914.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.-F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adamas, M.D., Venter, J.C., Hamilton, S.R., Petersen, G.M., Watson, P., Lynch, H.T., Peltomäki, P., Mecklin, J.-P., de la Chapelle, A., Kinzler, K.W., Volgelstein, B. (1994). Mutation of a *mutL* homolog in hereditary colon cancer. Science *263*, 1625-1629.

Papadopoulos, N., Nicolaides, N.C., Liu, B., Parsons, R., Lengauer, C., Palombo, F., D'Arrigo, A., Markowitz, S., Willson, J.K.V., Kinzler, K.W., Jiricny, J., and Vogelstein, B. (1995). Mutations of GTBP in genetically unstable cells. Science *268*, 1915-1917.

Parada, L.F., Land, H., Weinberg, R.A., Wolg, D., and Rotter, V. (1984). Cooperation between gene encoding p53 tumor antigen and ras in cellular transformation. Nature *312*, 649-651.

Parsons, R., Li, G.-M., Longley, Fang, W.-H., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K.W., Vogelstein, B., and Modrich, P. (1993). Hypermutability and mismatch deficiency in RER^+ tumor cells. Cell *75*, 1227-1236.

Pavletich, N., Chambers, K.A., and Pabo, C.O. (1993). The DNA binding domain of p53 contains the four conserved regions and the major mutation hotspots. Genes Dev. *7*, 2556-2564.

Pinhasi-Kimhi, O., Michalovitz, D., Ben-Ze'ev, A., and Oren, M. (1986). Specific interaction between the p53 cellular tumor antigen and major heat shock proteins. Nature *320*, 182-185.

Pinilla, C., Appel, J.R., Blanc, P., and Houghten, R.A. (1992). Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. Biotechniques *13*, 901-905.

Potter, P.M., Wilkinson, M.C., Fitton, J., Carr, F.J., Brennand, J., Cooper, D.P., and Margison, G.P. (1987). Characterization and nucleotide sequence of *ogt*, the O^6-alkylguanine-DNA-alkyltransferase gene of *E. coli.* Nucl. Acids Res. *15*, 9177-9193.

Prosser, J., Thompson, A.M., Cranston, G., Evans, H.J. (1990). Evidence that *p53* behaves as a tumor suppressor gene in sporadic breast tumors. Oncogene *5*, 1573-1579.

Protić, M., Hirschfeld, S., Tsang, A.P., Wagner, M., Dixon, K., and Levine, A.S. (1989). Induction of a novel damage-specific DNA binding protein correlates with enhancing DNA repair in primate cells. Mol. Toxicol. *2*, 255-270.

Protić-Sabljić, M., Tuteja, N., Munson, P.J., Hauser, J., Kraemer, K.H., and Dixon, K. (1986). UV light-induced cyclobutane pyrimidine pyrimidine dimers are mutagenic in mammalian cells. Mol. Cell. Biol. *6*, 3349-3356.

Raycroft, L., Wu, H., and Lozano, G. (1990). Transcriptional activation by wild-type but not transforming mutants of the *p53* anti-oncogene. Science *249*, 1049-1051.

Reardon, J.T., Nichols, A.F., Keeney, S., Smith, S.A., Taylor, J.S., Linn, S., and Sancar, A. (1993). Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts–t[c,s]t,t[t,s]t,t[6-4]t, and t[dewar]t. J. Biol. Chem. *268*, 21301-21308.

Reed, J.C. (1994). Bcl-2 and the regulation of programmed cell death. J. Cell Biol. *124*, 1-6.

Robins, P., Jones C.J., Biggerstaff, M., Lindahl, T., and Wood, R.D. (1991). Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. EMBO J. *10*, 3913-3921.

Ronai, Z.A., Gradia, S., Peterson, L.A., and Hecht, S.S. (1993). G to A transitions and G to T transversions in codon 12 of the Ki-*ras* oncogene isolated from mouse lung tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and related DNA methylating and pyridyloxobutylating agents. Carcinogenesis *14*, 2419-2422.

Rossi, S.C., Conrad, M., Voigt, J.M., and Topal, M.D. (1989). Excision repair of O^6-methylguanine synthesized at the rat H-*ras* *N*-methyl-*N*-nitrosourea activation site and introduced into *Escherichia coli*. Carcinogenesis *10*, 373-377.

Ruggeri, B., DiRado, M., Zhang, S.Y., Bauer, B., Goodrow, T., and Klein-Szanto, A.J.P. (1993). Benzo(*a*)pyrene-induced murine skin tumors exhibit frequent and characteristic G to T mutations in the *p53* gene. Proc. Natl. Acad. Sci. USA *90*, 1013-1017.

Saffran,W.A., Welsh, J.T., Knobler, R.M., Gasparro, F.P., Contor, C.R., and Edelson, R.L. (1988). Preparation and characterization of biotinylated psoralens. Nucl. Acids Res. *16*, 7221-7231.

Salmon, S.E., Lam, K.S., Lebl, M., Kankola, A., Khattri, P.S., Wade, S., Pátek, M., Kocis, P., Krchnák, V., Thorpe, D., and Felder, S. (1993). Discovery of biologically active peptides in random libraries: solution-phase testing after staged orthogonal release from resin beads. Proc. Natl. Acad. Sci. USA *90*, 11708-11712.

Samson, L., and Cairns, J. (1977). A new pathway for DNA repair in *Escherichia coli*. Nature 267, 281-283.

Samson, L., Thomale, J., and Rajewsky, M.F. (1988). Alternative pathways for the in vivo repair of O^6-alkylguanine and O^4-alkylthymine in *Escherichia coli*: the adaptive response and nucleotide excision repair. EMBO J. *7*, 2261-2267.

Sancar, A. (1994). Structure and function of DNA photolyase. Biochemistry *33*, 2-9.

Sancar, A., and Rupp, W.D. (1983). A novel repair enzyme: UVRABC excision nuclease of *Escherichia coli* cuts a DNA strand on both sides of the damaged region. Cell *33*, 249-260.

Sancar, A., and Tang, M.S. (1993). Nucleotide excision repair. Photochem. Photobiol. *57*, 905-921.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463-5467.

Santhanam, U., Ray, A., and Sehgal, P.B. (1991). Repression of the interleukin 6 gene promoter by *p53* and the retinoblastoma susceptibility gene product. Proc. Natl. Acad. Sci. USA *88*, 7605-7609.

Sarnow, P., Ho, Y.S., Williams, J., and Levine, A.J. (1982). The adenovirus E1b-58K tumor antigen and SV40 large tumor antigen are physically associated with the same 54K cellular protein. Cell *28*, 387-394.

Sassanfer, M., and Szostak, J.W. (1993). An RNA motif that binds ATP. Nature

364, 550-553.

Sato, T., Hanada, M., Bodrug, S., Irie, S., Iwama, N., Boise, L., Thompson, C., Golemis, E., Fong, L., Wang, H.-G., and Reed, J.C. (1994). Interactions among members of the *bcl-2* protein family analyzed with a yeast two-hybrid system. Proc. Natl. Acad. Sci. USA *91*, 9238-9242.

Satoh, M., and Lindahl, T. (1992). Role of poly(ADP-ribose) formation in DNA repair. Nature *356*, 356-358.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H.J., Chambon, P., and Egly, J.-M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science *260*, 58-63.

Schaeffer, L., Moncollin, V., Roy, R., Staub, A., Mezzina, M., Sarasin, A., Weeda, G., Hoeijmakers, J.H.J., and Egly, J.-M. (1994). The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. EMBO J. *13*, 2388-2392.

Scheffner, M., Werness, B.A., Huibregtse, J.M., Levine, A.J., and Howley, P.M. (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell *63*, 1129-1136.

Schneider, D., Tuerk, C., and Gold, L. (1992). Selection of high affinity RNA ligands to the bacteriophage R17 coat protein. J. Mol. Biol. *228*, 862-869.

Scott, J.K., and Smith, G.P. (1990). Searching for peptide ligands with an epitope library. Science *249*, 386-390.

Seto, E., Usheva, A., Zambetti, G.P., Momand, J., Horikoshi, N., Weinman, R., Levine, A.J., and Shenk, T. (1992). Wild-type p53 binds to the TATA-binding protein and represses transcription. Proc. Natl. Acad. Sci. USA *89*, 12028-12032.

Shaulsky, G., Goldfinger, N., Peled, A., and Rotter, V. (1991). Involvement of wild-type p53 protein in the cell cycle requires nuclear localization. Cell Growth Differ. *2*, 661-667.

Shaw, P., Bovey, R., Tardy, S., Sahli, R., Sordat, B., and Costa, J. (1992). Induction of apoptosis by wild-type p53 in a human colon tumor-derived cell line. Proc. Natl. Acad. Sci. USA *89*, 4495-4499.

Shibutani, S., Takeshita, Grollman, A.P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. Nature *349*, 431-434.

Shigenaga, M.K., Gimeno, C.J., and Ames, B.N. (1989). Urinary 8-hydroxy-2'-deoxyguanosine as a biological marker of in vivo oxidative DNA damage. Proc. Natl. Acad. Sci. USA *86*, 9697-9701.

Shiio, Y., Yamamoto, T., and Yamaguchi, N. (1992). Negative regulation of Rb expression by the *p53* gene product. Proc. Natl. Acad. Sci. USA *89*, 5206-5210.

Shopland, D.R., Eyre, H.J., and Pechacek, T.F. (1991). Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States? J. Natl. Cancer Inst. *83*, 1142-1148.

Sibghat-Ullah, and Sancar, A. (1990). Substrate overlap and functional competition between human nucleotide excision repair and *Escherichia coli* photolyase and (A)BC excision nuclease. Biochemistry *29*, 5711-5718.

Smith, C.A., and Taylor, J.-S. (1993). Preparation and characterization of a set of deoxyoligonucleotides 49-mers containing site-specific *cis-syn, trans-syn*-I, (6-4), and Dewar photoproducts of thymidyldylyl(3'→5')-thymidine. J. Biol. Chem. *268*, 11143-11151.

Spruck, C.H., Rideout, W.M., and Jones, P.A. (1993). DNA methylation and cancer. In DNA Methylation: Molecular Biology and Biological Significance, J.P.Jost and H.P.Saluz, eds. (Basel, Switzerland: Birkhauser Verlag), pp. 487-509.

Soussi, T., Caron de Fromentel, and May, P. (1990). Structural aspects of the p53 protein in relation to gene evolution. Oncogene *5*, 945-952.

Srivastava, S., Zou, Z.Q., Pirollo, K., Blattner, W., and Chang, E.H. (1990). Germ-line transmission of a mutated *p53* gene in a cancer-prone family with Li-Fraumeni syndrome. Nature *348*, 747-749.

Stürzbecher,H.W., Maimets, T., Chumakov, P., Brain, R., Addison, C., Simanis, V., Rudge, K., Philp, R., Grimaldi, M., Court, W., and Jenkins, J.R. (1990). p53 interacts with $p34^{cdc2}$ in mammalian cells: implications for cell cycle control and oncogenesis. Oncogene *5*, 795-801.

Stürzbecher, H.W., Brain, R., Addison, C., Rudge, K., Remm, M., Grimaldi, M., Keenan, E., and Jenkins, J.R. (1992). A C-terminal alpha helix plus basic motif is the major structural determinant of p53 tetramerization. Oncogene 7, 1513-1523.

Subler, M.A., Martin, D.W., and Deb, S. (1992). Inhibition of viral and cellular promoters by human wild-type p53. J. Virol. *66*, 4757-4762.

Svoboda, D.L., Taylor, J.-S., Hearst, J.E., and Sancar, A. (1993). DNA repair by eukaryotic nucleotide excision nuclease: removal of thymine dimer and psoralen monoadduct by HeLa cell-free extract and of thymine dimer by *Xenopus laevis* oocytes. J. Biol. Chem. *268*, 1931-1936.

Szostak, J.W. (1992). In vitro genetics. Trends Biochem. Sci. *17*, 89-93.

Takahashi, T., Nau, M.M., Chiba, I., Birrer, M.J., Rosenberg, R.K., Vinocour, M., Levitt, M., Pass, H., Gazdar, A.F., and Minna, J.D. (1989). *p53*: a frequent target for genetic abnormalities in lung cancer. Science *246*, 491-494.

Takao, M., Abramic, M., Moos, M., Otrin, V.R., Wootton, J.C., McLenigan, M., Levine, A.S., and Protić, M. (1993). A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime-mold protein. Nucl. Acids Res. *21*, 4111-4118.

Tamura, G., Kihana, T., Nomura, K., Terada, M., Sugimura, T., Hirohashi, S. (1991). Detection of frequent *p53* gene mutations in primary gastric cancer by cell sorting and polymerase chain reaction single-strand conformation polymorphism analysis. Cancer Res. *51*, 3056-3058.

Tanaka, K., and Wood, R.D. (1994). Xeroderma pigmentosum and nucleotide excision repair of DNA. Trends Biochem. Sci. *19*, 83-86.

Taylor, J.-S., Garrett, D.S., Brockie, I.R., Svoboda, D.L., and Telser, J. (1990). 1H NMR assignment and melting temperature study of *cis-syn* and *trans-syn* thymine dimer containing duplexes of d(CGTATTATGC)·d(GCATAATACG). Biochemistry *29*, 8858-8866.

Thomas, D.C., Okumoto, D.S., Sancar, A., and Bohr, V.A. (1989). Preferential DNA repair of (6-4) photoproducts in the dihydroflate reductase gene of Chinese hamster ovary cells. J. Biol. Chem. *264*, 18005-18010.

Tidd, D.M., Warenius, H.M. (1989). Partial protection of antisense oligonucleotides by terminal methyl phosphonate groups. Br. J. Cancer *60*, 343-350.

Toguchida, J., Yamaguchi, T., Dayton, S.H., Beauchamp, R.L., Herrera, G.E., Ishizaki, K., Yamamuro, T., Meyers, P.A., Little, J.B., Sasaki, M.S., Weichselbaum, R.R., and Yandell, D.W. (1992). Prevalence and spectrum of germline mutations of the *p53* gene among patients with sarcoma. New Engl. J. Med. *326*, 1301-1308.

Toney, J.H., Donahue, B.A., Kellett, P.J., Bruhn, S.L., Essigmann, J.M., and Lippard, S.J. (1989). Isolation of cDNA encoding a human protein that binds selectively to DNA modified by the anticancer drug *cis*-diamminedichloroplatinum(II). Proc. Natl. Acad. Sci. USA *86*, 8328-8332.

Treiber, D.K. (1993). The role of damaged DNA recognition proteins in the genotoxicities of ultraviolet light and the anticancer drug. Ph. D. Dissertation, Massachusetts Institute of Technology.

Treiber, D.K., Chen, Z.-H., and Essigmann, J.M. (1992). An ultraviolet light-damaged DNA recognition protein absent in xeroderma pigmentosum group E cells binds selectively to pyrimidine (6-4) pyrimidone photoproducts. Nucl. Acids Res. *20*, 5805-5810.

Treiber, D.K., Zhai, X.-Q., Jantzen, H.-M., and Essigmann, J.M. (1994). Cisplatin-DNA adducts are molecular decoys for the ribosomal RNA transcriptional factor hUBF (human upstream binding factor). Proc. Natl. Acad. Sci. USA *91*, 5672-5676.

Tsai, D.E., Kenan, D.J., and Keene, J.D. (1992). In vitro selection of an RNA epitope immunologically cross-reactive with a peptide. Proc. Natl. Acad. Sci. USA *89*, 8864-8868.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505-510.

Tuerk, C., MacDougal, S., and Gold, L. (1992). RNA pseudoknots that inhibits human immunodeficiency virus type 1 reverse transcriptase. Proc. Natl. Acad. Sci. USA *89*, 6988-6992.

Uhlmann, E., and Peyman, A. (1990). Antisense oligonucleotides: a new therapeutic principle. Chem. Rev. *90*, 543-584.

Umar, A., Boyer, J.C., Thomas, D.C., Nguyen, D.C., Risinger, J.I., Boyd, J., Ionov, Y., Perucho, M., and Kunkel, T.A. (1994). Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. J. Biol. Chem. *269*, 14367-14370.

Unger, T., Nau, M.M., Segal, S., and Minna, J.D. (1992). p53: a transdominant regulator of transcription whose function is ablated by mutations occurring in human cancer. EMBO J. *11*, 1383-1390.

van Assendelft, G.B., Rigney, E.M., and Hickson, I.D. (1993). Purification of a HeLa cell nuclear protein that binds selectively to DNA irradiated with ultraviolet light. Nucl. Acids Res. *21*, 3399-3404.

Wang, K.Y., Krawczyk, S.H., Bichofberger, Norbert, Swaminathan, S., and Bolton, P.H. (1993). The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. Biochemistry *32*, 11285-11292.

Wang, S.Y. (1976). Pyrimidine bimolecular photoproducts. In Photochemistry and Photobiology of Nucleic Acids, S.Y. Wang, eds. (New York: Academic Press), Vol. 1, pp. 295-356.

Wang, Y., Reed, M., Wang, P., Stenger, J.E., Mayr, G., Anderson, M.E., Schwedes, J.F., and Tegtmeyer, P. (1993). p53 domains: identification and characterization of two autonomous DNA binding regions. Genes Dev. *7*, 2575-2586.

Wang, Y., and Eckhart, W. (1992). Phosphorylation sites in the amino-terminal of mouse p53. Proc. Natl. Acad. Sci. USA *89*, 4231-4235.

Weintraub, H., Hauschka, S., and Tapscott, S.J. (1991). The *MCK* enhancer contains a p53 responsive element. Proc. Natl. Acad. Sci. USA *88*, 4570-4574.

Werness, B.A., Levine, A.J., and Howley, P.M. (1990). Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science *248*, 76-79.

Wood, M.L., Dizdaroglu, M., Gajewski, E., and Essigmann, J.M. (1990). Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. Biochemistry *29*, 7024-7032.

Wood, R.D. (1989). Repair of pyrimidine dimer ultraviolet light photoproducts by human cell extracts. Biochemistry *28*, 8287-8292.

Wood, R.D., Robins, P., and Lindahl. T. (1988). Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. Cell *53*, 97-106.

Wu, L., Bayle, J.H., Elenbaas, B., Pavletich, N.P., and Levine, A.J. (1995). Alternatively spliced forms in the carboxy-terminal domain of the p53 protein regulates its ability to promote annealing of complementary single strands of nucleic acids. Mol. Cell. Biol. *15*, 497-504.

Wu, X., Bayle, H.J., Olson, D., and Levine, A.J. (1993). The p53-mdm-2

autoregulatory feedback loop. Genes Dev.*7*, 1126-1132.

Wyatt, J.R., Vickers, T.A., Roberson, J.L., Buckheit, R.W., Jr., Klimkait, T., DeBaets, E., Davis, P.W., Rayner, B., Imbach, J.L., and Ecker, D.J. (1994). Combinatorially selected guanosine-quartet structure is a potent inhibitor of human immunodeficiency virus envelope-mediated cell fusion. Proc. Natl. Acad. Sci. USA *91*, 1356-1360.

Yang, J.L., Chen, R.H., Maher, V.M., and McCormick, J.J. (1991). Kinds and location of mutations induced by (±)-7β,8α-dihydroxy-9α,10α-epoxy-7,8,9,10-teterahydro-benzo(*a*)pyrene in the coding region fo the hypoxanthine (guanine) phosphoribosyltransferase gene in diploid human fibroblasts. Carcinogenesis *12*, 71-75.

Yin, X.-M., Oltvai, Z.N., and Korsmeyer, S.J. (1994). BH1 and BH2 domains of Bcl-2 are required for inhibition of apoptosis and heterodimerization with Bax. Nature *369*, 321-333.

Yonish-Rouach, E., Resnitzky, D., Lotem, J., Sachs, L., Kimchi, A., and Oren, M. (1991). Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6. Nature *352*, 345-347.

Zambetti, G.P., and Levine, A.J. (1993). A comparison of the biological activities of wild-type and mutant p53. FASEB J. *7*, 855-865.

Zambetti, G.P., Bargonetti, J., Walker, K., Prives, C., and Levine, A.J. (1992). Wild-type p53 mediates positive regulation of gene expression through a specific DNA sequence element. Genes Dev. *6*, 1143-1152.

Zauberman, A., Barak, Y., Ragimov, N., Levy, N., and Oren, M. (1993). Sequence-specific DNA binding by p53: identification of target sites and lack of binding to p53-MDM2 complexes. EMBO J. *12*, 2799-2808.

Zuckermann, R.N., Martin, E.J., Spellmeyer, D.C., Stauber, G.B., Shoemaker, K.R., Kerr, J.M., Figliozzi, G.M., Goff, D.A., Siani, M.A., Simmon, R.J., Banville, S.C., Brown, E.G., Wang, L., Richter, L.S., and Moos, W.H. (1994). Discovery of nanomolar ligands for 7-transmembrane G-protein-coupled receptors from a diverse *N*-(substituted)glycine peptoid library. J. Med. Chem. *37*, 2678-2685.

BIOGRAPHY

Zheng-Huan Chen was born on August 26, 1964 in Nanjing, a Chinese city famous for its richness in Chinese culture and history. He spent most of his childhood in a suburban district of the city, and attended his elementary and junior high schools there. During his high school years at the Nanjing No. 10 High School (also known as the Jin-Ling High School), he developed a strong interest in biological sciences. He entered the Fudan University in Shanghai, and earned a Bachelor of Sciences degree in biology in 1987. Upon graduation he enrolled in the graduate program at the Shanghai Institute of Cell Biology, Academia Sinica. In 1989, he transferred to the Division of Toxicology at the Massachusetts Institute of Technology, and joined the laboratory of Dr. John Essigmann, a professor of chemistry and toxicology at MIT. His doctoral work focused on developing a novel approach, called Fatal Engineering, of designing effective anticancer drugs. He completed his studies in August of 1995. He was the recipient of the Whitaker Fellowship in 1993. Upon graduation, he pursued postdoctoral work in the laboratory of Dr. Rudolf Jaenisch, a member of the Whitehead Institute for Biomedical Research and a professor of biology in the Department of Biology at MIT.