

60

# Total Synthesis of a Furan-Side Psoralen-Thymidine Monoadduct and its Biochemical Applications

by

William Rudolf Kobertz

B.S. Chemistry  
University of California, Berkeley

submitted to the Department of Chemistry  
in partial fulfillment of the requirements  
for the degree of

Doctor of Philosophy

at the  
Massachusetts Institute of Technology  
February 1998

© 1998 Massachusetts Institute of Technology  
All rights reserved

Signature of Author \_\_\_\_\_

Department of Chemistry  
December 3, 1997

Certified by \_\_\_\_\_

Professor John M. Essigmann  
Thesis Supervisor

Accepted by \_\_\_\_\_

Professor Dietmar Seyferth  
Chairman, Departmental Committee on Graduate Studies

MAR 03 1998

Science



### Thesis Committee

Professor Stephen L. Buchwald \_\_\_\_\_ Chairperson

Professor John M. Essigmann \_\_\_\_\_ Thesis Supervisor

Professor JoAnne Stubbe \_\_\_\_\_





*to Mom,  
a scientist at heart*



## Total Synthesis of a Furan-Side Psoralen-Thymidine Monoadduct and its Biochemical Applications

by

William Rudolf Kobertz

*Submitted to the Department of Chemistry on December 3, 1997 in partial fulfillment of  
the requirements for the degree of a Doctor of Philosophy in Chemistry.*

### Abstract

The first total chemical synthesis of a *cis*-syn furan-side photoproduct between a psoralen derivative and thymidine is described. The key step in the synthesis was an intramolecular [2+2] photocycloaddition, which directed the stereochemical course of the reaction to afford a product equivalent to that formed when a psoralen molecule is allowed to react at a 5'-TpA-3' site in DNA. A model system consisting of a simple benzofuranyl acid tethered to the 5' hydroxyl of thymidine showed that it was possible to bias the stereochemical outcome of the photochemical reaction in favor of the desired *cis*-syn product. Further refinement of the model system allowed for the elaboration of a benzofuran-thymidine photoproduct into a *cis*-syn 2-carboxypsoralen furan-side thymidine monoadduct. In addition, a facile synthesis of a *cis*-syn 2-carboxypsoralen-thymidine furan-side monoadduct was accomplished using a 2'-carboxypsoralen tethered to the 5'-hydroxyl of thymidine.

The site-specific incorporation of a 2-carboxypsoralen-thymidine furan-side monoadduct into several oligonucleotides was accomplished using phenoxyacetyl protected phosphoramidites. In > 90% yield, this oligonucleotide formed a crosslink with a complementary DNA strand by irradiation at 366 nm. The crosslinked duplex was reversible either with 254 nm light or by heating in aqueous base. The chemically synthesized psoralen containing oligonucleotides possessed the same crosslinking abilities as those produced by prior methods, but with the flexibility to place the adduct in any sequence context.

A 2-oxopentamide derivative of the 2-carboxypsoralen-thymidine monoadduct was synthesized and incorporated into an oligonucleotide. Treatment of these ketone bearing oligonucleotides with aminoxy 2-phenylindole derivatives formed an oligonucleotide containing an adduct that bound the estrogen receptor protein specifically. Subsequent incorporation of the 2-phenylindole psoralen and ketone psoralen containing oligonucleotides into circular double-stranded DNA plasmids afforded good substrates for nucleotide excision repair using human whole cell extracts.

Finally, the *cis*-syn benzofuran-thymidine monoadduct was used to alter the target selectivity of the DNA damaging agent aflatoxin B<sub>1</sub> epoxide, a DNA damaging agent formed by metabolic activation of the liver carcinogen, aflatoxin B<sub>1</sub>. DNA duplexes were prepared consisting of a target strand containing multiple potentially reactive guanines and a non-target strand containing a *cis*-syn thymidine-benzofuran photoproduct. Aflatoxin B<sub>1</sub> epoxide treatment

of a DNA duplex containing a *cis*-syn thymidine-benzofuran photoproduct in the intercalation site 5' to a specific guanine greatly reduced adduct formation at that site. Using this approach it was possible to simplify the production of site-specifically modified oligonucleotides containing AFB<sub>1</sub> adducts in the sequence context of a p53 mutational hotspot.

Thesis Supervisor: Professor John Essigmann

Title: Professor of Chemistry and Toxicology

## Acknowledgments

The relationship between a thesis advisor and graduate student is rarely defined as a friendship. I am very fortunate to be able to call my thesis advisor John Essigmann, a friend. John has taught me the most important aspect of being a research scientist, the ability to clearly present one's research. John has single-handedly improved my writing and communication skills by an order of magnitude. I also want to thank John and Ellen for their many invitations to their incredible cabin in Maine. I want to thank Dr. Gerald Wogan, my co-advisor, for also contributing significantly to my graduate education. Dr. Wogan was responsible for planting the intercalation seed into my mind that eventually grew into Chapter 4 of this dissertation. However, the person ultimately responsible for the completion of this thesis was Carolyn Bertozzi. Carolyn taught me, as undergraduate, how to design and solve interesting and exciting scientific problems. Most importantly, Carolyn instilled her unmatched enthusiasm for science into me. I am extremely grateful to have Carolyn as a friend and mentor.

It is seldom that one performs all of the experimental work in their entire dissertation. David Nauman synthesized all of the 2-phenylindole derivatives in Chapter 3 and provided me with boat-loads of important photo-precursors over the last 2 years. I hope I am as valuable to Young Dave Nauman's graduate education as he was to mine. All of the work described in Chapter 4 was performed with my co-worker David Wang. Dave was a cardinal part in the design and execution of every experiment in this chapter. Although Stanford alum, Dave and I have become friends, labmates, housemates and co-workers during our years at MIT. Dave has had a great effect on my scientific thinking and recreational activities. My downhill skiing, volleyball, and ice hockey skills have all improved thanks to Dave. Oh, I think I know how to set up a biochemical experiment now thanks to Dave too! Lastly, I would like to thank Isa Kuraoka. Isa taught me the nucleotide excision repair assay while I was working in Dr. Richard Wood's lab. All of the repair reactions performed in Chapter 3 were due to our collaboration. I also thank Rick and Isa for their generous hospitality while I was staying in England.

The Essigmann lab has been a wonderful place to do research primarily due to the members of the lab. The "older" students were a valuable resource to pull from during my first few years in the lab. I would like to thank Lisa Bailey for being patient and explaining aflatoxin lability and mutagenesis to me over and over again. Jill Mello was the best baymate that I could have ever asked for. I thoroughly enjoyed our conversations about "bad hair days," "Portuguese Princesses" and "winter coats." It was nice to work late in the lab with someone as fun as Jill. I would also like to thank Dan Treiber, Kevin Yarema, and Stacia Rink for being great teammates in different Essigmann Lab team sports. There are very few pleasures greater than teaching and the "younger" students in the lab, Maria Kartalou, Maryann Smela and Zoran Zdraveski allowed me to share with them my small knowledge base of DNA damage and repair. Clearly, the best Essigmann lab experience I've had has been my membership in the Fantastic Four. Jim Delaney, Debbie Kreutzer, Dave Wang and I joined the lab the same year and are responsible for the unparalleled levels of sarcasm that the lab now enjoys. I still love to reminisce about our first year trip up to John's cabin in Maine; however, I never want to listen to Genesis' ABACAB CD or play "I'm going on a picnic" ever again. I will always remember our 5+ years of graduate school together.

The friendships made in the lab will last a career; the friendships made while eating will last forever. I have had the pleasure of eating at least  $\frac{1}{3}$  of my meals in New England with Evan Powers, Martha Rook, Ben Turk, Jeff Eckert, Matt Bogyo and Dave Wang. As accurate as a Swiss watch, we have convened in the basement at 11:45 AM every weekday to make the trek to our final lunchtime destination. From Star Trek (the next generation of course) and the X-files to Chemistry League volleyball strategies and Indie rock sellouts, the topics for discussion were every bit as eclectic as the lunchtime crowd. In addition to our dining rituals, I've enjoyed our holiday excursions to Martha's Cabin in Vermont for fishing and leaf peeping, to Matt's Cabin at the Cape for sailing and bocce and to each other's houses for the season premieres and finales of the show du jour. In the future, we will clearly have to fly in for these important events.

Most importantly, I want to thank my mother and father for supporting me in all of my crazy endeavors I have undertaken during my entire lifetime. From bowling all over the west coast to whitewater rafting while paralyzed, they have always advised me what to do, but not told me what to do and I love them both for that. Lastly, I would like to thank Marjie, who is my best friend and love. I am truly fortunate to be able to share all of my thoughts, ideas and ambitions with someone as understanding and supportive as Marjie. I cannot thank her enough for her wonderful smile that can always bring a smile to my face.

## Table of Contents

Title Page .....	1
Committee Page .....	3
Abstract .....	7
Acknowledgments .....	9
Table of Contents .....	11
List of Figures .....	13
List of Tables .....	15
List of Abbreviations .....	17
<b>Introduction</b> .....	<b>19</b>
<b>Chapter 1 Synthesis of a <i>Cis</i>-Syn Furan-Side Psoralen-Thymidine Monoadduct</b> .....	<b>23</b>
Background.....	23
Synthetic Strategy .....	26
Synthesis of Benzofuran-Thymidine Photoproducts .....	30
Synthesis of a <i>Cis</i> -Syn Furan-Side Psoralen-Thymidine Monoadduct .....	35
A More Efficient Synthesis .....	40
Conclusions .....	42
<b>Chapter 2 Synthesis of Oligonucleotides Containing a Site-Specific Psoralen Adduct</b> .....	<b>45</b>
Introduction .....	45
Psoralen-DNA Adducts .....	49
NMR Solution Structure of a Psoralen Containing Oligonucleotide .....	50
Repair of Psoralen Adducts .....	52
Mutagenicity and Cytotoxicity of Psoralen Adducts .....	54
Biochemical Applications of Site-Specifically Psoralenated Oligonucleotides .....	55
Site-Specific Synthesis of an Oligonucleotide Containing a Furan-Side Psoralen Monoadduct .....	57
Conclusions .....	66
<b>Chapter 3 Nucleotide Excision Repair of Psoralen-Thymidine Adducts</b> .....	<b>67</b>
Introduction .....	67
Attachment of a Protein Recognition Domain .....	70
Estrogen Receptor Binding of 2-Phenylindole Containing Oligonucleotides .....	82
Nucleotide Excision Repair Experiments .....	84
Discussion .....	92
Future Directions .....	98

<b>Chapter 4 Design of an Intercalation Inhibitor</b> .....	101
Introduction .....	101
Intercalation Inhibition of Aflatoxin B <sub>1</sub> Epoxide .....	107
A Model System .....	110
Aflatoxin B <sub>1</sub> Epoxide Inhibition at Adjacent Guanines .....	112
Synthesis of Aflatoxin B <sub>1</sub> Adducts in a p53 Mutational Hotspot Sequence .....	113
Discussion .....	115
Conclusions .....	120
<b>Experimental Section</b> .....	123
Chapter 1 .....	125
Chapter 2 .....	145
Chapter 3 .....	151
Chapter 4 .....	171
References .....	177
Biography .....	187



## List of Figures

Figure 1.1.	Structures of the naturally occurring psoralens .....	24
Figure 1.2.	Structures of some of the common synthetic psoralen derivatives .....	25
Figure 1.3.	Structures of the isolated psoralen-thymidine monoadducts.....	26
Figure 1.4.	Structures of the eight possible furan-side monoadducts .....	28
Figure 1.5.	Synthesis of a <i>cis</i> -anti pyrone-side monoadduct .....	29
Figure 1.6.	Synthesis of a <i>cis</i> -syn “abbreviated” thymidine dimer .....	30
Figure 1.7.	Retrosynthetic route to a <i>cis</i> -syn 2-carboxypsoralen-thymidine furan-side monoadduct .....	31
Figure 1.8.	Synthesis of a <i>cis</i> -syn benzofuran-thymidine monoadduct .....	32
Figure 1.9.	NMR and CD spectra of benzofuran-thymidine monoadducts .....	34
Figure 1.10.	Synthesis of 5-formyl-6-methoxymethyl-2-benzofuran carboxylic acid .....	36
Figure 1.11.	Synthesis of a <i>cis</i> -syn 2-carbomethoxypsoralen-thymidine furan-side monoadduct .....	37
Figure 1.12.	NMR and CD spectra of the 2-carbomethoxypsoralen-thymidine furan-side monoadduct .....	39
Figure 1.13.	Synthesis of 2-carboxypsoralen .....	41
Figure 1.14.	Facile synthesis of a <i>cis</i> -syn 2-carbomethoxypsoralen-thymidine furan-side monoadduct .....	42
Figure 1.15.	NMR and CD spectra of the 2-carbomethoxypsoralen-thymidine furan-side monoadduct .....	44
Figure 2.1.	The four major strategies for synthesizing site-specifically modified oligonucleotides .....	47
Figure 2.2.	Schematic diagram of the synthesis of psoralen-DNA adducts using high intensity lasers .....	49
Figure 2.3.	NMR derived models of a DNA duplex crosslinked with a psoralen derivative.....	51
Figure 2.4.	The <i>E. Coli</i> error-free repair of a psoralen crosslink .....	53
Figure 2.5.	Model for the <i>E. Coli</i> RNA polymerase transcription elongation complex .....	57
Figure 2.6.	Structures of carboxypsoralen-thymidine monoadducts .....	59
Figure 2.7.	Scheme for solid-phase DNA synthesis .....	60
Figure 2.8.	HPLC traces of stability studies performed on a 2-carboxypsoralen- thymidine monoadduct .....	61
Figure 2.9.	Synthesis of a 2-carbomethoxypsoralen-thymidine phosphoramidite .....	62
Figure 2.10.	HPLC trace of an enzymatic digestion of a psoralen containing oligonucleotide .....	63
Figure 2.11.	Photoreactions of the furan-side monoadduct containing oligonucleotide .....	65
Figure 2.12.	Hand-held lamp irradiations of a furan-side monoadduct containing oligonucleotide .....	66

Figure 3.1.	The design of anticancer drugs that block DNA repair selectively in cancer cells .....	68
Figure 3.2.	Synthesis of a ketone benzofuran-thymidine monoadduct .....	72
Figure 3.3.	Synthesis of 2-phenylindole aminoxy derivatives .....	74
Figure 3.4.	HPLC trace of the reaction between a benzofuran-thymidine ketone nucleoside and a 2-phenylindole aminoxy compound.....	75
Figure 3.5.	Synthesis of a ketone benzofuran-thymidine phosphoramidite .....	76
Figure 3.6.	HPLC analysis of the reaction between a 2-phenylindole aminoxy compound and a benzofuran ketone bearing oligonucleotide .....	77
Figure 3.7.	Synthesis of ketone psoralen-thymidine monoadduct .....	78
Figure 3.8.	HPLC trace of the reaction between a psoralen-thymidine ketone nucleoside and a 2-phenylindole aminoxy compound.....	79
Figure 3.9.	Synthesis of a ketone benzofuran-thymidine phosphoramidite .....	80
Figure 3.10	HPLC trace of an enzymatic digestion of a psoralen ketone containing oligonucleotide .....	81
Figure 3.11.	HPLC analysis of the reaction between a 2-phenylindole aminoxy compound and a psoralen ketone bearing oligonucleotide .....	82
Figure 3.12	Estrogen receptor binding competition between estradiol and oligonucleotides containing 2PI-photolesions .....	83
Figure 3.13.	Construction of an internally labeled plasmid containing a site-specifically placed DNA lesion .....	84
Figure 3.14.	Scheme of an <i>in vitro</i> nucleotide excision repair assay .....	85
Figure 3.15.	Scheme of the repair blocking mechanism using an <i>in vitro</i> NER assay .....	85
Figure 3.16.	Construction of internally radiolabeled plasmids containing a single psoralen lesion .....	86
Figure 3.17.	An agarose gel demonstrating the presence of the psoralen lesion in a closed circular duplex DNA plasmid .....	87
Figure 3.18.	Kinetics of removal of psoralen adducts using HeLa cell extracts .....	88
Figure 3.19.	Effects of chloride ion and sodium orthovanadate concentration on the excision repair of psoralen-DNA adducts .....	90
Figure 3.20.	Effects of the estrogen receptor on the excision repair of psoralen lesions .....	91
Figure 4.1.	Structures of aflatoxin B <sub>1</sub> , its epoxide and the aflatoxin B <sub>1</sub> -induced DNA adducts .....	103
Figure 4.2.	Construction of aflatoxin B <sub>1</sub> containing oligonucleotides .....	104
Figure 4.3.	Experimental scheme for inhibition of intercalation .....	106
Figure 4.4.	Synthesis of a 2-carbomethoxybenzofuran-thymidine phosphoramidite .....	109
Figure 4.5.	Intercalation inhibition in a model system .....	111
Figure 4.6.	Intercalation inhibition with adjacent guanines .....	112
Figure 4.7.	Nearest neighbor inhibition effects .....	113
Figure 4.8.	Synthesis of aflatoxin B <sub>1</sub> adducts in an oligonucleotide derived from the human p53 sequence .....	114

## List of Tables

Table 2.1.	Comparison of the protecting groups and deprotection conditions for solid-phase DNA synthesis .....	58
Table 3.1.	Relative binding affinities of nitrogen mustard 2-phenylindole conjugates .....	71
Table 4.1.	Oligonucleotide sequences used in Chapter 4 .....	108
Table 4.2.	Relative monoadduct yield as a percentage of total monoadducts .....	121



## List of Abbreviations

2PI	2-phenylindole
2PI-C <sub>5</sub> -ONH <sub>2</sub>	2-phenylindole <i>N</i> -pentane hydroxylamine
2PI-C <sub>6</sub> -ONH <sub>2</sub>	2-phenylindole <i>N</i> -hexane hydroxylamine
AFB <sub>1</sub>	aflatoxin B <sub>1</sub>
AFB <sub>1</sub> -N7-Gua	8, 9-dihydro-8-(N7-guanyl)-9-hydroxyaflatoxin B <sub>1</sub> adduct
AMB	2-(acetoxymethyl)benzoyl
AMT	3-aminomethyltrioxsalen
AP site	apurinic site
ATP	adenosine triphosphate
BSA	bovine serum albumin
CCC	covalently closed circular
CD	circular dichroism
Chloramine-T	<i>N</i> -chloro- <i>p</i> -toluenesulfonamide, sodium salt
CPG	controlled pore glass
DBU	1, 8-diazabicyclo[5.4.0]undec-7-ene
cis-DDP	cis-diamminedichloroplatinum(II)
dNTP	deoxynucleoside triphosphate
DMT	4, 4-dimethoxytrityl chloride
ds	double-stranded
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene-bis(oxyethylenenitrilo)tetraacetic acid
EI	electron impact
EtBr	ethidium bromide
ER	estrogen receptor
FAB	fast atom bombardment
FAPY	formamidopyrimidine
HEPES	<i>N</i> -(2-hydroxyethyl)piperazine- <i>N'</i> -(2-ethanesulfonic acid)
HMG	high mobility group
HMT	3-hydroxymethyltrioxsalen
HPLC	high pressure liquid chromatography
HRMS	high resolution mass spectrometry
K <sub>d</sub>	dissociation constant
lin	linear
MALDI	matrix assisted laser desorption ionization
MOM	methoxymethyl
MOP	methoxy-psoralen

MS	mass spectrum
NER	nucleotide excision repair
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
OC	open circular
PAC	phenoxyacetyl
PAGE	polyacrylamide gel electrophoresis
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
RBA	relative binding affinity
ss	single-stranded
TCA	trichloroacetic acid
TBE	tris borate EDTA
TEMPO	2, 2, 6, 6-tetramethyl-1-piperidinyloxy, free radical
TMP	trioxsalen
Tris	Tris(hydroxymethyl)aminoethane
UV	ultra violet

# Introduction

Genetic material continuously encounters endogenous and exogenous DNA damaging agents that make a host of different modifications. These lesions consist of structural damage to the bases, deoxyribose and phosphate residues. Cells must repair this DNA damage in order to prevent the occurrence of mutations and lethal events. Although the entire processes from DNA damage to lethality are far from being completely understood, advances are being made on several fronts to probe these complicated pathways. Organic chemists have contributed to these efforts by synthesizing nucleosides damaged by UV and ionizing radiation, chemical carcinogens, and chemotherapeutic agents.<sup>1</sup> Subsequent incorporation of these modified nucleosides at predefined sites in oligonucleotides enables the design of biochemical experiments that are able to address incisively questions on the events that immediately follow damage to the genome.

Oligonucleotides containing a single DNA lesion have been utilized to determine the proteins involved in the four major pathways of DNA repair: base excision repair, nucleotide excision repair, direct reversal of damage and recombination.<sup>2</sup> The effects of DNA lesions on the replication and transcriptional apparatus have also been examined by using site-specifically modified oligonucleotides.<sup>3</sup> Synthetic methodology has provided large quantities of site-specifically modified oligonucleotides, which has allowed the high-resolution structural study of DNA lesions by x-ray crystallography and NMR spectroscopy.<sup>4</sup> Single-lesion oligonucleotide substrates have also been utilized to study how DNA damage effects chromatin assembly.<sup>5</sup> These structural studies have helped to explain mutagenic properties of a particular type of DNA damage.<sup>6</sup> Likewise, the mutagenic and cytotoxic study of anticancer drugs that work by covalent

modification of DNA has provided a knowledge base with which one can, in principle, design new anticancer agents that have diminished mutagenic potential, while preserving the cytotoxic properties necessary for therapeutic activity.<sup>7</sup>

The work presented in this dissertation describes the synthesis of oligonucleotides containing a site-specific psoralen adduct. I also describe some of the biochemical applications of these oligonucleotides. Chapter 1 reports the first total synthesis of a *cis*-syn furan-side psoralen-thymidine monoadduct. Subsequent incorporation of this DNA lesion into oligonucleotides by solid-phase synthesis is described in Chapter 2. In this Chapter, the photocrosslinking abilities of the furan-side psoralen-thymidine adduct is presented. Moreover, the accessibility to the pyrone-side monoadduct as well as the interstrand crosslink is demonstrated. In Chapter 3, the psoralen-thymidine monoadduct is derivatized in a first attempt to explore the possibility of a custom designed DNA damaging agent that can selectively kill cancer cells that overexpress a specific protein. Chapter 4 examines the unique *cis*-syn stereochemistry of the psoralen-thymidine monoadduct and exploits this structural feature to alter the target selectivity of other DNA damaging agents. To this end, the synthesis of aflatoxin B<sub>1</sub> adducts in a p53 mutational hotspot, as well other sequence contexts, is described.

In addition to the experiments described in this dissertation, the *cis*-syn furan-side psoralen-thymidine monoadduct and its conversion to the intrastrand crosslink has been utilized for other applications. Many DNA damaging agents form crosslinks and it would be valuable to have a convenient source of site-specific crosslinks. The conversion of the furan-side monoadduct into the crosslink affords a ready source of site-specific interstrand crosslinks. The site-specific psoralen-thymidine adduct is currently being used in ongoing experiments to



---

examine the role of the human nucleotide excision repair apparatus role in the removal of interstrand crosslinks. In addition, the DNA sequence context freedom provided by this methodology will allow for the full utilization of the cross-linking properties of site-specific psoralen containing oligonucleotides in hybridization assays as well as triple helix forming oligonucleotides. Presently, the site-specific psoralen crosslink is being used to ascertain whether electrospray mass spectrometry can accurately determine an intrastand crosslink's sequence location within a DNA helix. This methodology would improve the current analytical tools used to characterize damaged DNA. These current experiments as well as the experiments described in this dissertation would have been impractical, if not impossible, without the design and synthesis of a furan-side psoralen-thymidine monoadduct and its a site-specific incorporation into oligonucleotides.



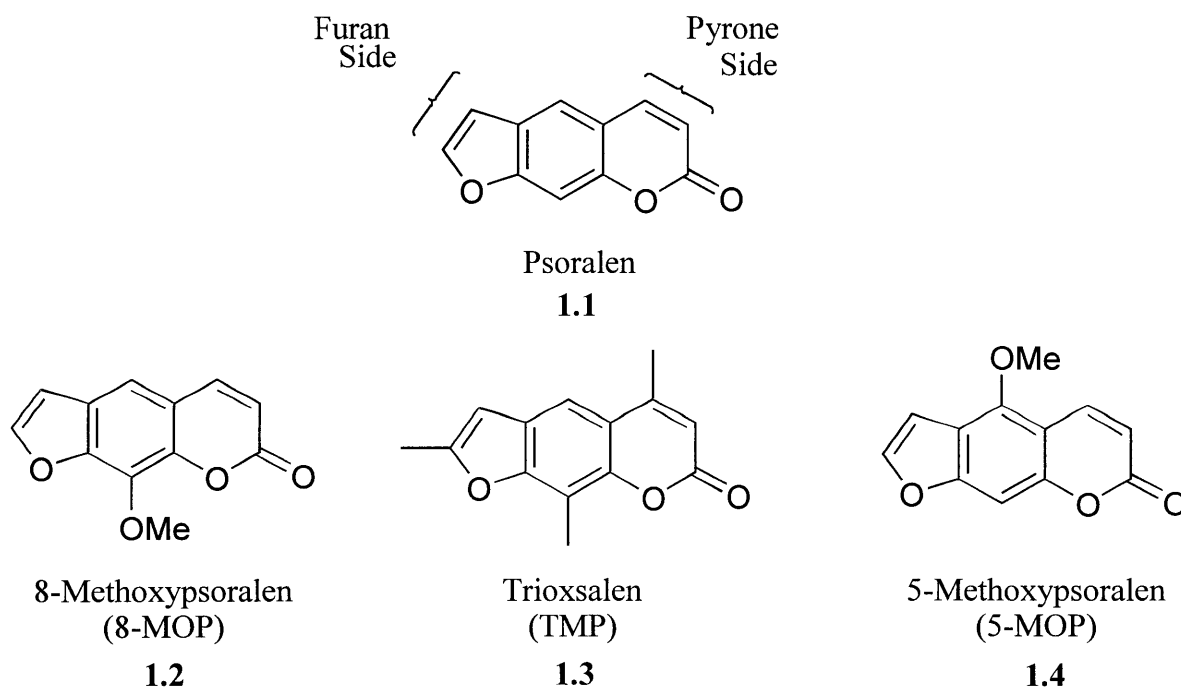
## Chapter 1

# Synthesis of a *Cis-Syn* Furan-Side Psoralen-Thymidine Monoadduct

### Background

The linear isomers of the furocoumarin family known as the psoralens were first recognized as therapeutic agents by the ancient Egyptians who employed fruit and vegetable extracts containing these agents for the treatment of vitiligo.<sup>8</sup> More recently, psoralens have been used medicinally for the treatment of psoriasis, eczema, and cutaneous T-cell lymphoma.<sup>9</sup> The cellular target for the psoralens is believed to be DNA. Cells treated with psoralen and long-wave UV light form adducts involving the pyrimidines in double-stranded nucleic acids. The mechanism of the photoreaction between psoralen and a nucleic acid helix involves several steps. The first step is the intercalation between the hydrophobic bases of double-stranded nucleic acids is also known as the dark reaction. Exposure to long-wave UV light (320-410 nm) results in either the pyrone or furan double bond of psoralen reacting via [2 + 2] cycloaddition with the 5,6-double bond of a pyrimidine to form an initial monoadduct. After a structural reorganization that takes approximately 1  $\mu$ s, a furan-side adduct can absorb a second photon and react with an adjacent pyrimidine on the opposite strand forming an interstrand crosslink.<sup>10</sup> By contrast, pyrone-side monoadducts cannot absorb long-wave UV light and therefore do not form crosslinks. When irradiated in the presence of model oligonucleotides, psoralens show sequence selectivity as evidenced by the 100-fold preference for reaction at 5'-TpA as compared to 5'-ApT or 5'-TpG sites.<sup>11</sup> All psoralen photoadducts can be reversed by 254 nm light;<sup>12</sup> crosslinks and furan-side adducts can also be reversed by treatment with base.<sup>13</sup>

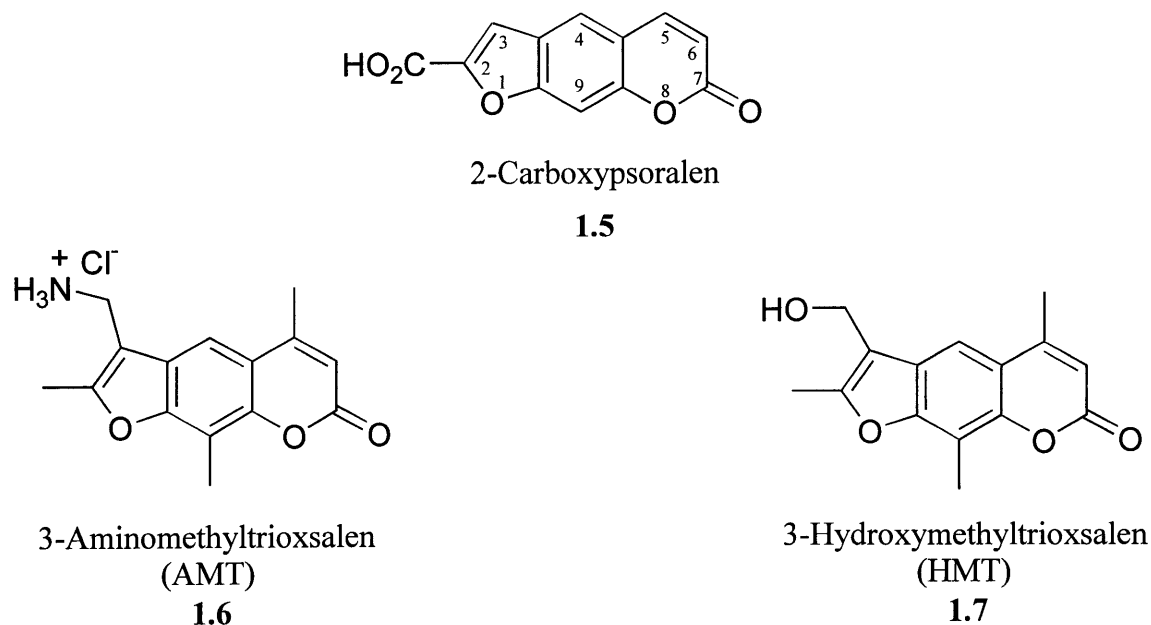
Psoralens have been isolated from plants typically from the Leguminosae, Umbelliferae and Rutaceae families.<sup>14</sup> Examples of some of the naturally occurring psoralens are shown in Figure 1.1. Synthetic efforts toward the synthesis of psoralens has been formidable with over 60 years of published experience in this area.<sup>15</sup> Isolated psoralens are typically hydrophobic and require aqueous/organic mixtures in order to solubilize them sufficiently to achieve effective therapeutic concentrations.<sup>16</sup> In order to improve their medicinal qualities, psoralen derivatives have been synthesized with the aim of enhancing water solubility and photoreactivity. Toward



**Figure 1.1.** Structures of the naturally occurring psoralens.

this end, the formal psoralen compound **1.1** has been derivatized at every non-quaternary carbon. Some of the more common synthetic derivatives are 2-carbomethoxypsoralen **1.5**, 3-aminomethyltrioxsalen **1.6** (AMT) and 3-hydroxymethyltrioxsalen (HMT) **1.7** (Fig 1.2).

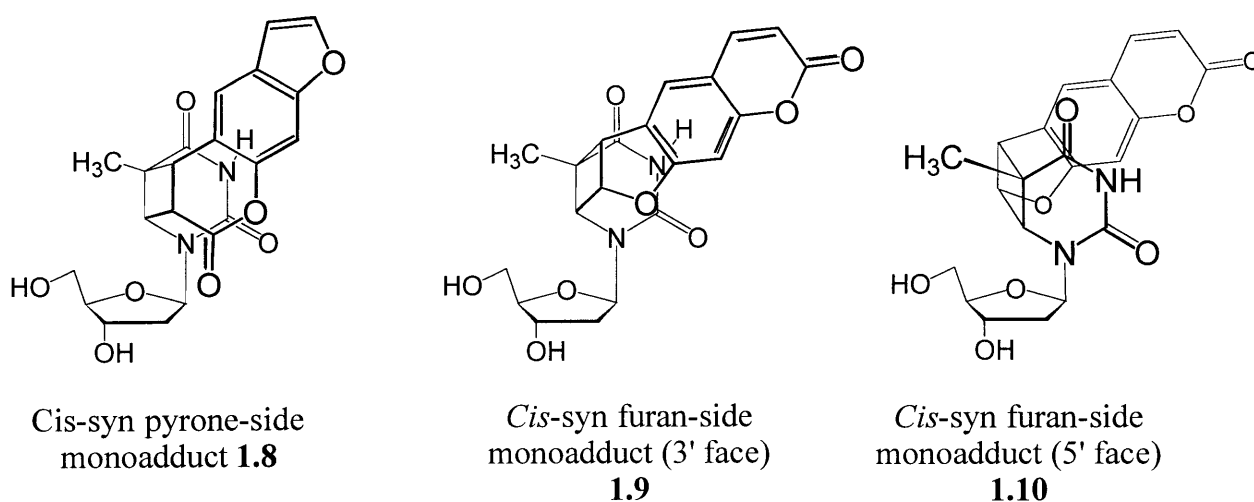
Typically, the numbering system of furocoumarins is based on the psoralen nucleus; however, the Ring Index nomenclature will be used because the numbering system remains constant when comparing different aromatic ring systems.



**Figure 1.2.** Structures of some of the common synthetic psoralen derivatives. The numbering system shown in 2-carboxypsoralen **1.5** is based on the Ring Index nomenclature and will be used herein.

Hearst and others have put much effort into the isolation and characterization of the psoralen-DNA adducts.<sup>12b, 17</sup> To date, three major monoadducts and one interstrand crosslink have been characterized. The structures and stereochemistries of the isolated adducts are shown in Figure 1.3. The monoadducts consist of a psoralen-thymidine *cis-syn* pyrone-side adduct **1.8** and both diastereomers of a *cis-syn* furan-side adduct **1.9** and **1.10**. The two furan-side adducts arise from reaction of the psoralen on either the 5' or the 3' face of thymidine. The interstrand crosslink is comprised of a furan and pyrone *cis-syn* monoadducts with thymidine. Psoralen

reaction with cytosines has been reported;<sup>18</sup> however, the limited reactivity between psoralen and cytosine has hampered the structural characterization of these lesions. Presumably they possess the same regio- and stereochemistry as determined for the thymidine adducts because furan-side cytidine monoadducts can be converted into crosslinks with subsequent treatment of long-wave UV light.



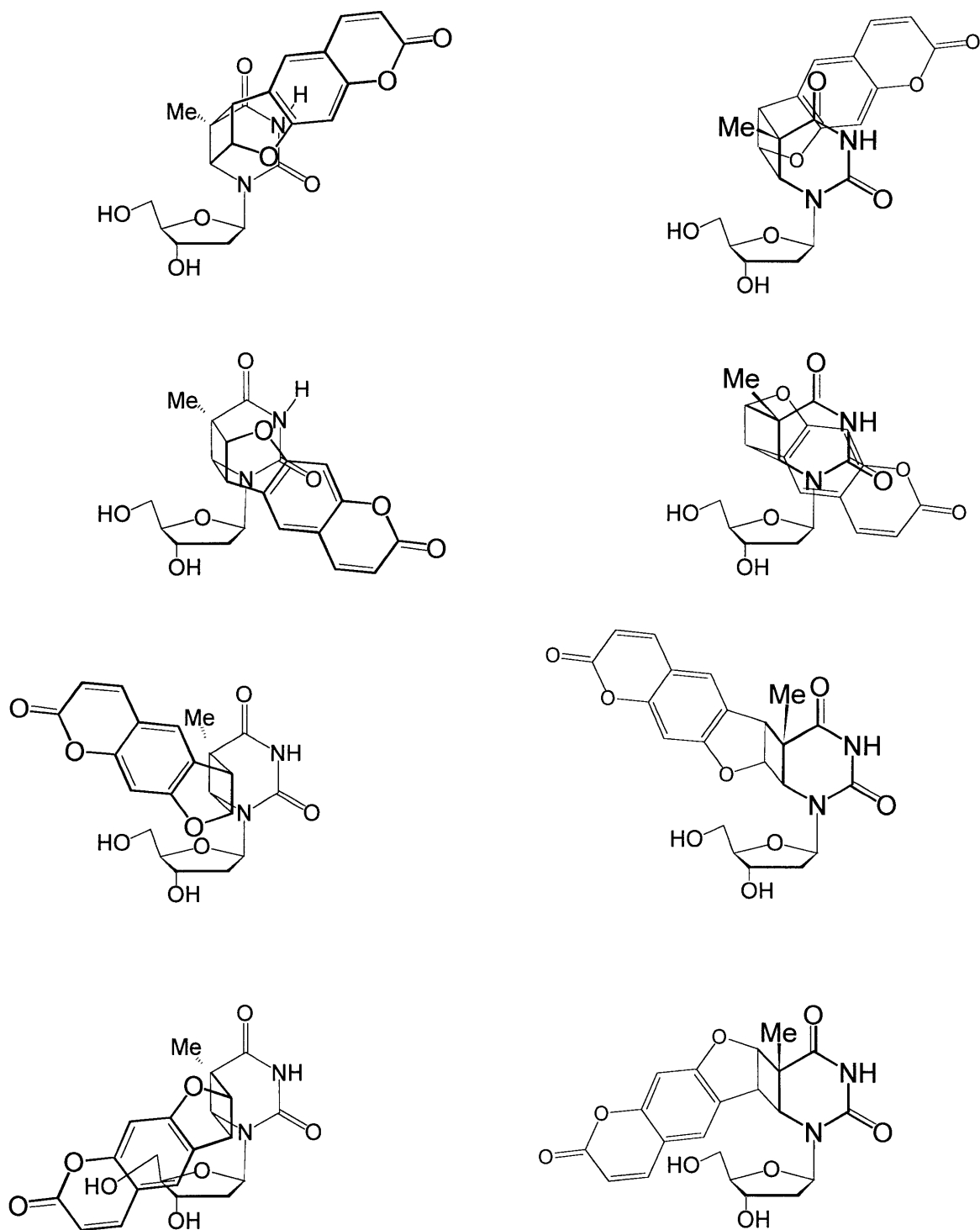
**Figure 1.3.** Structures of the isolated psoralen-thymidine monoadducts.

### Synthetic Strategy

Attempts to synthesize any *cis-syn* psoralen-thymidine adduct in the absence of DNA have met with modest success. DNA controls the stereochemistry of the photochemical reaction because the hydrophobic psoralen molecule intercalates between the nucleobases aligning the reactive double-bonds to afford *cis-syn* photoproducts. The stereoselectivity of the photoreaction in DNA is exceptional given that 64 monoadducts are theoretically possible between thymidine and psoralen; 48 of these adducts, however, do not form apparently because they would contain a highly strained and sterically hindered *trans* fusion about the pyrimidine or psoralen rings.

Elimination of these strained isomers still leaves 8 possible isomers for each monoadduct (Fig. 1.4). Of these isomers, the *cis-syn* adduct that occurs on the 3' face of the thymidine accounts for more than 99% of the total furan-side monoadducts.<sup>11</sup> Psoralen dimerization is another major synthetic obstacle because psoralen reacts with itself faster than with thymidine or thymine in water/organic mixtures.<sup>19</sup> The DNA helix prevents photodimerization by separating the psoralen molecules between the nucleobases. A successful total synthetic approach would have to duplicate the control the DNA helix exerts on the psoralen molecule to afford only *cis-syn* stereoisomers and to eliminate psoralen photodimerization.

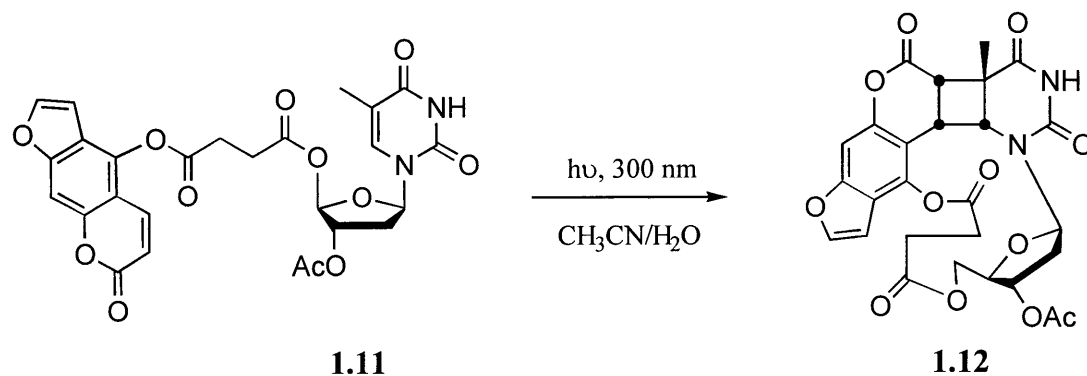
Some previous approaches to mimic the intercalation step involved irradiation of frozen aqueous solutions or evaporated thin films containing psoralen and thymine.<sup>20</sup> The rationale behind these experiments was that these solid state irradiations would minimize psoralen photodimerization and allow for the isolation of the desired *cis-syn* isomer. Unfortunately, numerous photoproducts, including photodimers of psoralen, were isolated and the desired *cis-syn* stereochemical isomer was purified in less than 3% yield. A more promising strategy by Lhomme and coworkers utilized succinic acid to attach a hydroxypsoralen derivative to the 5'-hydroxyl of thymidine in an attempt to control stereochemistry and eliminate psoralen photodimerization.<sup>21</sup> This method completely eliminated psoralen photodimerization and afforded only one psoralen-thymidine monoadduct **1.12** in 70% yield (Fig. 1.5). Unfortunately, the product isolated was the pyrone *cis-anti* isomer and not the desired *cis-syn* isomer. In addition, the removal of the succinyl linker was not reported because its removal would result in ring opening of the psoralen lactone. Closure of this molecule would result in side products due to the symmetry of the resulting catechol derivative. Furthermore, any attempts to use this



**Figure 1.4.** Structures of the eight possible furan-side monoadducts.



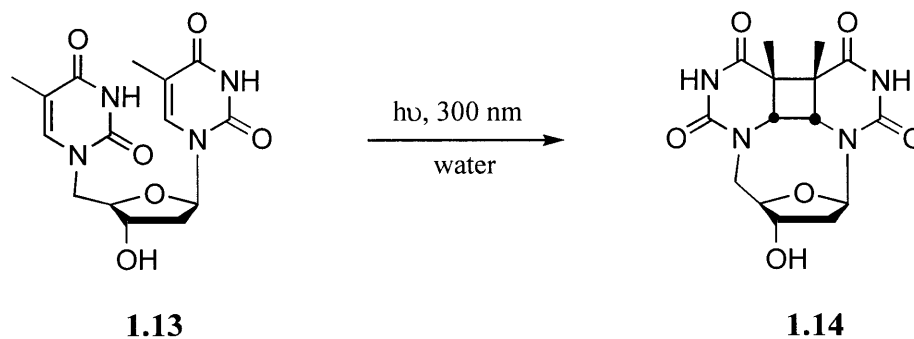
pyrone-side adduct in DNA synthesis would require a non-basic deprotection, which would involve the synthesis of specialized phosphoramidites for solid-phase oligonucleotide synthesis.



**Figure 1.5.** Synthesis of a *cis*-anti pyrone-side monoadduct **1.12**.<sup>21</sup>

Based upon the aforementioned discussion, there are several obstacles that must be addressed in order to synthesize a *cis*-syn psoralen-thymidine monoadduct. First, psoralen photodimerization would have to be eliminated. A suitably functionalized benzofuran derivative was chosen as a psoralen synthon because benzofuran is missing the photochemically more reactive pyrone ring of psoralen, thereby eliminating the possibility for psoralen dimerization.<sup>19,21</sup> Moreover, only furan-side photoproducts can form using a benzofuran derivative. Interestingly, the furan double bond has been shown to be the more reactive double bond when intercalated in DNA owing to its juxtaposition to the 5,6 thymidine double in a 5'-TpA site.<sup>22</sup> Control of the stereochemical outcome of the photoreaction to achieve the desired *cis*-syn isomer would be the second crucial aspect in the photochemical step. Lhomme et al. demonstrated that stereochemical control can be achieved by attachment of a psoralen derivative to the 5' hydroxyl of the sugar of thymidine; however, their flexible succinyl linker allowed for the synthesis of *cis*-

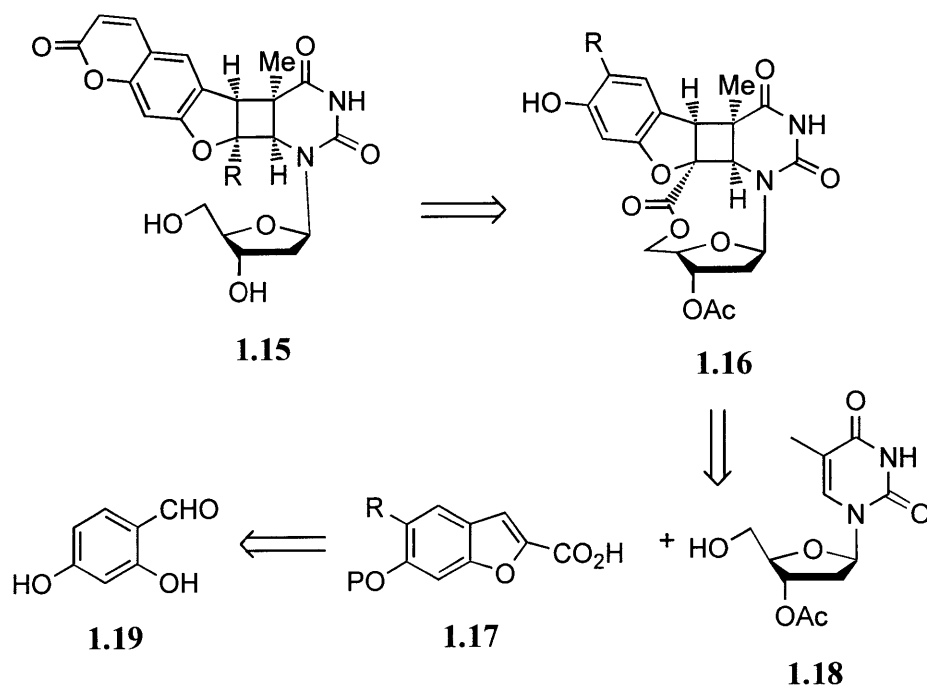
anti adducts.<sup>21</sup> Leonard and co-workers have utilized a ribose ring to control the stereochemical outcome of some "abbreviated" dinucleosides of thymidine.<sup>23</sup> These researchers synthesized two thymines held in close proximity by a 2-deoxyribose (Fig. 1.6). Irradiation of **1.13** afforded *cis-syn* stereoisomer **1.14** in 88% yield. The success of this approach is evidenced by the elimination of the anti isomers due to the geometrical restraint imposed by the attachment of both photoreactive groups to the single deoxyribose moiety. Based on this experimental evidence, a one carbon ester linkage was chosen in my work, to attach the benzofuran derivative to the 5'-hydroxyl of thymidine. The intramolecular photochemical reaction between benzofuran and thymidine double bonds should afford the desired *cis-syn* isomer. This monoadduct could then be further elaborated into the desired psoralen-thymidine monoadduct.



**Figure 1.6.** Synthesis of a *cis-syn* "abbreviated" thymine dimer **1.14**.<sup>23</sup>

### Synthesis of Benzofuran-Thymidine Photoproducts

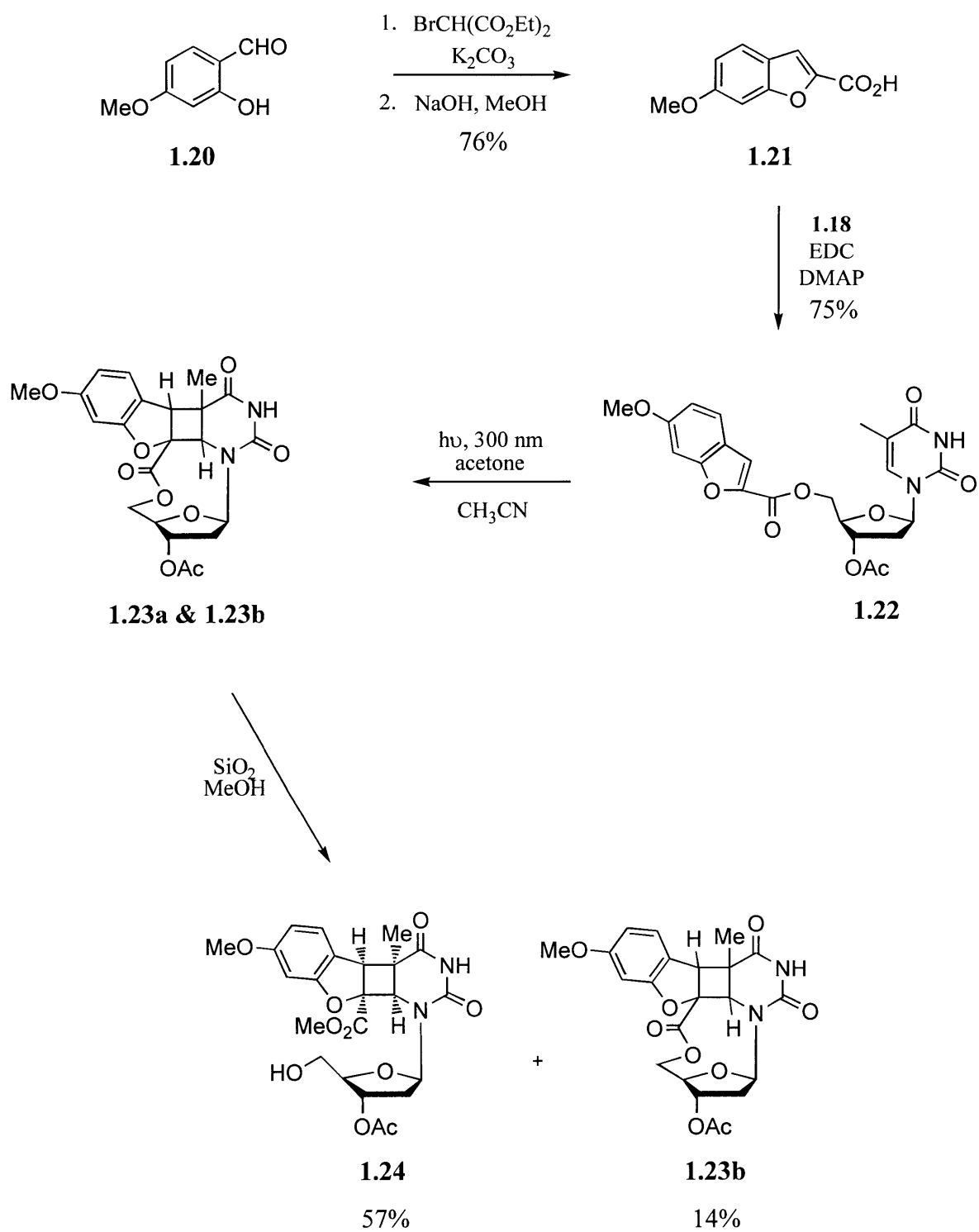
The retrosynthetic route of a *cis-syn* 2-carbomethoxypsoralen-thymidine furan-side adduct **1.15** is shown in Figure 1.7. Removal of the pyrone ring would provide benzofuran photoproduct **1.16**. Photoproduct **1.16** could be formed in a stereospecific manner from a benzofuranyl acid **1.17** linked to the 5'-hydroxyl of thymidine **1.18**.<sup>21</sup> Benzofuranyl acid **1.17** is



**Figure 1.7.** Retrosynthetic route to *cis-syn* 2-carbomethoxy-psoralen-thymidine furan-side monoadduct **1.15**.

readily obtained from salicylaldehyde **1.19**. The key step in the total synthesis would be the photochemical reaction with benzofuran. The photochemistry of benzofuran esters was not known, and therefore it was necessary to test this strategy in a model system. The primary goal of the model system was to direct the stereochemical outcome of the photochemical reaction to obtain the desired *cis-syn* stereochemistry. The successful application of this model is shown in Figure 1.8.

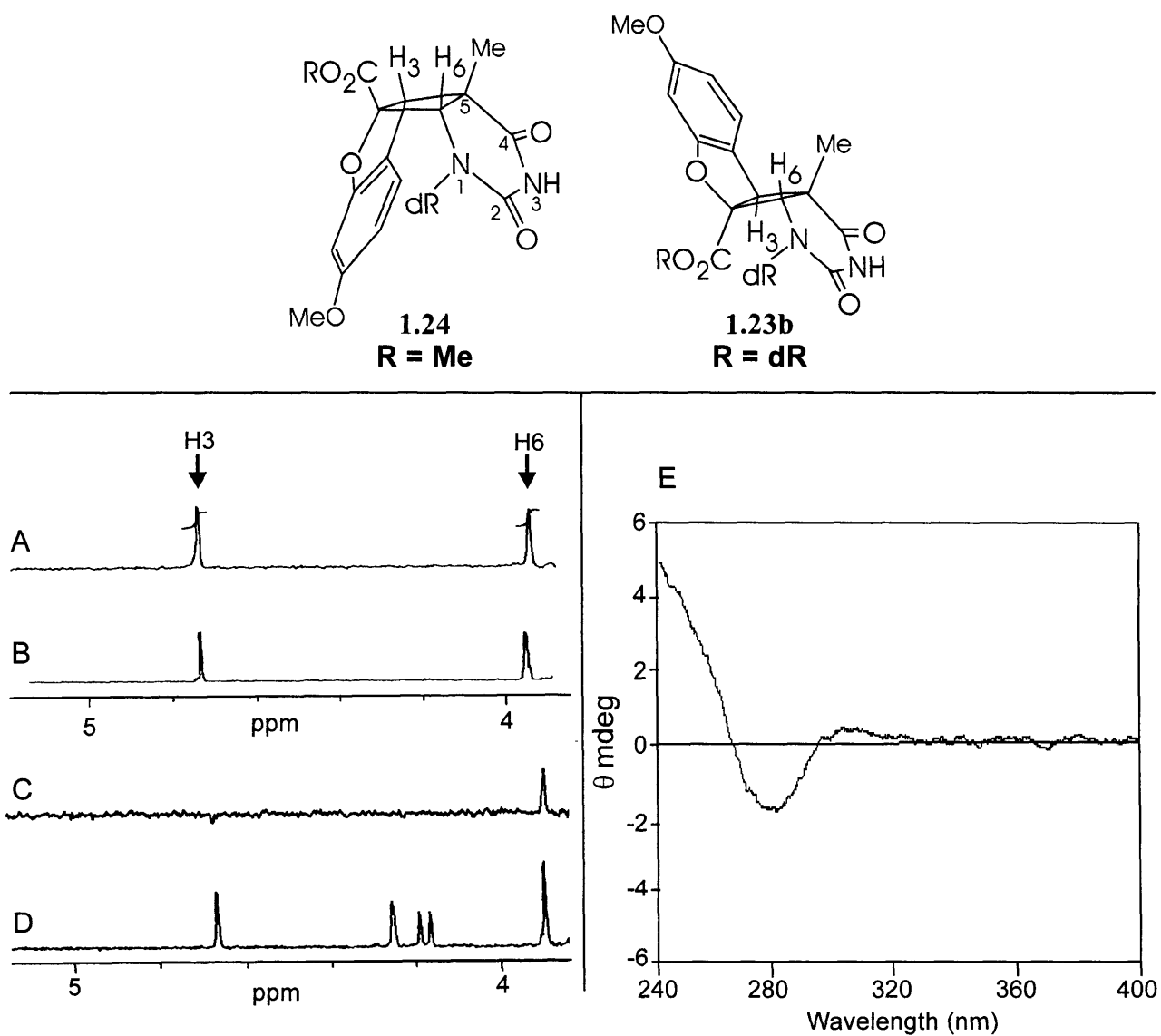
Commercially available salicylaldehyde **1.20** was allowed to react with diethyl bromomalonate to give a benzofuran ester, which was saponified to afford benzofuranyl acid **1.21**. Esterification of **1.21** with a 3'-protected thymidine, **1.18**, provided photoprecursor **1.22** in 75% yield. Irradiation of an acetone sensitized, argon degassed, dilute (< 2 mM) solution of **1.22**



**Figure 1.8.** Synthesis of *cis*-syn benzofuran-thymidine monoadduct **1.24**.

in acetonitrile afforded two diastereomers with a 4:1 ratio in 71% overall yield. Purification of photoproducts **1.23a** and **1.23b** by using standard silica gel chromatography with methylene chloride and methanol as an eluent, transesterified the 9-membered lactone of the major isomer **1.23a** to give ring opened photoproduct **1.24** containing a methyl ester. The minor isomer **1.23b**, did not transesterify with methanol and silica gel and could be purified as the lactone using methanol as the eluent.

NMR was used to determine both the regio- and stereo-chemistry of the isolated photoproducts **1.24** and **1.23b**. The regiochemistry of both photoproducts was determined by observing the coupling constant between the H6 proton of thymidine and the H3 proton of benzofuran. Only long range coupling (2 Hz) was observed for both photoproducts, which is indicative of syn stereochemistry.<sup>24</sup> To determine if the thymidine methyl group were on the same side of the four membered ring as the H3 and H6 protons, an NOE difference experiment was performed (Fig. 1.9). Irradiation of the thymidine methyl group of major isomer **1.24** gave a positive NOE for both ring protons, which confirmed that the major photoproduct possessed the desired *cis*-stereochemistry. Irradiation of the thymidine methyl group of minor isomer **1.23b** gave only one positive NOE for the H6 proton of thymidine (Figure 1.9C). Hence, the minor isomer is the *trans*-syn isomer. Attempts to crystallize the major photoproduct, to assign on which face of the thymidine ring the [2 + 2] cycloaddition occurred, have yet to yield a crystal of X-ray diffraction quality. However, Hearst et al. have shown that the CD spectrum of a 5'-TpA psoralen adduct has a negative ellipticity from 270 to 360 nm, whereas the 5'-ApT adduct has a positive ellipticity in the same range.<sup>11a</sup> It appears that the contribution of the deoxyribose ring affects the CD spectrum minimally and the two diastereomers behave as enantiomers in the

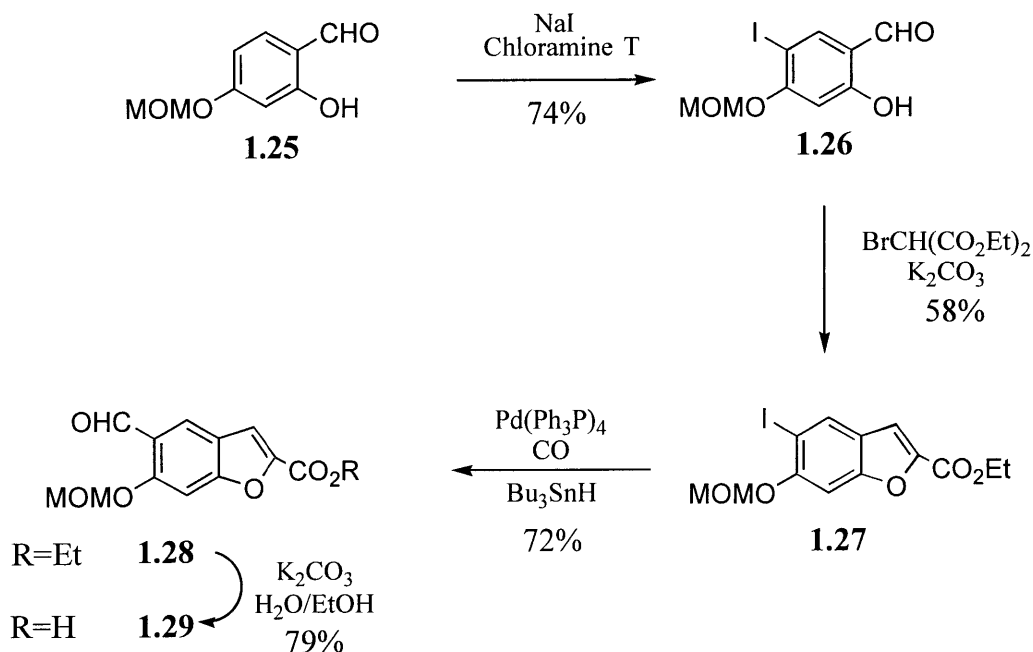


**Figure 1.9.** (Above) Structures of the benzofuran-thymidine photoproducts **1.24** and **1.23b**, where dR = 2'-deoxy-3'-O-acetylribose. (Below, left) Nuclear Overhauser enhancements for photoproducts **1.24** and **1.23b**. (A) difference spectrum between the major isomer **1.24** irradiated at the thymidine methyl resonance (1.78 ppm) and (B). (B) proton NMR spectrum of the major isomer **1.24** (4.0-5.0 ppm). (C) difference spectrum between the minor isomer **1.24** irradiated at the thymidine methyl resonance (1.28 ppm) and (D); (D) proton NMR spectrum of the minor isomer **1.23b** (4.0-5.0 ppm). (E) circular dichroism spectrum of the major isomer **1.24**.

presence of polarized light. Although the model compound did not contain the full coumarin ring, it seemed possible to ascertain the facial selectivity of the photoreaction by CD spectroscopy. The CD spectrum of the major isomer **1.24**, shown in Figure 1.9E, has a negative ellipticity from 260 to 290 nm suggesting the cycloaddition had occurred on the 3' face of thymidine. The benzofuran-thymidine model system had fulfilled the original goals of directing the stereochemical outcome of the photoreaction and eliminating photodimerization. In addition, the propensity of the *cis-syn* isomer to transesterify with methanol facilitated the rapid purification of the desired isomer.

### **Synthesis of a *Cis-Syn* Furan-Side Psoralen-Thymidine Monoadduct**

Having firmly established a methodology to construct a *cis-syn* photoproduct by using benzofuran as a psoralen guise, the total synthesis of a psoralen-thymidine adduct was within reach. Two minor modifications of the model system were necessary to convert it into a psoralen-thymidine adduct. First, a simple change of a protecting group on the phenolic alcohol would allow for a mild deprotection when coupled to the nucleoside. The second modification was to incorporate an aldehyde *ortho* to the methoxymethyl (MOM) protected phenol. Addition of a formyl group at this position would allow for a mild conversion of the benzofuran photoproduct into a psoralen-thymidine adduct.<sup>25</sup> Earlier formylations of 6-alkoxybenzofurans utilized DMF-POCl<sub>3</sub> in order to formylate the benzofuran directly. However, these conditions resulted in primarily 7-formylbenzofurans and afforded little 5-formyl product.<sup>16d</sup> Figure 1.10 depicts a two step approach utilizing halogenation followed by a palladium(0) catalyzed formylation<sup>26</sup> to attain a 5-formylbenzofuran. Treatment of selectively protected **1.25**<sup>27</sup> with

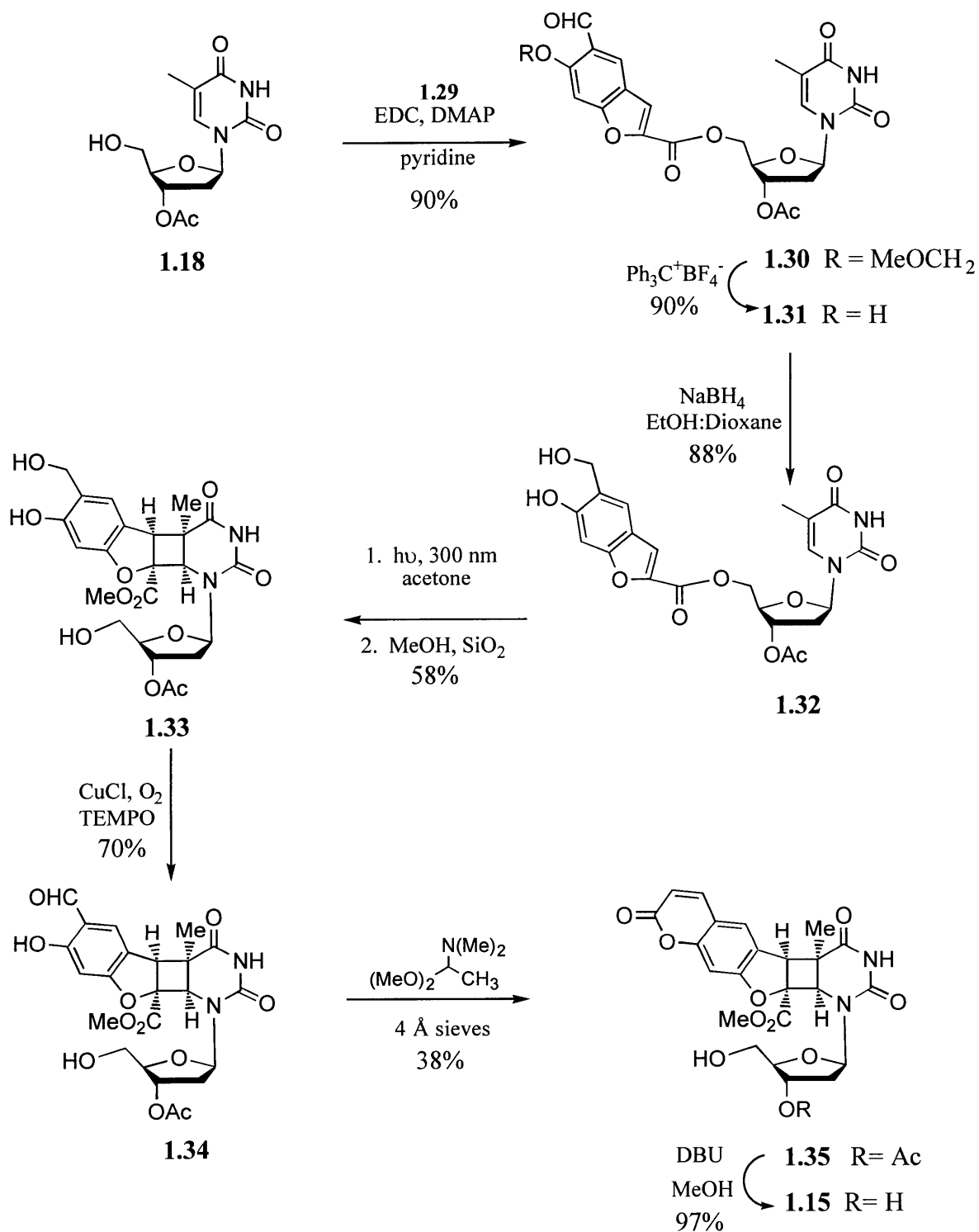


**Figure 1.10.** Synthesis of 5-formyl-6-methoxymethyl-2-benzofuran carboxylic acid **1.29**

sodium iodide and chloramine T gave primarily *para* iodinated **1.26**.<sup>28</sup> Cyclization of **1.26** to benzofuran **1.27** was accomplished with diethyl bromomalonate and potassium carbonate. The desired aldehyde was synthesized using a Stille coupling.<sup>26</sup> Palladium(0) coupling with **1.27**, carbon monoxide, and slow addition of tributyltin hydride gave 5-formylbenzofuran **1.28**. The overall two step yield for incorporation of the 5-aldehyde was 53%. Saponification of **1.28** yielded benzofuranyl acid **1.29**.<sup>29</sup>

With the benzofuran portion of the molecule in hand, the next task was to couple it to thymidine and produce a photoprecursor. The synthesis began with carbodiimide coupling of **1.29** with 3'-*O*-acetylthymidine **1.18** to give esterified **1.30**, as shown in Figure 1.11. Removal of the protecting group using trityl cation afforded phenol **1.31** in 90% yield.<sup>30</sup> Efforts to use HCl/methanol to remove the MOM group led to side products and lower yields. All attempts at

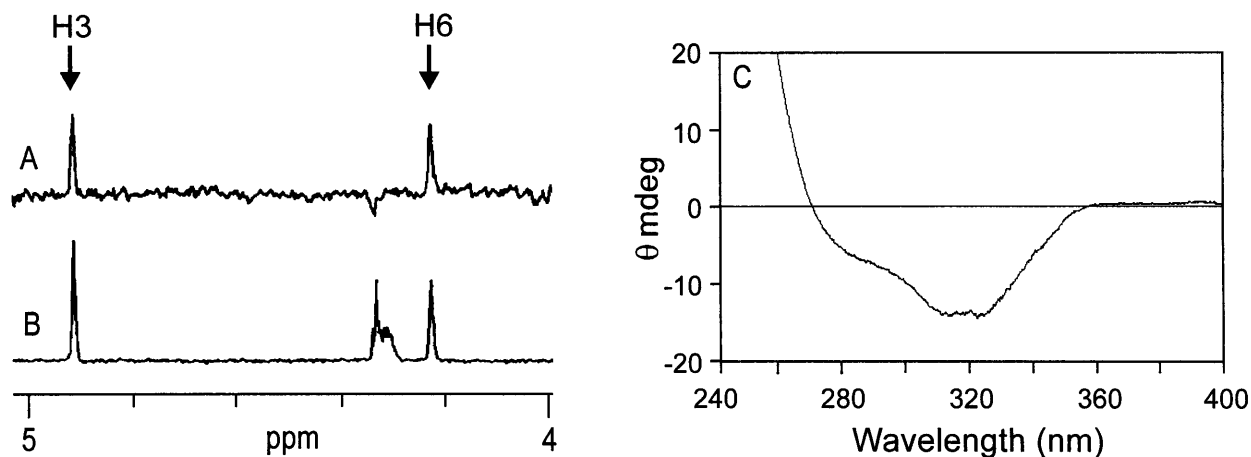




**Figure 1.11.** Synthesis of a *cis-syn* 2-carbomethoxypsoralen furan-side thymidine monoadduct **1.15**.

photochemistry with aldehyde **1.31** resulted in no reaction presumably owing to fluorescence quenching by the benzaldehyde moiety. To circumvent this problem, aldehyde **1.31** was protected by reduction with NaBH<sub>4</sub> to the readily oxidizable benzylic alcohol **1.32**. A dilute solution of photoprecursor **1.32** in 20:1 acetonitrile/acetone was irradiated with 300 nm light and afforded only one diastereomer, which was transesterified with silica gel and methanol to produce the ring opened photoproduct **1.33** in 58% yield. Selective oxidation of **1.33** with oxygen in the presence of CuCl and TEMPO afforded aldehyde **1.34**.<sup>31</sup> Typically harsh, basic conditions are needed to construct the pyrone ring from *ortho* formyl phenols.<sup>32</sup> Barton and others have used enediamines or commercially available acetamide acetals to effect this transformation under relatively mild conditions.<sup>25</sup> Treatment of aldehyde **1.34** with N',N'-dimethylacetamide dimethylacetal in the presence of 4 Å sieves at room temperature gave coumarin **1.35** in 38% yield. Although the yield for this step is moderate, the chemoselectivity of the reagent to form a carbon-carbon double bond in the presence of a free primary hydroxyl and imide is noteworthy. The protected 2-carbomethoxypsoralen-thymidine adduct was converted into the desired photoproduct **1.15** using DBU in methanol.

Determination of the stereochemistry of photoproduct **1.15** was identical to the model system. The coupling constant between the H6 thymidine proton and the H3 benzofuran proton of psoralen was small, 1.7 Hz, confirming the *syn* regioisomer. Figure 1.12 shows the difference NOE experiment where irradiation of the thymidine methyl shows an enhancement of both H6 and H3 protons, confirming the stereochemistry as *cis*. The final stereochemical determination was whether the 2-carbomethoxypsoralen was situated on the 5' or 3' face of thymidine. The CD spectrum of the synthesized adduct **1.15**, was nearly identical with the CD spectra of an isolated



**Figure 1.12.** (Left) Nuclear Overhauser enhancements for the 2-carbomethoxypsoralen-thymidine furan-side monoadduct **1.15**. (A) difference spectrum between sample irradiated at the thymidine methyl resonance (1.80 ppm) and (B). (B) 4.0-5.0 ppm region of the proton NMR spectrum. (C) Circular dichroism spectrum of the 2-carbomethoxypsoralen-thymidine furan-side monoadduct **1.15**.

psoralen-thymidine adduct even though the psoralen moieties are slightly different (Fig.

1.12C).<sup>11a, 17f</sup> The synthesized 2-carbomethoxypsoralen-thymidine adduct is the adduct equivalent to the reaction occurring at a 5'-TpA site in DNA; this is the hundred fold most prevalent adduct under normal conditions of DNA damage.

There were many key synthetic steps developed for the eventual synthesis of a *cis-syn* furan-side psoralen-thymidine monoadduct. Selective iodination of the salicylaldehyde derivative allowed for the incorporation of an aldehyde at the 5 position via a Stille formylation. The presence of the nucleosidic bond prevented the use of harsh mineral acids or bases, which eliminated many standard synthetic methodologies. Therefore, milder synthetic reactions were employed to maintain this important, labile bond. Using trityl cation for MOM-deprotection as well as pyrone ring annulation via Barton enediamines were just two of the functional group

manipulations that were critical in achieving this total synthesis. However, the success of this total synthesis was ultimately due the utilization of a benzofuran derivative in the key intramolecular photochemical step. This approach allowed for the sole formation of a *cis*-syn furan-side monoadduct without the complications of photodimerization.

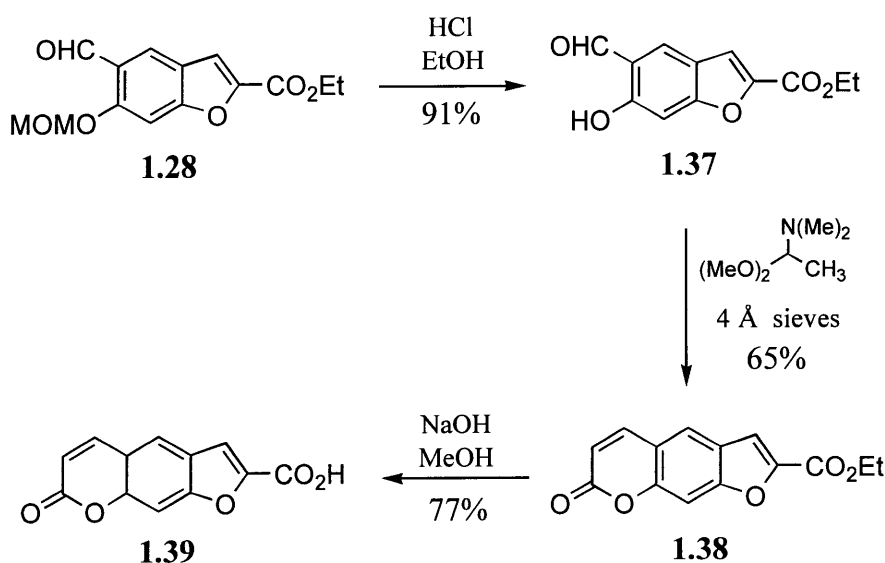
### A More Efficient Synthesis

The previous "two-stage" synthetic strategy eliminated photodimerization and yielded a single photoproduct containing a *cis*-syn geometry. While this approach gave the desired monoadduct in reasonable yield, it involved many steps with a delicate saturated pyrimidine. A more direct approach would utilize a full psoralen derivative in lieu of benzofuran; this approach is more appealing because it shortens the synthetic scheme and eliminates all but one reaction with a saturated pyrimidine. At the outset, a major concern with using an intact psoralen was that it could lead to self-dimerization instead of a thymidine-psoralen photoproduct.<sup>19,21</sup> The efficiency of the benzofuran-thymidine photochemical reactions suggested that a 2'-ester linkage might favor the intramolecular reaction between psoralen and thymidine over psoralen photodimerization. This section explores this highly effective linkage using an intact psoralen derivative.

The strategy involves linking a 2-carboxypsoralen to the 5'-hydroxyl of thymidine to control the stereochemistry of the photoreaction. Based on the earlier work with benzofuran-thymidine photoproducts, a new method was developed for the synthesis of 2'-carboethoxypsoralen shown in Figure 1.13. Removal of the MOM-protecting group from **1.28** with HCl in ethanol yielded phenol **1.37**. Cyclization of compound **1.37** with dimethylacetamide

dimethyl acetal in the presence of 4 Å sieves gave carboethoxypsoralen **1.38** in 65% yield.<sup>25</sup>

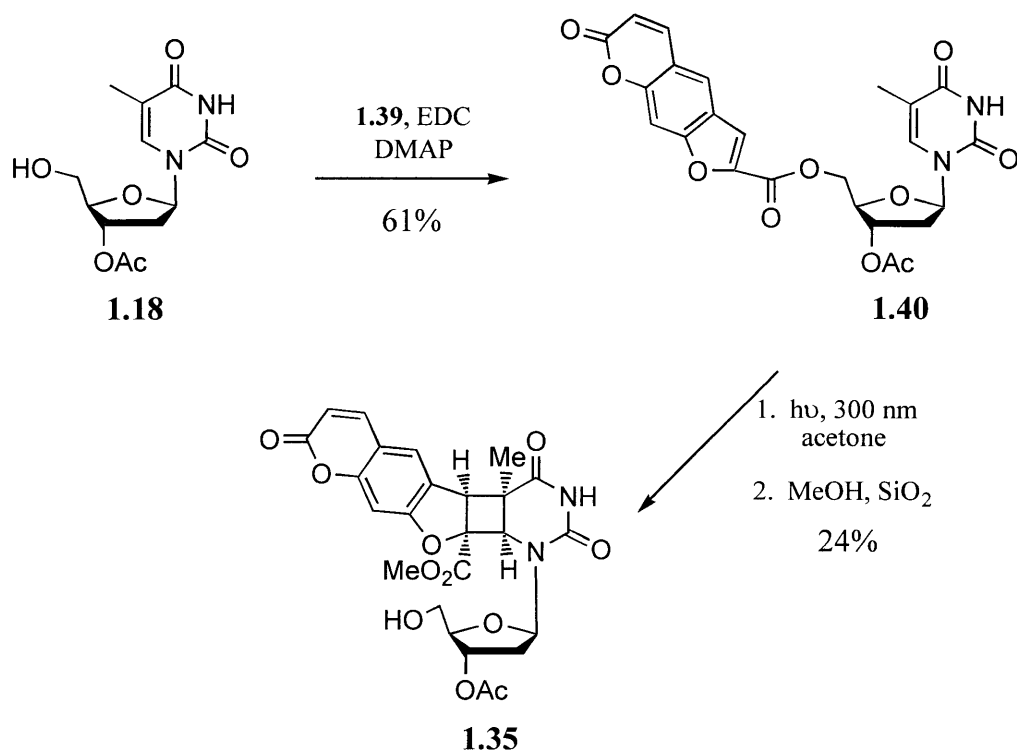
Again, the addition of sieves removed the methanol that is generated in this reaction and greatly improved the yield. Saponification of **1.38** afforded acid **1.39**. The 2-carboxypsoralen was then attached to a suitably protected thymidine.<sup>21</sup> As shown in Figure 1.14, coupling of psoralen **1.39**



**Figure 1.13.** Synthesis of 2-carboxypsoralen **1.39**.

with thymidine **1.18** yielded photoprecursor **1.40**. Attempts to effect the photochemical reaction by using long wave UV light (366 nm) produced only photodegradation of the starting material, compound **1.40**. However, the use of 300 nm light and acetone as a photosensitizer yielded one photoproduct, which underwent a transesterification with methanol and silica gel to afford **1.35** in 24% yield. The low yield was presumably due to some photodimerization even though the reactions were performed at low concentration (1 mM).

The regio- and stereochemistry of photoproduct **1.35** was again assigned by NMR and CD spectroscopy. The 1D NMR spectra of **1.35** was identical, including the regiochemistry, to



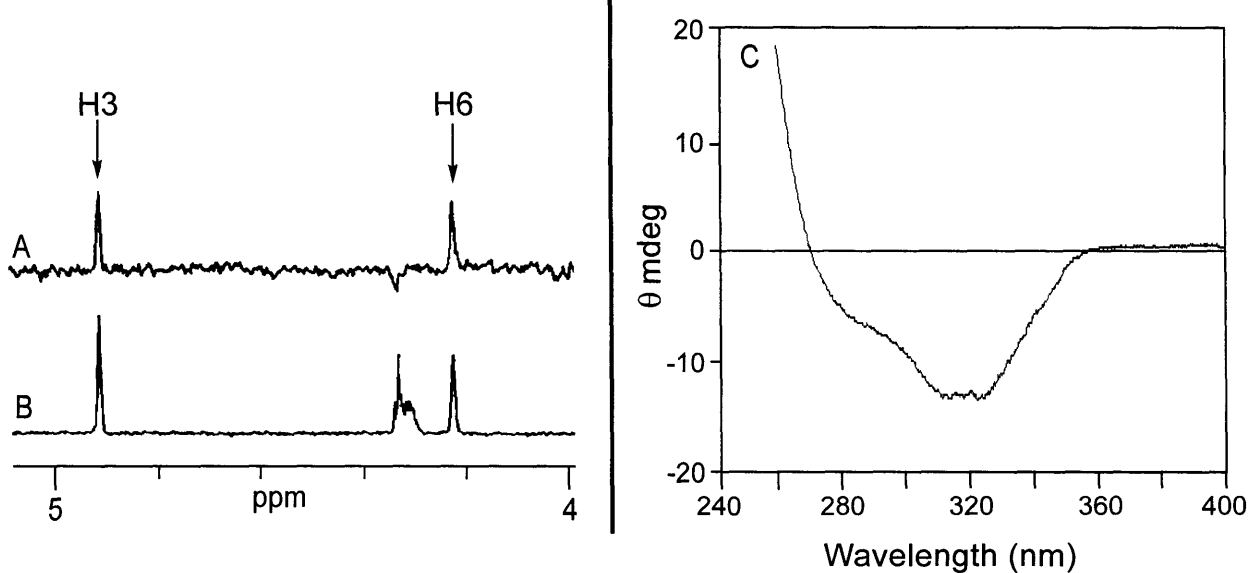
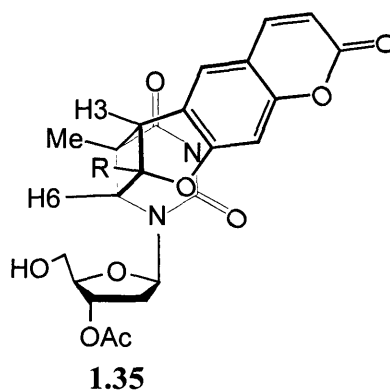
**Figure 1.14.** Facile synthesis of a *cis*-syn 2-carbomethoxypsoralen furan-side thymidine monoadduct **1.35**.

the previously synthesized furan-side adduct. A difference NOE experiment resulted in positive signals for both protons on the four-membered ring assigning the stereochemistry as *cis* (Fig. 1.15). Determination of the facial selectivity of the photoreaction was ascertained by CD spectroscopy (Figure 1.15). The CD spectrum of **1.35** was identical to the photoproduct generated by the earlier method and to an isolated psoralen-thymidine photoproduct.<sup>11a, 17f</sup>

## Conclusions

Two total syntheses of a *cis*-syn furan-side monoadduct between thymidine and a psoralen derivative have been described. Linking a benzofuranyl acid derivative or a 2-

carboxypsoralen to the 5' hydroxyl of thymidine was crucial in the control of the regio- and stereochemistry to give a *cis-syn* adduct. The key [2+2] cycloaddition occurs with high facial selectivity giving primarily or only the *cis-syn* 3' face adduct, which is the most prominent adduct found upon DNA-psoralen irradiation or a medicinal treatment. One possible synthetic organic transformation left to explore is the removal of the 2-carboxy group via a radical decarbonylation to provide the formal psoralen-thymidine adduct. However, the carboxylic acid group does provide a chemoselective handle that allows for the attachment of reporter groups at this site and thereby enhances the utility of the 2-carboxypsoralen. Another possible use of these structurally intriguing molecules is in the area of molecular scaffolding. The *cis-syn* adduct provides a molecular architecture where two sets of functional groups can be displayed in the same direction. This type of orientation is rare with small organic compounds and is usually observed in large macromolecules. With known lability of the nucleosidic bond and with the retainment of the *cis-syn* geometry shown by Hearst and co-workers,<sup>17e</sup> it seems possible to utilize the deoxyribose backbone to synthesize other *cis-syn* molecules that could be used to entertain some molecular recognition and scaffolding experiments.



**Figure 1.15.** (Above) Structure of photoproduct **1.35**, where  $R = \text{CO}_2\text{Me}$ . (Below, left) Nuclear Overhauser enhancements for photoproduct **1.35**. (A) Difference spectrum between sample irradiated at the thymidine methyl resonance (1.76 ppm) and spectrum B. (B) 4.0-5.0 ppm region of the proton NMR spectrum. (C) Circular dichroism spectrum of photoproduct **1.35**.



## Chapter 2

# Synthesis of Oligonucleotides Containing a Site-Specific Psoralen Adduct

### Introduction

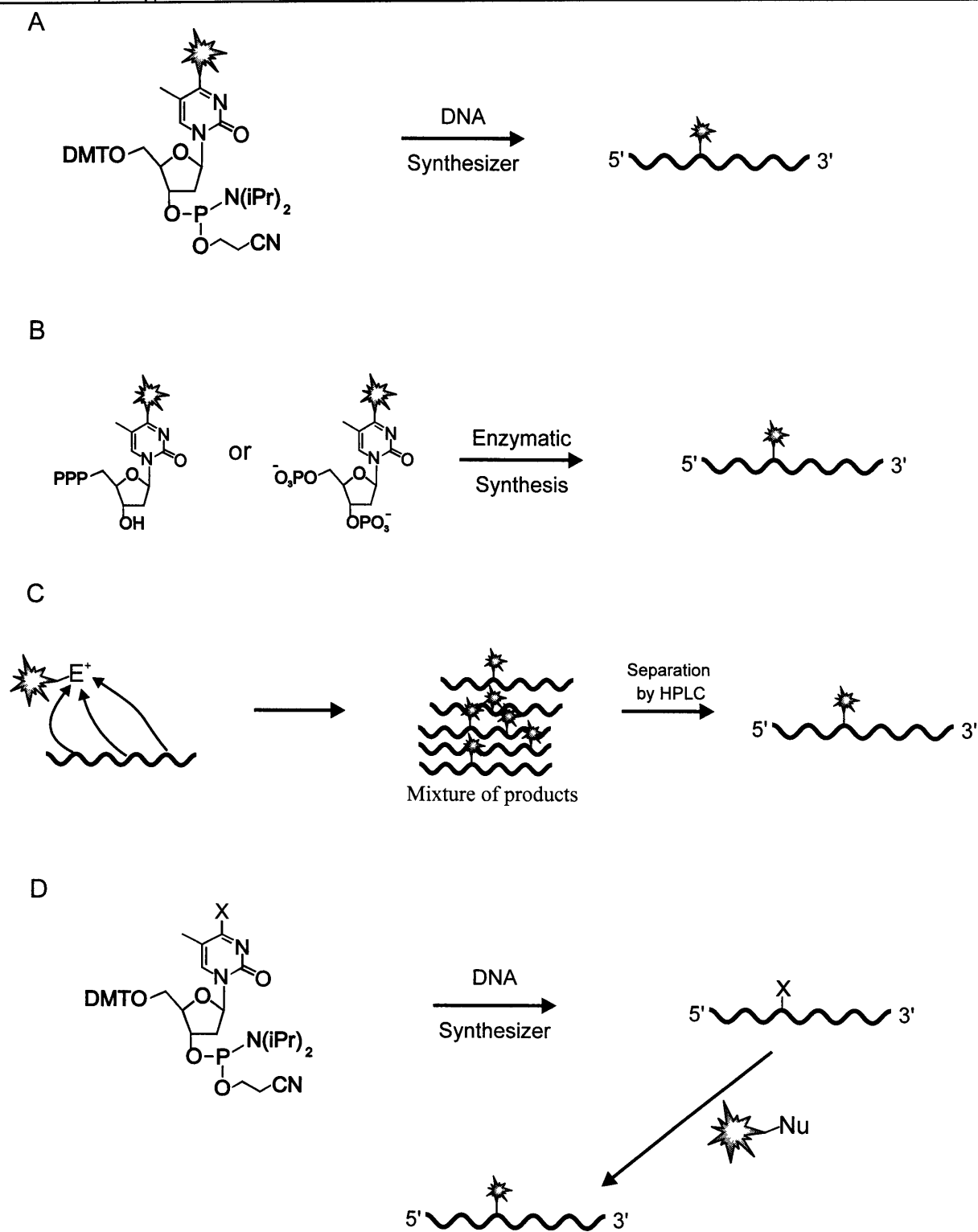
Oligonucleotides containing a modified nucleoside at a predefined site have proved to be valuable tools for understanding the mechanisms of DNA replication, repair and mutagenesis. Over the last 15 years many synthetic methodologies have been developed to construct oligonucleotides containing a site-specifically placed damaged nucleoside.<sup>1</sup> Four major approaches have been taken (Fig. 2.1): (1) The total synthetic approach that involves the synthesis of a damaged nucleoside phosphoramidite for its use in a solid-phase oligonucleotide synthesis. (2) The synthesis of a damaged mononucleotide for its use in an enzymatic synthesis of an oligonucleotide. (3) A direct treatment approach that involves the treatment of a unmodified oligonucleotide with a DNA damaging agent followed by purification of a singly-damaged oligonucleotide. (4) A “convertible” approach that involves the synthesis of oligonucleotide containing a nucleoside that can be converted into the desired DNA damage utilizing a reagent that chemoselectively modifies the site-specifically placed base. Each of these approaches has its advantages and disadvantages depending on the type of DNA damage to be site-specifically incorporated.

A successful total synthetic strategy is inherently the most powerful approach because once the methodology is developed the damaged nucleoside can be placed anywhere in any length oligonucleotide. Moreover, this approach affords enough material for the structural study of these oligonucleotide by NMR or x-ray crystallography. In addition, a synthetic approach can

provide access to the minor DNA adducts of a DNA damaging agent that, although low in yield, can be responsible for a disproportionately large fraction of a DNA damaging agent's biological consequences.<sup>32</sup> One limitation of this approach is that the damaged nucleoside of interest must be able to withstand the rigors of solid-phase synthesis and deprotection. Instability is problematic with modified oligonucleotide synthesis because chemical alteration of a natural nucleoside often increases its lability. Any contamination caused by chemical degradation may complicate or impede purification of the oligonucleotide rendering it useless for further biochemical studies. Additionally, site-specifically damaged oligonucleotide that are used for genetic studies must be of the utmost purity due to the possible enhanced signal a contamination may cause in these systems.

An alternative method to avoid some of the limitations of solid-phase oligonucleotide synthesis is to utilize an enzymatic synthesis incorporating the nucleoside adducts as triphosphates or 3',5' bisphosphates.<sup>33</sup> This approach has the advantage of mild reaction conditions, but suffers from low yields and it is somewhat constrained by DNA sequence context limitations.

The direct treatment approach is convenient and affords site-specific substrates to the synthetically challenged researcher.<sup>7,34</sup> This method has the advantage that it can provide oligonucleotide containing DNA lesions that are highly labile because the damaged oligonucleotide does not experience solid-phase oligonucleotide synthesis.<sup>1b</sup> Until recently, direct treatment of DNA with crosslinking agents was the only method to construct intrastand crosslinks in double-stranded DNA.<sup>35</sup> A clear disadvantage to this approach is that it requires DNA damaging agents that make predominately one type of adduct. For example, an



**Figure 2.1.** The four major strategies for synthesizing site-specifically modified oligonucleotides. (A) The total synthetic approach. (B) The enzymatic approach. (C) The direct treatment method. (D) The convertible nucleoside procedure.

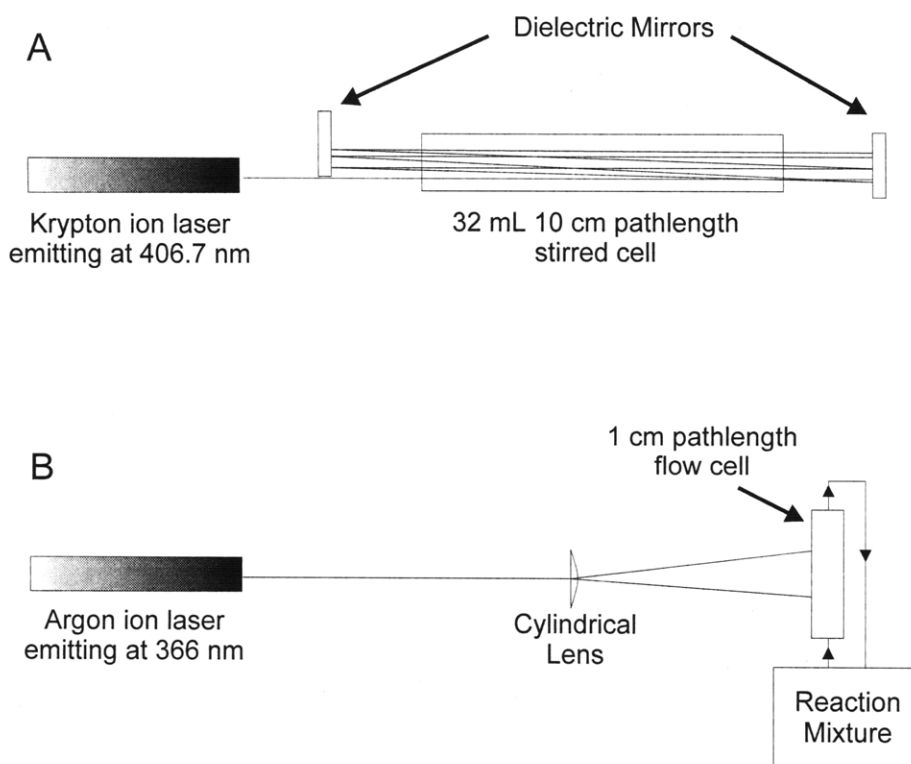
oligonucleotide treated with a simple alkylating agent or an oxidizing reagent would produce a wide range of DNA adducts and purification of an individual adduct would be impossible.

Another major drawback is that oligonucleotide sequence and length are limited. Typically short (8-12 nt) oligonucleotide with only one reactive nucleoside for that particular DNA damaging agent are used in order to minimize HPLC purification problems. These restrictions prevent the study of these lesions in oligonucleotide that have multiple potential sites of adduction. This is unfortunate because many of these sites have been shown to be mutational hotspots that occur in proto-oncogenes and tumor suppressor genes; such hotspot sites are thought to play a key role in neoplastic transformation. Despite these purification hurdles, some adducts have been site-specifically placed in oligonucleotide containing multiple reactive sites using the direct treatment approach.<sup>34</sup>

The fourth, and most recent approach, takes advantage of the chemoselective reaction of halo-purines and triazolo-pyrimidines with amines and alcohols.<sup>36</sup> An oligonucleotide is synthesized with a site-specifically placed “convertible” nucleoside. Depending on the protocol used, the resin-bound or free oligonucleotide can be treated with an amine or alcohol in the presence or absence protecting groups. This approach combines the DNA sequence independence of a total synthetic approach with the experimental ease of the direct treatment approach. This strategy does suffer from some side reactions, which complicates purification and limits oligonucleotide length. Another drawback is that the current methodology has been limited to adducts at the N<sup>2</sup> of guanine, N<sup>6</sup> of adenine, and O<sup>4</sup> of thymine or uracil.

### Psoralen-DNA Adducts

The synthesis of psoralen monoadducts and crosslinks have been thus far constructed by the direct treatment method and are limited to oligonucleotide with only one 5'-TA site.<sup>37</sup> Traditionally, Hanovia type lamps and blackray lights were utilized to synthesize psoralen adducts. This strategy provides reasonable amounts of crosslinked substrates; however, the separation and purification of the furan- and pyrone-side monoadducts has proved nearly impossible using denaturing polyacrylamide gel electrophoresis (PAGE) or by HPLC. Recently, Hearst et al. have synthesized oligonucleotides containing a site-specifically placed furan-side monoadduct and crosslink utilizing high intensity lasers (Fig. 2.2).<sup>38</sup> The photo-chemical

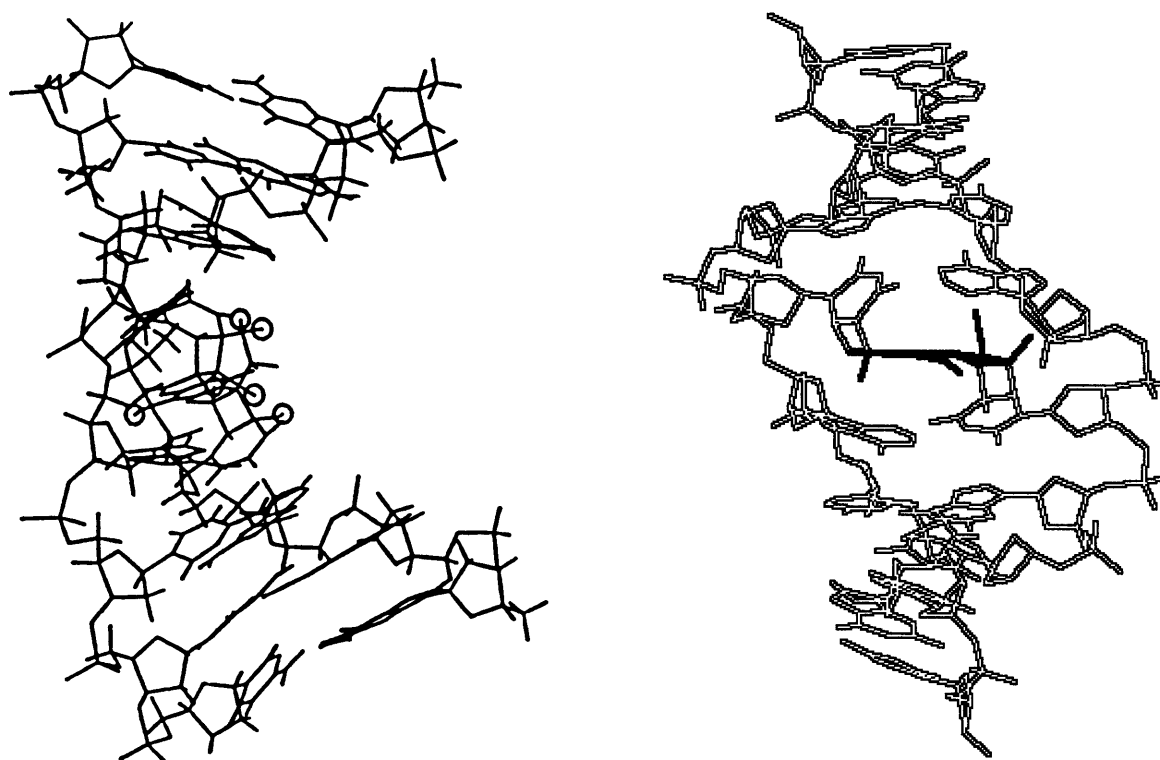


**Figure 2.2.** Schematic diagrams of the optical path during sample irradiation to form furan-side monoadducts (A) and interstrand crosslinks (B).<sup>38</sup>

selectivity of this method arises from the knowledge that psoralen molecules absorb light at wavelengths up to 410 nm whereas the absorption cutoff for furan-side monoadducts is ~395 nm. Therefore, irradiation at wavelengths greater than 395 nm would allow for the isolation of furan-side monoadducts without the conversion to crosslinks. The employment of a high-intensity monochromatic light source is necessary because the absorption coefficient of psoralen at wavelengths greater than 400 nm is only  $\sim 1\text{-}20$  ( $\text{mol}^{-1}, \text{cm}^{-1}$ ). This exotic methodology can provide large quantities of psoralen monoadducts; however, it is limited by sequence context because the target oligonucleotide can only contain one thymine base. It was this lack of oligonucleotide sequence freedom that prompted the total synthesis of a furan-side psoralen-thymidine monoadduct described in Chapter 1. Moreover, a total synthetic procedure allowing for the preparation of psoralen adducts in any sequence context is more desirable because recent studies show that context plays an important role in the replicative bypass, mutagenic, and genotoxic effect of DNA-damaging agents.<sup>6</sup>

### **NMR Solution Structure of a Psoralen Containing Oligonucleotide**

The 3-dimensional structural determinations of a psoralen containing oligonucleotide have come under some recent scrutiny. The first NMR-derived model of a DNA duplex crosslinked with AMT provided the first evidence that the psoralen moiety was indeed intercalated within the DNA helix because of the standard sequential connectivities in an unmodified oligonucleotide were interrupted by the psoralen lesion (Fig. 2.3).<sup>39</sup> Distance geometry methods proposed that the 3D structure of the psoralen cross-linked helix contained a



**Figure 2.3.** (Left) NMR-derived model for a DNA duplex crosslinked with AMT.<sup>39</sup> The methyl groups involved in the photolesion and the amino group of AMT are indicated by circles. (Right) NMR solution structure for the HMT interstrand crosslink.<sup>22</sup> The HMT is filled in with black.

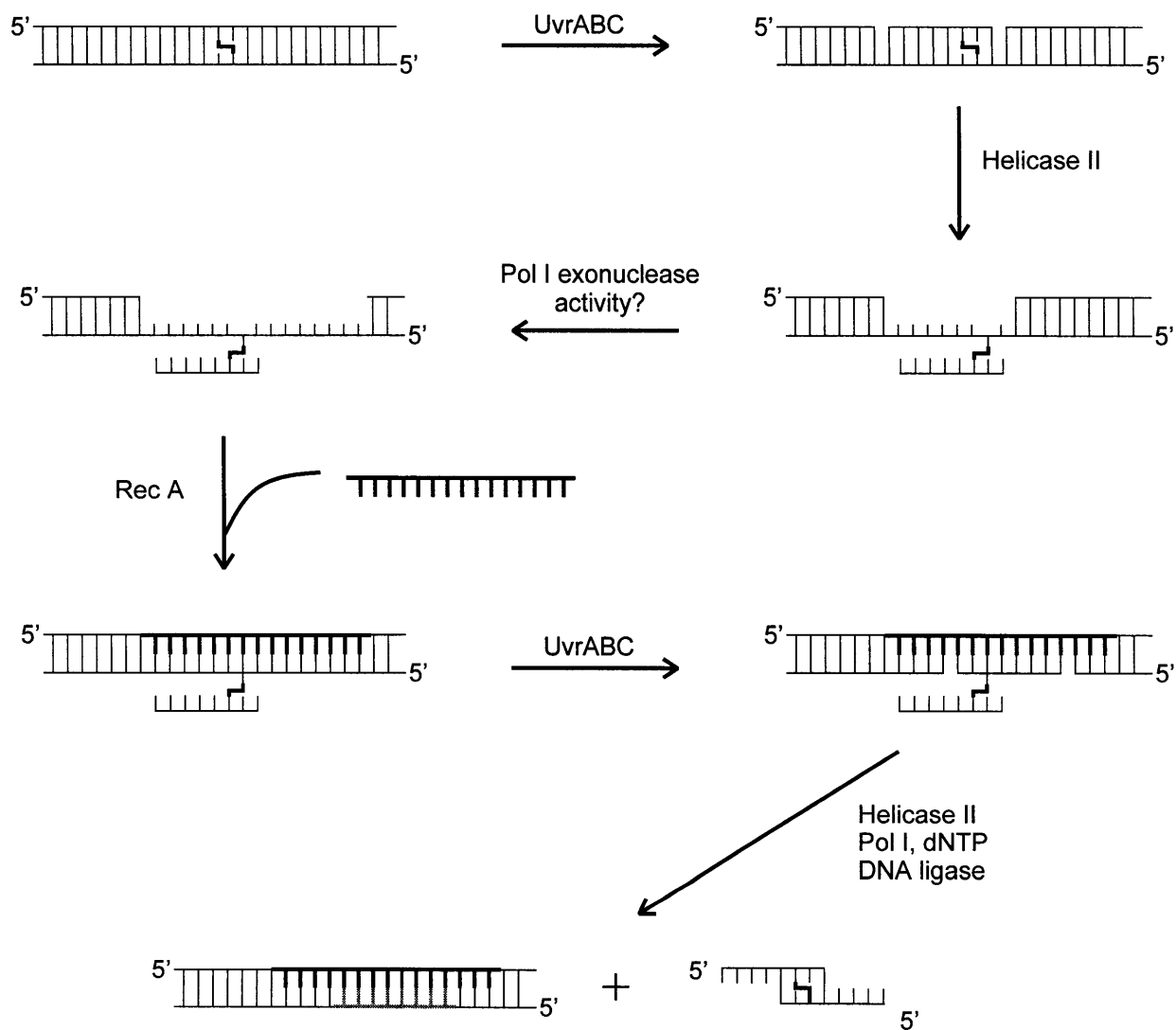
large kink ( $56^\circ$ ) in the phosphate backbone at the site of damage. In addition, the structure shows an asymmetry in the base-pairing stability adjacent to the crosslink, which is also detected by the temperature dependancies of the imino resonances. Recently, the NMR structure of both the monoadduct and the crosslink between HMT and DNA has been reported (Fig. 2.3).<sup>22</sup> Wemmer and co-workers show that this psoralen derivative is intercalated into the helix with the local DNA structure distorted, but returning to B-form DNA within three base pairs. However, no significant bend is seen in the helix with the monoadduct or the crosslink involving HMT. It

seems unlikely subtle difference of a primary amine (AMT) to a primary hydroxyl (HMT) could affect the DNA helix so dramatically. The discrepancy between the two NMR solution structures has been attributed to technical errors in the first structural determination suggesting that the latter structure is more representative of the structure of psoralen-DNA adducts.<sup>22</sup> Regardless of which structure best represents a psoralenated oligonucleotide, the psoralen moiety is intercalated between the bases of the DNA and it significantly alters the overall structure of the helical DNA.

### **Repair of Psoralen Adducts**

Due to the diverse types of chemical damage that can occur to DNA, cells have evolved a number of mechanisms to repair modified nucleosides. It is generally accepted that the structurally distorting psoralen adducts are substrates for the nucleotide excision repair (NER) pathway in both prokaryotes and eukaryotes.<sup>40</sup> In addition, recombination may also play a role in psoralen crosslink repair.<sup>41</sup> The NER pathway repairs bulky lesions by excising a small patch (13-32 nt) of DNA surrounding the lesion. DNA polymerase and ligase are used to complete the repair process. Oligonucleotides containing a site-specific psoralen crosslink have been used to determine NER's role in removal of intrastrand crosslinks. In *Escherichia coli* (*E. coli*), the UvrBC proteins first incise the DNA on the furan-side of the crosslink.<sup>42a</sup> If the lesion was not an interstrand crosslink, the gap would be filled in using DNA polymerase; however, the presence of the pyrone-side adduct in the template strand poses a block to DNA synthesis and a possible mutagenic site. Therefore, an error-free repair mechanism has been proposed from *in vitro* recombination experiments using the *E. coli*. RecA protein (Fig. 2.4).<sup>41</sup> One concern with





**Figure 2.4.** The *E. coli* error-free repair of a psoralen crosslink suggested by *in vitro* experiments with site-specifically modified oligonucleotides.<sup>41</sup>

this mechanism is that the gap required for recombination must be at least 12-13 nucleotides larger than the oligonucleotide gap made by UvrABC. The wider gap has been postulated to arise from the 5'→3' exonuclease activity of DNA polymerase I.<sup>42</sup> While the furan-side of the crosslink is incised first by UvrABC of *E. coli*, the pyrone-side of the crosslink is excised first using HeLa cell extracts.<sup>43</sup> In fact, it has been recently reported that DNA sequence context effects whether the furan or pyrone-side of a crosslink is repaired first by UvrABC.<sup>44</sup> In addition, Wood et al. have discovered the efficiency of human NER repair of psoralen furan-side monoadducts is affected by DNA sequence context as well.<sup>43</sup> The differences between eukaryotic and prokaryotic NER of psoralen adducts is common. *In vitro* studies show that the UvrABC proteins from *E. coli* repair the furan-side monoadduct more efficiently than the crosslink.<sup>18</sup> However, human cell extracts from HeLa cell lines repair the crosslink better than the furan-side monoadduct.<sup>18</sup> Psoralen adducts have also been utilized to study transcription coupled repair in CHO cells.<sup>45</sup> Hanawalt and co-workers have shown that psoralen crosslinks in transcribed strands are removed much more efficiently than in non-transcribed strands. Monoadduct repair, in contrast, is not influenced by adduct location in the genome. The ability to place a psoralen crosslink in double-stranded DNA without regard for sequence context should help clarify how cells process psoralen-DNA adducts.

### **Mutagenicity and Cytotoxicity of Psoralen Adducts**

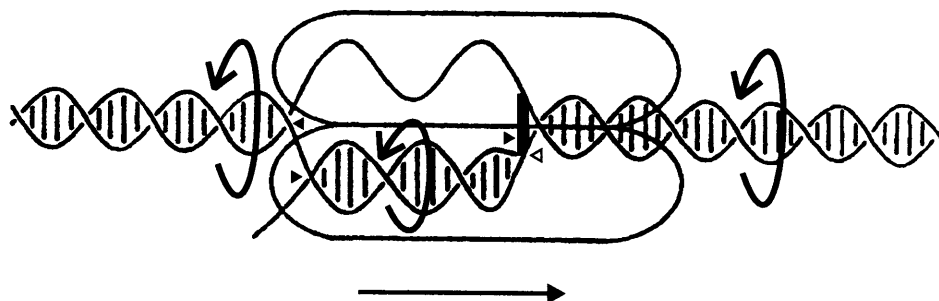
DNA damage that goes unrepaired can give rise to mutations during replication and eventually cause cell death. Psoralen-DNA crosslinks are an absolute block to DNA synthesis *in vitro*.<sup>46</sup> On the other hand, furan-side monoadducts have been shown to be bypassed during

replication on average 11% of the time with more replication errors occurring on the lagging-strand.<sup>47</sup> Crosslinks are considered to be both more mutagenic and lethal; however, the genotoxic differences have been correlated to species rather than the adduct type.<sup>48</sup> The mutational spectrum of HMT in the *lac* promoter region of M13mp10 phage DNA contains predominately T→G transversions and one-base deletions in runs of T and C.<sup>49</sup> A site-specific mutagenic approach utilizing the furan-side monoadduct and the crosslink has also confirmed that the major base mutation is a T→G.<sup>50</sup> In addition, this site-specific approach revealed that all deletions occur at the site of the psoralen adduct. In fact, all of the deletion mutants have been attributed solely to psoralen crosslinks.<sup>51</sup> Using 8-methoxypsoralen in *E. coli* NER deficient cells, all mutations are base substitutions, which were attributed to psoralen monoadducts because, without NER, crosslinks are lethal. In NER proficient cells, one-base deletions are also observed and have been postulated to occur during replication past the three-stranded psoralen lesion with DNA polymerase (Fig. 2.4). With the importance of sequence context has on mutagenesis and replicative bypass of many lesions,<sup>6</sup> a furan-side psoralen-thymidine phosphoramidite will be useful to help determine the effects of sequence context plays on these biological endpoints.

### **Biochemical Applications of Site-Specifically Psoralenated Oligonucleotides**

In spite of the arduous methods used to generate oligonucleotide containing site-specifically placed psoralen adducts, a substantial amount of biochemical work has been done utilizing oligonucleotide containing furan-side monoadducts. Furan-side psoralen-thymidine monoadducts can be used as hybridization probes that can covalently attach to the target

sequence by long-wave UV light. Hearst et al. have utilized psoralen photo-crosslinking oligonucleotide to identify unique sequences in the human genome, to enhance Southern blotting signal and to probe how secondary structure affects hybridization kinetics and equilibria.<sup>52</sup> The sequence context constraints imposed by the current methodologies to synthesize oligonucleotide containing site-specific furan-side monoadducts has minimized the general utility of these cross-linking probes. Site-specifically placed psoralen monoadducts have also been used to probe DNA-protein contacts.<sup>53</sup> Sastry et al. have utilized psoralen-protein photochemistry to crosslink proteins to DNA. Using an oligonucleotide containing a furan-side psoralen monoadduct, single-stranded DNA binding proteins such as T7 RNA polymerase, T4 gp32 and the *E. coli* proteins RecA and UvrB have been cross-linked to an oligonucleotide containing a furan-side monoadduct. This methodology has the added benefit that it uses long-wave UV light that is outside of the absorption maximum of DNA and most proteins. In theory, oligonucleotide containing a furan-side monoadduct could be used to probe any DNA-protein interaction where the DNA double helix is unwound by DNA tracking proteins. RNA-protein interactions could also be probed in this manner; however, the ability to generate experimental quantities of furan-side monoadducts in RNA has not been possible using any of the previously described techniques. Oligonucleotide containing site-specifically placed psoralen adducts have been used to arrest transcription and to generate a model for the ternary elongation complex between *E. coli* RNA polymerase, the DNA template, and the nascent RNA (Fig. 2.5).<sup>54</sup> A furan-side psoralen-thymidine phosphoramidite may increase the appeal of these crosslinkable substrates to probe DNA-protein interactions.



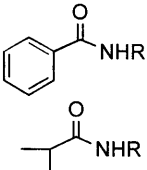
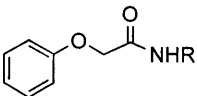
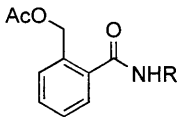
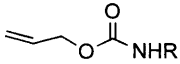
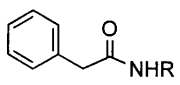
**Figure 2.5.** A model for the *E. coli* RNA polymerase transcription elongation complex.<sup>54</sup> The horizontal arrow indicates the direction of RNA synthesis. The curved arrows indicate the unwinding and rewinding of the DNA and DNA-RNA helix by the polymerase. The filled triangles denote the hypothetical unwindase and rewindase centers of the enzyme. The open triangle denotes the catalytic site of the polymerase, which is very close to the leading unwindase site. The thick, vertical bar indicates the site of the psoralen crosslink that was used to arrest RNA transcription. The footprint of the polymerase on the DNA is indicated by the length of the two lobes of the enzyme relative to the turns of the DNA helix.

## Site-Specific Synthesis of an Oligonucleotide Containing a Furan-Side Psoralen

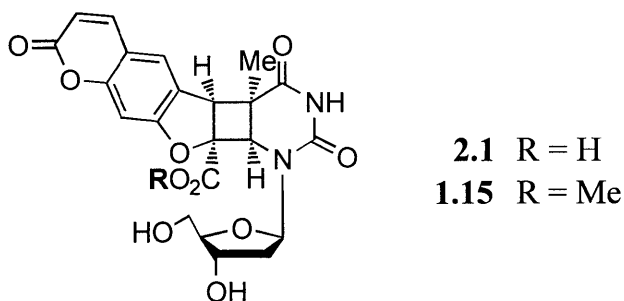
### Monoadduct

Synthesis of the furan-side monoadduct was chosen for two reasons. First, the furan-side adduct contains an intact coumarin chromophore, which possesses the aforementioned photocross-linking capabilities when irradiated with long wave UV light.<sup>12</sup> Second, the furan-side adduct could survive some basic deprotection conditions used in solid-phase DNA synthesis, whereas the unsaturated lactone of the pyrone-side monoadduct is known to transesterify readily in the presence of mild aqueous base.<sup>13</sup> It is also known that furan-side monoadducts can be reversed by treatment with strong base at elevated temperatures;<sup>13</sup> this consideration therefore prompted a stability study analysis on nucleoside **2.1** (Fig. 2.6).

Table 2.1. Comparison of the protecting groups and deprotection conditions for DNA synthesis.

Protecting Group	Structure(s)	Commercially Available	Deprotection Conditions
<b>Standard</b> {isobutyryl and benzoyl}		✓	NH <sub>3</sub> , 18 hr, 55° C.
<b>PAC</b> {phenoxyacetyl}		✓	NH <sub>3</sub> , 2 hr, rt. 0.1 M NaOH, 6 hr, rt. 10% DBU in EtOH, 24 hr, rt.
<b>AMB</b> {2-(acetoxymethyl)benzoyl}		✗	50 mM K <sub>2</sub> CO <sub>3</sub> in MeOH, 12 hr, rt.
<b>Alloc</b> {allyloxy}		✗	Pd(0), pH 5.5, 1 h, 55° C.
<b>Phenylacetyl</b>		✗	Penicillin G acylase, pH 7.0, 2 h, rt.

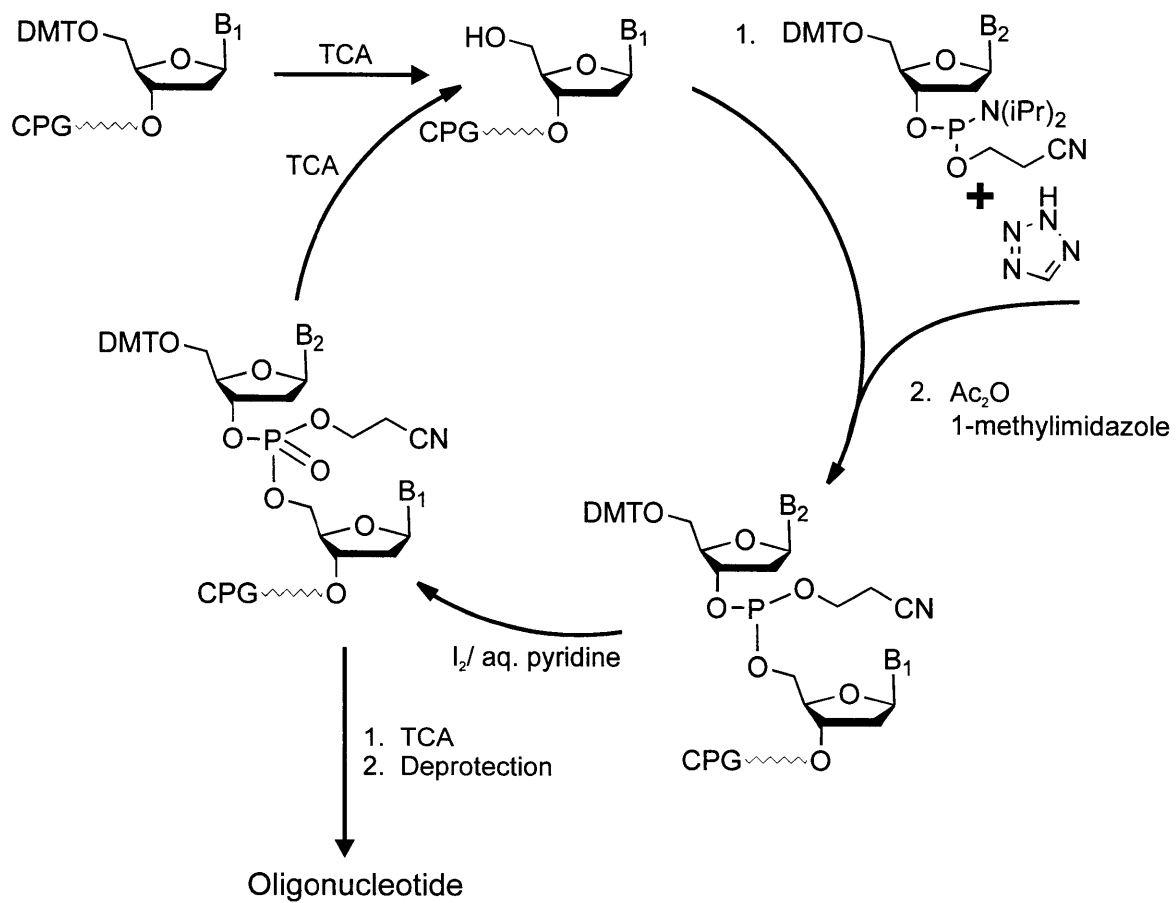
Unless otherwise stated, deprotection reactions are performed in water. rt, room temperature.



**Figure 2.6.** Structure of 2-carboxy- **2.1** and 2-carbomethoxypsoralen **1.15**.

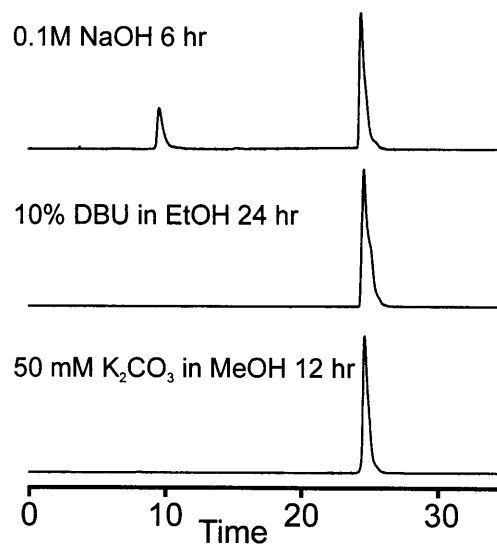
Nucleoside **2.1** was found to be stable for at least 30 min to every step in the DNA synthesis cycle shown in Figure 2.7. The last stability study to be performed was the deprotection step. Table 2.1 contains some of the protecting groups developed for the solid-phase synthesis of oligonucleotide. Oligonucleotide synthesized using standard phosphoramidites require treatment with concentrated aqueous ammonia at elevated temperatures. Ammonolysis of the lactone ring of **2.1** prevented the use of standard phosphoramidites. Therefore, nucleoside **2.1** was treated with deprotection conditions that have been reported to remove PAC or AMB protected oligonucleotide.<sup>19, 36c, d</sup> Sodium hydroxide treatment of **2.1** led to significant degradation; however, nucleoside **2.1** remained intact when treated with either 10% DBU in ethanol for 24 h at room temperature or the conditions required for AMB protecting group removal (Fig. 2.8).

Another issue was the observation that nucleoside **1.15** slowly hydrolyzes to the carboxylic acid derivative **2.1** when stored in neutral aqueous solutions. This hydrolysis did not affect the photo-crosslinking capabilities of the psoralen-thymidine monoadduct, but for purification purposes it was convenient to have a homogenous oligonucleotide. Therefore,



**Figure 2.7.** Scheme for solid-phase DNA synthesis.



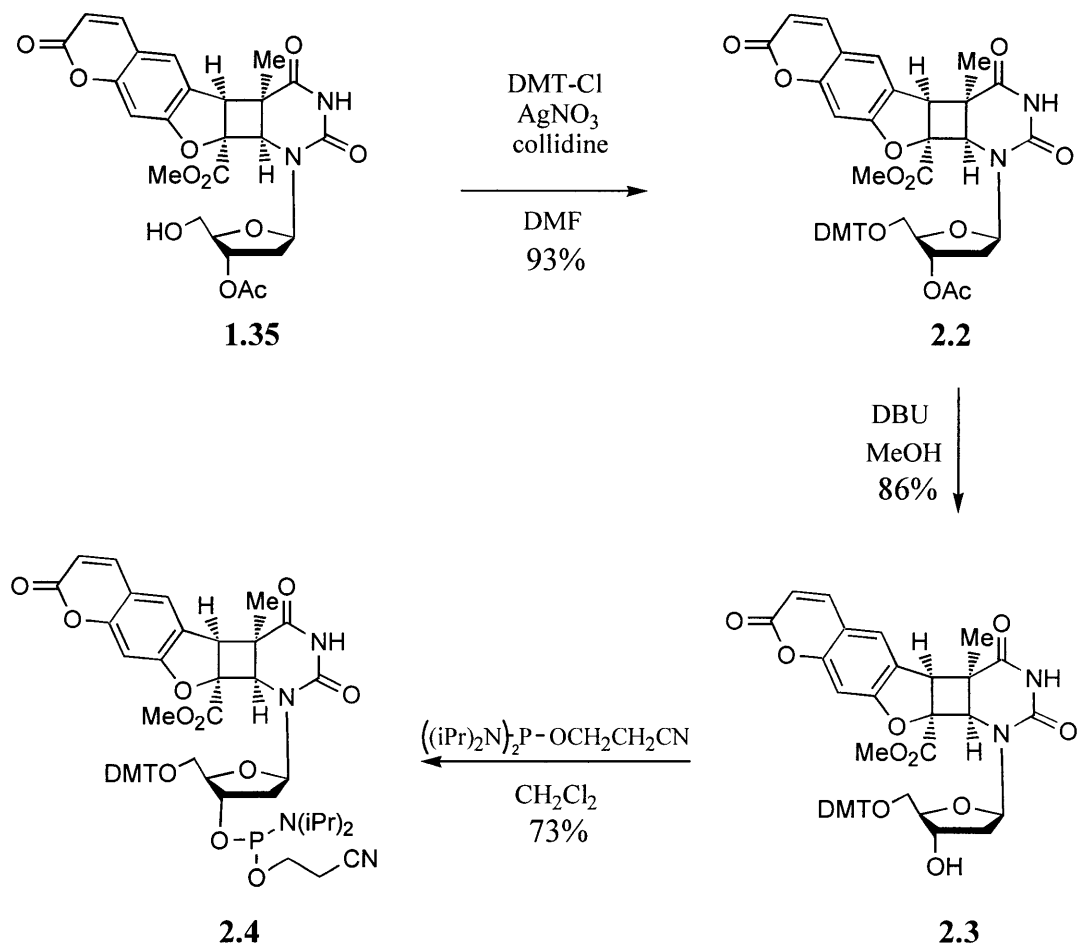


**Figure 2.8.** HPLC traces (330 nm) of the stability studies performed on nucleoside **2.1**. All reactions were performed at room temperature.

reaction conditions were developed to saponify the alkyl ester without destroying the coumarin ring and the photo-crosslinking ability of the monoadduct. Treatment of **2.1** with a 10 mM sodium carbonate solution at pH 9 for 12 h led to greater than 95% conversion to **1.15** alleviating purification problems.

The synthesis of the suitably protected psoralen-thymine phosphoramidite is shown in Figure 2.9. Protection of the 5' hydroxyl of **1.35** with DMT-Cl in the presence of silver nitrate led to rapid conversion to **2.2**.<sup>55</sup> Removal of the 3'-acetate with 5% DBU in freshly distilled methanol afforded **2.3**. Phosphitylation using standard conditions gave phosphoramidite **2.4**.

To demonstrate the flexibility of a site-specific approach, a prototypical human TATA box sequence was chosen to incorporate **2.1**. Runs of TA sequences are known to be mutational hotspots for psoralen<sup>56</sup> and it is possible that the therapeutic efficacy of psoralen might be due in part to the inability of transcriptional complexes to bind to modified TATA boxes. The modified

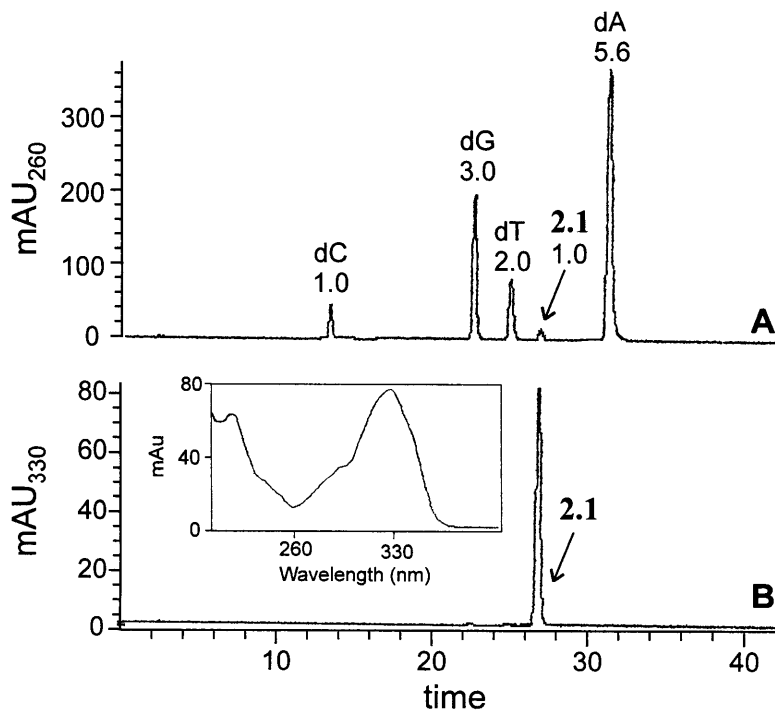


**Figure 2.9.** Synthesis of 2-carbomethoxypsoralen-thymidine phosphoramidite **2.4**

oligodeoxynucleotide 5'-AGCTA**2.1**AAAAGGT-3' **2.5** was synthesized on a 1  $\mu$ mol scale by using an automated DNA synthesizer. The coupling time for phosphoramidite **2.4** was extended to 15 min affording coupling yields of 90% based on DMT cation release. For unmodified phosphoramidites, yields are typically higher; the lower yield is presumably due to the increased bulk of the psoralen-thymidine phosphoramidite. The solid support was treated with 1 mL of 10% DBU in anhydrous ethanol in the presence of cetyltrimethyl ammonium bromide, which aided solubility.<sup>57</sup> After 24 h at room temperature, the deprotection solution was neutralized with

aqueous acetic acid and the salts were removed by using a Na<sup>+</sup> exchange column. The crude oligonucleotide was treated with 100 mM sodium carbonate (pH 9.0) solution for 12 h to saponify the alkyl ester into the carboxylic acid. Neutralization followed by purification by reversed phase HPLC afforded oligonucleotide **2.5**.

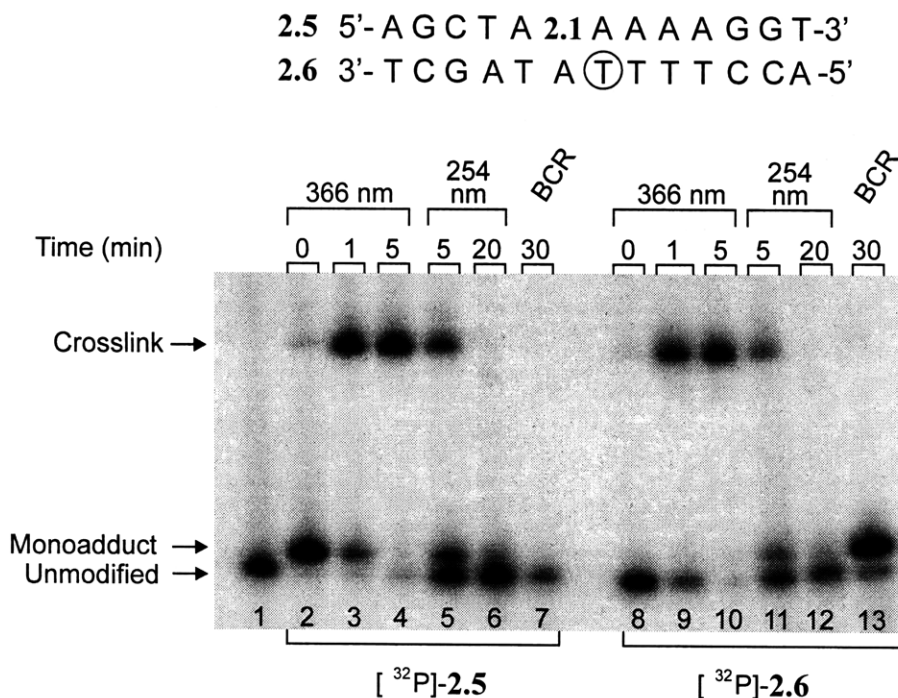
The integrity of oligonucleotide **2.5** was established by enzymatic digestion and electrospray mass spectrometry. Enzymatic digestion and HPLC analysis of **2.5** yielded nucleoside ratios which were within experimental error of the theoretical composition of oligonucleotide **2.5** (Fig. 2.10). The peak corresponding to the modified nucleoside had identical



**Figure 2.10.** HPLC profile of enzymatically digested **2.5**. Chromatograph A shows the corresponding nucleosides observed at 260 nm and the numbers represent the calculated peak ratio areas. Peak identities were determined by nucleoside standards. Chromatograph B shows the absorbance at 330 nm displaying only one peak corresponding to the psoralen-thymidine nucleoside **2.1**. Inset: UV spectrum of peak **2.1**, which is identical to the characteristic UV profile of a synthesized standard **2.1**.

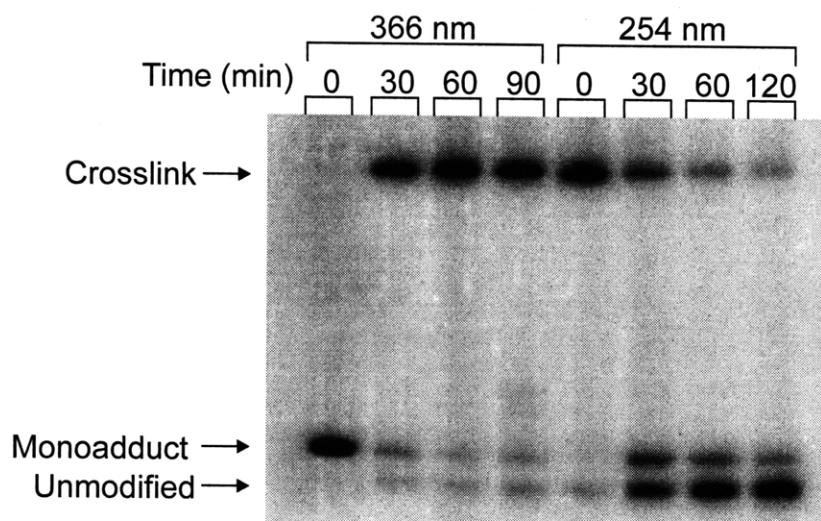
HPLC retention characteristics and UV spectrum as the synthesized standard **2.1**. Electrospray ionization mass spectrometry of **2.5** revealed the presence of ions at  $m/z$  1058.05, 846.5, 705.2 corresponding to 4<sup>-</sup>, 5<sup>-</sup>, 6<sup>-</sup> ions, respectively. The determined molecular weight was 4236.9, which agreed well with the calculated molecular weight of 4236.7.

One of the valuable applications of psoralen containing oligonucleotide is their ability to crosslink to a hybridized strand.<sup>12</sup> Figure 2.11 shows a photo-crosslinking experiments of oligonucleotide **2.5** with its complementary strand **2.6**. By <sup>32</sup>P-phosphate labeling the 5'-hydroxyl of only one of the two DNA stands in each experiment, the reactions of the two DNA stands can be independently analyzed by denaturing gel electrophoresis. Oligonucleotide **2.5** was labeled in lanes 2-7. Oligonucleotide **2.6** was labeled in lanes 8-13. The presence of the psoralen derivative caused oligonucleotide **2.5** to migrate slower than its unmodified counterpart **2.7** (lane 1 vs. 2). Irradiation of the duplex with 366 nm light (9.0 J/m<sup>2</sup>) afforded the slower moving interstrand crosslink in 5 min (lanes 4 and 10). A well known chemical property of psoralen-DNA crosslinks is their reversibility with 254 nm light<sup>12</sup> or their asymmetric conversion to the pyrone-side adduct by heating in the presence of base.<sup>13</sup> Treatment of the crosslinked duplex with 254 nm light (16.7 J/m<sup>2</sup>) for 20 min resulted in photoreversion to the monoadduct (lanes 5 and 11) and eventually complete or near complete reversal to the unmodified strand (lanes 6 and 12). Base catalyzed reversal results in conversion of the interstrand crosslink into a pyrone-side adduct, essentially transferring the original psoralen furan-side monoadduct in **2.5** to the complementary strand **2.6**. Treatment of a crosslinked duplex with 0.1 M NaOH at 90 °C for 30 min efficiently reversed the crosslink affording unmodified oligonucleotide (lane 7) and primarily a pyrone-side monoadduct (lane 13).<sup>13</sup> The crosslinking and photoreversion with a



**Figure 2.11.** Photoreactions of the furan-side monoadduct containing oligonucleotide and the photoreversion and base catalyzed reversal (BCR) of the crosslinked duplex. The photoreactive thymidine in oligonucleotide **2.6** is encircled. Oligonucleotide **2.5** and the complementary strand **2.6** were 5' phosphate labeled with  $^{32}\text{P}$  in lanes 2-7 and lanes 8-13, respectively: (Lane 1) Unmodified control oligonucleotide **2.7** labeled with  $^{32}\text{P}$  at the 5' terminus. (Lanes 2-4 & 8-10) Irradiation with 366 nm light. (Lanes 5-6 & 11-12) Photoreversion of the crosslink with 254 nm light. (Lane 7 & 13) Base catalyzed reversal of crosslink.

strong intensity lamp was so effective that a low-wattage UV lamp was used to effect the same transformations (Fig. 2.12). Irradiation with a hand-held 4 Watt long wave UV lamp afforded the crosslink in 84% yield in 90 min and a hand-held 254 nm lamp caused near complete reversal in 2 h. These experiments demonstrated that the synthesized psoralen oligonucleotide possesses the useful hybridization/crosslinking/reversal properties of psoralen containing oligonucleotide generated by traditional methods.



**Figure 2.12** Photoreactions of the furan-side monoadduct containing oligonucleotide **2.5** and photoreversal of the crosslinked duplex using a hand held UV lamp. Oligonucleotide **2.5** was 5' phosphate labeled with  $^{32}\text{P}$  at the 5' terminus in each lane.

## Conclusions

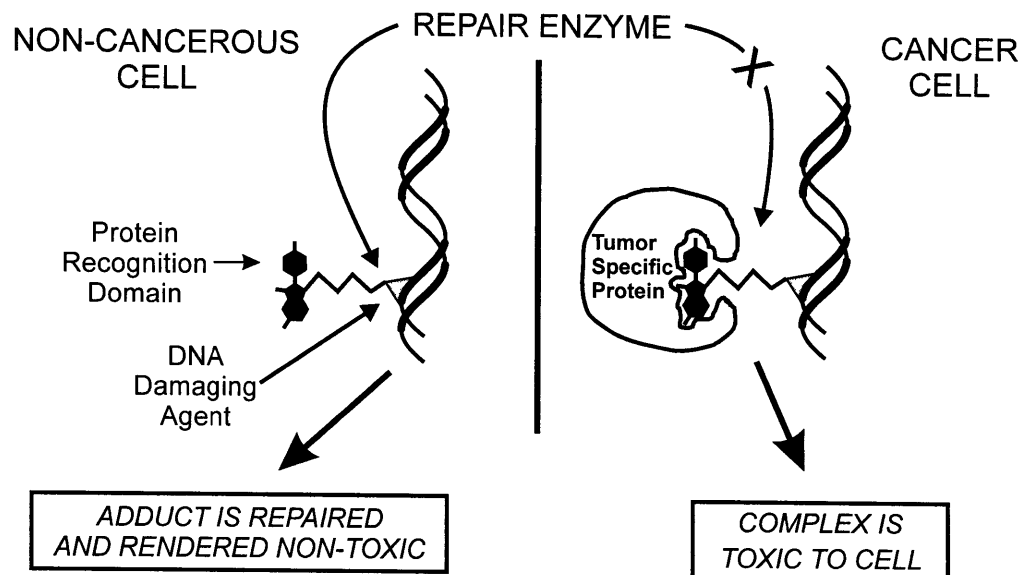
The successful site-specific incorporation of a *cis*-syn furan-side psoralen-thymidine monoadduct into oligonucleotide will enable the study of this therapeutic agent in DNA sequence contexts that contain multiple sites of reactivity. Such sites are known to be biologically important and, with this methodology, it will now be possible to design experiments to probe the details of how sequence context influences biological endpoints. In addition, this synthetic approach makes it feasible to utilize the valuable crosslinking properties of DNA and RNA oligonucleotide containing furan-side monoadducts in hybridization assays and other experiments where site-specific crosslinking is desired. In Chapter 3, the well-defined psoralen-thymidine monoadduct is derivatized with the goal of designing DNA damage that can selectively kill cancer cells overexpressing a tumor specific protein.

## Chapter 3

# Nucleotide Excision Repair of Psoralen-Thymidine Adducts

### Introduction

The development of new cancer therapies is an urgent challenge for medicinal chemists. Several approaches to cancer treatment have been investigated over the last few decades, with the development of DNA-binding drugs at the forefront. Cisplatin (*cis*-DDP) is among the most effective anticancer drugs in clinical use.<sup>58</sup> Cisplatin reacts with DNA to form adducts with specific nucleotide pairs.<sup>59</sup> Work from this lab in collaboration with the Lippard lab has demonstrated that the chemotherapeutically effective DNA adducts formed by *cis*-DDP are recognized by HMG box proteins.<sup>60</sup> These proteins are thought to block the repair enzymes required for adduct removal and hence cell survival.<sup>61</sup> These observations suggested a general mechanism for the design of anticancer drugs that act similarly to *cis*-DDP.<sup>62</sup> Bifunctional compounds that covalently modify DNA and bind tumor expressed or overexpressed proteins could block the repair enzymes that are essential for cell viability (Fig. 3.1). The DNA lesions formed by the bifunctional drug would be toxic in cells that possessed the overexpressed protein, whereas cells without the protein would survive because the DNA damage would be accessible to the repair enzymes and removed. If this repair blocking mechanism proves to be a viable method for targeting specific cell types, its generality could be applied to many forms of cancer. Many breast and ovarian tumors overexpress the estrogen receptor (ER) at high levels.<sup>63</sup> The ER and its known ligand binding domain provide a handle to eradicate these cancer cells selectively.<sup>64</sup> Already the Essigmann group has synthesized bifunctional compounds that contain



**Figure 3.1.** The design of anticancer drugs that block DNA repair selectively in cancer cells. The tumor expressed or overexpressed proteins in the cancer cell block the repair enzymes from removing the DNA lesions whereas the lesions in the non-cancerous cell are readily repaired.

a DNA damaging agent, a nitrogen mustard, attached to several different ER binding ligands.<sup>62</sup> These DNA damaging agents exhibited selective toxicity in ER+ cell lines compared with ER- cell lines; however, it was not possible to determine definitively if the cause of selectivity were due to the ER blocking DNA repair because a cell-based assay was used.<sup>62</sup>

A cell-free assay that allows for the visualization of DNA repair would be better suited to test the repair blocking mechanism directly. In fact, all of the repair blocking experiments performed with cisplatin have been observed *in vitro* using a cell-free nucleotide excision repair (NER) assay developed by Richard Wood and co-workers, or by a derivative of that assay.<sup>40d, 61</sup> In order to test the repair blocking mechanism using a cell-free assay, a lesion that can be repaired by NER would have to be site-specifically incorporated into a DNA plasmid. To date,



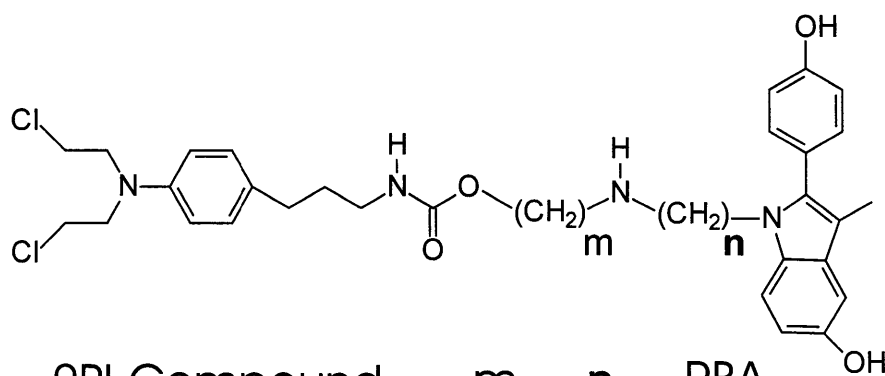
DNA damaging agents such as *cis*-DDP, *trans*-DDP, N-acetoxy-2-acetyl-2-aminofluorene (AAF), thymidine dimers, (6-4) photoproducts, psoralen and benzo[*a*]pyrene have been shown to be repaired by human NER in cell-free systems.<sup>65</sup> Of these adducts, psoralen and cisplatin were the two logical choices for the DNA lesion because both compounds are considered drugs. Psoralen was chosen as the DNA damaging agent because any *in vitro* results could be further examined in a cell-based assay because of the enormous literature precedent demonstrating that most psoralen derivatives still retain their DNA damaging properties.<sup>16</sup> Cisplatin adducts were avoided because they would likely attract HMG-box proteins, which would potentially confuse the results.

It was the lack of a convenient method for the site-specific incorporation of derivatized psoralen adducts into oligonucleotides that spawned the synthesis of the furan-side psoralen-thymidine monoadduct described in Chapters 1 and 2. In fact, the same 2-carboxy group that was crucial for the synthesis of the *cis*-syn furan-side psoralen-thymidine adduct was envisioned to provide a convenient site of attachment for the protein recognition group. This chapter describes the development of a method to attach small molecules chemoselectively to a psoralen-thymidine adduct site-specifically placed in an oligonucleotide. Using this methodology, oligonucleotides containing a site-specific psoralen-thymidine derivative were synthesized that bind specifically to the ER. In addition, the repair blocking hypothesis was tested using these oligonucleotides in the aforementioned *in vitro* NER assay in the presence or absence of recombinant ER.

### Attachment of a Protein Recognition Domain

The ER protein was chosen to test the repair blocking mechanism because it possess a high affinity (< nM) small molecule recognition site that provides the essential protein recognition domain. In addition, purified ER is readily available<sup>66</sup> as are ER+ and ER- human cell lines.<sup>67</sup> Extracts from these cell lines could be used in the *in vitro* NER assay. Although estradiol is the natural ligand for the ER, many non-steroidal derivatives have been synthesized that bind the ER with similar high affinity.<sup>68</sup> 2-Phenylindoles (2PI) are compounds that interact well with the ER and are simple to prepare.<sup>68</sup> Our lab has made a series of bifunctional compounds consisting of a 2PI ligand covalently attached to a nitrogen mustard.<sup>62</sup> The length of the linker between the nitrogen mustard and 2PI group was varied to assess the affinity of the compounds for the ER (Table 3.1). The bifunctional compounds that competed best with [<sup>3</sup>H]estradiol for the ER possessed at least a five (Mustard-C<sub>5</sub>), preferably six carbon (Mustard-C<sub>6</sub>) linker attached to the N2 position of the 2PI molecule. Because linker length greatly affected the relative binding affinity between the 2PI nitrogen mustard compounds and the ER, a modular synthesis was developed in which a protein recognition domain could be readily attached to an oligonucleotide containing a site-specifically placed psoralen-thymidine monoadduct. The chemoselective reaction between ketones and nucleophilic weak bases such as aminoxy, hydrazide or 1,2 aminothiols provides a means of attaching the 2PI derivative to a psoralen adduct within an oligonucleotide.<sup>69</sup> Such carbonyl addition reactions have been recently revitalized for the synthesis of peptides and other biopolymers because they can be performed chemoselectively in water in the absence of protecting groups.<sup>70</sup> This convergent strategy would

Table 3.1. Relative binding affinities of nitrogen mustard 2-phenylindole conjugates.

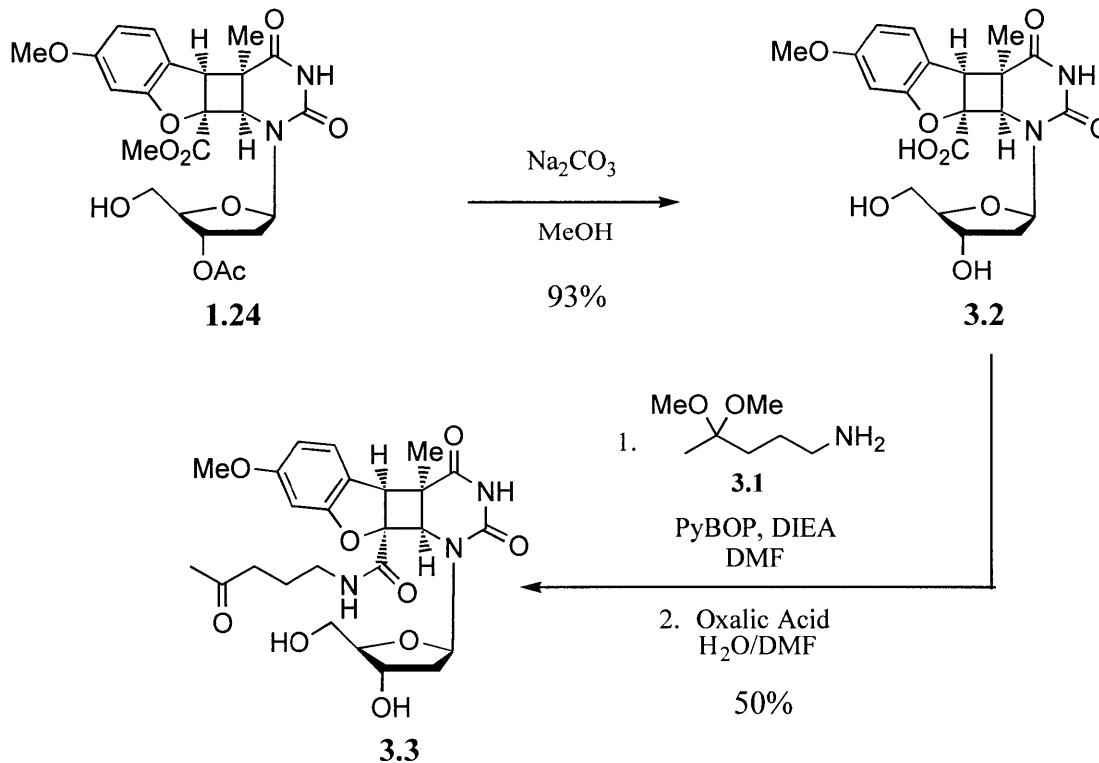


2PI-Compound	m	n	RBA
Mustard-C <sub>3</sub>	3	3	0
Mustard-C <sub>5</sub>	3	5	0.6
Mustard-C <sub>6</sub>	2	6	7.1

Relative binding affinity (RBA) for estradiol was set at 100.  
Adapted from Rink et al.<sup>63</sup>

allow for the synthesis of only one psoralen-thymidine phosphoramidite, yet have the flexibility to couple different protein recognition domains with various linkers.

The first step in the development of this method would be the attachment of a ketone onto the psoralen-thymidine monoadduct. The benzofuran-thymidine monoadduct was used as a model because it is easier to prepare while being structurally similar to the psoralen-thymidine monoadduct. Facile synthesis involved amide bond formation with an aminoketone and the 2-carboxybenzofuran-thymidine photoproduct, though protection of the ketone as a ketal would be necessary in order to avoid Schiff base formation. The synthesis of aminodimethyl ketal **3.1** has been previously described by Drueckhammer and co-workers.<sup>71</sup> The benzofuran-thymidine monoadduct **1.24** was saponified with carbonate to afford acid **3.2** (Fig. 3.2). Treatment of **3.2**

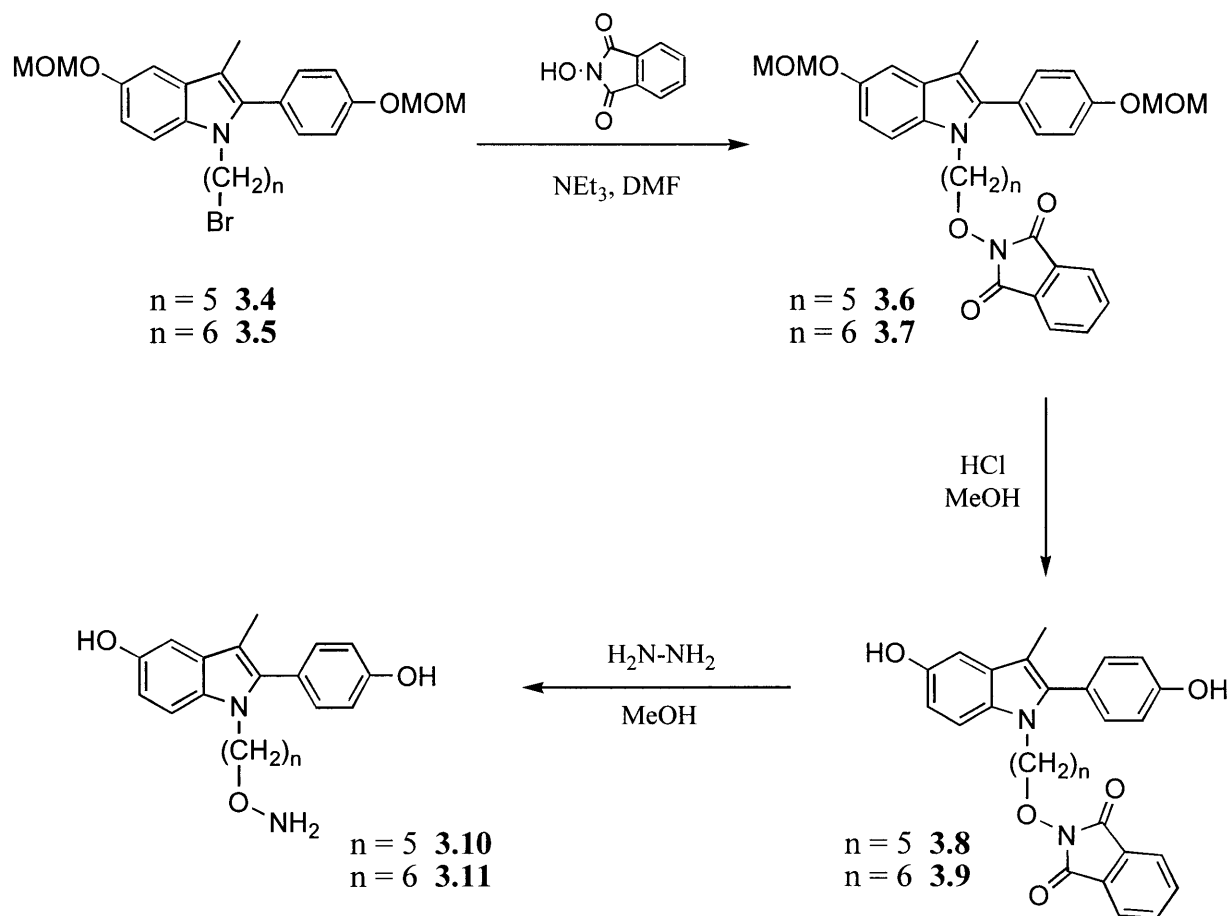


**Figure 3.2.** Synthesis of ketone benzofuran-thymidine monoadduct **3.3**.

and amino dimethylketal **3.1** with PyBOP and diisopropylethylamine selectively formed the amide bond. Immediate removal of the dimethyl ketal protecting group with aqueous oxalic acid afforded ketone **3.3**.

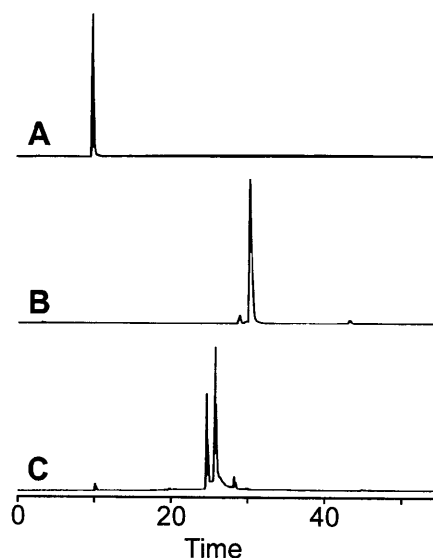
With the ketone “handle” in place, the next step was to determine which chemoselective nucleophile would be suitable for this strategy. Commercially available biotin hydrazide and benzyloxyamine were tested to see if they would react with ketone **3.3** in mildly acidic aqueous solutions. Biotin hydrazide failed to react with ketone **3.3** under several pH conditions. On the other hand, benzyloxyamine reacted quantitatively with ketone **3.3** yielding the oxime within 3 h in 0.1 M sodium phosphate buffer (pH 5.6). Two peaks, presumably the *cis* and *trans* isomers of the resulting oxime, were observed in the chromatogram.

These results suggested that oxime formation could be used to attach the 2PI moiety to the psoralen-thymidine ketone. Previous work in the lab had demonstrated that attachment of at least a five carbon linker to the N2 position of 2PI is required in order to retain ER binding (Table 3.1). The bromo-*N*-alkyl 2-phenylindoles shown in Figure 3.3 have been previously synthesized in the lab.<sup>62</sup> Displacement of the bromine with *N*-hydroxyphthalimide affords the protected aminoxy derivatives. Deprotection of the phenol protecting groups was accomplished with concentrated HCl in methanol. Removal of the phthalimide protecting group with hydrazine afforded the aminoxy 2PIs, 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11**. Figure 3.4C shows the reaction between 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and nucleoside ketone **3.3**. The reaction was nearly complete after 21 h at room temperature. The newly formed peaks (*cis* and *trans*) were isolated and their mass corresponded to oxime formation.



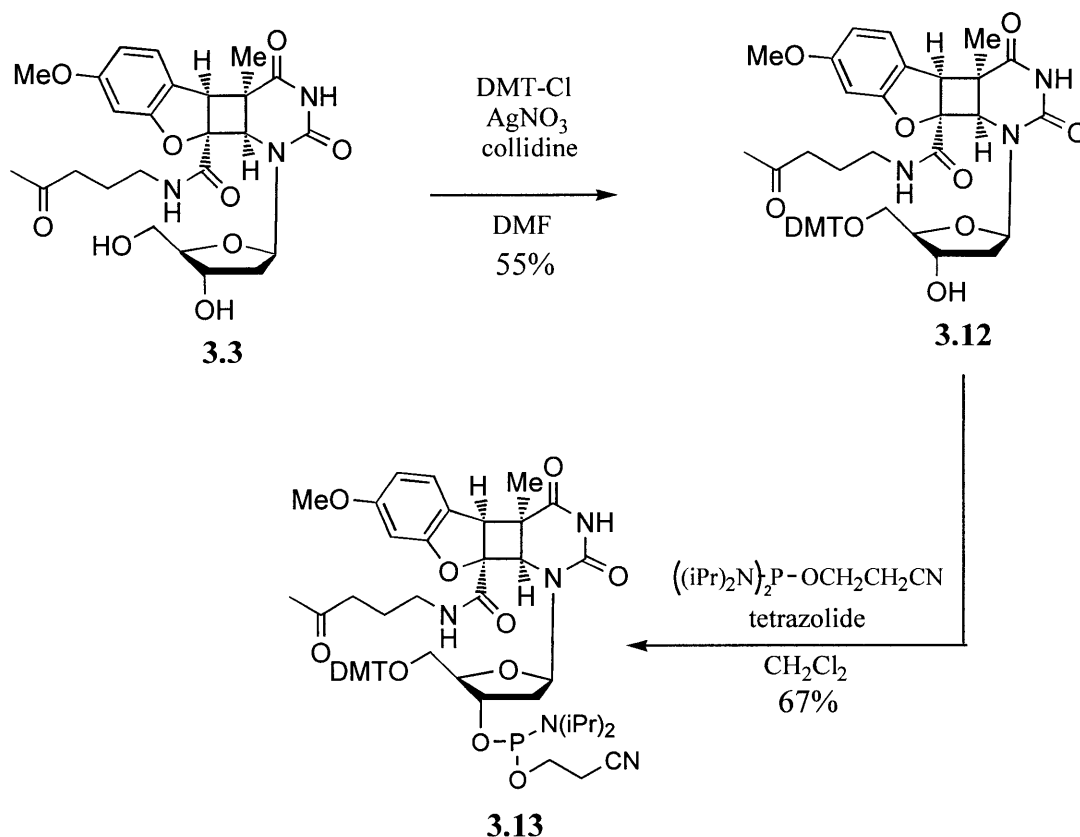
**Figure 3.3.** Synthesis of 2-phenylindole aminoxy derivatives **3.10** and **3.11**.

In order to determine if the chemoselective reaction would occur when the nucleoside ketone was within an oligonucleotide, the synthesis of a suitably protected phosphoramidite of **3.3** was needed (Fig. 3.5). The attachment of the ketone onto the benzofuran-thymidine adduct had the added benefit of not requiring protection during DNA synthesis. Silver nitrate activated DMT protection provided 5' protected **3.12**.<sup>55</sup> Phosphitylation under standard conditions afforded **3.13**. Phosphoramidite **3.13** was incorporated into a 13 base oligonucleotide



**Figure 3.4.** HPLC traces (260 nm) of (A) benzofuran-thymidine ketone **3.3**, (B) 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and (C) the reaction between **3.3** and **3.10** at room temperature (pH 5.6) for 21 h.

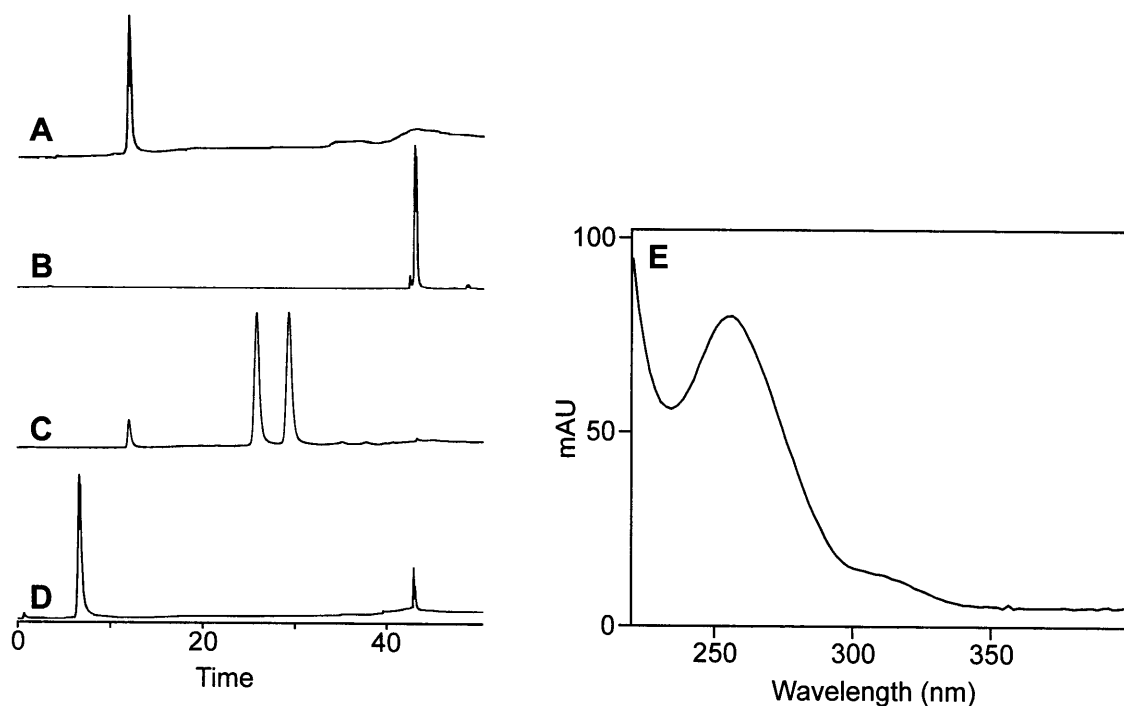
5'-AGCTATAAAAGGT using PAC-protected phosphoramidites. To conserve precious phosphoramidite, a 0.05 M solution of **3.13** was used in conjunction with a 15 min extended coupling time for the modified base. Although these non-ideal conditions afforded a coupling yield of only 40%, they provided an ample amount of material for these model studies. Deprotection for 24 h with 10% DBU in anhydrous ethanol, followed by neutralization and removal of the salts by sodium ion exchange gave the oligonucleotide **3.14**, which was further purified by reversed phase HPLC. Treatment of oligonucleotide **3.14** with 2 equivalents of either 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** or 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** in 0.2 M sodium phosphate (pH 5.6) for 24 h afforded two products **3.15** or **3.16** that were separable by reverse phase HPLC (Fig. 3.6C). All attempts to identify the nucleoside adduct by enzymatic digestion methods resulted in incomplete digestion presumably due to the presence of the added 2PI group. The modified oligonucleotides **3.15** and **3.16** contained an absorbance at 300 nm, which is outside of the absorbance of



**Figure 3.5.** Synthesis of ketone benzofuran-thymidine phosphoramidite **3.13**.

unmodified DNA and is the absorbance maxima for the 2PI (Fig. 3.6E). In addition, electrospray ionization mass spectrometry of either isomer of **3.15** gave a mass of 4607.0, which is consistent with the addition of the 2PI group. Likewise, the molecular weight of either isomer of oligonucleotide **3.16** was 4621.0 corresponding to an addition of a six carbon linked 2PI moiety. One plausible side reaction involves nucleophilic attack of hydroxylamine or methoxyamine at the N<sup>4</sup> position of cytosine. Although such reactions have been observed, they require harsh conditions of 1 M solutions of methoxyamine at low pH and elevated temperatures.<sup>72</sup> Moreover, the yield of modification under these conditions was less than 1% for 24 hr. To test the stability

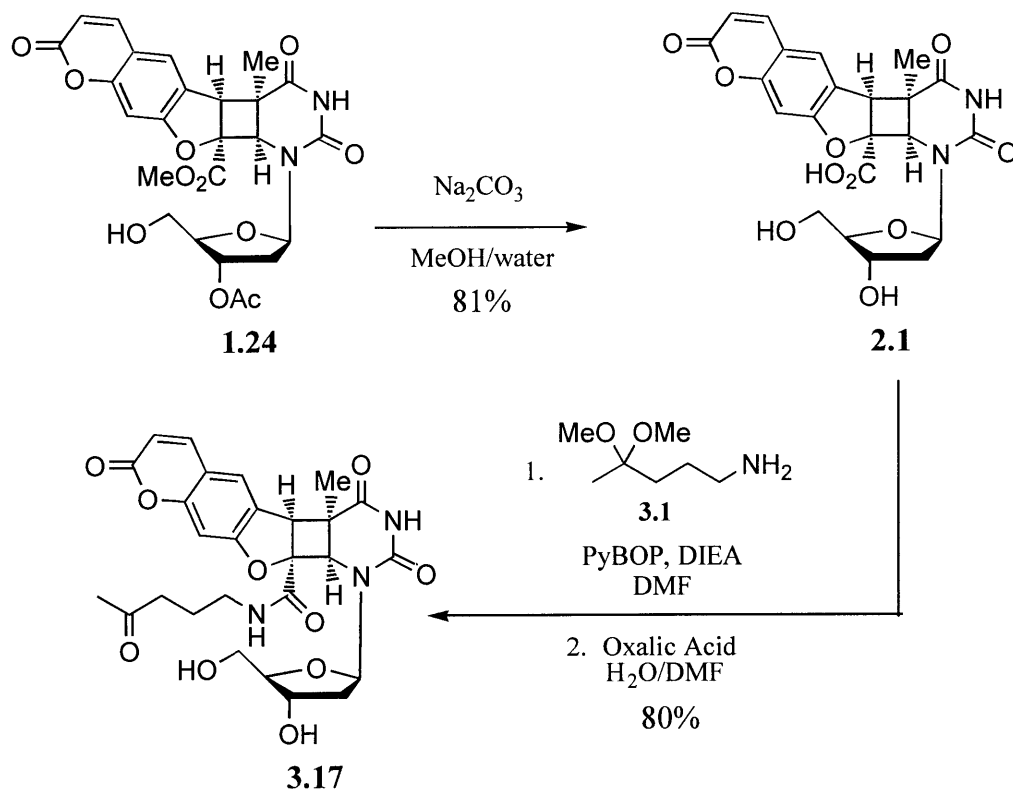




**Figure 3.6.** HPLC analysis of the reaction between aminoxy 2PI compounds and oligonucleotide **3.14**. Data for the reaction with 2PI-C<sub>5</sub>-ONH<sub>2</sub> is shown. HPLC traces (260 nm) of (A) oligonucleotide **3.14**, (B) 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and (C) the reaction between **3.14** and **3.10** at room temperature (pH 5.6) for 24 h. (D) The HPLC trace showing that no reaction occurred after 90 h between 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and oligonucleotide **2.5**, which is missing the ketone “handle”. (E) UV spectrum of oligonucleotide **3.15**. The shoulder at 300 nm is due to the 2PI moiety.

of cytosine under the more milder conditions, aminoxy **3.10** and oligonucleotide **2.5**, which is identical to oligonucleotide **3.14** except it lacks the ketone ‘handle’, were allowed to react under the same conditions. Even after 90 h at room temperature, no reaction products were observed (Fig. 3.6D). These results suggest that oxime formation is selective between oligonucleotide **3.14** and the aminoxy 2PIs.

Having developed a methodology to attach a 2PI group at the site of the photolesion via oxime chemistry in the model system, the next step was to synthesize psoralen-thymidine monoadduct with a ketone at the 2-carboxy position for its incorporation into an oligonucleotide.

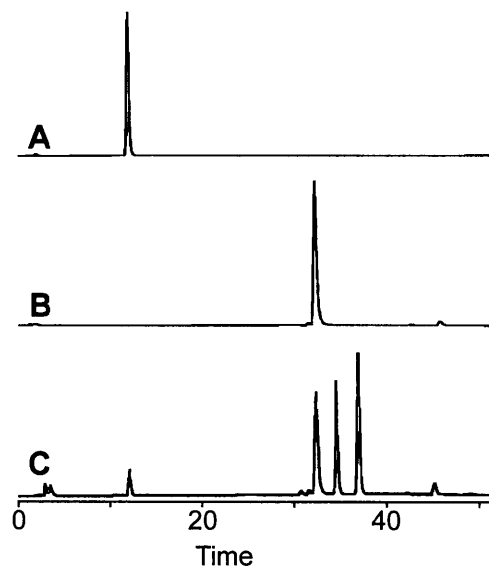


**Figure 3.7.** Synthesis of ketone psoralen-thymidine monoadduct **3.17**.

Synthesis of the psoralen derivative proceeded as designed in the model system (Fig. 3.7).

Saponification of psoralen photoproduct **1.24** with 50 mM sodium carbonate (pH 9.0) in methanol/water yielded carboxylate **2.1** without lactone ring rupture. PyBOP coupling of acid **2.1** with amine **3.1** followed by hydrolysis of the ketal with aqueous oxalic acid gave psoralen ketone **3.17**. Oxime formation with **3.17** and 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** occurred under the same conditions as the model system (Fig. 3.8). The two oxime isomers were collected and the addition of the 2PI group was confirmed by proton NMR.

The next step in the synthesis was to incorporate the psoralen-thymidine nucleoside into an oligonucleotide. Psoralen-ketone **3.17** was converted into a suitably protected

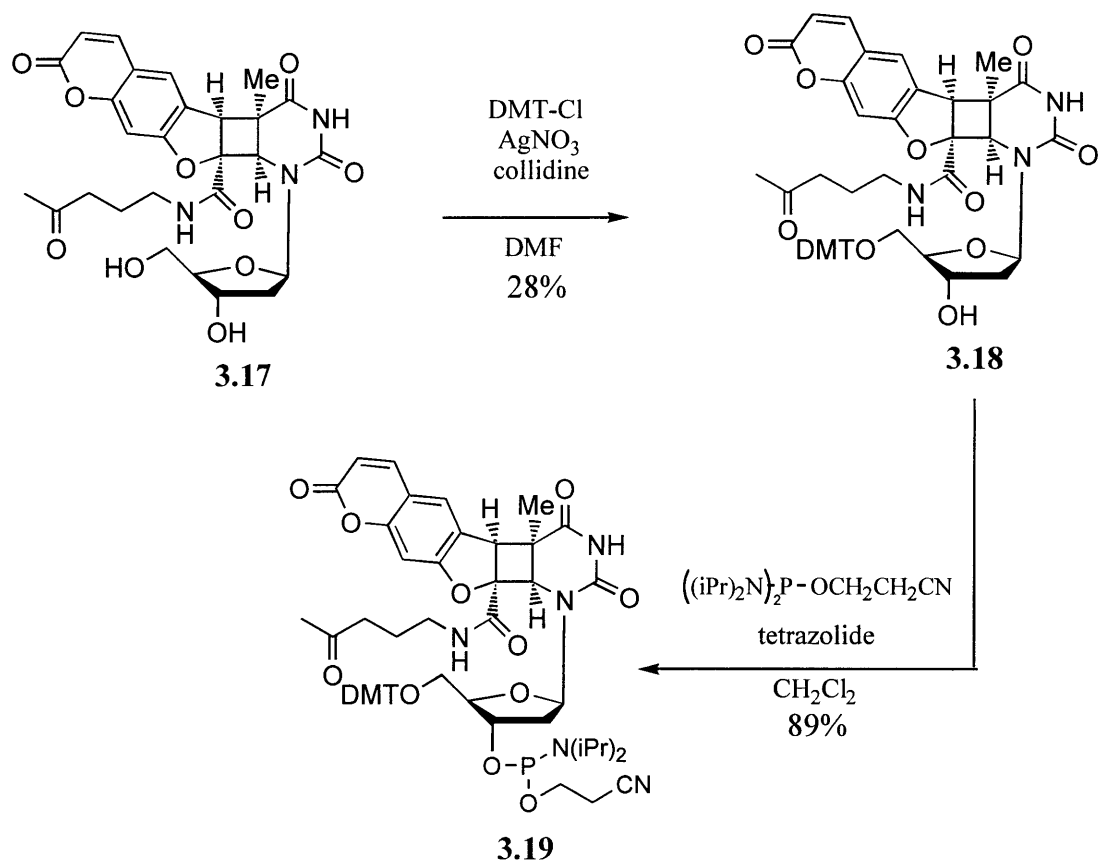


**Figure 3.8.** HPLC traces (260 nm) of (A) psoralen-thymidine ketone **3.17**, (B) 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** and (C) the reaction between **3.17** and **3.11** at room temperature (pH 5.6) for 20 h.

phosphoramidite as shown in Figure 3.9. Protection of the 5' hydroxyl with 4,4'-dimethoxytrityl chloride (DMT-Cl) in the presence of silver nitrate led to rapid conversion to **3.18**.<sup>55</sup>

Phosphitylation using standard procedures afforded phosphoramidite **3.19**.

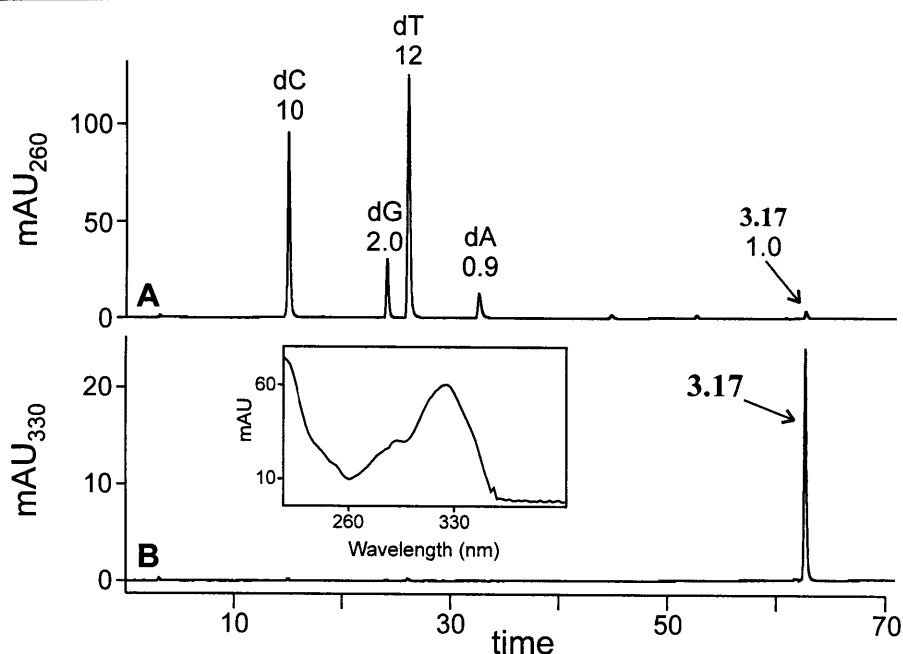
The oligonucleotide sequence 5'-CC TCT TCT TCT GXG CAC TCT TCT TCT **3.20** was chosen for incorporation of phosphoramidite **3.19**. This sequence is complementary to the plasmids used in the Dr. Richard Wood's laboratory with whom we were in collaboration. The Wood lab has shown that this particular oligonucleotide sequence provides the strongest nucleotide excision repair signal for the 2-carboxypsoralen adduct.<sup>43</sup> Oligonucleotide **3.20** was synthesized on a 1  $\mu$ mol scale using PAC-protected phosphoramidites. The modified phosphoramidite **3.19** was hand coupled for 15 min in order to maintain a 0.1 M phosphoramidite concentration. The coupling efficiency for **3.19** was 85%. PAC-group deprotection, neutralization and salt removal afforded crude oligonucleotide **3.20**. The length of



**Figure 3.9.** Synthesis of ketone psoralen-thymidine phosphoramidite **3.19**.

the oligonucleotide necessitated purification by strong anion exchange HPLC using a sodium chloride gradient. Enzymatic digestion established the integrity of oligonucleotide **3.20**. Figure 3.10 shows the HPLC trace of digested **3.20** that yielded nucleoside ratios that were within experimental error of the theoretical composition. The peak corresponding to the modified nucleoside had identical HPLC retention characteristics and UV spectrum as the synthesized standard **3.17**.

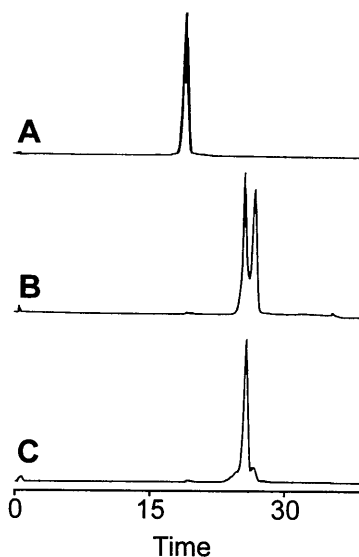
The next step was attachment of the 2PI derivative to the oligonucleotide containing the psoralen-ketone oligonucleotide. Oligonucleotide **3.20** and two equivalents of 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** were allowed to react in 100 mM phosphate buffer (pH 5.6) overnight (Fig. 3.11). Attempts



**Figure 3.10.** HPLC profile of enzymatically digested **3.20**. Chromatograph A shows the corresponding nucleosides observed at 260 nm and the numbers represent the calculated peak ratio areas. Peak identities were determined by nucleoside standards. Chromatograph B shows the absorbance at 330 nm displaying only one peak corresponding to the psoralen-thymidine nucleoside **3.17**. Inset: UV spectrum of peak **3.17**, which is identical to the characteristic UV profile of a synthesized standard.

to separate the two resulting oximes under many different HPLC conditions proved futile. However, it was discovered that the oligonucleotide oximes were in equilibrium with the starting materials and thus the isolated material would be contaminated with starting material. The oxime instability was surprising because it had not been observed in the benzofuran-thymidine model system. Reduction of the crude oxime with  $\text{NaCNBH}_3$  under acidic conditions afforded a stable secondary alkoxyamine **3.21**.<sup>73</sup> Reduction of the oxime also eliminated the geometrical restraints imposed by the carbon-nitrogen double bond and thus resulted in the isolation of only one peak (Fig. 3.11C). Presumably the reduction was not stereoselective and the isolated material was a mixture of inseparable diastereomers. The UV and MALDI mass spectrum of the

isolated material are consistent with the addition of the 2PI group. Most importantly, oligonucleotide **3.21** was stable in aqueous solution and could be used to test the repair blocking mechanism.

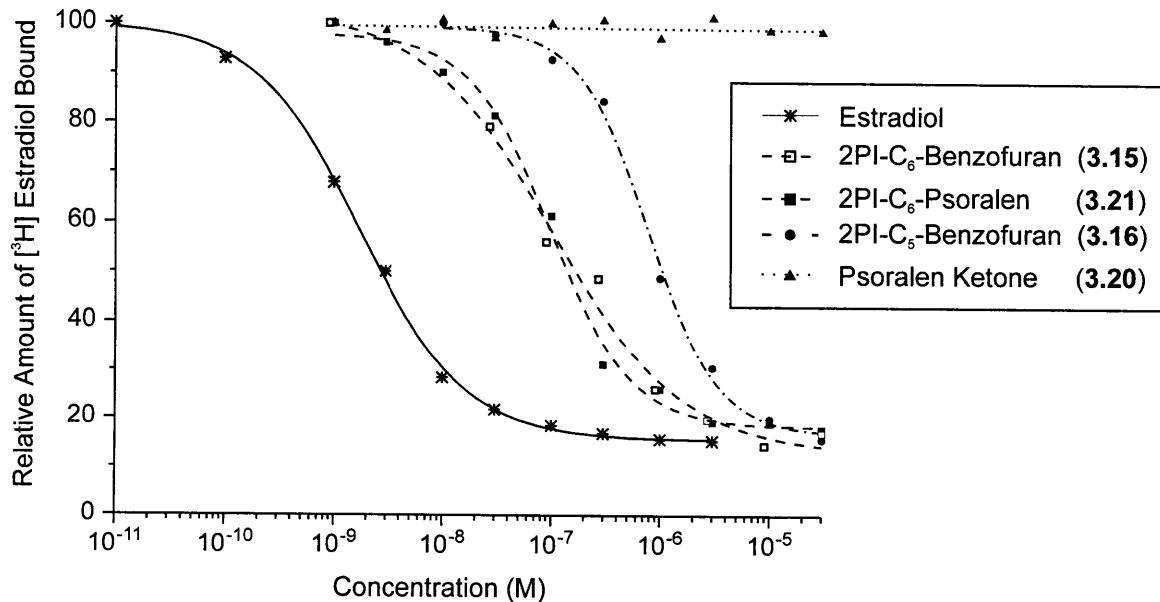


**Figure 3.11.** HPLC traces (260 nm) of (A) oligonucleotide **3.20** and (B) the reaction between **3.20** and 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** at room temperature (pH 5.6) for 24 h. (C) HPLC trace of the reduction of the oxime reaction with NaBH<sub>4</sub> at room temperature for 15 min to afford oligonucleotide **3.21**.

### ER Binding of 2PI Containing Oligonucleotides

Previous work in the Essigmann lab had demonstrated that 2PI nitrogen mustard compounds conjugated to oligonucleotides were capable of specific binding to the ER.<sup>62</sup> In order for the psoralen-2PI oligonucleotides to be blocked from repair, they would have to exhibit similar specific binding to the ER. Consequently, oligonucleotides containing the 2PI-psoralen lesion were assayed for ER binding affinity. The relative affinities of all 2PI containing substrates for the calf uterine ER were measured by using a competitive binding assay with 17-β-

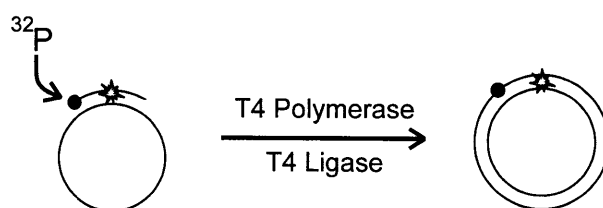
[<sup>3</sup>H]estradiol.<sup>74</sup> Figure 3.12 shows that all 2PI containing oligonucleotides, including the model substrates, competed with estradiol for binding to the ER. The relative binding affinity (RBA) of estradiol to remove 50% receptor-bound radioactivity was set at 100. Approximately a 10-fold improvement in binding affinity was observed with the substrates that possessed a six carbon linker (RBA = 1.4 - 1.8) versus a five (RBA = 0.3). DNA containing only the psoralen ketone damage did not compete with estradiol for the ER demonstrating that the interaction was specific for the 2PI moiety. These data established that the oligonucleotides containing a psoralen-2PI lesion can bind specifically to the ER.



**Figure 3.12.** Oligonucleotides containing the 2PI-photolesions compete with estradiol for the ER. The modified or control oligonucleotides were added to a calf uterine extract along with [<sup>3</sup>H]estradiol and the ability of each oligonucleotide to compete with estradiol was determined.

### Nucleotide Excision Repair Experiments

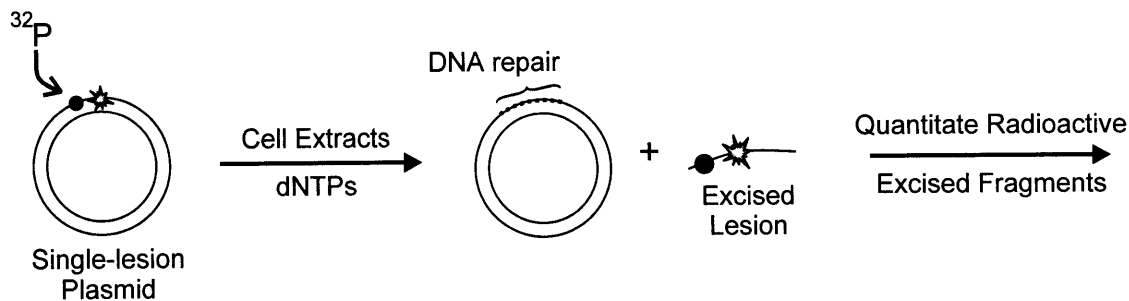
The cell-free nucleotide excision repair assay involves constructing a damaged DNA plasmid and treating it with whole cell extracts. NER competent extracts can excise a small patch (28-32 nucleotides) of DNA around the site of the lesion. The resulting gap is filled with dNTPs by DNA polymerase and ligated in place with DNA ligase.<sup>65</sup> There are currently two methods to detect NER using this assay. DNA repair synthesis can be visualized and quantitated by using <sup>32</sup>P labeled dNTPs. This approach requires the constructed covalently closed circular plasmids to be extremely pure because any nick in the plasmid will generate a repair signal and increase the background noise dramatically. Consequently, visualization of the excision step of NER using an internally radiolabeled single-lesion plasmid is the method of choice.



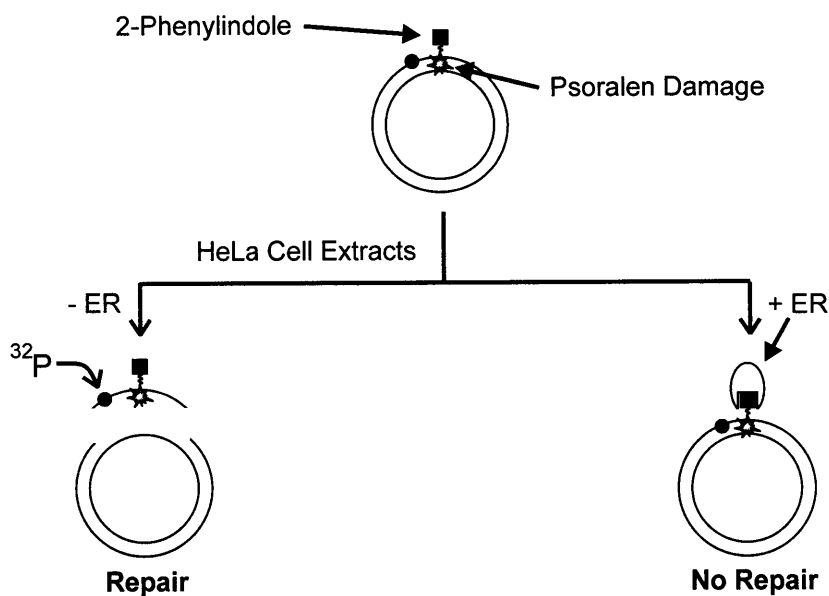
**Figure 3.13.** Construction of a plasmid utilizing a <sup>32</sup>P labeled oligonucleotide containing a site-specifically placed DNA lesion. Priming single-stranded circular DNA with a radiolabeled oligonucleotide followed by DNA synthesis and ligation with T4 DNA polymerase and ligase affords a single-lesion internally <sup>32</sup>P labeled plasmid.

Construction of the labeled plasmid is accomplished by 5'-phosphate labeling a site-specifically modified oligonucleotide followed by DNA synthesis and ligation (Fig. 3.13). Treatment of this substrate with cell extracts affords radioactive excised fragments that can be quantitated (Fig. 3.14). In order to perform the repair blocking experiments using this *in vitro* NER assay, an internally radiolabeled plasmid containing a site-specifically placed psoralen-2PI lesion would





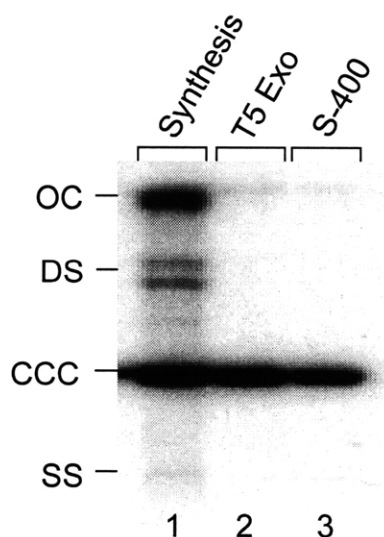
**Figure 3.14.** The *in vitro* nucleotide excision repair (NER) assay utilizing a internally  $^{32}\text{P}$  labeled single-lesion plasmid as the DNA substrate. The internal label allows for the observation and quantitation of the incision step of NER.



**Figure 3.15.** Testing of the repair blocking hypothesis using an *in vitro* NER assay.

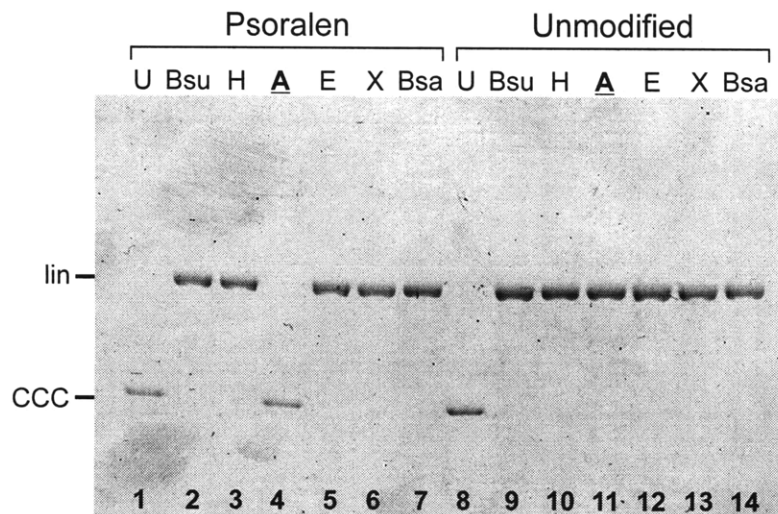
have to be constructed. Treatment of these single-lesion substrates with human whole cell extracts in the presence or absence of the targeted ER protein should determine if the psoralen-2PI lesion can be shielded from repair (Fig. 3.15).

Oligonucleotides containing three different psoralen derivatives were  $^{32}\text{P}$  phosphate labeled: psoralen ketone **3.20**, 2PI-psoralen **3.21** and 2-carboxypsoralen **3.22**. Covalently closed circular duplex DNA was produced by priming single stranded circular DNA with a radiolabeled oligonucleotide containing a psoralen lesion followed by DNA synthesis and ligation with T4 DNA polymerase and ligase (Fig. 3.13).<sup>43, 75</sup> As seen in Figure 3.16, this procedure afforded covalently closed circular (CCC) DNA; however, open circular (OC) species and linear DNA were present as well (Fig. 3.16, lane 1). Previous methods for removal of these side products were accomplished by using CsCl/EtBr density gradient centrifugation and visualization of the CCC DNA with long wave UV light.<sup>75</sup> Plasmids containing furan-side psoralen-thymidine



**Figure 3.16.** Construction of  $^{32}\text{P}$  internally labeled plasmids containing a single psoralen lesion. (Lane 1) Synthesis of the single-lesion substrate. (Lane 2) Digestion of open circular (OC), double-stranded (DS) and single-stranded (SS) DNA with T5 exonuclease. (Lane 3) The covalently closedcircular (CCC) plasmid after two size exclusion spin columns.

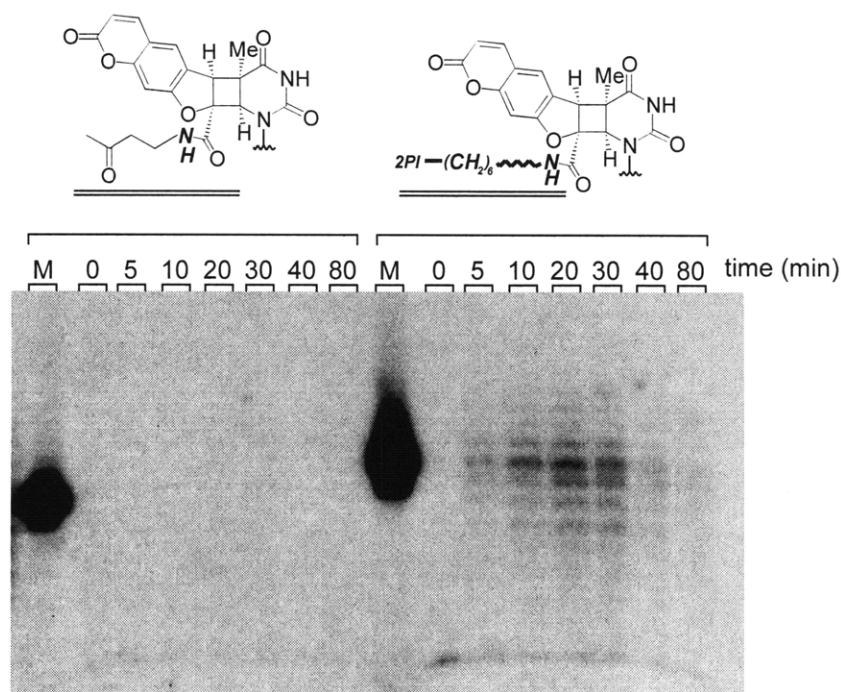
monoadducts cannot be purified in this manner because the UV light would convert them into the crosslink. Fortunately, Isa Kuraoka, in Dr. Wood's laboratory, developed a facile procedure for the removal of the side products of DNA synthesis without the use of arduous CsCl gradients and long-wave UV light.<sup>43</sup> Addition of T5 exonuclease for 3 h selectively degrades all DNA species except for CCC DNA into small oligonucleotide fragments (Fig. 3.16, lane 2). Removal of the oligonucleotide fragments using a size exclusion spin column afforded apparently homogeneous internally <sup>32</sup>P labeled single lesion plasmids, as determined by autoradiography. All of the plasmids constructed contained the psoralen adduct within the unique recognition sequence of the restriction enzyme *Apa*LI (Fig. 3.17). Closed circular plasmids containing the psoralen lesion were resistant to cleavage by *Apa*LI (Fig. 3.17, lane 4), while unmodified plasmid was completely linearized (Fig. 3.17, lane 11). Restriction enzymes with unique sites in DNA



**Figure 3.17.** An agarose gel (0.8%) demonstrating the presence of the 2-carboxypsoralen-thymidine monoadduct in a closed circular duplex DNA plasmid. *Lane 1*, uncut psoralen containing plasmid. *Lanes 2-7*, digestion of psoralen containing plasmid with *Bsu*36I (Bsu), *Hind*III (H), *Apa*LI (A), *Eco*RI (E), *Xho*I (X), and *Bsa*HI (Bsa). *Lane 8*, uncut unmodified plasmid. *Lanes 9-14*, digestion of unmodified plasmid with the same enzymes. The mobilities of covalently closed circular (CCC) and linear (lin) DNA are indicated.

sequences flanking the psoralen adduct linearized both psoralen and unmodified substrates (Fig. 3.17). These data confirm the presence of a lesion at the anticipated site of the site-specific psoralen monoadduct.

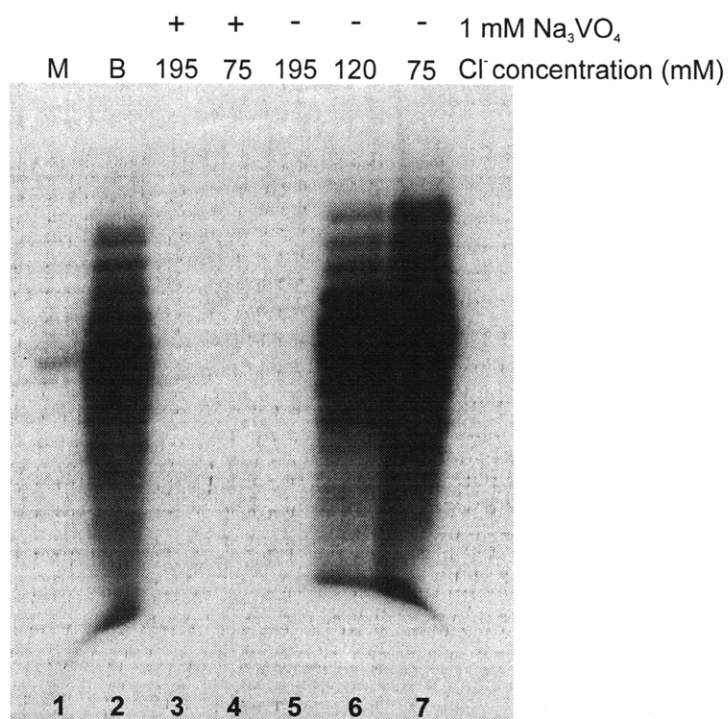
The newly synthesized psoralen ketone and 2PI-psoralen lesions were assayed for NER. A time course was performed to obtain the optimal repair signal for the new psoralen adducts. The plasmids were treated with HeLa whole cell extracts<sup>76</sup> for varying amounts of time and the excised oligonucleotides were separated by electrophoresis in a denaturing polyacrylamide gel (Fig. 3.18). The NER complex removes 28-32 nucleotide patches giving rise to a ladder of products. The smaller fragments in Figure 3.18 are presumably due to the exonuclease activity in



**Figure 3.18.** Kinetics of removal of psoralen adducts using HeLa cell extracts. An autoradiograph of an excision gel. The psoralen ketone and 2PI-psoralen substrates were incubated with extracts at 30 °C for the times indicated. No repair signal was observed for the psoralen ketone adduct. The best repair signal for the 2PI-psoralen adduct was between 20 and 30 min.

the whole cell extract.<sup>61,65</sup> The 2PI-psoralen adduct gave the best signal between 20 and 30 min after starting the repair reaction and the amount of excision ranged from 30-40%. In stark contrast, the psoralen ketone substrate afforded no repair signal (Fig. 3.18, left lanes). All attempts to improve or enhance a possible repair signal for the psoralen ketone lesion failed. The lack of a repair signal of the psoralen-ketone adduct was not detrimental for testing the repair blocking mechanism because the 2-carboxypsoralen derivative could serve as a negative control in these experiments. Nonetheless, the absence of a repair signal for the psoralen ketone adduct was quite intriguing and possible experiments to probe this result will be entertained in the Discussion section.

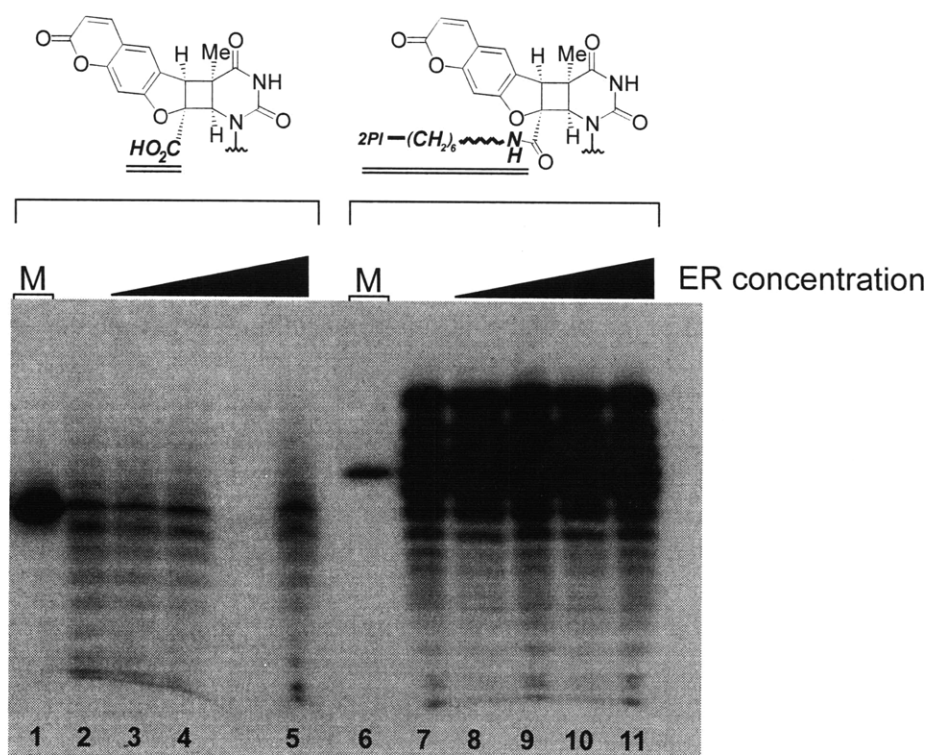
The NER assay was next used to test the repair blocking hypothesis (Fig. 3.15). One advantage of the *in vitro* repair assay is that varying amounts of shielding protein can be added to the reactions as long as the storage buffer of the purified protein is compatible with the repair reactions. The Wood lab has observed that high chloride ion concentrations and certain protein phosphatase inhibitors (PP2A-type) completely inhibit the repair reactions.<sup>77</sup> Unfortunately, purified, recombinant ER is typically stored in 500 mM KCl to minimize aggregation and 1 mM sodium orthovanadate to prevent ER inactivation by dephosphorylation during purification.<sup>66</sup> Due to the obvious incompatibility between the ER storage buffer and the *in vitro* repair reactions, a set of repair reactions was performed to determine a buffer that would not inhibit the repair reactions. As expected, the general phosphatase inhibitor sodium orthovanadate (1 mM) completely inhibited repair (Fig. 3.19, compare lanes 4 and 7) whereas decreasing the overall chloride ion concentration to 75 mM enhanced the repair signal remarkably (Fig. 3.19, lanes 5-7). The highest chloride ion concentration tested in this set of experiments represented the total



**Figure 3.19.** The effect of chloride ion and sodium orthovanadate concentration on the excision repair of the 2PI-psoralen lesion using HeLa cell extracts. *Lane 1*, marker. *Lane 2*, repair reaction using the standard repair buffer. *Lanes 3-4*, repair reaction in the presence of sodium orthovanadate. *Lane 3* represents the addition of 3  $\mu$ L of ER storage buffer in lieu of the standard buffer used in lane 2. *Lanes 5-7*, repair reaction in the presence of varying concentrations of chloride in the absence of sodium orthovanadate.

chloride concentration if a 3  $\mu$ L aliquot of ER storage buffer were to be added to a repair reaction. Based on these results, the ER protein was dialyzed into a sodium orthovanadate free buffer that contained 100 mM chloride ion, which, unfortunately, might compromise ER solubility. Dialysis was accomplished by using two successive size exclusion (G25) spin columns that were pre-equilibrated with the optimized buffer. A rapid salt exchange was preferred over a slow membrane dialysis because of the known aggregation problems with ER in low salt buffers. Successful dialysis was monitored by the recovery of NER for the psoralen lesions.

Now that the ER was in a buffer that could be added to, but not inhibit, the repair reactions, the repair blocking experiments could be performed. The blocking experiments were performed as close to the previously described conditions for the HMG box-cisplatin shielding experiments as possible.<sup>61</sup> The ER was dialyzed just prior to use and immediately added to the reaction mixture. Heat inactivated dialyzed ER was used to maintain the dialyzed buffer concentration in order to prevent any loss of signal due to incomplete buffer exchange. Plasmids containing a single psoralen adduct with or without the 2PI group were incubated at 30 °C for 30 minutes with increasing concentrations of ER. HeLa whole cell extracts were added and the repair reactions were stopped after 30 minutes. Figure 3.20 shows that no blocking had occurred with either substrate up to what was believed to be an ER concentration of 600 nM. The lack of



**Figure 3.20.** Effect of the estrogen receptor (ER) on the excision repair of the 2-carboxypsoralen and 2PI-psoralen lesions using HeLa cell extracts. Denaturing 14% polyacrylamide gel showing no inhibition of either adduct by addition of ER.

repair blocking prompted the determination of the activity of the commercial source of ER. The amount of active ER was determined by quantitation of a  $^3\text{H}$ -estradiol receptor complexes using a hydroxyl apatite batch assay.<sup>78</sup> The activity of the commercial source of ER was found to be only 1/6th of the reported 3  $\mu\text{M}$  concentration. This lower activity would reduce the final concentration of active ER to 100 nM in these experiments. In addition, quick spin dialysis may have also partially inactivated the ER additionally lowering the active concentration. Although repair blocking was not seen in this experiment, the ER activity complications may have been the primary cause for its lack of success. Clearly, these experiments are inconclusive at best; however, the materials and methods have been developed to determine whether or not DNA damage can be designed to be selectively repaired *in vitro*. Now that the general procedures have been established, future experiments with an ER preparation of better integrity will be used to evaluate the repair blocking hypothesis.

## Discussion

The combination of the cisplatin-HMG box shielding experiments<sup>61</sup> and the recent design of the selectively toxic DNA damaging agents in the Essigmann lab<sup>62</sup> has suggested a general strategy to design DNA damage that can be shielded from repair in the presence of a targeted protein (Fig. 3.1). In an attempt to observe repair blocking directly in a cell-free assay, site-specifically damaged oligonucleotides capable of binding the ER were synthesized utilizing a chemoselective oxime reaction. A very preliminary experiment suggested that repair blocking does not occur at protein concentrations up to  $\sim 100$  nM (Fig. 3.20), but the questionable activity



of the commercial available ER protein and the limited experimental data warrants further investigation of this system before any conclusions can be made.

In order to screen protein recognition domains rapidly and optimize linker length a general strategy was developed to attach small molecules bearing an aminoxy group to ketone photolesions site-specifically placed within an oligonucleotide. Oxime formation was used to attach the 2PI group to the nucleoside chemoselectively. The high barrier of rotation around the oxime bond facilitated the separation of the syn and anti isomers. It was surprising that biotin hydrazide and ketone nucleoside **3.3** failed to react under similar conditions; however, hydrazone formation with ketones typically requires at least a 100-fold excess of hydrazide to afford a 50% conversion.<sup>79</sup> Reaction of 2 equivalents of an aminoxy 2PI with the benzofuran-thymidine containing oligonucleotides afforded oligonucleotides with the 2PI group at the site of photoproduct damage. The syn and anti oximes of these oligonucleotides could also be separated by HPLC. The chemoselectivity of the reaction was confirmed by the lack of reaction between 2PI-C<sub>5</sub>-ONH<sub>2</sub> **3.10** and oligonucleotide **2.5** (Fig. 3.6D), which is missing the ketone “handle.” Reaction of the 2PI-C<sub>6</sub>-ONH<sub>2</sub> **3.11** with psoralen ketone oligonucleotide **3.20** afforded products that could not be separated by HPLC and more importantly were in equilibrium with the starting materials. The major difference between these two oligonucleotides was the sequence context surrounding the photolesion. In the model system, the benzofuran-ketone oligonucleotide was surrounded by adenines; however, the psoralen-ketone oligonucleotide was flanked by guanines. The possible sequence dependent stability of the oxime bond was not explored because it could be reduced to the secondary alkoxyamine to provide a 2PI-psoralen oligonucleotide that was stable. There is a possibility that the O<sup>4</sup> of the psoralen-thymidine adduct maybe susceptible to

reduction because it has been shown that  $\text{NaBH}_4$  can reduce the cyclobutane dimer of thymidine.<sup>80</sup> Although the molecular weight of the 2PI-psoralen oligonucleotide determined by MALDI mass spectroscopy was within experimental error, the additional 2 protons due to  $\text{O}^4$  reduction would be within this error. If this contamination does exist, it would have little to no effect on whether the psoralen adduct can be blocked from repair or not. However, any future experiments that require the integrity of the psoralen-thymidine adduct to be fully maintained may need to seek an alternative method to conjugate a psoralen containing oligonucleotide to a small molecule protein recognition element.

2PI containing oligonucleotides competed with [ $^3\text{H}$ ]-estradiol for binding to the ER (Fig. 3.12). It was discovered that these oligonucleotides competed with relative binding affinities from 0.3% to 1.8% of that of estradiol (Figure 3.12). Oligonucleotides containing a six carbon linker between the N2 of 2PI and the aminoxy group bound 10-fold higher to the ER than compounds with a five carbon chain. The fraction of adducts expected to exist in a complex with the ER in the *in vitro* repair reactions can be estimated from the  $K_d$ .<sup>81</sup> From the relative binding data, the apparent dissociation constant for the 2PI-psoralen adduct for the ER was estimated to be 1.8% that of estradiol ( $K_d = 0.35$  nM), implying a  $K_d$  of ~19 nM for the 2PI-psoralen lesion. Based on these rough calculations, ~ 84% of the 2PI adducts should have been bound at 100 nM ER which, in theory, should translate into repair blocking. One possible source of error for these numbers was in the calculation of the relative binding affinities for the 2PI-psoralen adducts. The ER binding assays were performed on single-stranded substrates whereas the repair experiments were conducted in double-stranded plasmids. Given the secondary structure of the B-DNA helix, it is possible that the hydrophobic planar 2PI molecule could intercalate between

the nucleobases and become available for ER binding. Previous work in the Essigmann lab has suggested that the double-stranded interstrand 2PI-nitrogen mustard crosslink binds 10-fold less than the alkylated single-stranded adducts.<sup>62</sup> However, this result was complicated by the fact that the exact structural nature of the single-stranded species is uncertain because complementary oligonucleotides were used in those experiments. The relative binding affinity of the 2PI-psoralen lesion should be determined in a double-stranded oligonucleotide to answer these questions.

The specific binding of the ER to the 2PI-psoralen oligonucleotides prompted study of the susceptibility of these lesions to NER. The site-specific incorporation of the 2PI-psoralen lesion into circular duplex plasmids allowed for the direct visualization of the incision step of NER. The side products of DNA synthesis were removed with T5 exonuclease since the psoralen adducts are UV sensitive (Fig. 3.16). This method was highly efficient and also eliminated the tedious CsCl/EtBr gradients that were previously required for plasmid purification.<sup>75</sup> All three psoralen adducts were assayed for NER with HeLa whole cell extracts. The 2PI-psoralen adduct was repaired to the greatest extent (Fig. 3.20, compare lanes 2 and 7). This result correlates well with the fact that nucleotide excision repair typically recognizes bulky, structurally distorting lesions. The presence of the 2PI group presumably distorts the structure of B-DNA more than the 2-carboxypsoralen adduct. A trend in the recent NER literature suggests that the more structurally distorting the DNA lesion is, the more efficiently it is removed.<sup>82</sup> In fact, a growing body of evidence has recently revealed that the repair difference between the “repair shielded” 1,2-intrastrand cisplatin-DNA crosslink and the free 1,3-intrastrand adduct can be attributed to the amount of structural distortion each lesion makes. Using a reconstituted NER system in the

absence of shielding HMG-box proteins, Lippard and co-workers demonstrate the 1,3-d(GpTpG) adduct is repaired at least 3-fold better than 1,2-(GpG) adduct.<sup>61c</sup> However, these researchers are quick to point out that the presence of 1  $\mu$ M HMG-box protein can partially reduce repair of the 1,2-(GpG). It remains to be determined if HMG-box proteins are present at these levels in cancer cells that are responsive to cisplatin. Recently, Wood et al. have shown that if the 1,2-d(GpG) adduct paired opposite thymidines it is repaired as well as the 1,3-d(GpTpG) adduct.<sup>83</sup> They speculate that the mismatches make the lesion more distorted and hence a more favorable target for NER. These results suggest that DNA distortion, not shielding, may be responsible for the difference in repair efficiencies for the two cisplatin lesions.

Taking all of this structural distortion into perspective, it is interesting to note that no repair was observed for the psoralen-ketone adduct (Fig. 3.18). Previous work in collaboration with the Wood lab has shown that the structurally similar 2-carboxypsoralen group is repaired by NER. The presence of the psoralen-ketone adduct in the plasmid was suggested by its resistance to cleavage with *Apa*LI. One possible explanation for the absence of a repair signal is that the psoralen-ketone adduct was actually repaired, but the repair signal was lost during the experiment. The excised oligonucleotides containing the ketone moiety maybe forming a Schiff base with the plasmid DNA or proteins in the cell extract preventing the oligonucleotide fragment from entering the gel and eliminating the repair signal. A different incision pattern by NER may also be responsible for the lack of repair signal because the 5' internal radiolabel was placed 14 bases away from the adduct. A 5' incision that occurred less than 14 bases away from the psoralen-ketone adduct would afford oligonucleotide fragments that would not be radiolabeled; however, this incision pattern would be entirely novel because 5' incisions for all

reported adducts to date, including the two other psoralen lesions, occur at least 16 bases from the site of the adduct.

Alternatively, the psoralen-ketone adduct may not be repaired by NER *in vitro*. There are few explanations that may account for its lack of repair. The first and most simple explanation is that the psoralen-ketone adduct is in fact less structurally distorting than the 2-carboxypsoralen adduct and therefore not recognized as a lesion by NER. Another possibility is that the ketone group is forming a Schiff base with a protein in the whole cell extract and thereby shielding itself from being repaired. One possible strategy to pinpoint the role of the ketone group in repair inhibition would be to synthesize isosteres of the psoralen-ketone adduct and monitor the repair of these adducts. If the ketone were indeed forming a Schiff base with a protein in the cell extracts, these isosteres should give a repair signal. Additionally, it should be possible to reduce the Schiff base between the protein and the psoralen-ketone adduct and trap this adduct.<sup>84</sup> A restriction digest of the protein-plasmid species followed by PAGE analysis would easily detect a protein-DNA adduct by retardation of the bands due to the added protein. One set of particularly enticing proteins that could be forming a Schiff base with the psoralen-ketone adduct are the damage recognition proteins of the NER. The psoralen-ketone adduct maybe acting as a pseudo-suicide substrate for NER trapping the proteins crucial for DNA damage recognition. If this scenario is indeed the case, reduction of the psoralen-ketone adduct using a reconstituted NER system would identify the repair protein that was forming the Schiff base. In any case, the lack of a repair signal for the psoralen-ketone adduct is quite intriguing.

The lack of any observable repair blocking with 2PI-psoralen adducts in the presence of ER was not surprising due to the limited amount of experimentation performed (Fig. 3.20). The

incompatibility between the ER storage buffer and repair reaction conditions hampered these experiments (Fig. 3.19). In addition, the lower than reported activity of the commercially available source of ER made the addition of  $\sim 1 \mu\text{M}$  concentrations of ER to the repair reaction infeasible. Although the high affinity of the 2PI compounds for the ER warrants further investigation of this protein for future experiments, a different approach may be needed to determine if the repair blocking hypothesis is a viable strategy leading predictably to selective cytotoxins. An alternative to using purified recombinant ER is to make whole cell extracts from ER+ MCF-7 cells. Repair reactions performed with these extracts could alleviate the ER concentration and storage buffer complications. This approach is somewhat limited because ER activity would be dependent on the MCF-7 cell line; however, the availability of the ER to bind the 2PI-psoralen adduct could be eliminated by adding saturating amounts of competitor estradiol to the repair reactions. While the experiments described in this chapter have not provided any evidence for or against a repair blocking mechanism *in vitro*, the substrates synthesized and methods utilized provide a high degree of flexibility in the design of future experiments.

### **Future Directions**

The uncertainties in this chapter provide many opportunities to design new experiments in order to probe these unanswered questions. The chemoselective conjugation chemistry developed has great potential as a general strategy for selectively attaching small molecules to DNA. For further development of this method it is necessary to determine the source of the apparent sequence context instability. The current oxime formation and reduction system is suitable for testing the repair blocking mechanism *in vitro* and can easily accommodate a change

of the small molecule protein recognition domain and shielding. Other protein/small ligand targets may be more suitable for the testing of the repair blocking hypothesis. Although not medicinally interesting, the high affinity interaction between biotin and avidin could replace the current ER and 2PI duo. If repair blocking is not observed with this strong small molecule-protein interaction, the likelihood of identifying other working combinations is dire. The recent correlation between structural distortion and *in vitro* repair efficiency suggests that therapeutically effective DNA damaging anticancer agents may make stable, non-distorting lesions. In this vein, the psoralen-ketone substrate could serve as a model candidate. The simple addition of the 2-pentanone group to the 2-carboxypsoralen moiety appears to harbor the adduct from damage recognition. Understanding the basis of this phenomenon may be the first step of designing a new generation of DNA damaging agents that can elude the DNA repair machinery and kill rapidly dividing cancer cells.





## Chapter 4

# Design of an Intercalation Inhibitor

### Introduction

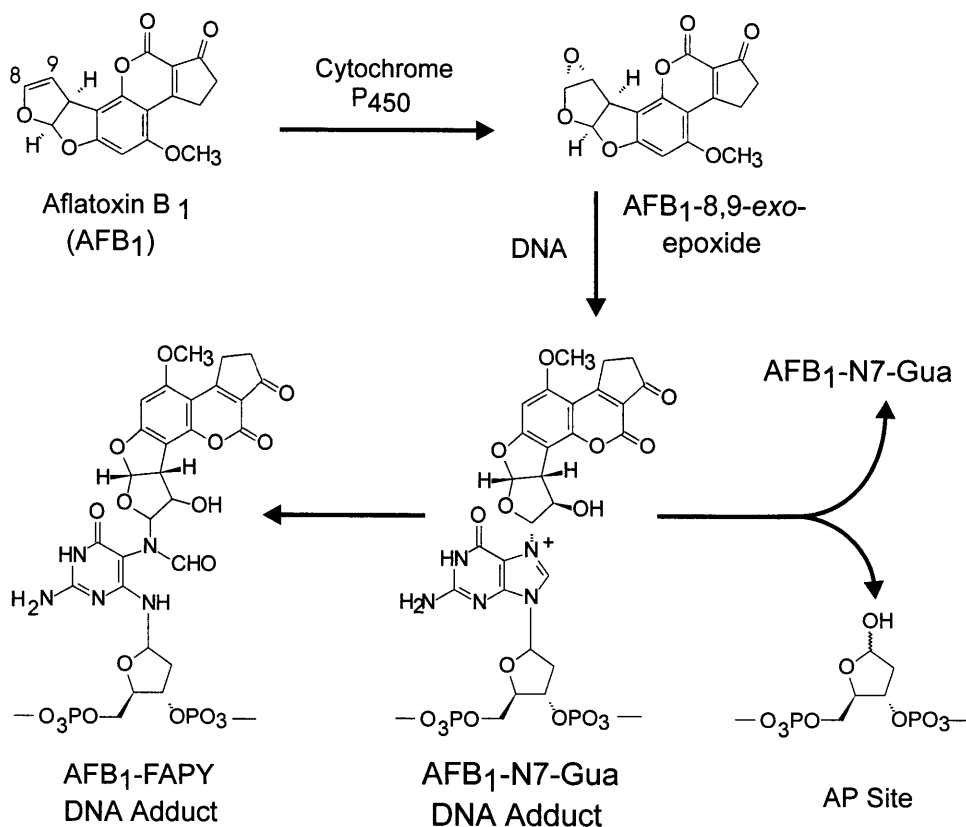
The benefits of working with oligonucleotides containing a site-specifically placed DNA adduct have been well documented in the literature and throughout this dissertation. The total synthesis of a psoralen-thymidine monoadduct illustrates the flexibility provided by using a total synthetic approach in the design of single-lesion substrates; however, there are DNA adducts that have not been or cannot be made using a total synthetic approach. DNA damaging agents that make N7-guanine adducts such as the aflatoxins, nitrogen mustards, and some alkylating agents are not suitable for a total synthetic approach because these lesions would decompose during DNA synthesis and/or deprotection.<sup>1b, 35</sup> Other lesions (benzo[*a*]pyrene, MeIQx, and aromatic amines) are deemed too synthetically challenging and are therefore synthesized by directly treating an oligonucleotide with an activated DNA damaging agent.<sup>34</sup> Although the direct treatment procedure is quite facile, the DNA sequence of the target oligonucleotide must be chosen carefully to avoid multiple products that would complicate purification. Unfortunately, this sequence restriction prevents the study of many biologically important lesions that occur in mutational hotspots in proto-oncogenes and tumor suppressor genes that are thought to play a key role in neoplastic transformation. It is desirable to have a method that combines the flexibility of a total synthetic approach and the simplicity of a direct treatment procedure.

One strategy to achieve site-specific placement of adducts exploits a key mechanistic step exhibited during the reaction of many hydrophobic alkylating agents with DNA -- intercalation

into the helix. Synthesis of a compound that can occupy an intercalation site in a DNA duplex could be used to direct an alkylating agent selectively to an unoccupied intercalation site. To explore the potential of this methodology, a system was developed using aflatoxin B<sub>1</sub> (AFB<sub>1</sub>) epoxide as the alkylating agent: DNA duplexes were prepared consisting of a target strand containing multiple potentially reactive guanines and a nontarget strand containing a *cis*-syn benzofuran-thymidine photoproduct. Because the covalently linked benzofuran moiety physically occupies an intercalation site, that site should be rendered inaccessible to AFB<sub>1</sub> epoxide. By strategic positioning of this intercalation inhibitor, it should be possible to alter the target selectivity of aflatoxin B<sub>1</sub> epoxide to afford singly-adducted oligonucleotides that have multiple sites of reactivity.

Aflatoxin B<sub>1</sub> (AFB<sub>1</sub>) is a hydrophobic toxic metabolite of the fungus *Aspergillus flavus*.<sup>85</sup> The fungus thrives in climates with high heat and humidity producing AFB<sub>1</sub> that has been detected during growth, harvest, and storage of foods such as peanuts, corn and rice.<sup>86</sup> Exposure to AFB<sub>1</sub> and infection by hepatitis B virus are known risk factors for hepatocellular carcinoma in developing areas of the world.<sup>87</sup> Although, the molecular mechanism of human hepatocellular carcinogenesis is poorly defined, AFB<sub>1</sub> is known to be a potent mutagen and hence may induce the mutations that appear in end stage tumors. Upon metabolic activation to the *exo*-8,9-epoxide, AFB<sub>1</sub> reacts almost exclusively with the N7 of guanine to form primarily guanine DNA adducts (AFB<sub>1</sub>-N7-Gua; Fig. 4.1).<sup>88</sup> The N7 adduct is quite unstable due to the positive charge on the imidazole ring. The half life of the intact N7 adduct at neutral pH has been determined to be 12-50 h *in vitro*, depending on the adduction levels of AFB<sub>1</sub>.<sup>89</sup> As with most positively charged N7 adducts, the AFB<sub>1</sub>-N7-Gua can depurinate giving rise to an apurinic site. Under mildly basic

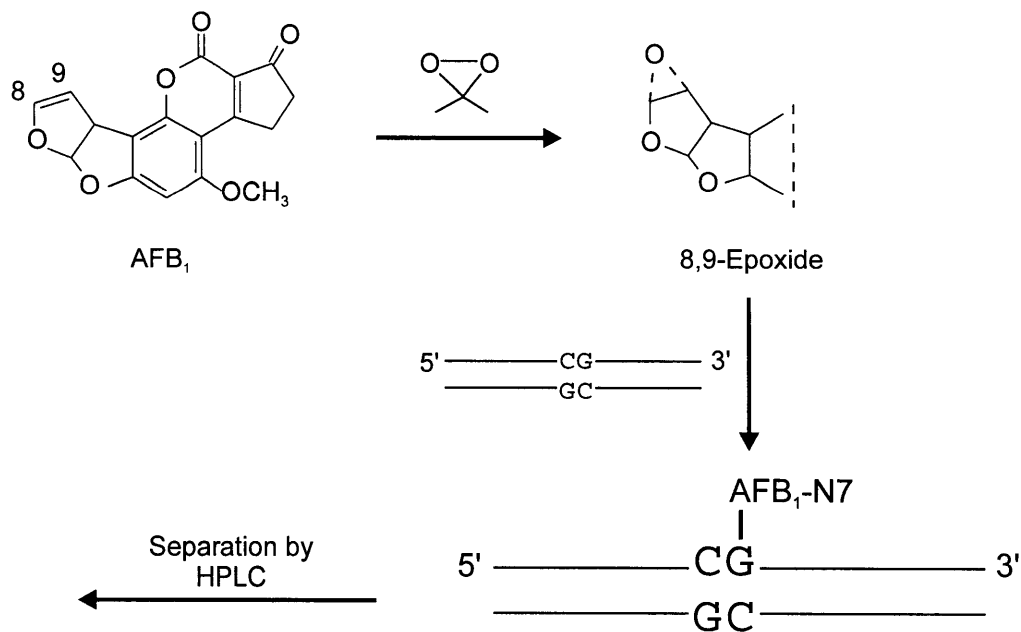
conditions, water can add to the C8 position and open the imidazole ring to form the chemically more stable AFB<sub>1</sub>-formamidopyrimidine adduct.<sup>90</sup> Additionally, AFB<sub>1</sub>-N7-Gua adducts have also been observed to reverse to guanine and free AFB-diol *in vitro*.<sup>91</sup>



**Figure 4.1.** AFB<sub>1</sub> is activated by cytochrome P<sub>450</sub> to generate its *exo* epoxide. The epoxide reacts with DNA to form the primary AFB<sub>1</sub> adduct, AFB<sub>1</sub>-N7-Gua, which can undergo depurination to an AP site or opening of the imidazole ring to form AFB<sub>1</sub>-FAPY.

Analysis of bacterial mutational spectra and phage genomes containing a specific AFB<sub>1</sub>-N7-Gua adduct reveals G→T substitutions as the predominant mutagenic event.<sup>92</sup> Interestingly, approximately 50% of hepatocellular carcinomas in portions of eastern Asia and sub-Saharan Africa, where exposure to AFB<sub>1</sub> through contaminated food is a frequent event,

possess G:C→T:A transversions in the third position of codon 249 of the p53 tumor suppressor gene.<sup>93</sup> The construction of oligonucleotides and genomes containing AFB<sub>1</sub>-N7-Gua adducts at specific sites would facilitate understanding of the molecular origin of the observed p53 mutational hotspot in hepatocellular carcinoma. A recent advance in chemical synthesis has facilitated the production of AFB<sub>1</sub> epoxide, which provides access to oligonucleotides containing site-specific aflatoxin B<sub>1</sub> adducts by simple treatment of an oligonucleotide duplex with AFB<sub>1</sub> epoxide (Fig. 4.2).<sup>1b</sup> However, the formidable obstacle of resolving and purifying highly labile AFB<sub>1</sub> adducts from complex mixtures of multiple reaction products has limited the synthesis of site-specifically modified AFB<sub>1</sub> adducts to targets containing a single guanine.<sup>92,94</sup> This synthetic limitation has hampered the study of AFB<sub>1</sub> adducts in the p53 mutational hotspot because of the presence of numerous guanines in the nucleotide sequence surrounding codon 249.



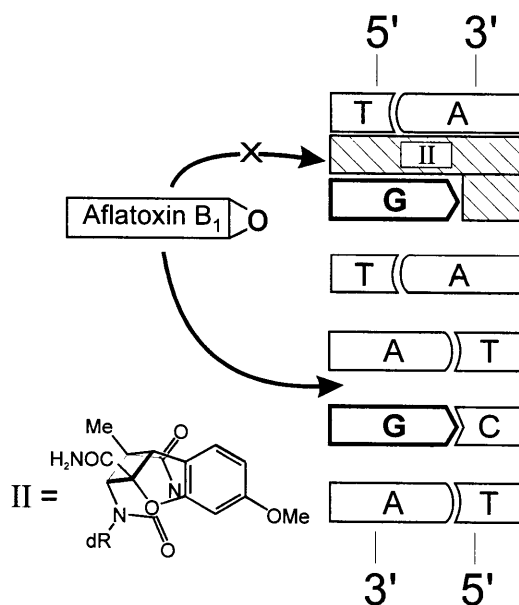
**Figure 4.2.** Construction of AFB<sub>1</sub> containing oligonucleotides using the direct treatment method developed by Harris and co-workers.<sup>1b</sup> Oxidation with dimethyl dioxirane affords AFB<sub>1</sub> epoxide. Treatment of a DNA duplex containing only one reactive guanine with AFB<sub>1</sub> epoxide and separation by HPLC yields an oligonucleotide containing a single AFB<sub>1</sub> adduct.

Current evidence suggests that formation of AFB<sub>1</sub>-N7-Gua DNA adducts proceeds through a transition state in which the AFB<sub>1</sub> epoxide is intercalated on the 5' side of the target guanine. This hypothesis is supported by the observation that the reactivity of AFB<sub>1</sub> epoxide with double stranded B-form DNA is greatly enhanced as compared to single stranded DNA or alternative duplex structures.<sup>95</sup> From NMR structural studies on the adduct, it is speculated that 5' intercalation facilitates adduct formation by positioning the epoxide for in-line nucleophilic attack by the guanine N7.<sup>96</sup> Additional evidence for a 5' intercalation event arises from stoichiometric studies with the two self-complementary hexamers ATCGAT and ATGCAT. Reaction with excess AFB<sub>1</sub> epoxide yields only a 1:1 ratio of AFB<sub>1</sub> to the duplex oligonucleotide (ATCGAT)<sub>2</sub>, where the two guanines share the same 5' intercalation site; in contrast, in the case where distinct 5' intercalation sites exist in each strand, a 2:1 AFB<sub>1</sub>: (ATGCAT)<sub>2</sub> ratio is observed.<sup>94</sup> Furthermore, modifications of the ring systems in AFB<sub>1</sub> that decrease planarity and therefore intercalation ability, a situation that exists with aflatoxin G<sub>1</sub>, decrease affinity for DNA and result in lower reactivity.<sup>97</sup> Finally, addition of the intercalating agent ethidium bromide to target DNA prior to treatment with AFB<sub>1</sub> epoxide greatly reduces guanine reactivity.<sup>95a</sup> These experiments, coupled with stopped-flow kinetic analysis, support a model in which the intercalated intermediate provides a kinetic and entropic advantage for productive reaction with DNA over hydrolysis.<sup>98</sup> Moreover, in the absence of this favorable interaction, AFB<sub>1</sub> epoxide can be readily hydrolyzed to the inactive AFB<sub>1</sub> diol.

Another agent that intercalates prior to its reaction with DNA and is the cornerstone of this thesis is psoralen.<sup>22</sup> Detailed structural analysis of psoralen-thymidine adducts reveal that both the *cis*-syn furan-side monoadduct and the crosslink are intercalated in the DNA duplex.<sup>22</sup>

The structurally analogous model compound, a *cis*-syn benzofuran-thymidine photoproduct, also possesses the same stereo- and regiochemistry as the *cis*-syn psoralen furan-side thymidine monoadduct; these constraints dictate that the benzofuran moiety occupies an intercalation site when situated in duplex DNA, raising the possibility that such a molecule could be used to inhibit subsequent intercalation by other molecular species.

In light of the above observations suggesting that intercalation of AFB<sub>1</sub> epoxide precedes covalent bond formation, it seemed reasonable that physical occupation of an intercalation site with a thymidine-benzofuran photoproduct would inhibit intercalation by AFB<sub>1</sub> epoxide and thereby prevent the formation of an AFB<sub>1</sub> adduct at the proximal guanine (Fig. 4.3). Removal of the inhibiting strand would provide singly-adducted AFB<sub>1</sub> oligonucleotides. The sequence



**Figure 4.3.** Experimental scheme. Occupation of an intercalation site 5' to a guanine should prevent intercalation by AFB<sub>1</sub> epoxide and reduce reactivity at the protected site. The covalently attached intercalation inhibitor (designated II), a *cis*-syn thymidine-benzofuran photoproduct, is positioned in the complementary strand. The benzofuran moiety is positioned on the 3' side of the T component of the *cis*-syn thymidine-benzofuran adduct.

context freedom of the benzofuran-thymidine monoadduct and the simplicity of treating duplex DNA with AFB<sub>1</sub> epoxide provides a means of combining the best attributes of both methodologies.

The ideas for this work arose from the many discussions with Drs. Gerald Wogan, Lisa Bailey and Marjorie Solomon. However, this work would not of come to fruition without my co-worker David Wang. Dave was an integral part in the design and execution of all of the experiments to be described. In this chapter, we present evidence that the reaction of DNA with AFB<sub>1</sub> epoxide is modulated by inhibiting intercalation. Moreover, judicious placement of the intercalation inhibitor in the complementary strand of a human p53 gene derived sequence results in simplification of the local adduct spectrum and improves the yield of individual adducts in the p53 mutational hotspot.

### **Intercalation Inhibition of AFB<sub>1</sub> Epoxide**

Evidence accumulated over the past several years suggests that intercalation of AFB<sub>1</sub> epoxide greatly enhances the reactivity of the potent carcinogen with DNA.<sup>94-96</sup> Accordingly, if it were possible to reduce the likelihood of intercalation at a given site, one would expect to diminish substantially the adduct yield at that site (Fig. 4.3). A series of duplex oligonucleotides was designed to examine this possibility (Table 4.1). In each case, one strand contained two or more guanines that were potential adduction sites. In the complementary strand, a *cis-syn* thymidine-benzofuran intercalation inhibitor was situated to occupy the intercalation site immediately 5' to one of the guanines. Chemical synthesis of the phosphoramidite of the intercalation inhibitor (Fig. 4.4) would enable the control of precisely the number and position of

Table 4.1. Oligonucleotides used in this chapter.

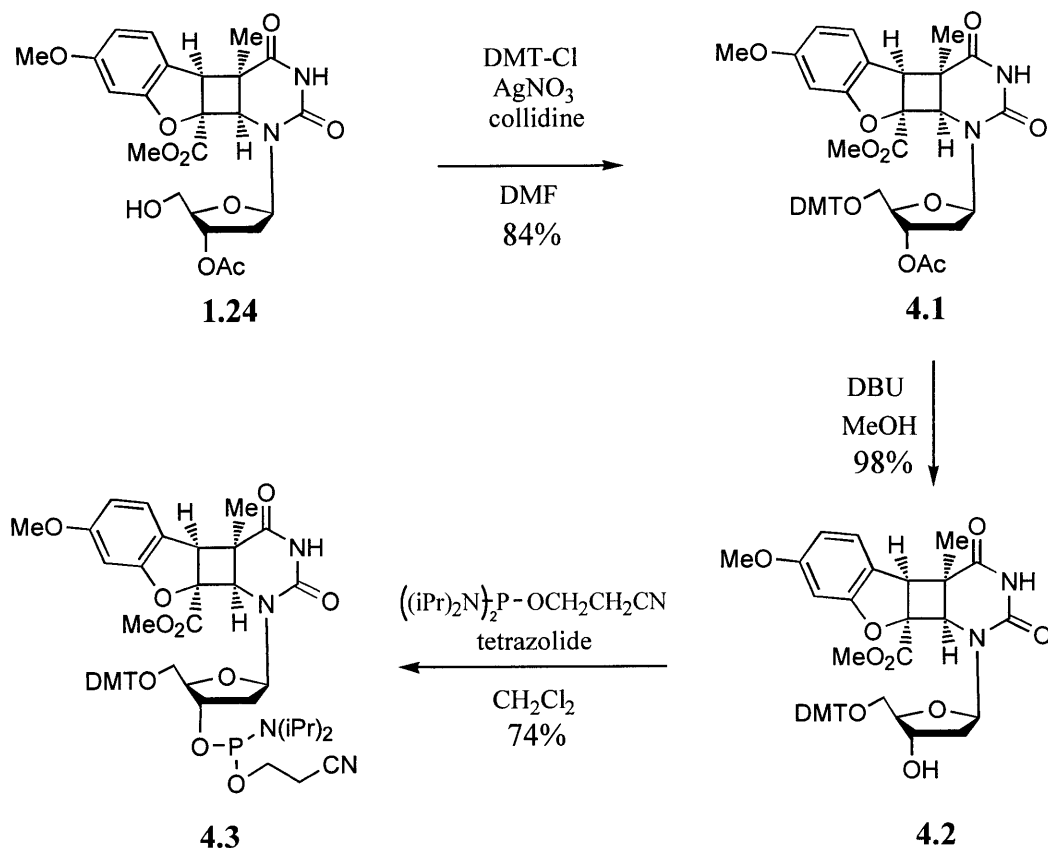
---

H	5' - ATAGATTGTA - 3'
QC	3' - TATCTAACAT - 5'
QT	3' - TATTTAACAT - 5'
Q8	3' - TAXTTAACAT - 5'
Q7	3' - TATXTAACAT - 5'
Q6	3' - TATTXAACAT - 5'
Q3	3' - TATCTAAXAT - 5'
GG	5' - TATAGGTTAT - 3'
R0	3' - ATATCCAATA - 5'
R6	3' - ATATXCAATA - 5'
R5	3' - ATATCXAATA - 5'
P53	5' - CCGGAGGCC - 3'
S0	3' - T77TCTCC7 - 5'
S6	3' - T77XCTCC7 - 5'
S5	3' - T77CXTCC7 - 5'

---

7=7-deaza-dG; X=thymidine-benzofuran  
photoproduct (intercalation inhibitor)





**Figure 4.4.** Synthesis of 2-carbomethoxybenzofuran-thymidine phosphoramidite **2.4**

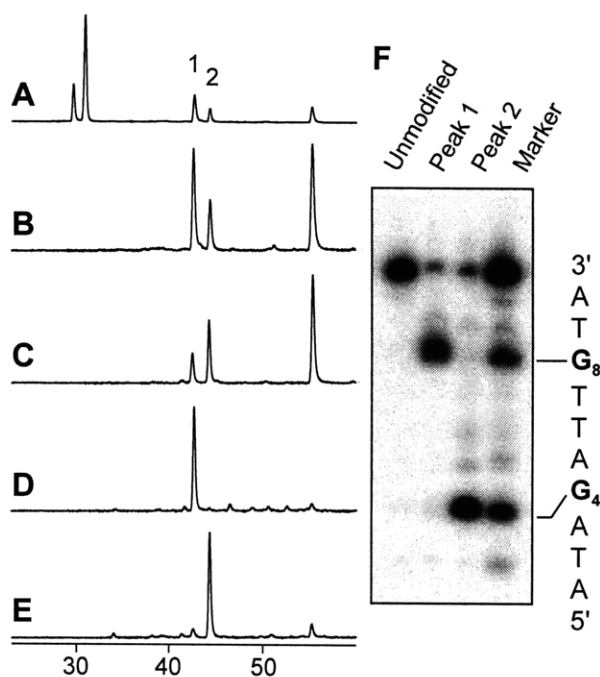
the intercalation inhibitors in a target duplex. The synthesis of deoxynucleoside **1.24** was described in Chapter 1. Treatment of **1.24** with 4,4'-dimethoxytrityl chloride in the presence of silver nitrate led to rapid protection of the 5' hydroxyl.<sup>55</sup> Removal of the acetate protecting group with 5% 1,8-diazabicyclo[5.4.0]undec-7-ene in freshly distilled methanol afforded the free 3' hydroxyl. Phosphitylation using standard conditions gave phosphoramidite **4.3**. Upon reacting the various DNA duplexes with AFB<sub>1</sub> epoxide, HPLC analysis was utilized to separate and isolate the reaction products. The standard elution conditions were sufficient to denature the duplexes, resolving the target strands from their complements. By this method, the complementary strands, which contained the unnatural thymidine-benzofuran moieties, could be

removed and the single-stranded site-specifically AFB<sub>1</sub> modified oligonucleotides could be isolated.

### *A Model System*

The simplest model system consisted of a 10-bp oligonucleotide duplex containing only two G residues in one strand separated by three nucleotides (Table 4.1, H/QC). In the control sequence, reaction with AFB<sub>1</sub> epoxide afforded two peaks corresponding to AFB<sub>1</sub> monoadducts and one peak representing the diadduct, as determined by the ratio of their UV absorbances at 360 nm and 260 nm (Fig. 4.5A and 4.5B). The two monoadduct peaks were isolated, desalted and then 5'-<sup>32</sup>P end-labeled by using polynucleotide kinase. The known alkaline lability of AFB<sub>1</sub>-N7-Gua adducts<sup>99</sup> was utilized to determine the identity of each peak. Electrophoresis of the piperidine treated samples revealed that peak 1 was the G8 AFB<sub>1</sub> adduct while peak 2 was the G4 adduct (Fig. 4.5F). It is known that the reaction of AFB<sub>1</sub> epoxide with guanines in duplex DNA is not random, but rather displays sequence specificity. Although the basis for this discrimination is not well understood, a systematic investigation by Loechler and coworkers has empirically determined reactivity rules based on the immediate 5' and 3' neighboring bases.<sup>100</sup> In the first model system, the two adducts formed in a ratio of 2.4:1, in close agreement to the predicted 2:1 ratio.

Upon altering the complementary strand to position an intercalation inhibitor in the 5' intercalation site of G4 (H/Q7), the ratio of observed products was dramatically shifted; essentially no monoadducts derived from G4 or diadducts were observed (Fig. 4.5D). Conversely, placement of the intercalation inhibitor opposite G8 (H/Q3) greatly reduced the

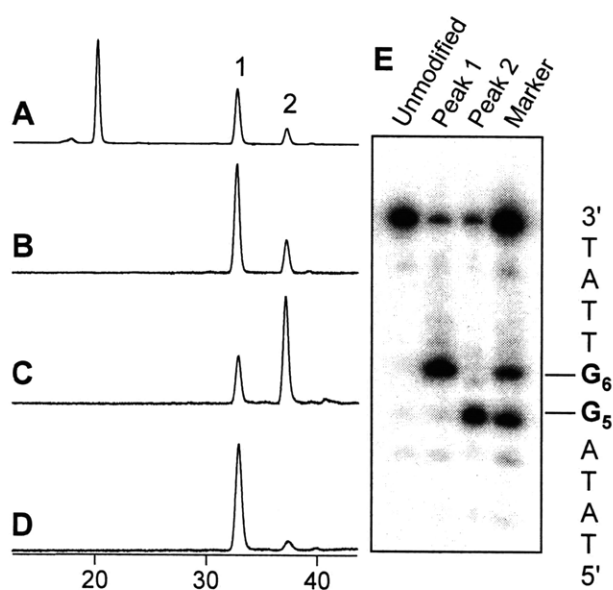


**Figure 4.5.** Intercalation inhibition in duplex (H/Q) series. Only molecules containing AFB<sub>1</sub> have UV absorbance at 360 nm. (A) 260 nm HPLC trace of the AFB<sub>1</sub> reaction with H/QC. The peaks near 30 min are the unreacted H and QC oligonucleotides, respectively. Peaks 1 and 2 are AFB<sub>1</sub> monoadducts, and the peak near 55 min is an AFB<sub>1</sub> diadduct. (B) 360 nm trace of H/QC. (C) 360 nm trace of H/QT reaction. (D) 360 nm trace of H/Q7. (E) 360 nm trace of H/Q3. (F) Electrophoretic analysis: piperidine cleavage of purified peaks 1 and 2 and comparison to markers generated by the Maxam and Gilbert G-reaction. Piperidine cleavage of the untreated duplex is shown in the lane labeled “unmodified.”

extent of reaction at G8 such that only 11% of the monoadducts arose from G8 (Fig. 4.5E). The asymmetry in the protection afforded is likely due to the inherently higher reactivity of G8, making total abolition of reactivity at that site more difficult.

Because positioning the intercalation inhibitor to occupy the intercalation site 5' to a target guanine necessarily forms a G-T mismatch at that site, the effect of a mismatch alone (H/QT) on adduct distribution was examined. Interestingly, substitution of a G-T base pair for a G-C actually resulted in increased reactivity of the mismatched G (compare Fig. 4.5B to 4.5C). Thus, the observed diminished reactivity upon introduction of an intercalation inhibitor can be

attributed to the presence of an intercalated moiety and not merely distortions resulting from a mismatched base pair.



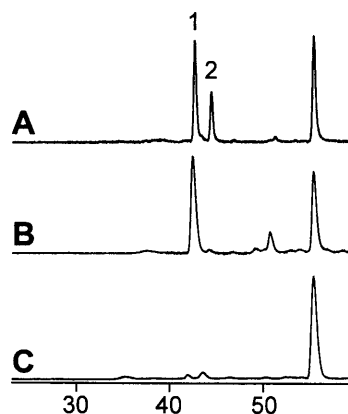
**Figure 4.6.** Intercalation inhibition in duplexes GG/R. (A) 260 nm HPLC trace of reaction with GG/R0. Large peak near 20 min is R0. Peaks 1 and 2 are AFB<sub>1</sub> monoadducts. (B) 360 nm trace of GG/R0. (C) 360 nm trace of GG/R5. (D) 360 nm trace of GG/R6. (E) Piperidine cleavage of purified peaks 1 and 2.

#### *AFB<sub>1</sub> Inhibition at Adjacent Guanines*

In the second model system, the reactivity and inhibition in oligonucleotides containing two adjacent guanines (GG) was examined. In the absence of an intercalation inhibitor (GG/R0), two peaks were observed in the HPLC trace, reflecting the two different monoadducts (Fig. 4.6A and 4.6B). Loechler's rules<sup>100</sup> predict that the two adducts should form in a 3.25 to 1 ratio, which agrees well with the observed product ratio of 3.3:1 (Fig. 4.6B). Interestingly, there were no diadduct peaks present in the trace. This observation is consistent with the neighbor exclusion principle that has been postulated to govern the ability of molecules to intercalate at adjacent sites.<sup>101</sup> Initial intercalation and reaction of an AFB<sub>1</sub> epoxide molecule at one guanine results in

inhibition of intercalation at the other adjacent guanine, thereby preventing the formation of a second adduct. Nevertheless, placement of the intercalation inhibitor at either of the two guanine sites (GG/R5 and GG/R6) resulted in an increase in the yield of adduct at the non-protected guanine (Fig. 4.6C and 4.6D).

Because the AFB<sub>1</sub> adducts seemed to exhibit neighbor exclusion properties, the benzofuran moiety was tested to see if it possessed similar characteristics. Duplexes were prepared with an intercalation inhibitor occupying the intercalation sites either 3' to or 5' to the actual target intercalation site (Fig. 4.7). Positioning the intercalation inhibitor 3' to the target site provided some inhibition (Fig. 4.7B), whereas placement of the intercalation inhibitor to the 5' side had no effect (Fig. 4.7C).



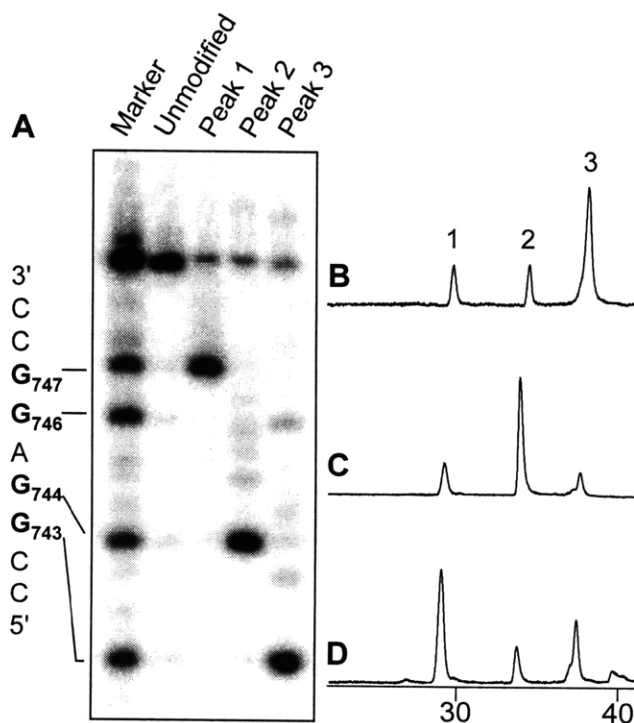
**Figure 4.7.** Nearest neighbor inhibition in duplex (H/Q) series. (A) 360 nm trace of H/QC reaction. Peaks 1 and 2 are AFB<sub>1</sub> monoadducts, and the peak near 55 min is an AFB<sub>1</sub> diadduct. (B) 360 nm trace of H/Q6 reaction. (C) 360 nm trace of H/Q8 reaction.

#### *Synthesis of AFB<sub>1</sub> Adducts in a p53 Mutational Hotspot*

Finally a sequence representing nucleotides 741 to 749 of the human p53 cDNA, 5' CCGGAGGCC (P53), was tested (Fig. 4.8). In the complementary strands, guanine bases were

substituted with 7-deaza-dG to prevent formation of additional AFB<sub>1</sub> adducts in this strand; such adducts would complicate analysis of the adduct spectrum. Treatment of a control duplex (P53/S0) revealed three HPLC peaks that corresponded to AFB<sub>1</sub> monoadducts (Fig. 4.8B). After <sup>32</sup>P labeling and alkali cleavage, it was determined that peak 1 contained exclusively the G747 adduct, peak 2 contained the G744 adduct, while peak 3 was a mixture of the G743 and G746 (Fig. 4.8A). Under certain conditions, peak 3 could actually be resolved to afford two peaks.

Experiments carried out with two intercalation inhibitors present within several bp of each other yielded reaction profiles consistent with single-stranded DNA, indicating that too



**Figure 4.8.** Effects of introducing an intercalation inhibitor into duplexes derived from the human p53 sequence. For clarity, all HPLC traces are at 360 nm and show only the monoadduct region. (A) Electrophoretic analysis: piperidine cleavage of peaks 1, 2 and 3 and comparison to markers. The sequence of oligonucleotide P53 is shown (left). Numbers refer to the nucleotide position of each base in the human p53 cDNA. (B) AFB<sub>1</sub> reaction with P53/S0. Peaks 1, 2 and 3 are monoadducts. (C) Reaction with duplex P53/S6. (D) Reaction with duplex P53/S5.

much structural distortion was imposed by multiple intercalators or mismatches to maintain duplex structure. Thus, the effects of placing a single intercalation inhibitor at different positions in the DNA was determined. The first experiment examined protection of G743 (P53/S6), the nucleotide predicted to have the greatest inherent reactivity. Protection of that site altered the product profile significantly: peak 3 was greatly reduced; peak 2 increased by approximately 4-fold; and peak 1 was relatively unchanged (Fig. 4.8C). In contrast, protection of G744 (P53/S5) resulted in diminished yield of peak 3 while enhancing peak 1 by 3-fold (Fig. 4.8D). The presence of intact AFB<sub>1</sub> monoadduct in each sample was confirmed by electrospray MS.

## **Discussion**

Given that the half life of AFB<sub>1</sub> epoxide in water is approximately one second,<sup>102</sup> it has been proposed that intercalation of AFB<sub>1</sub> epoxide provides a kinetic and entropic advantage for productive reaction with DNA versus unproductive hydrolysis.<sup>97,98</sup> In order to examine directly the effects of intercalation on AFB<sub>1</sub> epoxide reactivity with DNA, a means was devised to occupy intercalation sites with covalently linked benzofuran derivatives (Fig. 4.3). These results support the hypothesis that intercalation in DNA by AFB<sub>1</sub> epoxide facilitates adduct formation, as evidenced by the observation that abolishing intercalation at a given site greatly diminished reactivity at that site. Furthermore, application of this principle to the problem of synthesizing specific AFB<sub>1</sub>-N7-Gua adducts within a region of the p53 gene and increased the yield of isolatable adducts by several fold.

The first set of experiments clearly revealed that intercalation contributes significantly to the ability of AFB<sub>1</sub> epoxide to react with DNA (Fig. 4.5). The relative adduct yields could be

shifted from almost exclusively one adduct to almost exclusively the other depending upon the location of the intercalation inhibitor. In addition, diadduct formation was almost completely abolished. There remained, however, some residual adduction at the protected site (Fig. 4.5D and 4.5E). This observation was not unexpected since both single stranded DNA and mononucleotides are known to react, albeit to a much lesser extent, with AFB<sub>1</sub> epoxide<sup>95b</sup> indicating that intercalation is not an obligate step for reaction with guanine.

The susceptibility of AFB<sub>1</sub> adducts to piperidine cleavage enabled the assignment of the HPLC peaks to specific guanine adducts by comparison to Maxam-Gilbert G-reaction markers (Fig. 4.5F). No cross contamination of the two different adducts was observed. A minor band (10%) of similar mobility to the unmodified target oligonucleotide was observed in the piperidine treated samples for every monoadduct (Fig. 4.5F, 4.5E and 4.5A). This band most likely arises during manipulation of the samples from chemical reversal of the AFB<sub>1</sub>-N7-Gua adduct to give free AFB<sub>1</sub>-diol and guanine, as has been previously observed *in vitro*.<sup>91</sup> It is unlikely to be a contamination from the HPLC purification process as the band persisted, even after multiple HPLC purifications.

An intriguing finding is the increased reactivity of DNA with AFB<sub>1</sub> epoxide upon substituting T for C in the complementary strand (Fig. 4.5C), suggesting that the diminished reactivity observed after introduction of an intercalation inhibitor is due to the presence of the intercalated benzofuran moiety and not merely to distortions induced by a G-T mismatch. One possible explanation is that the guanine is more favorably positioned for nucleophilic attack on the epoxide when it is engaged in a wobble pairing characteristic of G-T base pairs.<sup>103</sup>

Alternatively, the mismatch itself, or the presence of the additional hydrophobic methyl group of



T, might facilitate intercalation. As CpG groups in mammalian DNA are often methylated, methylation status could play a significant role in the reaction of AFB<sub>1</sub> epoxide with DNA *in vivo*. Of particular relevance is recent evidence that CpG methylation significantly alters the distribution of DNA adducts of another intercalative carcinogen, benzo[*a*]pyrene diol epoxide.<sup>104</sup>

The second model system tested the reactivity of AFB<sub>1</sub> epoxide with a target DNA containing two adjacent guanines (Fig. 4.6). One interesting observation was that essentially no diadducts were observed with this duplex. The absence of diadducts can be rationalized on the basis of the nearest neighbor exclusion principle, which postulates that occupation of a single intercalation site prevents subsequent intercalation both immediately 5' and 3' to the initial site.<sup>101</sup> In this scenario, initial modification at one guanine results in occupancy of an intercalation site, effectively blocking intercalation at the adjacent guanine target site and thus diminishing reactivity of that guanine. Nevertheless, addition of intercalation inhibitors in the complementary strand had the expected effects. In both cases (GG/R5 and GG/R6), reaction at the protected guanine diminished while the yield of the other adduct increased (Fig. 4.6C and 4.6D). From these experiments, it is unclear whether the benzofuran moiety possesses neighbor excluding properties and to what extent this principle affects the reaction. If the neighbor exclusion principle is equally applicable to all intercalators, then the expectation upon placing an intercalation inhibitor at one of the two sites would be that all reactivity should be abolished. This was not the case, however, suggesting that the neighbor exclusion principle is not strictly applicable. Indeed, other violations of the neighbor exclusion principle have been reported.<sup>105</sup>

Based on the observations with the GG/R0 duplex, the intercalation inhibitor was tested for nearest neighbor exclusion properties in a simpler system using oligonucleotide H (Fig. 4.7).

Duplexes containing an intercalation inhibitor shifted one base to either the 5' or 3' side of the target were treated with AFB<sub>1</sub> epoxide. These experiments revealed an asymmetry in the protection afforded by these constructs. Placing the intercalation inhibitor immediately 3' to the target resulted in some protection, but the presence of diadducts suggests that this site is somewhat accessible to the AFB<sub>1</sub> epoxide. However, when placed to the 5' side, the reactivity at the target provided exclusively diadduct. The lack of consistently observable neighbor exclusion by this particular intercalation inhibitor may stem from the size of the intercalation inhibitor or the specific helix-distortion imposed. Expansion of the benzofuran moiety to a bulkier molecule that completely spans the helix may fill the entire intercalation space and serve as a more efficient inhibitor of intercalation. In any case, examination of the basis of the neighbor exclusion principle warrants further investigation.

Having established the feasibility of altering chemical reactivities by intercalation inhibition, this methodology was applied to help understand the conundrum presented by an observed AFB<sub>1</sub>-related p53 mutational hotspot. In greater than 50% of hepatocellular carcinomas in regions of the world with exposure to AFB<sub>1</sub> and hepatitis B, a G to T mutation at G747 (the third position of codon 249) is observed.<sup>93</sup> Based on Loechler's predictions, G747 should not be exceptionally prone to adduction; the four potential target guanines in the p53 mutational hotspot sequence (Fig. 4.8E) are expected to have relative reactivities of 3:1.4:1:1.4.<sup>100</sup> Why, then, are so many mutations observed at G747? Although selective pressures certainly play a role in shaping the mutations observed in end-stage tumors, selection alone is unlikely to explain why G747 mutations are so abundant. One approach towards understanding the factors and processes subsequent to adduct formation that lead to a mutation is the use of oligonucleotides containing

adducts at specific sites.<sup>106</sup> By synthesizing a series of oligonucleotides each containing an AFB<sub>1</sub> adduct at a different site, one can begin to study the relative ability of each adduct to induce a mutation. In addition, site-specific methods can address the identity of the actual mutagenic species. AFB<sub>1</sub>-N7-Gua adducts are highly labile, undergoing either depurination to afford an abasic site or imidazole ring opening to form the formamidopyrimidine derivative. Insertion of defined substrates into an appropriate vector will enable the first site-specific examination of the ability of each of the AFB<sub>1</sub> adducts or their decomposition products to induce mutations in the p53 sequence. However, to date, no study of the synthesis of site-specific AFB<sub>1</sub> adducts in such a sequence or any oligonucleotide containing multiple guanines has been reported. The ultimate goal of our strategy was to utilize intercalation inhibitors to facilitate the synthesis and purification of specific AFB<sub>1</sub> adducts in such a sequence. In studies of the GC-rich p53 sequence between nucleotides 741-749 (Fig. 4.8), 7-deaza-dG, which lacks a nitrogen at position 7, was used in the place of dG in the complementary strand to simplify the reaction products. In the absence of an intercalation inhibitor (P53/S0), three peaks corresponding to monoadducts were obtained with peak 3 the dominant reaction product (Fig. 4.8B). Electrospray MS confirmed that the isolated products were of the expected molecular weight corresponding to intact AFB<sub>1</sub> monoadducts. As anticipated, placement of two intercalation inhibitors within several bases of each other in a short oligonucleotide appeared to prevent hybridization of the duplex. Thus, the effects of a single intercalation inhibitor on the product profile were examined. Upon protection of the theoretically most reactive site, G743, the yield of the other peaks increased significantly, with the G744 adduct comprising greater than 50% of the monoadducted species (Fig. 4.8C). Similarly, protection of G744 lead to a large increase in the yield of the G747 adduct (Fig. 4.8D).

This is consistent with the observed asymmetry of protection afforded by the intercalation inhibitor, which appears to inhibit intercalation at both the site itself as well as the 5' neighboring intercalation site. Thus, depending upon the location of the intercalation inhibitor, it was possible to alter the adduct distribution such that any of the three monoadduct peaks represented the major reaction product (Table 4.2).

### **Conclusions**

This work demonstrates clearly that inhibiting intercalation of AFB<sub>1</sub> epoxide can significantly alter the resulting adduct spectrum and that introduction of intercalation inhibitors into DNA duplexes provides a facile means of manipulating product yields (Table 4.2). Even in complex sequence environments, such as the p53 mutational hotspot sequence surrounding codon 249, a significant effect on adduct distribution with a single intercalation inhibitor can be demonstrated (Table 4.2). In addition, it should be possible to use these novel molecules to gain a better understanding of the phenomenon of intercalation. Structure-function investigations with intercalation inhibitors may shed more light on factors that contribute to the neighbor exclusion principle, and second generation intercalation inhibitors can be designed with superior inhibitory or neighbor excluding capabilities. Moreover, intercalation inhibitors can be used to probe the intercalation requirements of other carcinogens, such as benzo[*a*]pyrene and aromatic amines, and facilitate the site-specific synthesis of carcinogen-DNA adducts.

Table 4.2. Relative monoadduct yield as a percentage of total monoadducts.

Duplex	% Peak 1	% Peak 2	% Peak 3
H/QC	70	30	N/A
H/QT	34	66	N/A
H/Q7	>95	<5	N/A
H/Q3	11	89	N/A
GG/R0	77	23	N/A
GG/R6	93	7	N/A
GG/R5	33	67	N/A
P53/S0	17	14	69
P53/S6	21	60	18
P53/S5	51	14	35

Peak areas determined by integration of  $A_{360}$ .

N/A, not applicable



## Experimental Section

**General Procedures.** Unless otherwise specified, materials were obtained from commercial suppliers and were used without further purification. Tetrahydrofuran (THF) was distilled from potassium/benzophenone ketyl. Methanol (MeOH) was distilled from sodium metal and used immediately after distillation. Anhydrous solvents were otherwise obtained from Aldrich. Column chromatography was performed on Merck silica gel (230-400 mesh). <sup>1</sup>H NMR spectra were recorded at 500 or 300 MHz on superconducting FT spectrometers. <sup>13</sup>C NMR spectra were proton decoupled and were recorded at 125 or 75 MHz. <sup>1</sup>H NMR data are tabulated in order: multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad; app, apparent), number of protons, coupling constant(s) in Hertz. Melting points (Pyrex capillary) are uncorrected. Fast atom bombardment (FAB+) and electron impact (EI) were recorded at the MIT Mass Spectral Laboratory. All high-pressure liquid chromatography (HPLC) separations were conducted with a Rainin analytical or preparative instrument with Rainin Dynamax or Beckmann ODS columns.

**Synthesis of Modified Oligonucleotides.** All oligonucleotides were synthesized using an Applied Biosystems Model 391 DNA synthesizer on a 1 μmol scale. The modified base was typically dissolved to 0.1 M concentration in anhydrous acetonitrile. The modified base was added to the reaction column by either utilizing the optional port for the modified phosphoramidite or manually added by attaching two 1 mL syringes containing the modified base and tetrazole to the DNA synthesis column. The coupling time for the addition of any non-commercial modified base was always extended to 15 min. The coupling yields for the modified phosphoramidites were determined by trityl release. Oligonucleotides containing the modified base 7-deaza-dG (Glen Research) were oxidized by using a 0.5 M solution of (1S)-(+)-(10-camphorsulfonyl)oxaziridine in anhydrous acetonitrile to prevent degradation of the 7-deazaguanine residues. The non-aqueous oxidation step was extended to 5 min to ensure complete oxidation. Oligonucleotides were deprotected by using "standard conditions", NH<sub>4</sub>OH for 18 hr at 55°C, and then lyophilized. Otherwise, oligonucleotides were deprotected with a 1:9

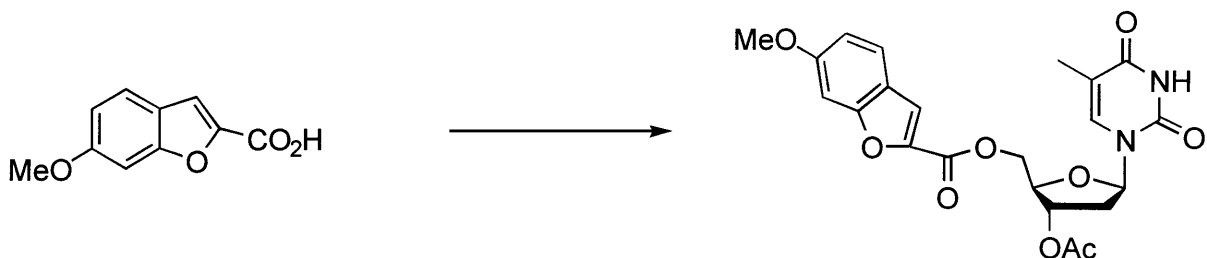
(v/v) DBU:ethanol solution containing 10 mg of cetyltrimethyl ammonium bromide in a sealed Eppendorf tube. After 24 h at room temperature, this reaction was quenched with 100  $\mu\text{L}$  of a 1:1 acetic acid:water solution. The neutralized solution was immediately placed onto a  $\text{Na}^+$  Dowex resin column and eluted with water to remove the cetyltrimethyl ammonium and DBU salts. The fractions containing the oligonucleotide were collected and lyophilized. Oligonucleotides were purified by HPLC and desalted on a Sep-Pak C18 cartridge eluting with 1:1 acetonitrile:water.

**Enzymatic Digestion Conditions.** To a solution of 10  $\mu\text{g}$  of oligonucleotide in 10  $\mu\text{L}$  of water were added 1  $\mu\text{L}$  of 1M NaOAc at pH 5.0, 1  $\mu\text{L}$  of 20 mM  $\text{ZnCl}_2$ , and 1  $\mu\text{L}$  (1.7 U/ $\mu\text{L}$ ) of nuclease P1 (Sigma). After 1 h at 37  $^\circ\text{C}$ , the reaction was cooled on ice and 1.3  $\mu\text{L}$  of 1 M Tris $\cdot\text{HCl}$  (pH 9), 1.0  $\mu\text{L}$  of snake venom phosphodiesterase (ICN, 0.05 U/ $\mu\text{L}$ ) and 1  $\mu\text{L}$  of alkaline phosphatase (Sigma, 25 U/ $\mu\text{L}$ ) were added. After 90 min at 37 $^\circ\text{C}$ , the entire reaction mixture was analyzed by C18 reversed phase HPLC using a photodiode array detector. Solvent A: 0.1 M  $\text{NH}_4\text{OAc}$  in water; solvent B: 0.1 M  $\text{NH}_4\text{OAc}$  in 1:1 water: $\text{CH}_3\text{CN}$ ; gradient 0-8.2% B over 35 min, 8.2-30% B over 25 min. The nucleoside ratios were quantitated by integration of the peak areas at 260 nm.

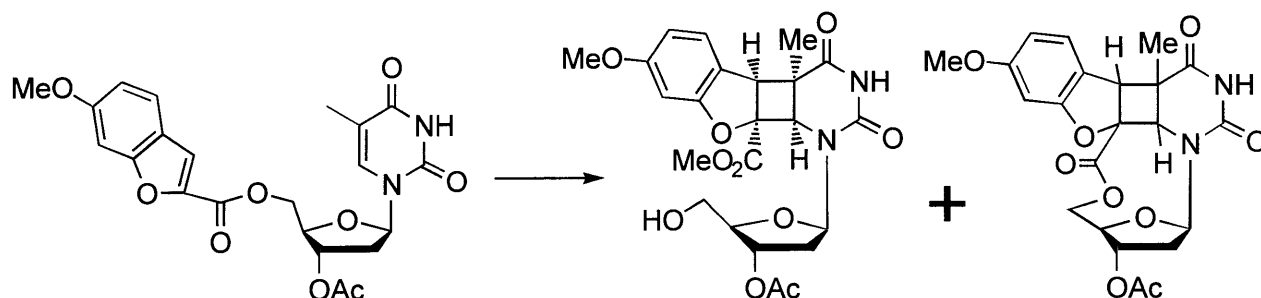




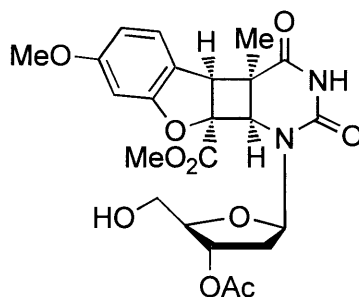
**6-Methoxy-2-benzofuran carboxylic acid (1.21).** To a solution of 5.74 g (37.7 mmol) of 4-methoxysalicylaldehyde **1.20** in 400 mL of dry 2-butanone were added 9.70 mL (56.6 mmol) of diethyl bromomalonate and 15.6 g (113 mmol) of oven dried K<sub>2</sub>CO<sub>3</sub>. The solution was stirred vigorously and heated to reflux for 7 hr, cooled and filtered. The solvents were removed and the solids diluted in 100 mL of methanol and 200 mL of 1N NaOH solution. The solution was then heated to reflux for 1 hr. After cooling to room temperature, the reaction mixture was extracted once with ethyl acetate and then acidified to pH 1. The tan precipitate that formed was collected, rinsed with 1N HCl, and dried to afford 5.53 g (76%) for two steps: <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 3.82 (s, 3 H), 6.96 (dd, 1 H, *J* = 2.0, 8.3), 7.28 (d, 1 H, *J* = 1.5), 7.58 (s, 1 H), 7.63 (d, 1 H, *J* = 8.3); <sup>13</sup>C NMR δ 55.62, 95.81, 101.70, 113.60, 119.89, 123.18, 145.24, 156.37, 159.86, 159.95; HRMS (EI) calcd for C<sub>10</sub>H<sub>8</sub>O<sub>4</sub> (M)<sup>+</sup> 192.0423, found 192.0424.



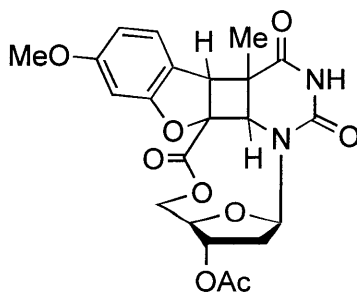
**5'-O-(6-Methoxy-2-benzofuranyl)-3'-O-acetylthymidine (1.22).** To a solution of 0.38 g (1.52 mmol) of compound **1.21** in 5 mL of dry pyridine was added 0.46 g (1.60 mmol) of 3'-O-acetylthymidine **1.18**, 0.35 g (1.82 mmol) of 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) and 20 mg (0.15 mmol) of DMAP. The solution was stirred at room temperature under an argon atmosphere for 16 hr. The reaction mixture was concentrated, dissolved in  $\text{CH}_2\text{Cl}_2$ , washed with 1 N HCl and saturated  $\text{NaHCO}_3$  and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with 100:1 dichloromethane/methanol to afford 0.54 g (75%) of a white foam:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.63 (app s, 3 H), 2.12 (s, 3 H), 2.38 (ddd, 1 H,  $J = 6.2, 8.8, 14$ ), 2.50 (app dd, 1 H,  $J = 5.6, 14$ ), 3.84 (s, 3 H), 4.36 (app d, 1 H,  $J = 1.6$ ), 4.59 (dd, 1 H,  $J = 2.5, 12$ ), 4.66 (dd, 1 H,  $J = 2.9, 12$ ), 5.35 (app d, 1 H,  $J = 6.2$ ), 6.44 (dd, 1 H,  $J = 5.6, 8.8$ ), 6.92-6.96 (m, 2 H), 7.26-7.54 (m, 3 H), 9.54 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  12.18, 20.78, 37.52, 55.71, 64.55, 74.89, 82.35, 84.89, 95.48, 111.52, 114.52, 115.44, 119.98, 123.21, 134.76, 143.80, 150.49, 157.21, 158.61, 161.11, 163.53, 170.35;  $\text{UV}_{\text{max}}$  ( $\text{CH}_3\text{CN}$ ) 312, 268, 246 nm; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{22}\text{H}_{22}\text{O}_9\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  459.1404, found 459.1404.



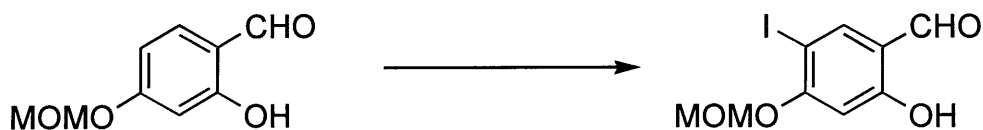
**5'-O-(6-Methoxy-2-benzofuranyl)-3'-O-acetylthymidine photoproduct.** A solution of 75 mg (0.16 mmol) of photoprecursor **1.22** in 85 mL of dry  $\text{CH}_3\text{CN}$  was deaerated with argon bubbling for 30 min. Acetone (4.25 mL) was added and the solution was irradiated with 300 nm light in a sixteen bulb Rayonet photoreactor for 5 hr at room temperature. The solvent was removed *in vacuo* and the crude material was dissolved in 25 mL of methanol with 150 mg of silica gel and stirred at room temperature for 30 min to effect transesterification. The methanol was removed and the absorbed material was directly added to the top of a silica gel column. Elution with a gradient of 100:1 to 50:1 dichloromethane/methanol afforded 45 mg (57%) of a major isomer **1.24** and 10 mg (14%) of a minor isomer **1.23b**.



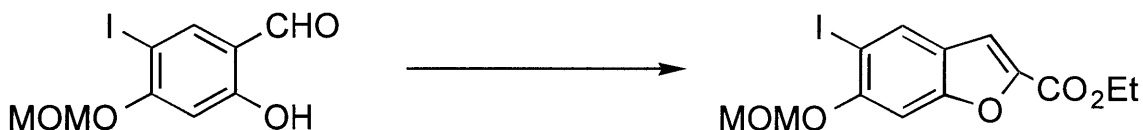
**Major isomer *cis-syn* (1.24):**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  1.78 (s, 3 H), 2.07-2.12 (m, 1 H), 2.08 (s, 3 H), 2.37 (ddd, 1 H,  $J = 2.0, 5.4, 14$ ), 3.76 (s, 3 H), 3.79-3.83 (m, 2 H), 3.86-3.88 (m, 1 H), 3.87 (s, 3 H), 3.95 (d, 1 H,  $J = 2.0$ ), 4.73 (d, 1 H,  $J = 2.0$ ), 5.19-5.22 (m, 1 H), 5.95 (dd, 1 H,  $J = 5.4, 9.3$ ), 6.46-6.48 (m, 2 H), 6.99 (d, 1 H,  $J = 7.8$ ), 7.06 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  21.00, 22.71, 36.12, 46.34, 53.44, 55.52, 56.05, 57.85, 62.58, 74.48, 84.25, 85.42, 86.72, 96.49, 108.75, 115.30, 126.81, 150.69, 161.70, 162.33, 169.67, 170.65, 170.69;  $\text{UV}_{\text{max}}$  ( $\text{CH}_3\text{CN}$ ) 284 nm; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{23}\text{H}_{26}\text{O}_{10}\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  491.1666, found 491.1664.



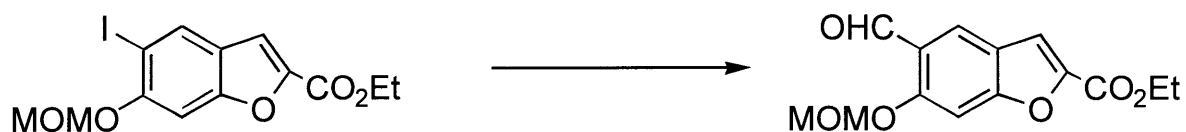
**Minor isomer *trans-syn* (1.23b):**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  1.28 (s, 3 H), 2.10 (s, 3 H), 2.24 (app d, 1 H,  $J = 7.3, 15$ ), 2.84 (ddd, 1 H,  $J = 5.9, 8.8, 15$ ), 3.80 (s, 3 H), 3.89 (s, 3 H), 4.18 (app d, 1 H,  $J = 12$ ), 4.25 (app d, 1 H,  $J = 2.0$ ), 4.66 (s, 1 H), 5.19 (dd, 1 H,  $J = 2.0, 12$ ), 5.48 (app d, 1 H,  $J = 5.9$ ), 6.35 (app t, 1 H,  $J = 8.8$ ), 6.55-6.58 (m, 2H), 7.15 (d, 1 H,  $J = 8.3$ ), 7.50 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  20.98, 22.49, 32.25, 48.00, 55.61, 65.59, 67.57, 77.20, 77.85, 82.60, 87.86, 91.55, 97.13, 108.35, 114.93, 126.97, 151.24, 161.12, 161.73, 165.55, 170.58, 171.77; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{22}\text{H}_{22}\text{O}_9\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  459.1404 found 459.1398.



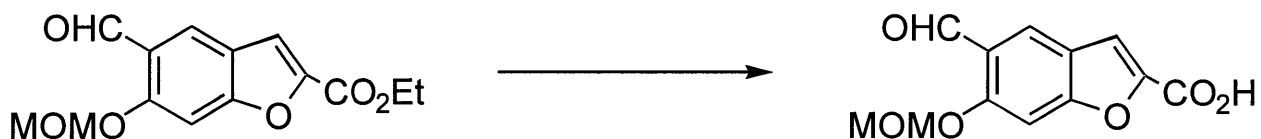
**2-Hydroxy-5-iodo-4-(methoxymethyl)benzaldehyde (1.26).** Chloramine T hydrate (4.54 g, 20.6 mmol) in 15 mL of DMF was added over a 10 min period to a solution of 3.13 g (17.2 mmol) **1.25** and 3.09 g (20.6 mmol) of sodium iodide in 50 mL of DMF. The solution was stirred for an additional 30 min, diluted with ethyl acetate, washed with 1 N HCl, and 5% sodium thiosulfate. The organic layer was then extracted twice with 1N NaOH and the aqueous extractions separated. The aqueous layers were acidified with 1N HCl in the presence of ethyl acetate. The organic layer was dried over MgSO<sub>4</sub>. Purification of the crude product by silica gel chromatography eluting with 15:1 hexanes/ethyl acetate afforded 3.92 g (74%) of a white flaky solid: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 3.34 (s, 3 H), 5.12 (s, 2 H), 6.48 (s, 1 H), 7.72 (s, 1H), 9.52 (s, 1 H), 11.13 (s, 1 H); <sup>13</sup>C NMR δ 56.74, 74.47, 94.81, 102.55, 117.70, 143.81, 162.15, 164.08, 193.6; HRMS (EI) calcd for C<sub>9</sub>H<sub>9</sub>O<sub>4</sub>I (M)<sup>+</sup> 307.9546, found 307.9545; Anal. calcd for C<sub>9</sub>H<sub>9</sub>O<sub>4</sub>I: C, 35.09; H, 2.94; I, 41.19. Found: C, 35.24; H, 3.04; I, 41.52; mp 77-79° C.



**2-Carboethoxy-5-iodo-6-(methoxymethyl)benzofuran (1.27).** To a solution of 3.89 g (12.6 mmol) of compound **1.26** in 200 mL of dry 2-butanone was added 3.25 mL (18.9 mmol) of diethyl bromomalonate and 5.25 g (37.8 mmol) of oven dried  $K_2CO_3$ . The solution was stirred vigorously and heated to reflux for 24 hr, cooled and filtered. Removal of the solvents *in vacuo* and purification by silica gel chromatography eluting with 10:1 hexanes/ethyl acetate afforded 2.76 g (58%) of a fluffy white solid:  $^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  1.39 (t, 3 H,  $J = 7.1$ ), 3.49 (s, 3 H), 4.39 (q, 2 H,  $J = 7.1$ ), 5.26 (s, 2 H), 7.31 (s, 1 H), 7.36 (s, 1 H), 8.05 (s, 1 H);  $^{13}C$  NMR  $\delta$  14.23, 56.39, 61.39, 82.87, 95.30, 98.59, 112.47, 123.26, 132.27, 145.81, 155.31, 156.51, 159.09; HRMS (EI) calcd for  $C_{13}H_{13}O_5I$  ( $M$ ) $^+$  375.9808, found 375.9807; Anal. calcd for  $C_{13}H_{13}O_5I$ : C, 41.51; H, 3.48; I, 33.74. Found: C, 41.42; H, 3.48; I, 33.93; mp 90-91 ° C.

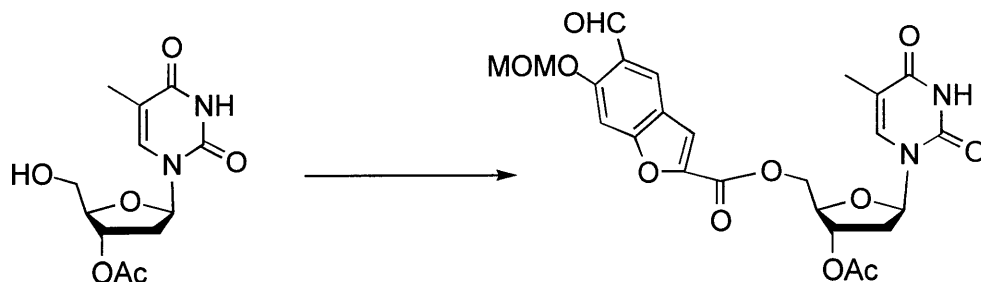


**2-Carboethoxy-5-formyl-6-(methoxymethyl)benzofuran (1.28).** A solution of 1.00 g (2.66 mmol) of compound **1.27** and 231 mg (0.20 mmol) of Pd(Ph<sub>3</sub>P)<sub>4</sub> in 10 mL of dry THF was heated to 50 °C under an CO atmosphere (1 atm). A solution of 0.79 mL (2.93 mmol) of tributyltinhydride in 20 mL of dry THF was added to the solution over a 6 hr period. The solution was cooled to room temperature and the solvents were removed *in vacuo*. The crude material was purified by a short silica gel column eluting with chloroform to separate the tin impurities from the product. Purification of the tin free product eluting with a gradient of 10:1 hexanes/ethyl acetate to 5:1 hexanes/ethyl acetate afforded 0.53 g (72%) of yellow-white solid: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 1.35 (t, 3 H, *J* = 7.2), 3.48 (s, 3 H), 4.36 (q, 2 H, *J* = 7.2), 5.30 (s, 2 H), 7.33 (s, 1 H), 7.42 (s, 1 H), 8.09 (s, 1 H), 10.43 (s, 1 H); <sup>13</sup>C NMR δ 14.10, 56.39, 61.43, 95.04, 98.36, 113.99, 121.14, 123.38, 123.42, 146.62, 158.73, 159.47, 159.49, 188.76; HRMS (EI) calcd for C<sub>14</sub>H<sub>14</sub>O<sub>6</sub> (M)<sup>+</sup> 278.0790, found 278.0792; Anal. calcd for C<sub>14</sub>H<sub>14</sub>O<sub>6</sub>: C, 60.43; H, 5.07. Found: C, 60.34; H, 5.07; mp 70-71 ° C.

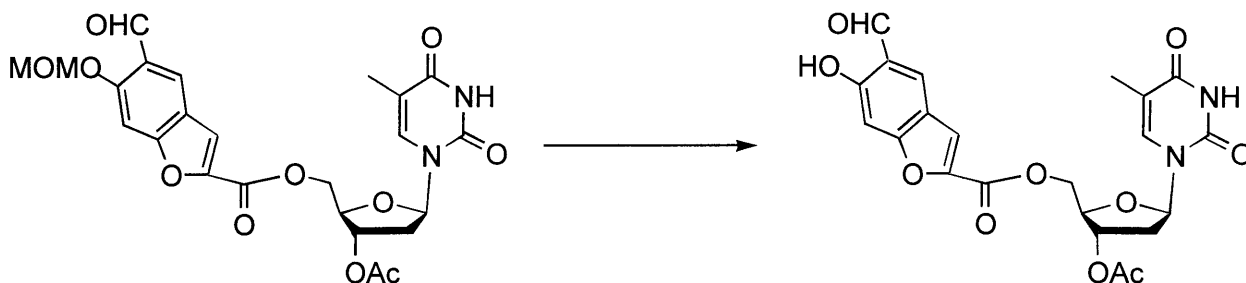


**5-Formyl-6-methoxymethyl-2-benzofuran carboxylic acid (1.29).** A solution of 0.49 g (1.77 mmol) of compound **1.28** and 7.5 mL of saturated aqueous K<sub>2</sub>CO<sub>3</sub> in 30 mL of ethanol was heated to reflux for 30 min. The solvents were removed, and the crude product was dissolved in water and the aqueous layer was washed twice with ethyl acetate. The aqueous layer was acidified to pH 3 with 0.1 M H<sub>2</sub>SO<sub>4</sub>. The resulting precipitate was filtered, washed with cold pH 3 water and dried overnight to afford 0.35 g (79%) of a white solid: <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 3.46 (s, 3 H), 5.43 (s, 2H), 7.52 (s, 1 H), 7.67 (s, 1 H), 8.13 (s, 1 H), 10.39 (s, 1 H); <sup>13</sup>C NMR δ 56.31, 94.91, 98.64, 114.11, 121.14, 123.17, 123.44, 147.23, 158.84, 159.02, 159.62, 188.74; HRMS (FAB<sup>+</sup>, glycerol) calcd for C<sub>12</sub>H<sub>10</sub>O<sub>6</sub> (M+H)<sup>+</sup> 251.0556, found 251.0552.

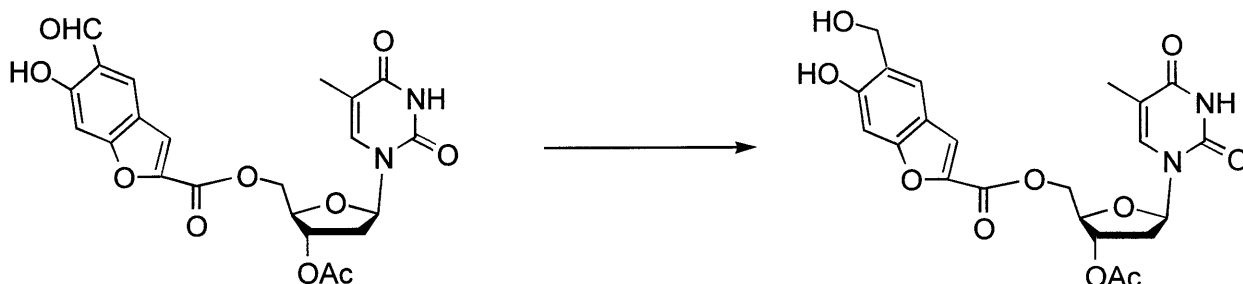




**5'-O-[5-Formyl-6-(methoxymethyl)-2-benzofuranyl]-3'-O-acetylthymidine (1.30).** To a solution of 0.30 g (1.18 mmol) of compound **1.29** in 4 mL of dry pyridine were added 0.36 g (1.25 mmol) of 3'-O-acetylthymidine **1.18**, 0.29 g (1.50 mmol) of 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) and a catalytic amount of DMAP. The solution was stirred at room temperature under an argon atmosphere for 24 hr. The solvent was concentrated, dissolved in CH<sub>2</sub>Cl<sub>2</sub>, washed with 1 N HCl and saturated NaHCO<sub>3</sub> and then dried over MgSO<sub>4</sub>. The crude product was purified by silica gel chromatography eluting with 100:1 dichloromethane/methanol to afford 0.50 g (90%) of a white foam: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 1.67 (d, 3 H, *J* = 1.2), 2.12 (s, 3 H), 2.37 (ddd, 1 H, *J* = 6.4, 9.0, 14), 2.51 (ddd, 1 H, *J* = 1.3, 5.5, 14), 3.52 (s, 3 H), 4.34-4.37 (m, 1 H), 4.59 (dd, 1 H, *J* = 2.7, 12), 4.67 (dd, *J* = 3.1, 12), 5.32-5.37 (m, 1 H), 5.34 (s, 2 H), 6.42 (dd, 1 H, *J* = 5.5, 9.0), 7.32 (s, 1 H), 7.44 (d, 1 H, *J* = 1.2), 7.62 (s, 1 H), 8.19 (s, 1 H), 9.46 (br. s, 1 H), 10.48 (s, 1 H); <sup>13</sup>C NMR δ 12.31, 20.83, 37.43, 56.58, 64.91, 74.70, 82.13, 84.87, 95.14, 98.31, 111.51, 115.75, 120.98, 123.85, 123.94, 134.71, 145.64, 150.45, 158.13, 159.59, 160.00, 163.52, 170.43, 188.76; HRMS (FAB<sup>+</sup>, 3-NBA) calcd for C<sub>24</sub>H<sub>24</sub>O<sub>11</sub>N<sub>2</sub>(M+H)<sup>+</sup> 517.1458, found 517.1458.

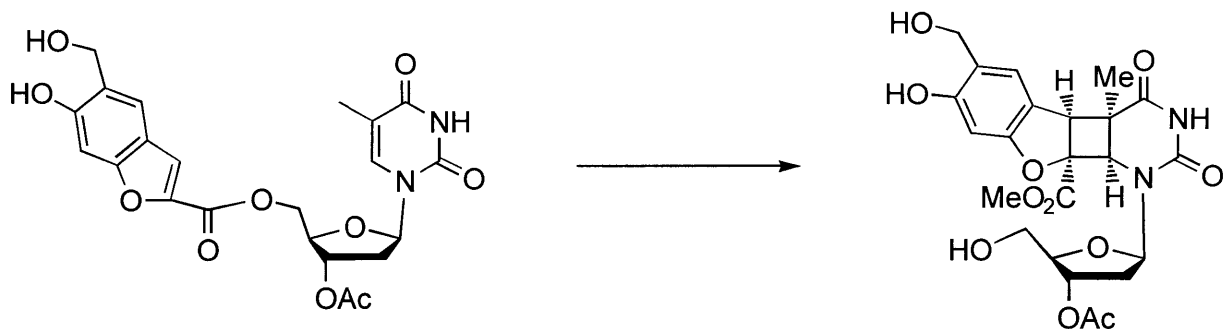


**5'-O-(5-Formyl-6-hydroxy-2-benzofuranyl)-3'-O-acetylthymidine (1.31).** To a solution of 0.44 (0.86 mmol) of compound **1.30** in 20 mL of dry  $\text{CH}_2\text{Cl}_2$  was added 0.32 g (0.95 mmol) of trityl tetrafluoroborate. The solution was stirred at room temperature under an argon atmosphere for 30 min. Completion of the reaction was monitored by the product's green fluorescence under long wave UV light (366 nm). The solution was diluted with  $\text{CH}_2\text{Cl}_2$ , washed with water and dried over  $\text{Na}_2\text{SO}_4$ . Purification of the crude product by silica gel chromatography eluting with a gradient of 100:1 to 50:1 dichloromethane/methanol afforded 0.50 g (90%) of a white foam:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.68 (d, 3 H,  $J = 1.0$ ), 2.12 (s, 3 H), 2.37 (ddd, 1 H,  $J = 6.4, 8.8, 14$ ), 2.53 (ddd, 1 H,  $J = 1.3, 5.5, 14$ ), 4.35-4.38 (m, 1 H), 4.60 (dd, 1 H,  $J = 3.0, 12$ ), 4.67 (dd, 1 H,  $J = 3.4, 12$ ), 5.34 (app d, 1 H,  $J = 6.4$ ), 6.39 (dd, 1 H,  $J = 5.5, 8.8$ ), 7.02 (s, 1 H), 7.39 (d, 1 H,  $J = 1.0$ ), 7.61 (s, 1 H), 7.93 (s, 1 H), 9.58 (br s, 1 H), 9.95 (s, 1 H), 11.23 (s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  12.33, 20.83, 37.39, 64.93, 74.63, 82.16, 85.04, 99.88, 111.45, 115.20, 119.51, 120.08, 130.23, 134.66, 145.80, 150.45, 158.08, 160.13, 162.20, 163.60, 170.45, 195.79; HRMS (FAB<sup>+</sup>, 3-NBA) calcd for  $\text{C}_{22}\text{H}_{20}\text{O}_{10}\text{N}_2$  (M+H)<sup>+</sup> 473.1196, found 473.1202.

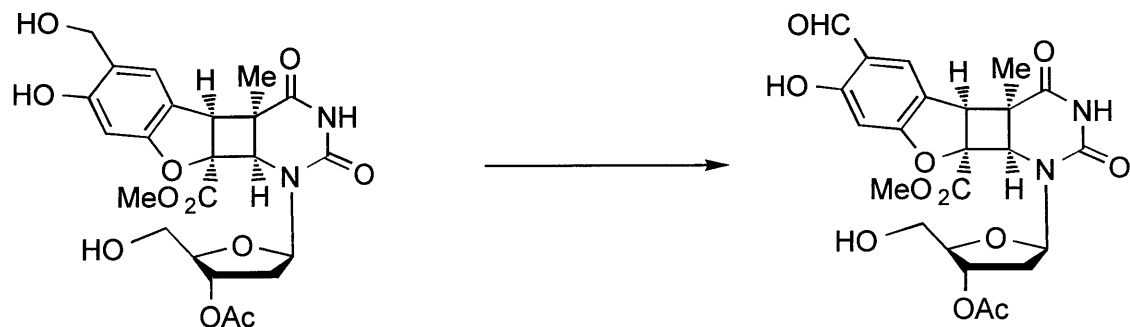


**5'-O-[6-Hydroxy-5-(hydroxymethyl)-2-benzofuranyl]-3'-O-acetylthymidine (1.32).**

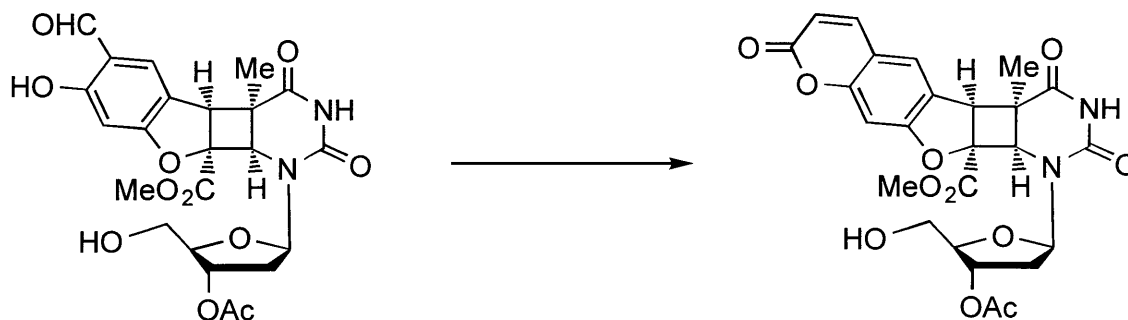
To a solution of 0.32 g (0.69 mmol) of compound **1.31** in 8 mL of a 1:1 mixture of ethanol/dioxane was added 10 mg (0.26 mmol) of  $\text{NaBH}_4$ . The solution was stirred at room temperature for 15 min, neutralized with saturated  $\text{NH}_4\text{Cl}$ , diluted with  $\text{CH}_2\text{Cl}_2$ , washed with water, and dried over  $\text{MgSO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 20:1 to 10:1 dichloromethane/methanol to afford 0.29 g (88%) of a white solid:  $^1\text{H NMR}$  (300 MHz, acetone- $d_6$ )  $\delta$  1.63 (d, 3 H,  $J = 1.8$ ), 2.10 (s, 3 H), 2.47 (ddd, 1 H,  $J = 6.4, 8.7, 14$ ), 2.55 (ddd, 1 H,  $J = 2.0, 5.9, 14$ ), 2.87 (br. s, 1 H), 4.40-4.42 (m, 1 H), 4.59 (dd, 1 H,  $J = 3.7, 12$ ), 4.67 (dd, 1 H,  $J = 4.0, 12$ ), 4.79 (d, 2 H,  $J = 2.0$ ), 5.42-5.45 (m, 1 H), 6.38 (dd, 1 H,  $J = 5.9, 8.7$ ), 7.03 (s, 1 H), 7.63 (d, 1 H,  $J = 2.6$ ), 7.71 (app s, 2 H);  $^{13}\text{C NMR}$  (DMSO- $d_6$ )  $\delta$  11.97, 20.74, 35.88, 58.37, 64.33, 74.15, 81.14, 84.03, 96.62, 109.97, 115.65, 118.32, 120.62, 128.04, 135.50, 142.75, 150.39, 155.53, 155.91, 158.36, 163.55, 170.02;  $\text{UV}_{\text{max}}$  ( $\text{CH}_3\text{CN}$ ) 314, 270, 249 nm; HRMS (FAB $^+$ , glycerol) calcd for  $\text{C}_{22}\text{H}_{22}\text{O}_{10}\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  475.1353, found 475.1355; Anal. calcd for  $\text{C}_{22}\text{H}_{22}\text{O}_{10}\text{N}_2$ : C, 55.70; H, 4.67; N, 5.91. Found: C, 55.48; H, 4.76; N, 5.63.



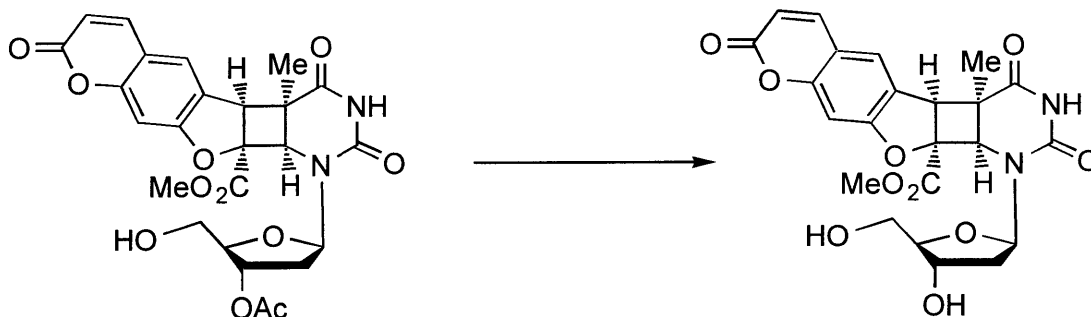
**2-Carbomethoxy-6-hydroxy-(5-methylalcohol)benzofuran 3'-O-acetylthymidine *cis-syn* photoproduct (1.33).** A solution of 0.25 g (0.52 mmol) of compound **1.32** in 350 ml of dry  $\text{CH}_3\text{CN}$  was deaerated with argon bubbling for 60 min. Acetone (17.5 mL) was added and the solution was irradiated with 300 nm light in a sixteen bulb Rayonet photoreactor for 3 hr at room temperature. Removal of the solvents and absorption of the crude material onto silica gel using methanol as the solvent effected the ring opening. The absorbed compound was loaded on top of a silica gel column and purified by eluting with a gradient of 30:1 to 15:1 dichloromethane/methanol to afford 0.15 g (58%) of a clear foam:  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.70 (s, 3 H), 2.00-2.11 (m, 2 H), 2.06 (s, 3 H), 3.69-3.71 (m, 2 H), 3.85 (s, 3 H), 3.88-3.90 (m, 1 H), 4.00 (d, 1 H,  $J = 1.0$ ), 4.04 (t, 1 H,  $J = 5.4$ , (-OH)), 4.55 (d, 1 H,  $J = 13$ ), 4.63 (d, 1 H,  $J = 13$ ), 4.84 (d, 1 H,  $J = 2.0$ ), 5.17-5.19 (m, 1 H), 6.18 (dd, 1 H,  $J = 5.4, 9.3$ ), 6.37 (s, 1 H), 6.94 (s, 1 H), 8.71 (br s, 1 H), 8.93 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  21.07, 22.95, 35.80, 47.35, 53.43, 56.76, 57.56, 61.68, 63.06, 75.56, 84.42, 84.84, 88.11, 98.32, 116.04, 122.28, 126.57, 152.55, 157.85, 162.83, 170.87, 170.93;  $\text{UV}_{\text{max}}$  ( $\text{CH}_3\text{CN}$ ) 287 nm; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{23}\text{H}_{26}\text{O}_{11}\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  507.1615, found 507.1611.



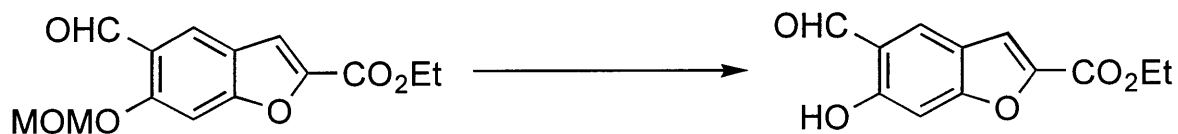
**(2-Carbomethoxy-5-formyl-6-hydroxybenzofuran) 3'-O-acetylthymidine *cis*-syn photoproduct (1.34).** To a solution of 0.16 g (0.31 mmol) of compound **1.33** in 2 mL of DMF were added 13 mg (0.08 mmol) of TEMPO and 8 mg (0.08 mmol) of CuCl. Oxygen was bubbled through the solution for 2 hr after which another 13 mg of TEMPO and 8 mg of CuCl were added. After stirring under a oxygen atmosphere for 6 hr, the reaction was diluted with ethyl acetate and washed with 1 N HCl, saturated NaHCO<sub>3</sub> and brine and then dried over MgSO<sub>4</sub>. Purification of the crude product by silica gel chromatography eluting with 50:1 dichloromethane/methanol afforded 0.11 g (70%) of a white foam: <sup>1</sup>H NMR (300 MHz, acetone-*d*<sub>6</sub>) δ 1.75 (s, 3 H), 1.96-2.11 (m, 2 H), 2.06 (s, 3 H), 3.70-3.73 (m, 2 H), 3.88 (s, 3 H), 3.89-3.91 (m, 1 H), 4.16 (app s, 1 H), 4.95 (d, 1 H, *J* = 1.8), 5.18-5.21 (m, 1 H), 6.19 (dd, 1 H, *J* = 5.6, 9.3), 6.49 (s, 1 H), 7.47 (d, 1 H, *J* = 0.7), 9.80 (d, 1 H, *J* = 0.7); <sup>13</sup>C NMR δ 21.07, 22.78, 35.84, 47.50, 53.70, 55.04, 57.55, 63.04, 75.59, 84.53, 84.77, 89.53, 98.98, 117.57, 119.14, 133.49, 152.30, 166.11, 169.34, 169.73, 170.59, 170.87, 196.10; HRMS (FAB<sup>+</sup>, 3-NBA) calcd for C<sub>23</sub>H<sub>24</sub>O<sub>11</sub>N<sub>2</sub> (M+H)<sup>+</sup> 505.1458, found 505.1455.



**(2-Carbomethoxypsoralen) 3'-O-acetylthymidine *cis-syn* photoproduct (1.35).** To a solution of 43 mg (0.09 mmol) of compound **1.34** in 0.4 mL of dry THF were added 4 Å sieves and 125  $\mu$ L (0.85 mmol) of N,N-dimethylacetamide dimethylacetal. The solution stirred for 2 hr at room temperature under an argon atmosphere. The reaction was diluted with ethyl acetate, washed with saturated  $\text{NH}_4\text{Cl}$ , saturated  $\text{NaHCO}_3$  and brine and then dried over  $\text{MgSO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 50:1 to 33:1 dichloromethane/methanol to afford 17 mg (38%) of a white foam:  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.76 (s, 3 H), 2.01-2.05 (m, 1 H), 2.07 (s, 3 H), 2.13 (ddd, 1 H,  $J = 2.0, 5.9, 14$ ), 3.71-3.73 (m, 2 H), 3.88 (s, 3 H), 3.90-3.92 (m, 1 H), 4.11 (t, 1 H,  $J = 5.4$ , (-OH)), 4.24 (s, 1 H), 4.97 (d, 1 H,  $J = 2.0$ ), 5.20-5.22 (m, 1 H), 6.20 (dd, 1 H,  $J = 5.9, 9.8$ ), 6.24 (d, 1 H,  $J = 9.3$ ), 6.90 (s, 1 H), 7.39 (s, 1 H), 7.90 (d, 1 H,  $J = 9.3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  20.10, 22.85, 35.80, 46.60, 53.68, 54.97, 57.28, 62.53, 74.27, 84.10, 84.94, 87.77, 99.35, 113.81, 114.41, 121.48, 125.81, 143.23, 150.43, 156.32, 160.47, 163.98, 169.20, 169.70, 170.69; UV (MeOH) 324, 293; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{25}\text{H}_{24}\text{O}_{11}\text{N}_2$  ( $\text{M}+\text{H}$ ) $^+$  529.1458, found 529.1456.



**2-Carbomethoxypsoralen thymidine *cis*-syn photoproduct (1.15).** To a solution of 9 mg (17  $\mu$ mol) of compound **1.35** in 0.5 mL of freshly distilled methanol was added 50  $\mu$ L of DBU. After stirring for 10 min at room temperature the reaction was diluted with ethyl acetate, washed with saturated  $\text{NH}_4\text{Cl}$ , saturated  $\text{NaHCO}_3$  and brine and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with 25:1 dichloromethane/methanol to afford 8 mg (97%) of a white foam:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.80 (s, 3 H), 1.87 (br s, 2 H), 2.02-2.11 (m, 1 H), 2.27 (ddd, 1 H,  $J = 3.1, 5.9, 14$ ), 3.70 (dd, 1 H,  $J = 3.8, 12$ ), 3.80-3.89 (m, 2 H), 3.89 (s, 3 H), 4.02 (s, 1 H), 4.39-4.43 (m, 1 H), 4.97 (d, 1 H,  $J = 1.7$ ), 5.96 (dd, 1 H,  $J = 5.9, 8.1$ ), 6.27 (d, 1 H,  $J = 9.5$ ), 6.91 (s, 1 H), 7.24 (s, 1 H), 7.40 (br s, 1 H), 7.60 (d, 1 H,  $J = 9.5$ );  $^{13}\text{C}$  NMR (methanol- $d_4$ )  $\delta$  22.92, 38.64, 48.44, 54.01, 56.16, 57.97, 63.36, 72.41, 85.16, 87.35, 89.97, 99.67, 113.97, 115.85, 124.40, 127.84, 145.90, 153.72, 157.68, 162.91, 166.21, 170.73, 172.26; UV (1:1 methanol:water) 332, 294; HRMS (FAB<sup>+</sup>, 3-NBA) calcd for  $\text{C}_{23}\text{H}_{22}\text{O}_{10}\text{N}_2$  ( $\text{M}+\text{H}$ )<sup>+</sup> 487.1353, found 487.1349.

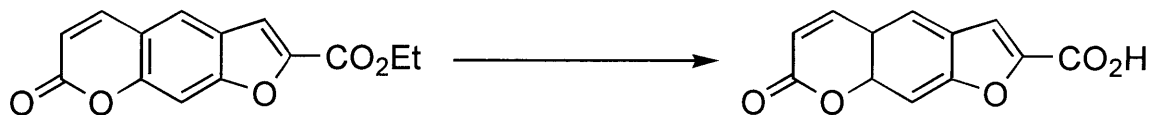


**2-Carboethoxy-5-formyl-6-hydroxybenzofuran (1.37).** A solution of 1.02 g (3.67 mmol) of compound **1.28**, 2 mL of concentrated HCl, and 100 mL of ethanol was heated at reflux for 1 h. After cooling to room temperature, the solution was neutralized with saturated NaCHO<sub>3</sub> and diluted with ethyl acetate. The layers were separated and the aqueous portion was washed twice with ethyl acetate. The combined organic washes were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. Removal of the solvents *in vacuo* and purification by silica gel chromatography eluting with a gradient of 10:1 to 2:1 hexanes/ethyl acetate afforded 0.78 g (91%) of a yellow-white solid. Recrystallization from hexanes/ethyl acetate afforded yellow plates: mp 158-159 °C; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 1.41 (t, 3 H, *J* = 7.2), 4.42 (q, 2 H, *J* = 7.2), 7.09 (s, 1 H), 7.49 (s, 1 H), 7.88 (s, 1 H), 9.95 (s, 1 H), 11.23 (s, 1 H); <sup>13</sup>C NMR δ 14.3, 61.7, 100.1, 113.7, 119.2, 120.5, 129.9, 147.0, 158.9, 160.2, 161.8, 195.9; HRMS (EI) calcd for C<sub>12</sub>H<sub>10</sub>O<sub>5</sub> (M)<sup>+</sup> 234.0528, found 234.0529. Anal. Calcd for C<sub>12</sub>H<sub>10</sub>O<sub>5</sub>: C, 61.54; H, 4.30. Found: C, 61.17; H, 4.11.

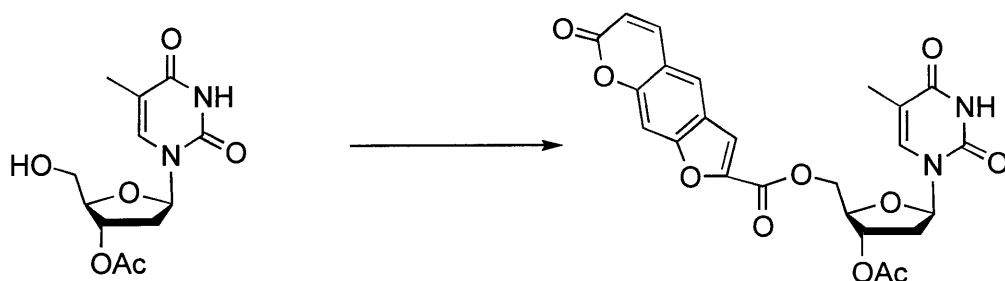




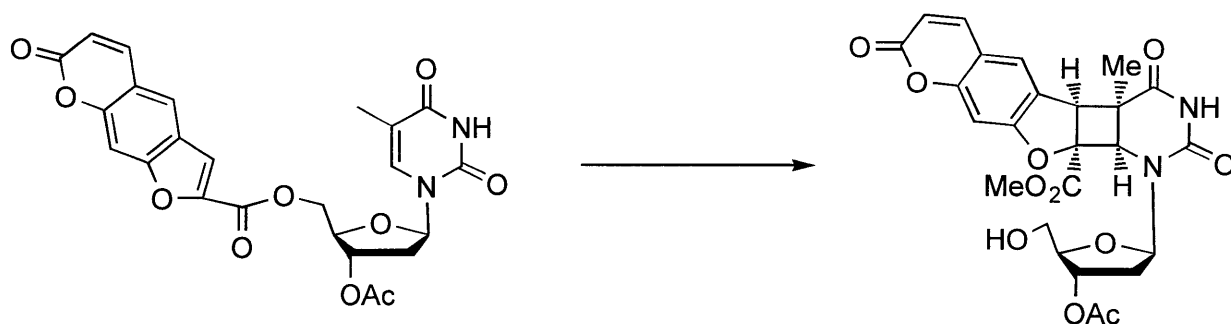
**2-Carboethoxypsoralen (1.38).** To a solution of 95 mg (0.41 mmol) of compound **1.37** in 4 mL of dry THF were added 4 Å sieves and 120  $\mu\text{L}$  (0.82 mmol) of *N,N*-dimethylacetamide dimethylacetal. The solution was stirred for 5 h at room temperature under an argon atmosphere. The reaction mixture was diluted with ethyl acetate, washed with saturated 1 N HCl, saturated  $\text{NaHCO}_3$  and brine and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 3:1 to 1:1 hexanes/ethyl acetate to afford 69 mg (65%) of a white solid. Recrystallization from hexanes/ethyl acetate afforded white needles: mp 221-222 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.44 (t, 3 H,  $J = 7.2$ ), 4.46 (q, 2 H,  $J = 7.2$ ), 6.42 (d, 1 H,  $J = 9.6$ ), 7.55 (s, 1 H), 7.57 (s, 1 H), 7.79-7.82 (m, 2 H);  $^{13}\text{C}$  NMR  $\delta$  14.3, 61.9, 100.6, 113.2, 115.5, 116.5, 121.8, 124.2, 143.5, 147.6, 153.7, 156.8, 158.9, 160.2; HRMS (EI) calcd for  $\text{C}_{14}\text{H}_{10}\text{O}_5$  ( $\text{M}$ ) $^+$  258.0528, found 258.0529. Anal. Calcd for  $\text{C}_{14}\text{H}_{10}\text{O}_5$ : C, 65.12; H, 3.90. Found: C, 65.05; H, 3.72.



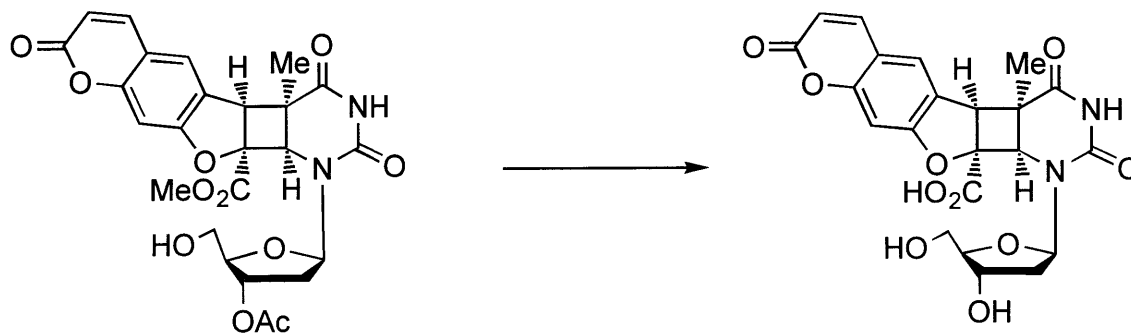
**2'-Carboxypsoralen (1.39).** A solution of 0.27 g (1.06 mmol) of compound **1.38** and 10 mL of aqueous 1 N NaOH in 10 mL of methanol was heated to reflux for 30 min. The solvents were removed, and the crude product was dissolved in water and acidified to pH 1 with 1 N HCl. The resulting precipitate was filtered, washed with cold 1 N HCl and ether and then dried overnight to afford 0.19 g (77%) of a yellowish-white solid: <sup>1</sup>H NMR (300 MHz, DMF-*d*<sub>7</sub>) δ 6.51 (d, 1 H, *J* = 9.6), 7.79 (s, 1 H), 7.83 (s, 1 H), 8.23 (s, 1 H), 8.26 (d, 1 H, *J* = 9.6); <sup>13</sup>C NMR δ 100.9, 114.4, 116.0, 117.8, 124.1, 125.7, 145.7, 149.7, 154.9, 157.8, 160.9, 160.9; HRMS (EI) calcd for C<sub>12</sub>H<sub>6</sub>O<sub>5</sub> (M)<sup>+</sup> 230.0215, found 230.0214.



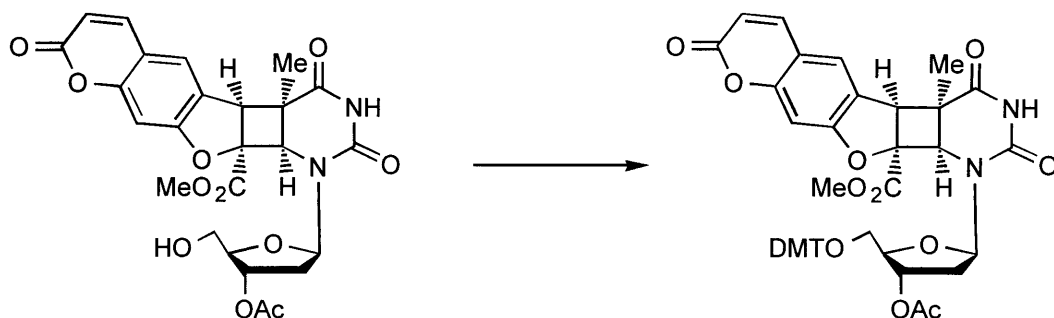
**5'-O-[2-psoralenyl]-3'-O-acetyl thymidine (1.40).** To a solution of 80 mg (0.35 mmol) of compound **1.39** in 1.5 mL of dry DMF and 1.5 mL of dry pyridine were added 150 mg (0.53 mmol) of 3'-O-acetylthymidine **1.18** and the suspension was sonicated for 5 min. To the milky white solution were added 170 mg (0.88 mmol) of 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) and 5 mg (0.35 mmol) of DMAP. The solution was stirred at room temperature under an argon atmosphere for 24 h. The solvent was concentrated, dissolved in  $\text{CH}_2\text{Cl}_2$ , washed with 1 N HCl and saturated  $\text{NaHCO}_3$  and then dried over  $\text{MgSO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 100:1 to 50:1 dichloromethane/methanol to afford 106 mg (61%) of a white foam: mp 264-266 °C (dec);  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.70 (d, 3 H,  $J = 1.0$ ), 2.14 (s, 3 H), 2.39 (ddd, 1 H,  $J = 6.6, 8.7, 14$ ), 2.55 (ddd, 1 H,  $J = 1.4, 5.5, 14$ ), 4.34-4.40 (m, 1 H), 4.64-4.71 (m, 2 H), 5.36 (app d, 1 H,  $J = 6.4$ ), 6.40 (dd, 1 H,  $J = 5.5, 8.7$ ), 6.43 (d, 1 H,  $J = 9.7$ ), 7.39 (d, 1 H,  $J = 1.2$ ), 7.48 (s, 1 H), 7.67 (s, 1 H), 7.80 (d, 1 H,  $J = 9.7$ ), 7.83 (s, 1 H), 9.01 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  12.4, 20.9, 37.4, 65.1, 74.6, 82.1, 85.1, 100.4, 111.5, 114.6, 115.9, 116.9, 122.2, 123.9, 134.7, 143.3, 146.5, 150.3, 154.2, 156.8, 158.1, 159.9, 163.3, 170.5;  $\text{UV}_{\text{max}}$  (MeOH) 334, 264 nm; HRMS (FAB<sup>+</sup>, NBA) calcd for  $\text{C}_{24}\text{H}_{20}\text{O}_{10}\text{N}_2$  (M+H)<sup>+</sup> 497.1196, found 497.1203.



**(2-Carbomethoxypsoralen)-3'-O-acetylthymidine *Cis-Syn* photoproduct (1.35).** A solution of 20 mg (0.04 mmol) of compound **1.40** in 40 mL of dry CH<sub>3</sub>CN was deaerated with argon bubbling for 30 min. Acetone (2 mL) was added and the solution was irradiated with 300 nm light in a sixteen bulb Rayonet photoreactor for 3 h at room temperature. Removal of the solvents and adsorption of the crude material onto silica gel using methanol as the solvent effected lactone ring opening. The absorbed compound was loaded on top of a silica gel column and purified by eluting with a gradient of 100:1 to 50:1 dichloromethane/methanol to afford 6 mg (24%) of a clear foam. *Attempts to perform this reaction on a scale greater than 50 mg afforded lower yields; however, multiple reaction vessels placed in a Rayonet photoreactor afforded yields in the 20-25% range:* <sup>1</sup>H NMR (300 MHz, acetone-*d*<sub>6</sub>) δ 1.76 (s, 3 H), 2.01-2.05 (m, 1 H), 2.07 (s, 3 H), 2.13 (ddd, 1 H, *J* = 2.0, 5.9, 14), 3.71-3.73 (m, 2 H), 3.88 (s, 3 H), 3.90-3.92 (m, 1 H), 4.11 (t, 1 H, *J* = 5.4, (-OH)), 4.24 (s, 1 H), 4.97 (d, 1 H, *J* = 2.0), 5.20-5.22 (m, 1 H), 6.20 (dd, 1 H, *J* = 5.9, 9.8), 6.24 (d, 1 H, *J* = 9.3), 6.90 (s, 1 H), 7.39 (s, 1 H), 7.90 (d, 1 H, *J* = 9.3); <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 20.1, 22.9, 35.8, 46.6, 53.7, 55.0, 57.3, 62.5, 74.3, 84.1, 84.9, 87.8, 99.3, 113.8, 114.4, 121.5, 125.8, 143.2, 150.4, 156.3, 160.5, 164.0, 169.2, 169.7, 170.7; UV (MeOH) 324, 293; HRMS (FAB<sup>+</sup>, 3-NBA) calcd for C<sub>25</sub>H<sub>24</sub>O<sub>11</sub>N<sub>2</sub> (M+H)<sup>+</sup> 529.1458, found 529.1456.

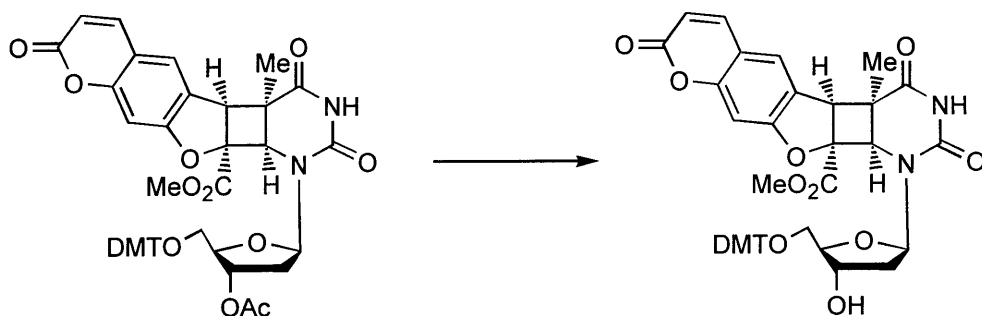


**2-Carboxypsoralen Thymidine *Cis-Syn* Photoproduct (2.1).** To a solution of 47 mg (89  $\mu\text{mol}$ ) of compound **1.35** in 2 mL of methanol was added 2 mL of 100 mM  $\text{Na}_2\text{CO}_3$  (pH 9.0) solution. After stirring for 1 h at room temperature, the reaction was neutralized with 1N HCl and adsorbed onto silica gel via lyophilization. The crude product was purified by silica gel chromatography eluting with 5:1 dichloromethane/methanol to afford 34 mg (81%) of a white solid:  $^1\text{H}$  NMR (500 MHz, methanol- $d_4$ )  $\delta$  1.76 (s, 3 H), 2.02-2.12 (m, 2 H), 3.62 (dd, 1 H,  $J = 5.9, 12$ ), 3.66 (dd, 1 H,  $J = 4.9, 12$ ), 3.77-3.79 (m, 1 H), 3.95 (s, 1 H), 4.22-4.25 (m, 1 H), 4.81 (d, 1 H,  $J = 1.5$ ), 6.21 (d, 1 H,  $J = 9.3$ ), 6.22-6.26 (m, 1 H), 6.85 (s, 1 H), 7.32 (s, 1 H), 7.84 (d, 1 H,  $J = 9.3$ );  $^{13}\text{C}$  NMR  $\delta$  22.80, 38.20, 48.40, 50.00, 56.30, 58.47, 63.78, 72.38, 85.33, 87.19, 99.48, 113.05, 115.31, 125.64, 127.54, 146.22, 153.95, 157.51, 163.34, 167.11, 173.16, 175.46; HRMS (FAB+, glycerol) calcd for  $\text{C}_{22}\text{H}_{20}\text{O}_{10}\text{N}_2$  ( $\text{M} + \text{H}$ ) $^+$  473.1196, found 473.1198.

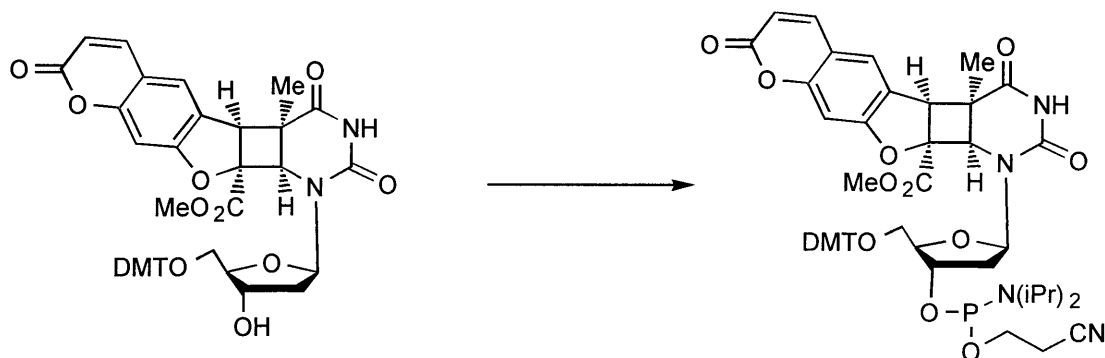


### 2-Carbomethoxypsoralen 5'-*O*-DMT-3'-*O*-acetylthymidine *Cis-Syn* Photoproduct

(2.2). To a solution of 42 mg (80  $\mu\text{mol}$ ) of compound **1.35** in 500  $\mu\text{L}$  of dry DMF were added 34 mg (100  $\mu\text{mol}$ ) of dimethoxytrityl chloride, 17 mg (100  $\mu\text{mol}$ ) of silver nitrate and 17  $\mu\text{L}$  (120  $\mu\text{mol}$ ) of collidine. The cloudy orange solution was stirred for 15 min at room temperature under an argon atmosphere. The reaction mixture was diluted with ethyl ether, washed with water, saturated  $\text{CuSO}_4$ , and brine, and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with 100:1 dichloromethane/methanol to afford 62 mg (93%) of a yellow foam:  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.69 (s, 3 H), 2.03-2.10 (m, 1 H), 2.05 (s, 3 H), 2.14-2.22 (m, 1 H), 3.24 (dd, 1 H,  $J = 4.4, 10$ ), 3.37 (dd, 1 H,  $J = 4.4, 10$ ), 3.78 (s, 3 H), 3.79 (s, 3 H), 3.80 (s, 3 H), 3.96 (app dd, 1 H,  $J = 4.3, 8.5$ ), 4.26 (s, 1 H), 4.79 (s, 1 H), 5.23-5.28 (m, 1 H), 6.21-6.26 (m, 2 H), 6.84 (d, 1 H,  $J = 8.7$ ), 6.90-6.95 (m, 4 H), 7.19 (d, 1 H,  $J = 8.8$ ), 7.23-7.42 (m, 7 H), 7.55 (d, 2 H,  $J = 7.7$ ), 7.90 (d, 1 H,  $J = 9.5$ ), 9.12, (s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  20.99, 23.17, 35.47, 47.80, 53.93, 55.57, 55.62, 55.66, 57.66, 64.09, 73.89, 81.94, 83.74, 87.11, 89.46, 99.25, 113.71, 114.01, 114.10, 115.13, 123.66, 127.47, 127.52, 128.81, 129.13, 130.15, 131.12, 136.89, 137.02, 144.77, 146.29, 152.25, 157.30, 159.78, 160.65, 165.70, 170.12, 170.33, 170.71; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{46}\text{H}_{42}\text{O}_{13}\text{N}_2$  ( $\text{M} + \text{H}$ ) $^+$  831.2765, found 831.2758.

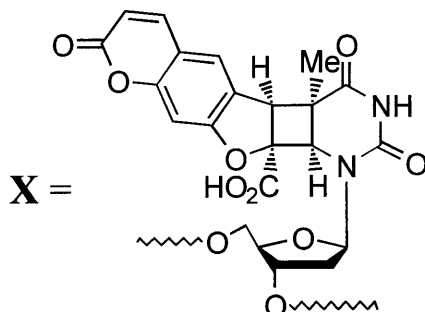


**2-Carbomethoxypsoralen 5'-*O*-DMT-thymidine *Cis*-Syn Photoproduct (2.3).** To a solution of 42 mg (50  $\mu\text{mol}$ ) of compound **2.2** in 2.5 mL of freshly distilled methanol was added 125  $\mu\text{L}$  of DBU under an argon atmosphere. After stirring for 10 min at room temperature the reaction was diluted with ethyl acetate, washed with saturated  $\text{NH}_4\text{Cl}$ , saturated  $\text{NaHCO}_3$ , and brine, and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with 50:1 dichloromethane/methanol to afford 34 mg (86%) of a yellow foam:  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.65 (s, 3 H), 1.84-1.93 (m, 1 H), 2.10 (ddd, 1 H,  $J = 4.7, 6.7, 14$ ), 3.20 (dd, 1 H,  $J = 5.0, 10$ ), 3.29 (dd, 1 H,  $J = 4.3, 10$ ), 3.73 (s, 3 H), 3.80 (s, 3 H), 3.80 (s, 3 H), 3.82-3.87 (m, 1 H), 4.24 (s, 1 H), 4.29-4.34 (m, 1 H), 4.47 (d, 1 H,  $J = 4.6$ , (-OH)), 4.75 (d, 1 H,  $J = 1.5$ ), 6.23-6.3 (m, 2 H), 6.90-6.96 (m, 5 H), 7.26 (d, 1 H,  $J = 7.3$ ), 7.33-7.43 (m, 7 H), 7.55 (d, 2 H,  $J = 7.2$ ), 7.89 (d, 1 H,  $J = 9.6$ ), 9.06, (s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  23.19, 38.21, 47.96, 53.69, 55.63, 57.57, 64.63, 71.40, 83.69, 84.97, 86.94, 89.73, 99.23, 113.96, 114.06, 115.09, 123.73, 127.43, 127.66, 128.73, 129.20, 131.14, 137.09, 144.74, 146.40, 152.16, 157.29, 159.77, 160.63, 165.71, 170.14, 170.38; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{44}\text{H}_{40}\text{O}_{12}\text{N}_2$  ( $\text{M} + \text{H}$ ) $^+$  789.2660, found 789.2663.



**2-Carbomethoxypsoralen 5'-O-DMT-3'-O-(isopropylamine-2-cyanoethyl phosphoramidite)-thymidine *Cis-Syn* Photoproduct (2.4).** To a solution of 25 mg (32  $\mu\text{mol}$ ) of compound **2.3** in 500  $\mu\text{L}$  of dry  $\text{CH}_2\text{Cl}_2$  were added 6 mg (32  $\mu\text{mol}$ ) of bisisopropylammonium tetrazolide and 20  $\mu\text{L}$  (64  $\mu\text{mol}$ ) of bis(isopropylamine)-2-cyanoethyl phosphoramidite. After stirring for 5 h at room temperature under an argon atmosphere the solvents were removed *in vacuo*. The crude product was purified by silica gel chromatography eluting with 1:1 hexanes/ethyl acetate. The purified product was dissolved in benzene and lyophilized to afford 23 mg (73%) of a white crunchy foam:  $^{31}\text{P}$  NMR  $\delta$  (121 MHz, benzene- $d_6$ ) 149.30, 149.65; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{53}\text{H}_{57}\text{O}_{13}\text{N}_4\text{P}$  ( $\text{M} + \text{H}$ ) $^+$  989.3738, found 989.3737.



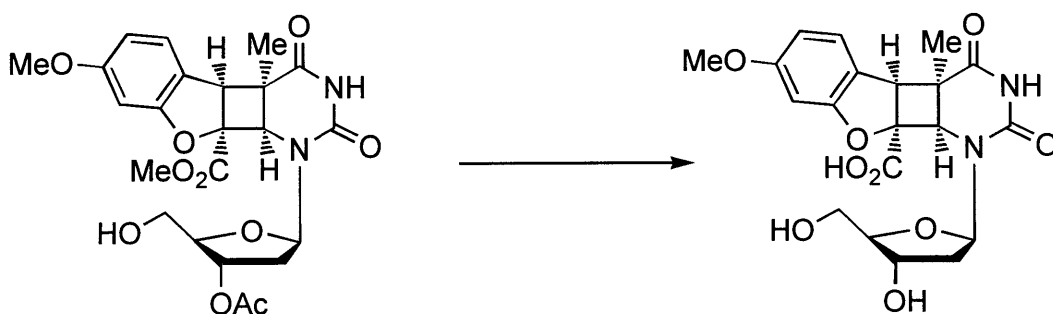


### 5'-AGCTAXAAAAGGT-3'

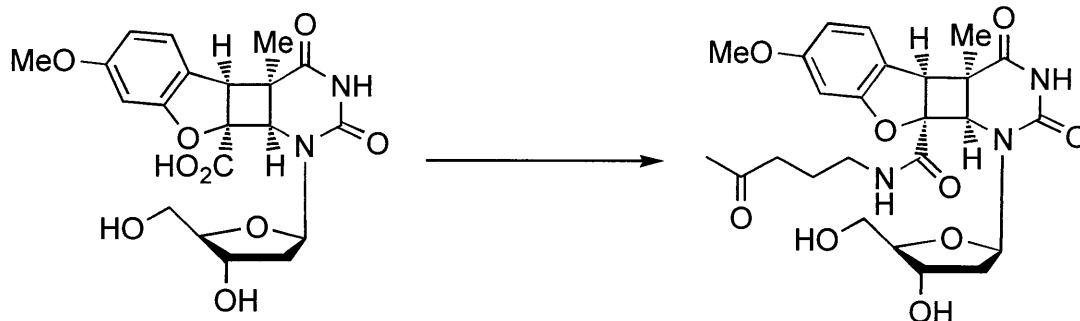
**Oligonucleotide (2.5).** A 0.1 M solution of the modified phosphoramidite **2.4** was added using the optional port. The coupling yield for the modified base was 85-92%. The oligonucleotide was deprotected using 10% DBU in ethanol and the 2'-carboethoxypsoralen was converted into the carboxylic acid by treatment with 1 mL of a 100 mM Na<sub>2</sub>CO<sub>3</sub> solution at pH 9.0 for 12 h. The reaction was neutralized with 50  $\mu$ L of a 50% acetic acid water solution and purified by C18 reversed phase HPLC. Solvent A: 0.1 M NH<sub>4</sub>OAc in water; solvent B: 0.1 M NH<sub>4</sub>OAc in 1:1 water:CH<sub>3</sub>CN; gradient 10-30% B over 30 min. Electrospray ionization mass spectrometry revealed the presence of ions at  $m/z$  1058.05, 846.5, 705.2 corresponding to 4<sup>-</sup>, 5<sup>-</sup>, 6<sup>-</sup> ions, respectively. The determined molecular weight was 4236.9, which agreed well with the calculated molecular weight of 4236.7. Enzymatic digestion and HPLC analysis yielded nucleoside ratios which were within experimental error of the theoretical composition of the oligonucleotide. The peak corresponding to the modified nucleoside had identical HPLC retention characteristics and UV spectrum as the synthesized standard **2.1**.

**Synthesis of the Psoralen-DNA Cross-link.** Oligonucleotides **2.5-2.7** (50 pmol) were phosphorylated with [ $\gamma$ - $^{32}$ P]ATP (New England Nuclear [6000 Ci/mmol]) and T4 polynucleotide kinase (New England Biolabs) followed by a chase of the kinase reaction with unlabeled ATP to achieve full 5' phosphorylation. The kinase was heat inactivated (65 °C, 15 min), and the appropriate unlabeled complementary strand was added. The oligonucleotides were annealed in 100 mM NaCl in a total volume of 100  $\mu$ L at 75 °C for 5 min and allowed to cool slowly to 4 °C. The labeled duplexes were placed in 6x50 mm pyrex tubes and were irradiated with 366 nm light using a 16 bulb Ray-O-Net photoreactor (9.0 J/m<sup>2</sup>) for 5 min at 4 °C. Hand-held irradiations were accomplished by placing a 366 nm UVP 4 watt lamp directly over the samples in opened Eppendorf tubes at 4 °C.

**Reversal of the Psoralen-DNA Cross-link.** (A) Photoreversal: Photocrosslinked oligonucleotides were photoreversed by irradiation of the samples with 254 nm light source (16.7 J/m<sup>2</sup>) for 20 min or by a 254 nm hand-held UVP 4 watt lamp placed directly over the samples for 2 h. (B) Base catalyzed reversal: The crosslink was treated with 0.1 M NaOH for 30 min at 90 °C. The reaction mixture was chilled on ice, neutralized with saturated NH<sub>4</sub>Cl, and the salts were removed by a C18 Waters Sep-Pak eluting with 1:1 CH<sub>3</sub>CN:water. All products were analyzed by electrophoresis on a denaturing 20% (19:1 mono:bis) polyacrylamide gel.

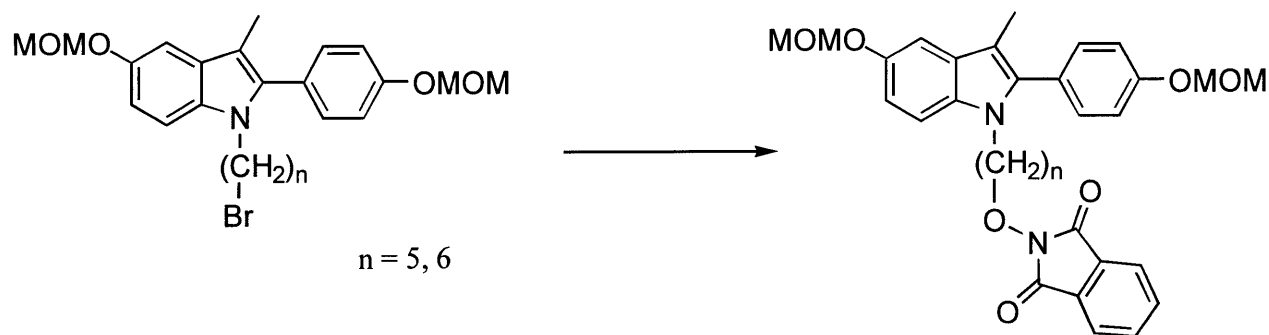


**6-Methoxy-2-carboxybenzofuran Thymidine *Cis-Syn* Photoproduct (3.2).** To a solution of 156 mg (0.32 mmol) of compound **1.24** in 15 mL of methanol was added 10 mL of a saturated Na<sub>2</sub>CO<sub>3</sub> solution. After stirring for 1 h at room temperature, the reaction was neutralized with 1N HCl and adsorbed onto silica gel via lyophilization. The crude product was purified by silica gel chromatography eluting with 5:1 ethyl acetate/methanol to afford 126 mg (93%) of a white solid: <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 1.75 (s, 3 H), 1.92-2.04 (m, 1 H), 2.16 (ddd, 1 H, *J* = 2.4, 5.9, 14), 3.63-3.75 (m, 2 H), 3.83 (s, 3 H), 3.88-3.94 (m, 1 H), 4.01 (s, 1 H), 4.31-4.39 (m, 1 H), 4.74 (s, 1 H), 6.27 (dd, 1 H, *J* = 5.5, 9.5), 6.56-6.63 (m, 2 H), 7.08 (d, 1 H, *J* = 8.3).



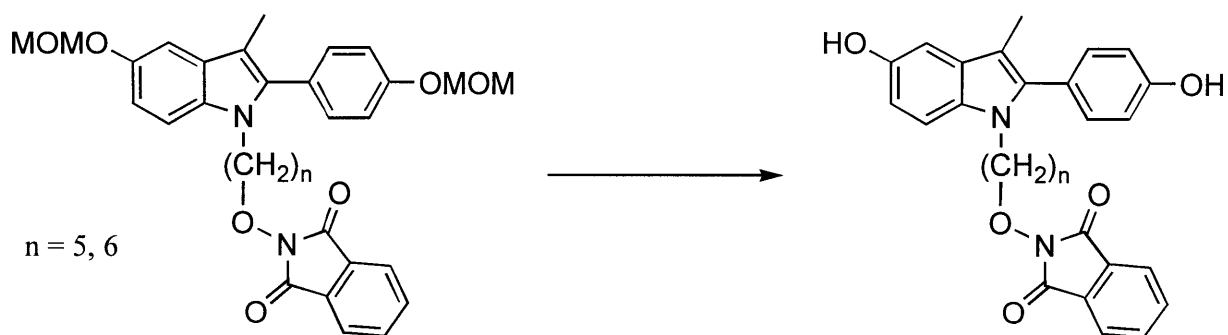
**6-Methoxy-2-(2-oxopentamide)benzofuran Thymidine *Cis-Syn* Photoproduct (3.3).**

To a solution of 67 mg (0.15 mmol) of compound **3.2** in 2 mL of dry DMF were added 66 mg (0.45 mmol) of 5-amino-2,2-dimethoxypropane **3.1**,<sup>71</sup> 88 mg (0.17 mmol) of PyBOP, and 108  $\mu$ L (0.62 mmol) of diisopropylethylamine. The solution was stirred at room temperature under an argon atmosphere for 2 h. The solvent was removed and the residual was dissolved in 2 mL of a 0.2 M solution of oxalic acid in water. After 5 min, the reaction was neutralized with saturated NaHCO<sub>3</sub> and the solvents were removed *in vacuo*. The residue was dissolved in methanol, adsorbed onto silica gel and directly added to the top of a silica gel column. Elution with a gradient of 20:1 to 10:1 dichloromethane/methanol afforded 39 mg (50%) of a yellow-white solid: <sup>1</sup>H NMR (500 MHz, acetone-*d*<sub>6</sub>)  $\delta$  1.67 (s, 3 H), 1.71-1.81 (m, 3 H), 2.01 (ddd, 1 H, *J* = 2.4, 5.9, 13), 2.07 (s, 3 H), 2.49 (app t, 2 H, *J* = 7), 3.15 (ddd, 1 H, *J* = 6.8, 13, 13), 3.35 (ddd, 1 H, *J* = 6.8, 14, 14), 3.59-3.60 (m, 2 H), 3.75 (s, 3 H), 3.76-3.79 (m, 1 H), 3.84 (s, 1 H), 4.28-4.30 (m, 1 H), 4.77 (d, 1 H, *J* = 1.9), 6.27 (dd, 1 H, *J* = 5.9, 8.8), 6.42 (d, 1 H, *J* = 2.5), 6.46 (dd, 1 H, *J* = 2.5, 7.8), 6.97 (d, 1 H, *J* = 7.8), 7.94 (br s, 1 H), 8.84 (br s, 1 H); <sup>13</sup>C NMR  $\delta$  22.27, 24.35, 38.08, 39.65, 41.05, 47.35, 55.97, 56.40, 56.69, 63.56, 72.16, 84.31, 87.05, 89.71, 97.65, 108.79, 117.99, 128.16, 152.72, 162.29, 163.52, 170.54, 171.16, 208.21.



**5-*N*-Hydroxyphthalimide 1-*N*-(2-(4-methoxymethylphenyl)-5-methoxymethylindole)pentane (3.6).** A solution of 0.22 g (0.46 mmol) of bromide **3.4**<sup>62</sup> and 0.11 g (0.65 mmol) of *N*-hydroxyphthalimide in 0.5 mL of dry DMF was heated to 60 °C under an argon atmosphere. Triethylamine (91  $\mu$ l, 0.65 mmol) was added and the solution was stirred overnight. The solution was diluted with ethyl ether, neutralized with  $\text{NH}_4\text{Cl}$ , and dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 5:1 to 3:1 hexanes/ethyl acetate to afford 0.22 g (86%) of a viscous yellow oil:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  1.35-1.38 (m, 2 H), 1.60-1.75 (m, 4 H), 2.20 (s, 3 H), 3.54 (s, 3 H), 3.55 (s, 3 H), 4.00-4.10 (m, 4 H), 5.23 (s, 2 H), 5.25 (s, 2 H), 6.97 (dd, 1 H,  $J = 2.4, 8.8$ ), 7.16 (d, 2 H,  $J = 8.8$ ), 7.23-7.25 (m, 2 H), 7.31 (d, 2 H,  $J = 8.8$ ), 7.71-7.74 (m, 2 H), 7.81-7.84 (m, 2 H);  $^{13}\text{C}$  NMR  $\delta$  9.23, 22.84, 27.64, 29.62, 43.66, 55.74, 55.81, 56.05, 78.00, 94.50, 95.95, 105.29, 108.36, 110.06, 112.87, 116.14, 123.36, 125.88, 128.96, 131.66, 132.39, 134.32, 138.03, 151.22, 156.99, 163.44.

**5-*N*-Hydroxyphthalimide 1-*N*-(2-(4-methoxymethylphenyl)-5-methoxymethylindole)hexane (3.7).** A solution of 46 mg (94  $\mu$ mol) of bromide **3.5**<sup>62</sup> and 23 mg (141  $\mu$ mol) of *N*-hydroxyphthalimide in 0.5 mL of dry DMF was heated to 60 °C under an argon atmosphere. Triethylamine (20  $\mu$ L, 141  $\mu$ mol) was added and the solution was stirred overnight. The solution was diluted with ethyl ether, neutralized with  $\text{NH}_4\text{Cl}$ , and dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 5:1 to 3:1 hexanes/ethyl acetate to afford 41 mg (76%) of a viscous yellow oil:  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  1.11-1.15 (m, 4 H), 1.48-1.56 (m, 4 H), 2.15 (s, 3 H), 3.54 (s, 3 H), 3.55 (s, 3 H), 3.94 (t, 2 H,  $J = 7$ ), 4.02 (t, 2 H,  $J = 6.7$ ), 5.23 (s, 2 H), 5.25 (s, 2 H), 6.75 (dd, 1 H,  $J = 2.4, 8.6$ ), 6.94 (d, 1 H,  $J = 2.4$ ), 6.99 (d, 2 H,  $J = 8.8$ ), 7.16 (d, 1 H,  $J = 8.6$ ), 7.22-7.26 (m, 2 H), 7.72-7.76 (m, 2 H), 7.82-7.87 (m, 2 H).

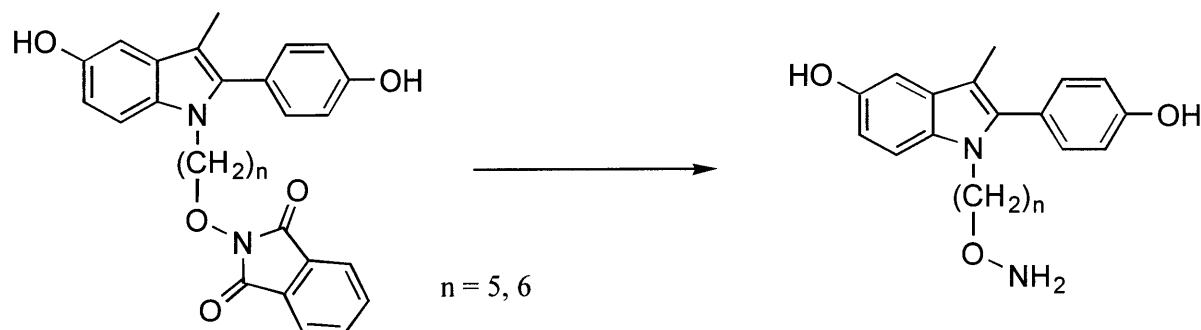


**5-*N*-Hydroxyphthalimide 1-*N*-(2-(4-hydroxyphenyl)-5-hydroxyindole) pentane (3.8).**

To a solution of 0.22 g (0.40 mmol) of **3.6** in 10 mL of methanol was added 0.2 mL of concentrated HCl. The solution was heated to reflux for 1 h, allow to cool to room temperature, and neutralized with saturated NaHCO<sub>3</sub>. The reaction mixture was diluted with ethyl acetate and the organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. Thin layer chromatography showed two spots suggesting that the phthalimide had partially degraded with the heat and acid. Purification of the crude mixture using silica gel chromatography eluting with a gradient of 5:1 to 2:1 hexanes/ethyl acetate afforded 0.10 g (54%) of a yellow oil: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 1.24-1.28 (m, 2 H), 1.52-1.63 (m, 4 H), 2.11 (s, 3 H), 3.95 (t, 2 H, *J* = 7), 4.03 (t, 2 H, *J* = 6.7), 5.26 (br s, 1 H), 6.14 (br s, 1 H), 6.79 (dd, 1 H, *J* = 2.4, 8.7), 6.94-6.97 (m, 3 H), 7.15-7.22 (m, 3 H), 7.70-7.72 (m, 2 H), 7.78-7.82 (m, 2 H).

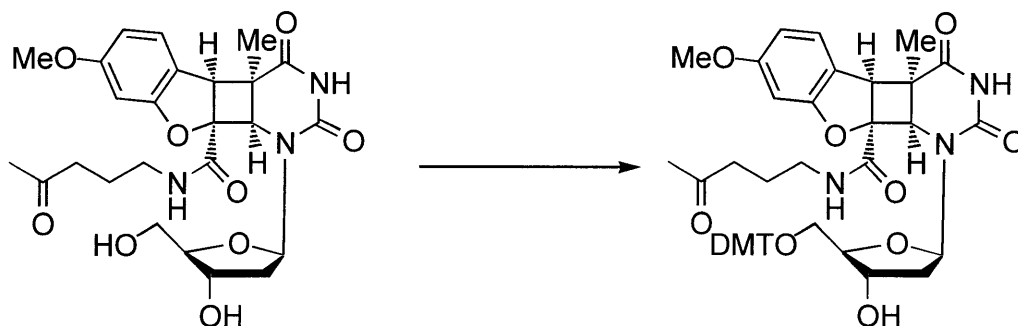
**5-*N*-Hydroxyphthalimide 1-*N*-(2-(4-hydroxyphenyl)-5-hydroxyindole) hexane (3.9).**

To a solution of 41 mg (71 μmol) of **3.7** in 2 mL of methanol was added 40 μL of concentrated HCl. The solution was heated to reflux for 1 h, allow to cool to room temperature, and neutralized with saturated NaHCO<sub>3</sub>. The reaction mixture was diluted with ethyl acetate and the organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. Thin layer chromatography showed two spots suggesting that the phthalimide had partially degraded with the heat and acid. Purification of the crude mixture using silica gel chromatography eluting with a gradient of 5:1 to 5:2 hexanes/ethyl acetate afforded 17 mg (50%) of a yellow oil: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.11-1.15 (m, 4 H), 1.48-1.56 (m, 4 H), 2.15 (s, 3 H), 3.94 (t, 2 H, *J* = 7), 4.02 (t, 2 H, *J* = 6.7), 6.75 (dd, 1 H, *J* = 2.4, 8.6), 6.94 (d, 1 H, *J* = 2.4), 6.99 (d, 2 H, *J* = 8.8), 7.16 (d, 1 H, *J* = 8.6), 7.22-7.26 (m, 2 H), 7.72-7.76 (m, 2 H), 7.82-7.87 (m, 2 H).



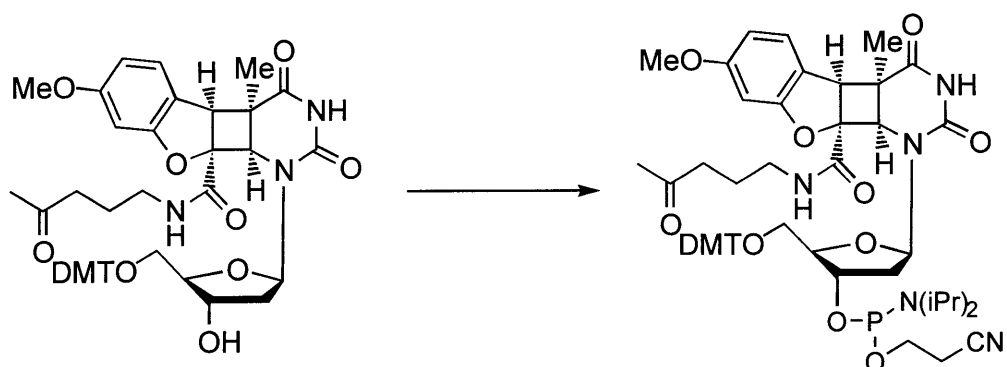
**5-*N*-Hydroxyamine-*N*-(2-(4-hydroxyphenyl)-5-hydroxyindole) pentane (3.10).** To a solution of 25 mg (50  $\mu$ mol) of **3.8** in 5 mL of methanol were added 150  $\mu$ L of hydrazine. After stirring for 20 min at room temperature, the solvents were removed *in vacuo*. The crude material was dissolved in methanol, adsorbed onto silica gel and was directly added to the top of a silica gel column. Elution with 25:1 dichloromethane/methanol afforded 14 mg (82%) of a yellowish oil:  $^1\text{H NMR}$  (300 MHz, methanol- $d_4$ )  $\delta$  1.06-1.16 (m, 2 H), 1.31-1.41 (m, 2 H), 1.44-1.56 (m, 2 H), 2.09 (s, 3 H), 3.44 (t, 2 H,  $J = 6.7$ ), 3.96 (t, 2 H,  $J = 7$ ), 6.69 (dd, 1 H,  $J = 2.4, 8.6$ ), 6.85-6.93 (m, 3 H), 7.11-7.19 (m, 3 H).

**5-*N*-Hydroxyamine-*N*-(2-(4-hydroxyphenyl)-5-hydroxyindole) hexane (3.11).** To a solution of 16 mg (33  $\mu$ mol) of **3.9** in 3 mL of methanol were added 90  $\mu$ L of hydrazine. After stirring for 30 min at room temperature, the solvents were removed *in vacuo*. The crude material was dissolved in methanol, adsorbed onto silica gel and was directly added to the top of a silica gel column. Elution with 25:1 dichloromethane/methanol afforded 5 mg (45%) of a yellowish oil:  $^1\text{H NMR}$  (500 MHz, methanol- $d_4$ )  $\delta$  0.99-1.11 (m, 4 H), 1.31-1.37 (m, 2 H), 1.42-1.49 (m, 2 H), 2.06 (s, 3 H), 3.47 (t, 2 H,  $J = 6.7$ ), 3.93 (t, 2 H,  $J = 7$ ), 6.65 (dd, 1 H,  $J = 2.4, 8.6$ ), 6.81-6.87 (m, 3 H), 7.10-7.15 (m, 3 H).



**6-Methoxy-2-(2-oxopentamide)benzofuran 5'-O-Dimethoxytritylthymidine *Cis-Syn* Photoproduct (3.12).** To a solution of 39 mg (75  $\mu\text{mol}$ ) of compound **3.3**<sup>71</sup> in 1 mL of dry DMF were added 28 mg (83  $\mu\text{mol}$ ) of dimethoxytrityl chloride, 14 mg (83  $\mu\text{mol}$ ) of silver nitrate and 15  $\mu\text{L}$  (113  $\mu\text{mol}$ ) of collidine. The cloudy orange solution was stirred for 15 min at room temperature under an argon atmosphere. The reaction mixture was diluted with ethyl ether, washed with water, saturated  $\text{CuSO}_4$ , and brine, then dried over  $\text{MgSO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 100:1 to 50:1 dichloromethane/methanol to afford 34 mg (55%) of a yellow foam:  $^1\text{H}$  NMR (500 MHz, acetone- $d_6$ )  $\delta$  1.57 (s, 3 H), 1.70-1.74 (m, 2 H), 1.88-1.92 (m, 1 H), 2.02 (s, 3 H), 2.02-2.06 (m, 1 H), 2.47 (app t, 2 H,  $J = 7$ ), 3.11-3.17 (m, 2 H), 3.29-3.37 (m, 2 H), 3.75 (s, 1 H), 3.75-3.79 (m, 1 H), 3.80 (s, 6 H), 3.86 (s, 1 H), 4.36-4.39 (m, 1 H), 4.74 (d, 1 H,  $J = 1.5$ ), 6.31 (app t, 1 H,  $J = 7$ ), 6.43 (d, 1 H,  $J = 2.0$ ), 6.47 (dd, 1 H,  $J = 2.4, 8.3$ ), 6.91-6.95 (m, 4 H), 6.98 (d, 1 H,  $J = 8.3$ ), 7.21-7.25 (m, 1 H), 7.34-7.37 (m, 2 H), 7.40-7.42 (m, 4 H), 7.54-7.56 (m, 2 H), 7.92 (br s, 1 H), 8.87 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  22.37, 24.38, 38.03, 39.57, 41.04, 47.51, 55.57, 55.93, 56.25, 57.02, 64.70, 71.32, 83.22, 84.96, 86.87, 89.87, 97.63, 108.72, 114.02, 114.03, 118.02, 127.51, 128.16, 128.72, 129.21, 131.14, 131.17, 137.11, 137.25, 146.47, 152.68, 159.64, 159.65, 162.23, 163.54, 170.33, 171.05, 208.05.

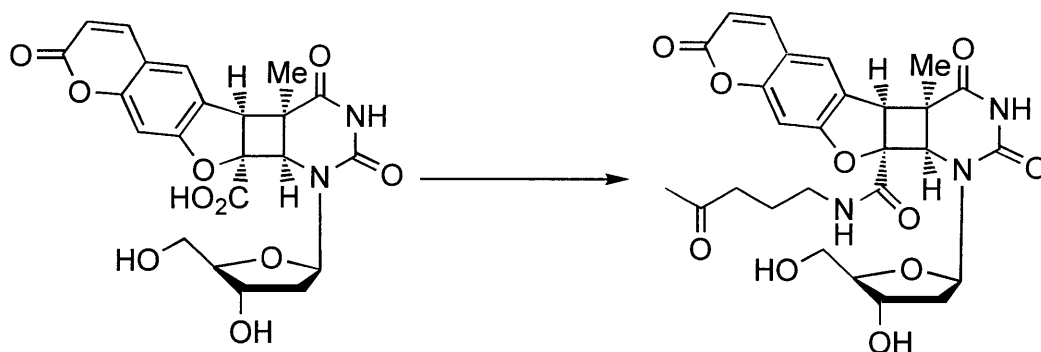




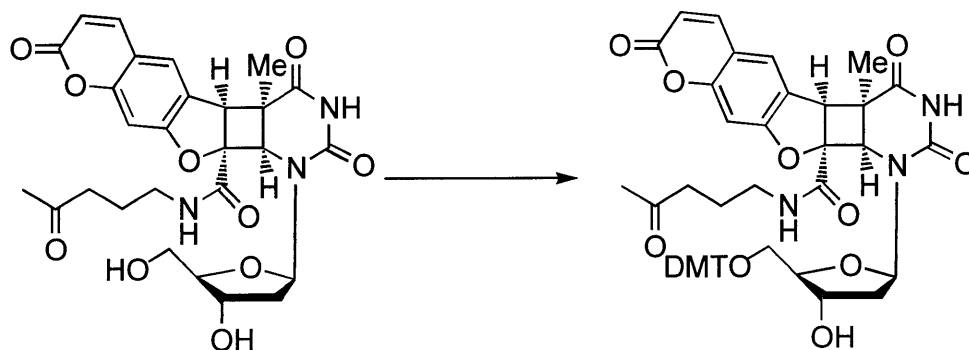
**6-Methoxy-2-(2-oxopentamide)benzofuran 5'-O-Dimethoxytrityl-3'-O-**

**(isopropylamine-2-cyanoethyl phosphoramidite)thymidine *Cis-Syn* Photoproduct (3.13).**

To a solution of 34 mg (41  $\mu\text{mol}$ ) of compound **3.12** in 400  $\mu\text{L}$  of dry  $\text{CH}_2\text{Cl}_2$  were added 6 mg (32  $\mu\text{mol}$ ) of bisisopropylammonium tetrazolide and 16  $\mu\text{L}$  (64  $\mu\text{mol}$ ) of bis(isopropylamine)-2-cyanoethyl phosphoramidite. After stirring for 3 h at room temperature under an argon atmosphere the solvents were removed *in vacuo*. The crude product was purified by silica gel chromatography eluting with 1:1 hexanes/ethyl acetate. The purified product was dissolved in benzene and lyophilized to afford 28 mg (67%) of a white crunchy foam:  $^{31}\text{P}$  NMR (121 MHz, benzene- $d_6$ )  $\delta$  147.89, 149.16.

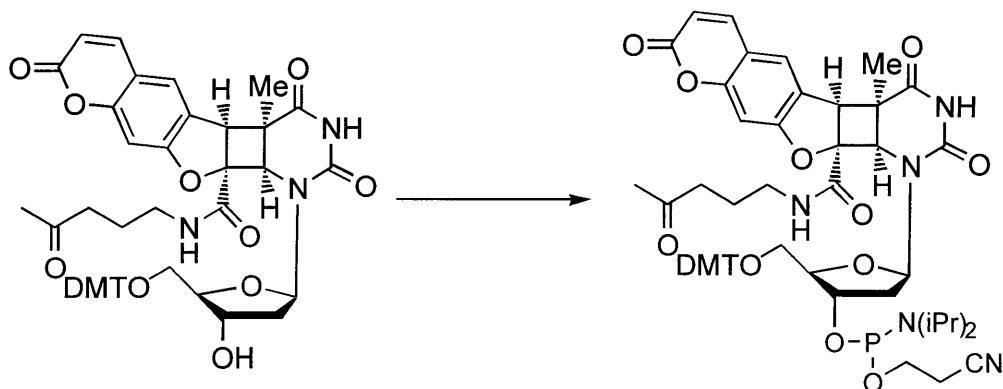


**2-(2-Oxopentamide)psoralen Thymidine *Cis*-Syn Photoproduct (3.17).** To a solution of 34 mg (72  $\mu\text{mol}$ ) of compound **2.1** in 1 mL of dry DMF were added 42 mg (290  $\mu\text{mol}$ ) of 5-amino-2,2-dimethoxypropane **3.1**, 56 mg (110  $\mu\text{mol}$ ) of PyBOP, and 70  $\mu\text{L}$  (400  $\mu\text{mol}$ ) of diisopropylethylamine. The solution was stirred at room temperature under an argon atmosphere for 2 h. The solvent was removed and the residual was dissolved in 3 mL of a 0.2 M solution of oxalic acid in water. After 5 min, the reaction was neutralized with saturated  $\text{NaHCO}_3$  and the solvents were removed *in vacuo*. The residue was dissolved in methanol, adsorbed onto silica gel and directly added to the top of a silica gel column. Elution with 20:1 dichloromethane/methanol afforded 32 mg (80%) of a yellow-white solid:  $^1\text{H}$  NMR (300 MHz, methanol- $d_4$ )  $\delta$  1.74 (s, 3 H), 1.64-1.87 (m, 3 H), 2.03 (ddd, 1 H,  $J = 3.0, 5.9, 13$ ), 2.09 (s, 3 H), 2.50 (app t, 1 H,  $J = 7.1$ ), 3.12-3.29 (m, 2 H), 3.59-3.62 (m, 2 H), 3.75-3.79 (m, 1 H), 4.04 (s, 1 H), 4.17-4.21 (m, 1 H), 6.23-6.28 (m, 1 H), 6.26 (d, 1 H,  $J = 9.5$ ), 6.98 (s, 1 H), 7.39 (s, 1 H), 7.85 (d, 1 H,  $J = 9.8$ );  $^{13}\text{C}$  NMR  $\delta$  22.23, 24.47, 29.94, 38.30, 40.25, 41.52, 47.49, 56.14, 57.49, 63.69, 72.43, 84.87, 87.37, 91.27, 100.30, 113.98, 116.02, 124.50, 127.94, 145.93, 153.89, 157.42, 162.92, 165.65, 170.72, 172.51, 211.28.

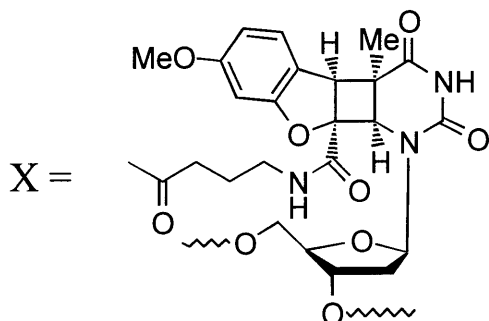


### 2-(2-Oxopentamide)psoralen 5'-O-Dimethoxytritylthymidine *Cis-Syn* Photoproduct

**(3.18).** To a solution of 32 mg (58  $\mu\text{mol}$ ) of compound **3.17** in 1 mL of dry DMF were added 22 mg (64  $\mu\text{mol}$ ) of dimethoxytrityl chloride, 11 mg (64  $\mu\text{mol}$ ) of silver nitrate and 12  $\mu\text{L}$  (87  $\mu\text{mol}$ ) of collidine. The cloudy orange solution was stirred for 15 min at room temperature under an argon atmosphere. The reaction mixture was diluted with ethyl ether, washed with water, saturated  $\text{CuSO}_4$ , and brine, and then dried over  $\text{MgSO}_4$ . The crude product was purified by silica gel chromatography eluting with a gradient of 100:1 to 50:1 dichloromethane/methanol to afford 14 mg (28%) of a yellow foam:  $^1\text{H}$  NMR (500 MHz, acetone- $d_6$ )  $\delta$  1.61 (s, 3 H), 1.70-1.75 (m, 2 H), 1.89-1.93 (m, 1 H), 2.02 (s, 3 H), 2.04-2.11 (m, 1 H), 2.47 (app t, 2 H,  $J = 7$ ), 3.14-3.18 (m, 2 H), 3.31-3.35 (m, 2 H), 3.77-3.80 (m, 1 H), 3.80 (s, 6 H), 4.06 (s, 1 H), 4.38-4.42 (m, 1 H), 4.82 (d, 1 H,  $J = 1.5$ ), 6.25 (d, 1 H,  $J = 9.8$ ), 6.30-6.33 (m, 1 H), 6.83 (s, 1 H), 6.92-6.95 (m, 4 H), 7.22-7.25 (m, 1 H), 7.35-7.38 (m, 2 H), 7.41-7.43 (m, 5 H), 7.55-7.57 (m, 2 H), 7.90 (d, 1 H,  $J = 9.8$ ), 8.07 (br s, 1 H), 8.98 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  22.39, 24.29, 29.13, 38.16, 39.72, 41.04, 47.95, 55.36, 55.58, 57.06, 64.60, 71.30, 83.23, 85.04, 86.90, 91.12, 99.73, 113.95, 114.04, 114.05, 115.24, 124.04, 127.53, 127.60, 128.73, 129.23, 131.16, 131.18, 137.12, 137.25, 144.82, 146.45, 152.48, 156.99, 159.66, 160.67, 165.21, 169.49, 170.80, 208.05; HRMS (FAB+, glycerol) calcd for  $\text{C}_{48}\text{H}_{47}\text{O}_{12}\text{N}_3$  ( $\text{M} + \text{H}$ ) $^+$  858.3238, found 858.3236.

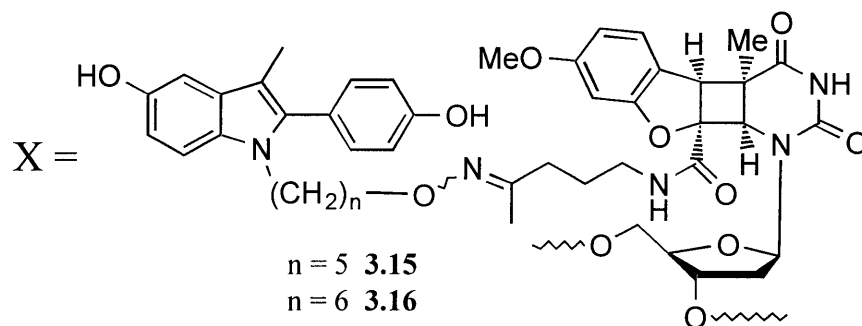


**2-(2-Oxopentamide)psoralen 5'-O-Dimethoxytrityl-3'-O-(isopropylamine-2-cyanoethyl phosphoramidite)thymidine *Cis-Syn* Photoproduct (3.19).** To a solution of 14 mg (16  $\mu\text{mol}$ ) of compound **3.18** in 200  $\mu\text{L}$  of dry  $\text{CH}_2\text{Cl}_2$  were added 3 mg (13  $\mu\text{mol}$ ) of bisisopropylammonium tetrazolide and 7  $\mu\text{L}$  (20  $\mu\text{mol}$ ) of bis(isopropylamine)-2-cyanoethyl phosphoramidite. After stirring for 3 h at room temperature under an argon atmosphere the solvents were removed *in vacuo*. The crude product was purified by silica gel chromatography eluting with 1:1 hexanes/ethyl acetate. The purified product was dissolved in benzene and lyophilized to afford 15 mg (89%) of a white crunchy foam:  $^{31}\text{P}$  NMR (121 MHz, benzene- $d_6$ )  $\delta$  147.75, 149.43.



### 5'-AGC TAX AAA AGG T-3'

**Oligonucleotide (3.14).** A 0.05 M solution of phosphoramidite **3.13** was added using the optional port on the DNA synthesizer. The coupling yield for the modified base was 40%. The oligonucleotide was deprotected using 10% DBU in ethanol and purified by C18 reversed phase HPLC. Solvent A: 0.1 M NH<sub>4</sub>OAc in water; solvent B: 0.1 M NH<sub>4</sub>OAc in 1:1 water:CH<sub>3</sub>CN; gradient 15-25% B over 30 min. Electrospray ionization mass spectrometry revealed the presence of ions at *m/z* 1426.9, 1070.7, 856.8, 714.0 corresponding to 3<sup>-</sup>, 4<sup>-</sup>, 5<sup>-</sup>, 6<sup>-</sup> ions, respectively. The determined molecular weight was 4286.8, which agreed well with the calculated molecular weight of 4281.8.



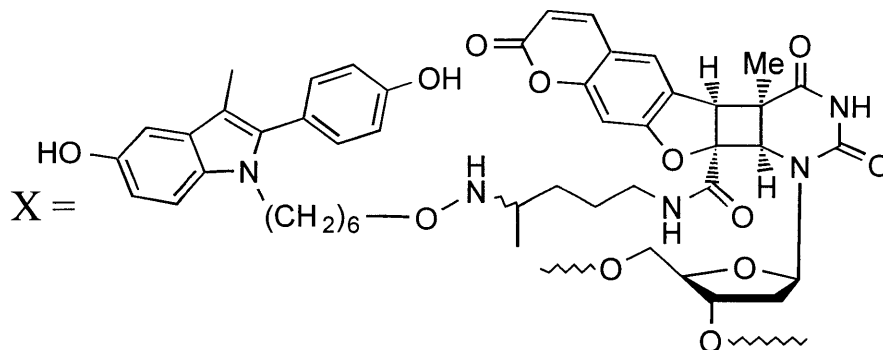
### 5'-AGC TAX AAA AGG T-3'

#### Oxime Formation Between Oligonucleotide (3.14) and 2-Phenylindole Aminoxy

**Compounds.** To a solution of 19 nmol of oligonucleotide **3.14** in 200  $\mu\text{L}$  of water were added 196  $\mu\text{L}$  of 0.2 M sodium phosphate buffer at pH 5.6, and 4  $\mu\text{L}$  (40 nmol) of a 10 mM solution of 2-phenylindole aminoxy compound in 1:1 DMF:water. After overnight incubation at room temperature the samples were analyzed and purified by C18 reversed phase HPLC. Solvent A: 0.1 M  $\text{NH}_4\text{OAc}$  in water; solvent B: 0.1 M  $\text{NH}_4\text{OAc}$  in 1:1 water: $\text{CH}_3\text{CN}$ ; gradient 15-25% B over 10 min, 25-40% B over 5 min, 40-60% B over 20 min, 60-100% B over 10 min.

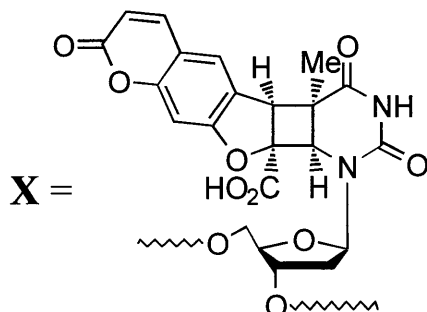
Oligonucleotides **3.15** and **3.16** were characterized by UV absorbance and electrospray ionization mass spectrometry. Electrospray ionization mass spectrometry of either peak of **3.15** revealed the presence of ions at  $m/z$  1534.0, 1150.9 and 920.7 corresponding to 3<sup>-</sup>, 4<sup>-</sup> and 5<sup>-</sup> ions, respectively. The determined molecular weight of **3.15** was 4607.0, which agreed well with the calculated molecular weight of 4604.5. Electrospray ionization mass spectrometry of either peak of **3.16** revealed the presence of ions at  $m/z$  1539.0 and 1154.5 corresponding to 3<sup>-</sup> and 4<sup>-</sup> ions, respectively. The determined molecular weight of **3.16** was 4621.0, which agreed well with the calculated molecular weight of 4222.0. The UV spectrum of every isolated product contained a shoulder at 300 nm corresponding to the addition of the 2PI group (Fig. 3.6E).



**5'-CC TCT TCT TCT GXG CAC TCT TCT TCT-3'**

**Oligonucleotide (3.21).** To a solution of 10 nmol of oligonucleotide **3.20** in 48.5  $\mu$ L of water were added 50  $\mu$ L of 0.2 M sodium phosphate buffer at pH 5.6, and 1.5  $\mu$ L (15 nmol) of a 10 mM solution of compound **3.11** in 1:1 DMF:water. After overnight incubation at room temperature, 20  $\mu$ L aliquots were reduced with 18  $\mu$ L of 1M NaCNBH<sub>3</sub> in water and 2  $\mu$ L of 1M HCl for 15 min. The reduced samples were purified by strong anion exchange HPLC. Solvent A: 0.025 M NH<sub>4</sub>OAc in 10% CH<sub>3</sub>CN; solvent B: 0.025 M NH<sub>4</sub>OAc, 1 M NaCl in 10% CH<sub>3</sub>CN; gradient 35-60% B over 30 min. Oligonucleotides were characterized by UV absorbance and MALDI mass spectrometry. The determined molecular weight was 8408.0, which agreed well with the calculated molecular weight of 8409.4.



**5'-CC TCT TCT TCT GXG CAC TCT TCT TCT-3'**

**Oligonucleotide (3.21).** The modified phosphoramidite **2.4** was added manually and the coupling yield for this step was 85%. The oligonucleotide was deprotected using 10% DBU in ethanol and the 2'-carboethoxypsoralen was converted into the carboxylic acid by treatment with 1 mL of a 100 mM Na<sub>2</sub>CO<sub>3</sub> solution at pH 9.0 for 12 h. The reaction was neutralized with 50  $\mu$ L of a 50% acetic acid water solution and purified by strong anion exchange HPLC. Solvent A: 0.025 M NH<sub>4</sub>OAc in 10% CH<sub>3</sub>CN; solvent B: 0.025 M NH<sub>4</sub>OAc, 1 M NaCl in 10% CH<sub>3</sub>CN; gradient 35-60% B over 30 min.

**Oxime Formation Between Ketone Nucleosides and 2-Phenylindole Aminoxy**

**Compounds.** To a solution of 250 nmol of a ketone nucleoside in 25  $\mu\text{L}$  of water were added 60  $\mu\text{L}$  of 0.2 M sodium phosphate buffer at pH 5.6, 80  $\mu\text{L}$  of DMF, and 35  $\mu\text{L}$  (350 nmol) of a 10 mM solution of 2-phenylindole aminoxy compound in 1:1 DMF:water. After overnight incubation at room temperature the samples were analyzed and purified by C18 reversed phase HPLC. Solvent A: 0.1 M  $\text{NH}_4\text{OAc}$  in water; solvent B: 0.1 M  $\text{NH}_4\text{OAc}$  in 1:1 water: $\text{CH}_3\text{CN}$ ; gradient 30-50% B over 10 min, 50-100% B over 30 min.

**Oxime of 2PI- $\text{C}_6$ - $\text{ONH}_2$  and Psoralen-Ketone.** The two isolated peaks had nearly identical proton NMR spectra.  $^1\text{H}$  NMR (500 MHz, acetone- $d_6$ )  $\delta$  1.10-1.21 (m, 4 H), 1.43-1.49 (m, 2 H), 1.53-1.57 (m, 2 H), 1.71 (s, 3 H), 1.74 (s, 3 H), 1.73-1.83 (m, 3 H), 2.01-2.05 (m, 1 H), 2.10 (s, 3 H), 2.13-2.16 (m, 2 H), 3.12-3.17 (m, 2 H), 3.61-3.62 (m, 2 H), 3.76-3.79 (m, 1 H), 3.87 (t, 2 H,  $J = 6.3$ ), 3.98-4.01 (m, 3 H), 4.31-4.33 (m, 1 H), 4.86 (d, 1 H,  $J = 1.5$ ), 6.24 (d, 1 H,  $J = 9.8$ ), 6.27 (dd, 1 H,  $J = 5.9, 8.8$ ), 6.74 (dd, 1 H,  $J = 2.4, 8.8$ ), 6.81 (s, 1 H), 6.91 (d, 1 H,  $J = 2.0$ ), 6.99 (d, 2 H,  $J = 8.8$ ), 7.21 (d, 1 H,  $J = 8.3$ ), 7.24 (d, 2 H,  $J = 8.8$ ), 7.36 (m, 2 H), 7.89 (d, 1 H,  $J = 9.8$ ), 8.05 (br s, 1 H).

**Relative Affinity of 2-Phenylindole Derivatives for the ER.** The relative affinities of the 2-phenylindole containing substrates for the calf uterine ER were measured by a competitive binding assay<sup>73</sup> with  $17\beta$ - $^3\text{H}$ estradiol (New England Nuclear, 85-115 Ci/mmol). A 1:500 dilution of the stock solution of  $^3\text{H}$ estradiol was made using 3:1 DMF:water. The 2PI derivatives were dissolved in 10  $\mu\text{L}$  of the  $^3\text{H}$ estradiol dilution solution. Calf uterine extract<sup>73</sup> (90  $\mu\text{L}$ ) was added, the samples were vortexed, and incubated overnight at 4  $^\circ\text{C}$ . After at least 8 hr, 100  $\mu\text{L}$  of a 0.5% dextran-activated charcoal solution in 10 mM Tris $\cdot\text{HCl}$  (pH 7.5) was added to each sample and vortexed. After a 1 hr incubation at 4  $^\circ\text{C}$ , the samples were spun for 10 min at 4  $^\circ\text{C}$  and 100  $\mu\text{L}$  of each sample was counted for radioactivity. By varying the ratios of the molar concentrations of two ligands for the ER, it was possible to determine the concentration of

the test substrate necessary to reduce receptor-bound radioactivity by 50%.

**Assay for Recombinant ER Activity.** The activity of the purchased human estrogen receptor was determined by the protocol described by Panvera.<sup>66</sup> Briefly, 5  $\mu\text{L}$  of 400 nM [ $^3\text{H}$ ]-estradiol (New England Nuclear, 40-60 Ci/mmol) was added to 95  $\mu\text{L}$  of ER in 10 mM Tris·HCl (pH 7.5), 10% glycerol, 1 mM DTT, and 95  $\mu\text{g}$  of BSA. The reactions were incubated overnight at 4  $^\circ\text{C}$ . A hydroxyl apatite slurry (100  $\mu\text{L}$ ) was added to each reaction and vortexed three times during a 15 min incubation on ice. A wash buffer (1 mL) containing 40 mM Tris·HCl (pH 7.4), 100 mM KCl, 1 mM EDTA and 1 mM EGTA was added, the solution was vortexed, and spun 5 min at 10,000 in a microcentrifuge. The supernatant was discarded and the wash step was repeated two more times. The hydroxyl apatite pellet was suspended in 200  $\mu\text{L}$  of ethanol and transferred into a scintillation vial. The reaction tube was washed with 200  $\mu\text{L}$  ethanol and added to the scintillation vial for quantitation. A non-specific binding (NSB) control was performed in parallel containing 300-fold excess of cold estradiol. The ER concentration was determined by using the following equation:

$$[\text{ER}] = \frac{((\text{cpm for ER}) - (\text{cpm of NSB})) \times (\text{dilution factor})}{(0.005 \text{ mL}) \times (\text{cpm/pmol for } [^3\text{H}]\text{-estradiol})}$$

**Construction of  $^{32}\text{P}$ -Internally Labeled Closed Circular Plasmids Containing a Site-Specific Psoralen Adduct.** In separate reactions, 31 pmol of oligonucleotides **3.20**, **3.21** or **3.22** were  $^{32}\text{P}$ -5'-phosphorylated with [ $\gamma$ - $^{32}\text{P}$ ] ATP (70  $\mu\text{Ci}$ ), in a total volume of 10  $\mu\text{L}$ , by incubation with T4 polynucleotide kinase (10 U) at 37  $^\circ\text{C}$  for 30 min. T4 polynucleotide kinase (2 U) and 1  $\mu\text{L}$  of 100  $\mu\text{mol}$  of unlabeled ATP were added and the reactions were again incubated for 30 min. The reactions were heat inactivated at 65  $^\circ\text{C}$  for 20 min. To the  $^{32}\text{P}$  labeled oligonucleotides were added 4 pmol of M13mp18GTGx<sup>74</sup> circular single-stranded DNA in

NEB2 (New England Biolabs) buffer in a total volume of 23  $\mu\text{L}$ . The solutions were heated to 70  $^{\circ}\text{C}$  for 5 min allowed to cool to 37  $^{\circ}\text{C}$  for 30 min, and then cooled to room temperature for 20 min. The annealed substrates were stored on ice while the DNA synthesis premix was prepared. The premix consisted of 10  $\mu\text{L}$  of 10X NEB1 (New England Biolabs) buffer, 1  $\mu\text{L}$  of 100X BSA (New England Biolabs), 2  $\mu\text{L}$  of 100 mM ATP (Pharmacia), 30  $\mu\text{L}$  of a solution that contained 2 mM of each dNTP, 22  $\mu\text{L}$  of water, 10  $\mu\text{L}$  of T4 DNA polymerase (3 U/ $\mu\text{L}$ ) (New England Biolabs), and 2  $\mu\text{L}$  of T4 DNA ligase (400 U/ $\mu\text{L}$ ) (New England Biolabs). The DNA synthesis premix (77  $\mu\text{L}$ ) was added to the DNA and incubated at 37  $^{\circ}\text{C}$  for 3 hr. The reaction was heat inactivated at 75  $^{\circ}\text{C}$  for 10 min. Digestion of the DNA synthesis side products was accomplished by addition of 1.5  $\mu\text{L}$  of T5 Exonuclease and incubation at 37  $^{\circ}\text{C}$  for 3 hr. The reaction was heat inactivated at 70  $^{\circ}\text{C}$  for 10 min and the plasmid was purified by size exclusion chromatography using two 1 mL Sephacryl-S400 spin columns (1700 rpm) pre-equilibrated with 10 mM Tris·HCl (pH 8.0) and 1 mM EDTA. The DNA substrates were stored at -80  $^{\circ}\text{C}$ .

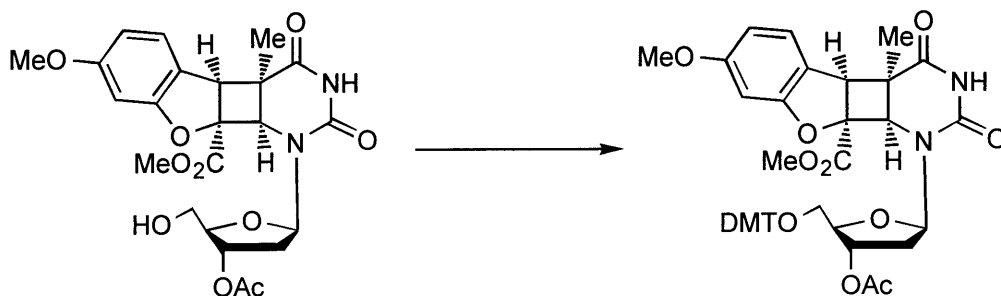
The purity of the covalently closed circular DNA was determined by electrophoresis on a 15 x 30 cm 0.8% agarose gel at 50 V overnight in TBE buffer. The gel was exposed to BioMax film at room temperature for 15 min. The site-specific placement of the psoralen adduct in the plasmid was confirmed by enzymatic digestion at unique sites in DNA sequences flanking and directly containing the psoralen lesion. Typically, 200-250 ng of plasmid was digested with 6 units of a restriction enzyme in a 10  $\mu\text{L}$  (NEB2 buffer) reaction at 37  $^{\circ}\text{C}$  for 90 min (Fig. 3.17). The digestion products were analyzed by electrophoresis on a 15 x 30 cm 0.8% agarose gel at 50 V overnight in TBE buffer. The gels were exposed to BioMax film at room temperature for 15 min.

***In vitro* Nucleotide Excision Repair Assays.** Reaction mixtures (10  $\mu\text{L}$ ) contained 150 ng of plasmid containing a site-specific psoralen adduct and 20-40  $\mu\text{g}$  of HeLa whole cell extract<sup>75</sup> in a buffer containing 50 mM HEPES-KOH (pH 7.8), 70 mM KCl, 10 mM MgCl<sub>2</sub>, 1.3 mM DTT, 0.4 mM EDTA, 2 mM ATP, 22 mM phosphocreatine (di-Tris salt), 2.0  $\mu\text{g}$  creatine phosphokinase (type I, Sigma), 6.8% glycerol, 0.02% NP40 and 18  $\mu\text{g}$  of BSA. HeLa cell

extract was preincubated with buffer at 30 °C for 5 min, DNA substrate added, and reactions incubated at 30 °C for the times indicated. The repair reactions were stopped by heating to 95 °C for 5 min. The incisions were analyzed by electrophoresis on a 14% 7 M urea polyacrylamide gel (0.4 mM) at 50 °C for 2.5 hr. The gels were exposed to BioMax film at - 80 °C for 12-24 hr. Amount of nucleotide excision was quantitated using a Molecular Dynamics PhosphorImager.

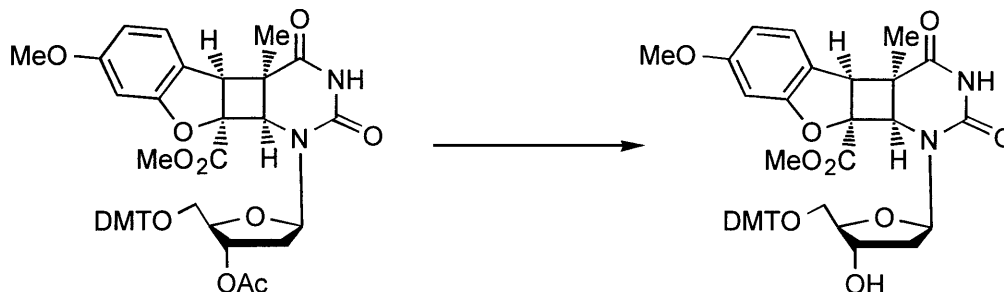
**Blocking of Nucleotide Excision Repair by Recombinant ER.** The ER storage buffer (PanVera) (100 µL) was exchanged by passing it through two G25 (1 mL) spin columns (1900 rpm) pre-equilibrated with 50 mM Tris·HCl (pH 7.5), 100 mM KCl, 1 mM EDTA, 2 mM DTT, 12 mM MgCl<sub>2</sub>, 0.067% NP40 and 10% glycerol. Half of the ER was heated to 95 °C for 5 min and spun (14,000) at 4 °C for 5 min in order to pellet the precipitated protein. The heat inactivated solution served as buffer for the repair reactions when less than 3 µL of dialyzed ER was used. Reaction mixtures (10 µL) contained 150 ng of plasmid containing a site-specific psoralen adduct, 0-3 µL of ER, and 20-40 µg of HeLa whole cell extract<sup>75</sup> in a buffer containing 42.5 mM HEPES-KOH (pH 7.8), 15 mM Tris·HCl (pH 7.5), 40 mM KCl, 10 mM MgCl<sub>2</sub>, 1.3 mM DTT, 0.4 mM EDTA, 2 mM ATP, 22 mM phosphocreatine (di-Tris salt), 2.0 µg creatine phosphokinase (type I, Sigma), 4.7% glycerol, 0.02% NP40 and 18 µg of BSA. The DNA substrate and 3 µL of a ratio of ER and heat inactivated ER were preincubated with buffer at 30 °C for 30 min. HeLa whole cell extract was added and the repair reaction was incubated at 30 °C for 30 min. The repair reactions were stopped by heating to 95 °C for 5 min. The incisions were analyzed by electrophoresis in a 14% 7 M urea polyacrylamide gel (0.4 mM) at 50 °C for 2.5 hr. The gels were exposed to BioMax film at - 80 °C for 12-24 hr.





#### 6-Methoxy-2-carbomethoxybenzofuran 5'-*O*-DMT-3'-*O*-acetylthymidine *Cis*-Syn

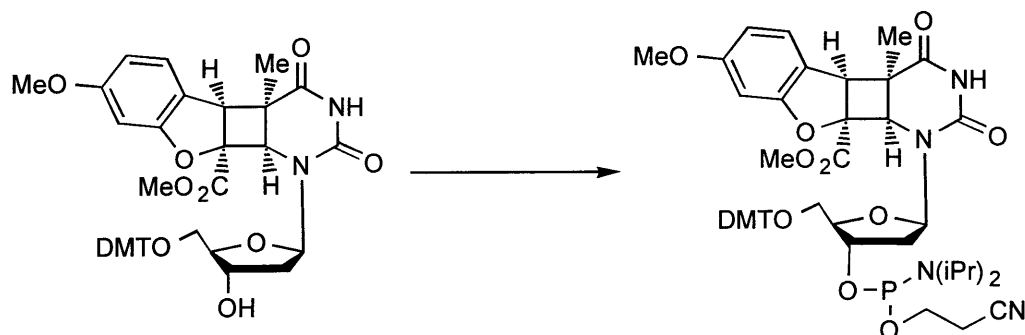
**Photoproduct (4.1).** To a solution of 0.25 g (0.50 mmol) of compound **1.24** in 2.5 mL of dry DMF were added 0.20 g (0.60 mmol) of dimethoxytrityl chloride, 0.10 g (0.60 mmol) of silver nitrate and 100  $\mu$ L (0.75 mmol) of collidine. The cloudy orange solution was stirred for 15 min at room temperature under an argon atmosphere. The reaction mixture was diluted with ethyl ether, washed with water, saturated  $\text{CuSO}_4$ , and brine, and then dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by silica gel chromatography eluting with 100:1 dichloromethane/methanol to afford 0.33 g (84%) of a yellow foam:  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.65 (s, 3 H), 2.03-2.14 (m, 2 H), 3.24 (dd, 1 H,  $J = 4.4, 10$ ), 3.37 (dd, 1 H,  $J = 4.4, 10$ ), 3.76 (s, 3 H), 3.77 (s, 3 H), 3.80 (s, 6 H), 3.93-3.97 (m, 1 H), 4.06 (s, 1 H), 4.73 (d, 1 H,  $J = 1.4$ ), 5.22-5.25 (m, 1 H), 6.22-6.27 (m, 1 H), 6.45-6.50 (m, 2 H), 6.93-7.00 (m, 5 H), 7.23-7.28 (m, 1 H), 7.35-7.43 (m, 6 H), 7.54-7.57 (m, 2 H), 9.03 (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  20.97, 23.15, 35.32, 47.34, 53.64, 55.58, 55.92, 56.62, 57.62, 64.18, 73.98, 81.85, 83.70, 87.06, 88.25, 96.97, 108.84, 114.05, 117.42, 127.65, 127.98, 128.76, 129.07, 131.07, 136.85, 136.99, 146.25, 152.42, 159.72, 162.53, 164.08, 170.56, 170.65, 170.94; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{44}\text{H}_{44}\text{O}_{12}\text{N}_2$  ( $\text{M} + \text{H}$ ) $^+$  792.2894, found 792.2892.



**6-Methoxy-2-carbomethoxybenzofuran 5'-O-DMT-thymidine *Cis-Syn* Photoproduct**

**(4.2).** To a solution of 0.33 g (0.42 mmol) of compound **4.1** in 10 mL of freshly distilled methanol was added 500  $\mu$ L of DBU under an argon atmosphere. After stirring for 10 min at room temperature the reaction was diluted with ethyl acetate, washed with saturated  $\text{NH}_4\text{Cl}$  and water, and then dried over  $\text{Na}_2\text{SO}_4$ . The yellow solid was one spot by TLC and used in the next reaction without further purification (98%):  $^1\text{H}$  NMR (300 MHz, acetone- $d_6$ )  $\delta$  1.60 (s, 3 H), 1.86-1.92 (m, 1 H), 2.03-2.07 (m, 1 H), 3.16-3.21 (m, 1 H), 3.25-3.30 (m, 1 H), 3.69 (s, 3 H), 3.74 (s, 3 H), 3.79 (s, 6 H), 3.77-3.85 (m, 1 H), 4.02 (s, 1 H), 4.28-4.30 (m, 1 H), 4.67 (d, 1 H,  $J = 2.3$ ), 6.27-6.32 (m, 1 H), 6.44 (dd, 1 H,  $J = 2.6, 8.4$ ), 6.47 (d, 1 H,  $J = 2.7$ ), 6.90-6.98 (m, 5 H), 7.23-7.26 (m, 1 H), 7.35-7.42 (m, 6 H), 7.53-7.56 (m, 2 H), 8.96, (br s, 1 H);  $^{13}\text{C}$  NMR  $\delta$  23.24, 38.07, 47.58, 53.49, 55.64, 55.98, 56.55, 57.48, 57.53, 64.68, 71.33, 83.59, 84.86, 86.93, 88.61, 97.04, 108.85, 114.05, 117.58, 127.69, 128.01, 128.78, 129.21, 131.16, 137.07, 137.12, 146.44, 152.43, 159.73, 162.56, 164.14, 170.71, 171.07; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{42}\text{H}_{42}\text{O}_{11}\text{N}_2$  (M + H) $^+$  750.2789, found 750.2790.





**6-Methoxy-2-carbomethoxybenzofuran 5'-O-DMT-3'-O-(isopropylamine-2-cyanoethyl phosphoramidite)-thymidine *Cis-Syn* Photoproduct (4.3).** To a solution of 0.31 g (0.41 mmol) of compound **4.2** in 4 mL of dry  $\text{CH}_2\text{Cl}_2$  were added 60 mg (0.33 mmol) of bisisopropylammonium tetrazolide and 160  $\mu\text{L}$  (0.50 mmol) of bis(isopropylamine)-2-cyanoethyl phosphoramidite. After stirring for 5 h at room temperature under an argon atmosphere the solvents were removed *in vacuo*. The crude product was purified by silica gel chromatography eluting with 2:1 hexanes/ethyl acetate. The purified product was dissolved in benzene and lyophilized to afford 0.29 (74%) of a white crunchy foam:  $^{31}\text{P}$  NMR (121 MHz, benzene- $d_6$ )  $\delta$  149.00, 149.47; HRMS (FAB $^+$ , 3-NBA) calcd for  $\text{C}_{51}\text{H}_{59}\text{O}_{12}\text{N}_4\text{P}$  ( $\text{M} + \text{H}$ ) $^+$  951.3945, found 951.3941.

**Oligonucleotides.** A 0.1 M solution of the modified phosphoramidite **4.3** was added using the optional port. The coupling yield for the intercalating inhibitor was 85-92%. Oligonucleotides containing the modified base 7-deaza-dG were oxidized as described in the general experimental section. Oligonucleotides were deprotected using standard conditions and purified by C18 reversed phase HPLC. Solvent A: 0.1 M NH<sub>4</sub>OAc in water; solvent B: 0.1 M NH<sub>4</sub>OAc in 1:1 water:CH<sub>3</sub>CN; gradient 10-30% B over 30 min. Oligonucleotides were annealed by heating for 5 min at 80 °C in 100 mM NaHPO<sub>4</sub> (pH 7.0), 100 mM NaCl. The samples were allowed to cool at room temperature for 1 hr and then at 4 °C for 1 hr.

**AFB<sub>1</sub> Epoxide Reaction with DNA.** AFB<sub>1</sub> epoxide was generated as reported.<sup>94</sup> Reactions contained 10 nmol duplex DNA in a total volume of 20 µl (0.5 mM DNA). In a 4 °C room, samples were treated with 1-2.5 molar equivalents of AFB<sub>1</sub> epoxide, mixed by vortexing for 5 min, diluted with buffer, and then the AFB<sub>1</sub>-diol was removed by extraction with CH<sub>2</sub>Cl<sub>2</sub>. Note: Everything contacting AFB<sub>1</sub> or AFB<sub>1</sub> epoxide solutions was treated with bleach to inactivate any residual toxin.

**HPLC Purification of AFB<sub>1</sub> Modified Oligonucleotides.** Samples were loaded onto a C18 reversed-phase analytical column and eluted with a gradient of 10-30% B over 60 min (A= 0.1 M NH<sub>4</sub>OAc in H<sub>2</sub>O; B= 0.1 M NH<sub>4</sub>OAc in 50% H<sub>2</sub>O/acetonitrile) with UV monitoring at both 260 nm and 360 nm. Samples were desalted on a Sep-pak C18 cartridge at 4 °C and eluted with 50% acetonitrile.

**<sup>32</sup>P-Labeling and Cleavage of AFB<sub>1</sub> Treated Oligonucleotides.** Purified oligonucleotides were 5'-end labeled with T4 polynucleotide kinase and γ-<sup>32</sup>P ATP (New England Nuclear, 6000 Ci/mmol) in 70 mM Tris-HCl (pH 7.6) for 5 min at 37 °C. Unincorporated label was removed by centrifugation of the sample through a Sephadex G25 column. Assays to identify the position of adducts formed were based upon the known lability of

AFB<sub>1</sub> adducts to conditions of alkali and heat.<sup>99</sup> One hundred  $\mu$ l of 10% piperidine was added and heated for 30 min at 90 °C. Piperidine was removed by lyophilization, and samples were subsequently electrophoresed through 20% 7 M urea polyacrylamide gels and visualized using a Molecular Dynamics PhosphorImager. The mobilities of bands from samples were compared to those of marker bands generated by the Maxam-Gilbert G-specific reaction.<sup>107</sup>



## References

1. (a) Basu, A. K.; Essigmann, J. M. *Chem. Res. Toxicol.* **1988**, *1*, 1. (b) Baertschi, S. W.; Raney, K. D.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **1988**, *110*, 7929. (c) Bodepudi, V.; Iden, C. R.; Johnson, F. *Nucleosides and Nucleotides* **1991**, *10*, 755. (d) Wood, M. L.; Esteve, A.; Morningstar, M. L.; Kuziemko, G. M.; Essigmann, J. M. *Nucleic Acids Res.* **1992**, *20*, 6023. (e) Sowers, L. C.; Beardsley, G. P. *J. Org. Chem.* **1993**, *58*, 1664. (f) Matray, T. J.; Greenberg, M. M. *J. Am. Chem. Soc.* **1994**, *116*, 6931. (g) Reddy, G. R.; Marnett, L. J. *J. Am. Chem. Soc.* **1995**, *117*, 5007. (h) Manchnda, R.; Dunham, S. U.; Lippard, S. J. *J. Am. Chem. Soc.* **1996**, *118*, 5144. (i) Iwai, S.; Shimizu, M.; Kamiya, H.; Ohtsuka, E. *J. Am. Chem. Soc.* **1996**, *118*, 7642. (j) Sugiyama, H.; Matsuda, S.; Kino, K.; Zhang, Q.-M.; Yonei, S.; Saito, I. *Tetrahedron Lett.* **1996**, *37*, 9067. (k) Chenna, A.; Singer, B. *Chem. Res. Toxicol.* **1997**, *10*, 165.
2. Friedberg, E. C.; Walker, G. C.; Siede, W. In *DNA Repair and Mutagenesis*; ASM Press: Washington, D. C., 1995.
3. (a) Banerjee, S. K.; Christensen, R. B.; Lawrence, C. W.; LeClerc, J. E. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8141. (b) Lawrence, C. W.; Borden, A.; Banerjee, S. K.; LeClerc, J. E. *Nucleic Acids Res.* **1990**, *18*, 2153. (c) Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.* **1991**, *60*, 477. (d) Lambert, I. B.; Napolitano, R. L.; Fuchs, R. P. P. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1310. (e) Rajagopalan, M.; Lu, C.; Woodgate, R.; O'Donnell, M.; Goodman, M. F.; Echols, H. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10777. (f) Horsfall, M. J.; Lawrence, C. W. *J. Mol. Biol.* **1994**, *235*, 465. (g) Napolitano, R. L.; Lambert, I. B.; Fuchs, R. P. P. *Biochemistry* **1994**, *33*, 1311.
4. (a) Fountain, M. A.; Krugh, T. R. *Biochemistry* **1995**, *34*, 3152. (b) Cosman, M.; de los Santos, C.; Fiala, R.; Hingerty, B. E.; Ibanez, V.; Luna, E.; Harvey, R.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1993**, *32*, 4145. (c) Eckel, L. M.; Krugh, T. R. *Biochemistry* **1994**, *33*, 13611. (d) Huang, H.; Zhu, L.; Reid, B. R.; Drobny, G. P.; Hopkins, P. B. *Science* **1995**, *270*, 1842. (e) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. *Nature* **1995**, *377*, 649. (f) Lipscomb, L. A.; Peek, M. E.; Morningstar, M. L.; Verghis, S. M.; Miller, E. M.; Rich, A.; Essigmann, J. M.; Williams, L. D. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 719.
5. Gaillard, P.-H. L.; Moggs, J. G.; Roche, D. M.J.; Quivy, J.-P.; Becker, P. B.; Wood, R. D.; Almouzni, G. *EMBO J.* **1997**, *16*, 6281.
6. Loechler, E. L. *Carcinogenesis* **1996**, *17*, 895.

7. (a) Yarema, K. J.; Lippard, S. J.; Essigmann, J. M. *Nucleic Acids Res.* **1995**, *23*, 4066. (b) Yarema, K. J.; Wilson, J. M.; Lippard, S. J.; Essigmann, J. M. *J. Mol. Biol.* **1994**, *236*, 1034.
8. (a) Pathak, M. A.; Fitzpatrick, T. B. *J. Photochem. Photobiol. B* **1992**, *14*, 3 (b) Scheel, L. D. In *Biochemistry of Some Foodborne Microbial Toxins*; MIT Press: Cambridge, MA, 1967; p 109. (c) Perone, V. B. *Microb. Toxins* **1972**, *8*, 71.
9. (a) Fitzpatrick, J. B.; Stern, R. S.; Parrish, J. A. "Proceedings of the Third International Symposium" in *Psoriasis*. Grune and Stratton, New York, **1982**, pp. 149-156. (b) Parrish, J. A.; Stern, R. S.; Pathak, M. A.; Fitzpatrick, T. B. "Photochemotherapy of skin diseases" in *The Science of Photomedicine*. Plenum Press, New York, **1982**, pp. 595-624. (c) Knobler, R. M.; Hönigsmann, H.; Edelson, R. L. "Psoralen phototherapies" in *Psoralen DNA Photobiology*. CRC Press, Boca Raton, **1988**, Vol. II, pp. 117-134. (d) Edelson, R. L. *Sci. Am.* **1988**, *259*, 68. (e) Scott, D. R.; Pathak, M. A.; Moh, G. R. *Mutat. Res.* **1976**, *39*, 29. (f) Parrish, J. A.; Fitzpatrick, T. B.; Tanenbaum, L.; Pathak, M. A. *N. Engl. J. Med.* **1974**, *291*, 206.
10. Johnston, B. H.; Kung, A. H.; Moore, C. B.; Hearst, J. E. *J. Am. Chem. Soc.* **1981**, *20*, 735.
11. (a) Gamper, H.; Piette, J.; Hearst, J. E. *Photochem. Photobiol.* **1984**, *40*, 29. (b) Tessman, J. W.; Isaacs, S. T.; Hearst, J. E. *Biochemistry* **1985**, *24*, 1669. (c) Gia, O.; Magno, S. M.; Garbesi, A.; Colonna, F. P.; Palumbo, M. *Biochemistry* **1992**, *31*, 11818.
12. (a) Shim, S. C.; Kim, Y. Z. *Photochem. Photobiol.* **1983**, *38*, 265. (b) Kanne, D.; Straub, K.; Hearst, J. E.; Rapoport, H. *J. Am. Chem. Soc.* **1982**, *104*, 6754. (c) Yeung, A. T.; Jones, B. K.; Chu, C. T. *Biochemistry* **1988**, *27*, 3204.
13. (a) Yeung, A. T.; Dinehart, W. J.; Jones, B. K. *Biochemistry* **1988**, *27*, 6332. (b) Shi, Y-B.; Spielmann, H. P.; Hearst, J. E. *Biochemistry*, **1988**, *27*, 5174.
14. Bisagni, E. *J. Photochem. Photobiol. B: Biol.* **1992**, *14*, 23.
15. (a) Späth, E. *Chem. Ber.* **1937**, *70A*, 83. (b) Dean, F. M. In *Naturally Occuring Oxygen Ring Compounds*; Butterworths: London, 1963; pp 176-220. (c) Mustafa, A. In *Furoprans and Furoprones*; Interscience: London, 1967; pp 14-101.
16. (a) Isaacs, S. T.; Shen, C. J.; Hearst, J. E.; Rapoport, H. *Biochemistry* **1977**, *16*, 1058. (b) Worden, L. R.; Kaufman, K. D.; Weis, J. A.; Schaaf, T. K. *J. Org. Chem.* **1969**, *34*, 2311. (c) Zubía, E.; Luis, F. R.; Massanet, G. M.; Collado, I. G. *Tetrahedron* **1992**, *48*, 4239. (d) Queval, P.; Bisagni, E. *Eur. J. Med. Chem.* **1974**, *9*, 335.

17. (a) Straub, K.; Kanne, D.; Hearst, J. E.; Rapoport, H. *J. Am. Chem. Soc.* **1981**, *103*, 2347. (b) Kanne, D.; Straub, K.; Rapoport, H.; Hearst, J. E. *Biochemistry* **1982**, *21*, 861. (c) Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E. *Ann. Rev. Biochem.* **1985**, *54*, 1151. (d) Hearst, J. E.; Isaacs, S. T.; Kanne, D.; Rapoport, H.; Straub, K. *Q. Rev. Biophys.* **1984**, *17*, 1. (e) Peckler, S.; Graves, B.; Kanne, D.; Rapoport, H.; Hearst, J. E.; Kim, S-H. *J. Mol. Biol.* **1982**, *162*, 157. (f) Hearst, J. E. *Chem. Res. Tox.* **1989**, *2*, 69.
18. Reardon, J. T.; Spielmann, P.; Huang, J.-C.; Sastry, S.; Sancar, A.; Hearst, J. E. *Nucleic Acids Res.* **1991**, *19*, 4623.
19. (a) Gervais, J.; Boens, N.; de Schryver, F. C. *Nouv. J. Chim* **1979**, *3*, 163. (b) Gervais, J.; de Schryver, F. C. *Photochem. Photobiol.* **1975**, *21*, 71.
20. (a) Land, E. J.; Rushton, F. A. P.; Beddoes, R. L.; Bruce, J. M.; Cernik, R. J.; Dawson, S. C.; Mills, O. S. *J. Chem. Soc. Chem. Commun.* **1982**, *22*. (b) Cadet, J.; Voituriez, L.; Gaboriau, F.; Vigny, P.; Negra, S. D. *Photochem. Photobiol.* **1983**, *37*, 363.
21. Décout, J.-L.; Lekchiri, Y.; Lhomme, J. *Photochem. Photobiol.* **1992**, *55*, 639.
22. (a) Spielmann, H. P.; Dwyer, T. J.; Sastry, S. S.; Hearst, J. E.; Wemmer, D. E. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2345. (b) Spielmann, H. P.; Dwyer, T. J.; Hearst, J. E.; Wemmer, D. E. *Biochemistry* **1995**, *34*, 12937.
23. Logue, M. W.; Leonard, N. J. *J. Am. Chem. Soc.* **1972**, *94*, 2842.
24. (a) Pretsch, E.; Clerc, T.; Seibl, J.; Simon, W. In *Tables of Spectral Data for Structure Determination of Organic Compounds*; Springer-Verlag: New York, **1989**; p H185. (b) Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. In *Spectrometric Identification of Organic Compounds*; John Wiley & Sons: New York, **1981**; p 209.
25. (a) Barton, D. H. R.; Hewitt, G.; Sammes, P. G. *J. Chem. Soc. (C)* **1969**, *16*. (b) Begley, M. J.; Crombie, L.; Slack, D. A.; Whiting, D. A. *J. Chem. Soc. Perkin Trans. I* **1977**, 2402. (c) Effenberger, F.; Maier, R. *Liebigs Ann. Chem.* **1969**, *729*, 246.
26. Baillargeon, V. P.; Stille, J. K. *J. Am. Chem. Soc.* **1986**, *108*, 452
27. Süsse, M.; Johne, S.; Hesse, M. *Helv. Chim. Acta* **1992**, *75*, 457.
28. Kometani, T.; Watt, D. S.; Ji, T. *Tetrahedron Lett.* **1985**, *26*, 2043.
29. Miknis, G. F.; Williams, R. M. *J. Am. Chem. Soc.* **1993**, *115*, 536.

30. Nakata, T.; Schmid, G.; Vranesic, B.; Okigawa, M.; Smith-Palmer, T.; Kishi, Y. *J. Am. Chem. Soc.* **1978**, *100*, 2933.
31. Semmelhack, M. F.; Schmid, C. R.; Cortés, D. A.; Chou, C. S. *J. Am. Chem. Soc.* **1984**, *106*, 3374.
31. Murray R. D. H.; Méndez, J.; Brown, S. A. In *The Natural Coumarins: Occurrence, Chemistry and Biochemistry*; Wiley: New York, **1982**; pp 131-161.
32. (a) Lakshman, M. K.; Sayer, J. M.; Yagi, H.; Jerina, D. M. *J. Org. Chem.* **1992**, *57*, 4585. (b) Lee, H.; Luna, E.; Hinz, M.; Stezowski, J. J.; Kiselyov, A. S.; Harvey, R. G. *J. Org. Chem.* **1995**, *60*, 5604. (c) Chaturvedi, S.; Lakshman, M. K. *Carcinogenesis* **1996**, *17*, 2747.
33. (a) Walker, G. C.; Uhlenbeck, O. C. *Biochemistry* **1975**, *14*, 817. (b) Brennan, C. A.; Gumpert, R. I. *Nucleic Acids Res.* **1985**, *13*, 8665. (c) Wood, M. L.; Dizdaroglu, M.; Gajewski, E.; Essigmann, J. M. *Biochemistry* **1990**, *29*, 7024.
34. (a) Belguise-Valladier, P.; Fuchs, R. P. *J. Mol. Biol.* **1995**, *249*, 903. (b) Mao, B.; Xu, J.; Li, B.; Margulis, L. A.; Smirnov, S.; Ya, N.-Q.; Courtney, S. H.; Geacintov, N. E. *Carcinogenesis* **1995**, *16*, 357.
35. (a) Millard, J. T.; Raucher, S.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2459. (b) Woo, J.; Sigurdsson, S. T.; Hopkins, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 1199. (c) Kirchner, J. J.; Hopkins, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 4682.
36. (a) Harris, C. M.; Zhou, L.; Strand, E. A.; Harris, T. M. *J. Am. Chem. Soc.* **1991**, *113*, 4328. (b) Kim, S. J.; Stone, M. P.; Harris, C. M.; Harris, T. M. *J. Am. Chem. Soc.* **1992**, *114*, 5480. (c) Xu, Y.-Z.; Zheng, Q.; Swann, P. F. *J. Org. Chem.* **1992**, *57*, 3839. (d) Tsarouhtsis, D.; Kuchimanchi, S.; DeCorte, B. L.; Harris, C. M.; Harris, T. M. *J. Am. Chem. Soc.* **1995**, *117*, 11013. (d) DeCorte, B. L.; Tsarouhtsis, D.; Kuchimanchi, S.; Cooper, M. D.; Horton, P.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **1996**, *9*, 630.
37. (a) Sinden, R. R.; Hagerman, P. J. *Biochemistry* **1984**, *23*, 6299. (b) Cimino, G. D.; Shi, Y.-B.; Hearst, J. E. *Biochemistry* **1986**, *25*, 3013.
38. (a) Spielmann, H. P.; Sastry, S. S.; Hearst, J. E. *Proc. Natl. Sci. USA* **1992**, *89*, 4514. (b) Sastry, S. S.; Spielmann, H. P.; Dwyer, T. J.; Wemmer, D. E.; Hearst, J. E. *J. Photochem. Photobiol. B: Biol.* **1992**, *14*, 65.
39. Tomic, M. T.; Wemmer, D. E.; Kim, S.-H. *Science* **1988**, *238*, 1722.



40. (a) Sancar, A.; Franklin, K. A.; Sancar, G.; Tang, M. S. *J. Mol. Biol.* **1985**, *184*, 725. (b) Yeung, A. T.; Jones, B. K.; Capraro, M.; Chu, T. *Nucleic Acids Res.* **1987**, *15*, 4957. (c) Jones, B. K.; Yeung, A. T. *Proc. Natl. Sci. USA* **1988**, *85*, 8410. (d) Wood, R. D.; Robins, P.; Lindahl, T. *Cell* **1988**, *53*, 97. (e) Ullah-Sibghat, Husain, I.; Carlton, W.; Sancar, A. *Nucleic Acids Res.* **1989**, *17*, 4471.
41. (a) Van Houten, B.; Gamper, H.; Holbrook, S. R.; Hearst, J. E.; Sancar, A. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 8077. (b) Cheng, S.; Sancar, A.; Hearst, J. E. *Nucleic Acids Res.* **1991**, *19*, 657. (c) Cheng, S.; Van Houten, B.; Gamper, H. B.; Sancar, A.; Hearst, J. E. *J. Biol. Chem.* **1988**, *263*, 15110.
42. Sladek, F. M.; Munn, M. M.; Rupp, W. D.; Howard-Flanders, P. *J. Biol. Chem.* **1989**, *264*, 6755.
43. Kuraoka, I.; Kobertz, W. R.; Essigmann, J. M.; Wood, R. D. unpublished results.
44. Jones, B. K.; Yeung, A. T. *J. Biol. Chem.* **1990**, *265*, 3489.
45. Islas, A. L.; Baker, F. J.; Hanawalt, P. C. *Biochemistry* **1994**, *33*, 10794.
46. Vos, J.-M., H. In *DNA Repair Mechanisms: Impact on Human Diseases and Cancer*; R. G. Landes Co.: Austin, TX, 1995; pp 187-218.
47. Thomas, D. C.; Svoboda, D. L.; Vos, J.-M., H.; Kunkel, T. A. *Mol. Cell. Biol.* **1996**, *16*, 2537.
48. Laquerbe, A.; Moustacchi, E.; Papadopoulo, D. *Mutat. Res.* **1995**, *346*, 173.
49. (a) Piette, J.; Decuyper-Debergh, D.; Gamper, H. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 7355. (b) Saffran, W. A.; Cantor, C. R. *J. Mol. Biol.* **1984**, *178*, 595.
50. Piette, J.; Gamper, H. B.; van de Vorst, A.; Hearst, J. E. *Nucleic Acids Res.* **1988**, *16*, 9961.
51. Yatagai, F.; Horsfall, M. J.; Glickman, B. W. *J. Mol. Biol.* **1987**, *194*, 601.
52. (a) Hearst, J. E. *Photobiochem. Photobiophys., Suppl.* **1987**, *23*. (b) Gamper, H. B.; Cimino, G. D.; Isaacs, S. T.; Ferguson, M.; Hearst, J. E. *Nucleic Acids Res.* **1986**, *14*, 9943. (c) Gamper, H. B.; Cimino, G. D.; Hearst, J. E. *J. Mol. Biol.* **1987**, *197*, 349.
53. (a) Sastry, S. S.; Spielmann, H. P.; Hoang, Q. S.; Phillips, A. M.; Sancar, A.; Hearst, J. E. *Biochemistry* **1993**, *32*, 5526. (b) Umlauf, S. W.; Cox, M. M.; Inman, R. B. *J. Biol. Chem.* **1990**, *265*, 16898.

54. (a) Shi, Y.-B.; Gamper, H.; Van Houten, B.; Hearst, J. E. *J. Mol. Biol.* **1988**, *199*, 277. (b) Shi, Y.-B.; Gamper, H.; Hearst, J. E. *J. Biol. Chem.* **1988**, *263*, 527. (c) Sastry, S. S.; Hearst, J. E. *J. Mol. Biol.* **1991**, *221*, 1091. (d) Sastry, S. S.; Hearst, J. E. *J. Mol. Biol.* **1991**, *221*, 1111.
55. Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K. *Can. J. Chem.* **1982**, *60*, 1106.
56. (a) Miller, S. S.; Eisenstadt, E. *J. Bacteriol.* **1987**, *169*, 2724. (b) Gunther, E. J.; Yeasky, T. M.; Gasparro, F. P.; Glazer, P. M. *Cancer Res.* **1995**, *55*, 1283. (c) Yang, S. C.; Lin, J. G.; Chiou, C. C.; Chen, L. Y.; Yang, J. L. *Carcinogenesis* **1994**, *15*, 201.
57. Xu, Y.-Z.; Swann, P. F. *Nucleic Acids Research* **1990**, *18*, 4061.
58. McEvoy G. K. In *AHFS Drug Information 94*; American Society of Hospital Pharmacists: 1994; pp 572-580.
59. Lippard, S. J. In *The Robert A. Welch Foundation 37th Conference on Chemical Research*; The Robert A. Welch Foundation: 1993; pp 49-60.
60. Toney, J. H.; Donahue, B. A.; Kellett, P. J.; Bruhn, S. L.; Essigmann, J. M.; Lippard, S. J. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8328.
61. (a) Huang, J.-C.; Zamble, D. B.; Reardon, J. T.; Lippard, S. J.; Sancar, A. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10394. (b) Zamble, D. B.; Lippard, S. J. *Trends Biochem. Sci.* **1995**, *20*, 435. (c) Zamble, D. B.; Mu, D.; Reardon, J. T.; Sancar, A.; Lippard, S. J. *Biochemistry* **1996**, *35*, 10004. (d) Szymkowski, D. E.; Yarema, K. Y.; Essigmann, J. M.; Lippard, S. J.; Wood, R. D. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10772.
62. Rink, S. M.; Yarema, K. J.; Solomon, M. S.; Paige, L. A.; Tadayoni-Rebek, B. M.; Essigmann, J. M.; Croy, R. G. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15063.
63. (a) Ferno, M.; Johansson, B. U.; Olsson, N. H.; Ryden, S.; Sellberg, G. *Acta Oncol.* **1990**, *29*, 129. (b) Slotman, B. J.; Rao, B. R. *Anticancer Res.* **1988**, *8*, 417.
64. (a) Katzenellenbogen, J. A. *J. Nucl. Med.* **1995**, *36*, Suppl. 6, 8s. (b) Katzenellenbogen, B. S.; Montano, M. M.; Le Goff, P.; Schodin, D. J.; Kraus, W. L.; Bhardwaj, B.; Fujimoto, N. *J. Steroid Biochem. Mol. Biol.* **1995**, *53*, 387.
65. (a) Wood, R. D.; Coverley, D. *Bioessays* **1991**, *13*, 447. (b) Wood, R. D. *Annu. Rev. Biochem.* **1996**, *65*, 135.

66. (a) Recombinant human estrogen receptor (ER) is commercially available from PanVera Corporation, 545 Science Dr., Madison, WI 53711. (b) Arnold, S. F.; Obourn, J. D.; Jaffe, H.; Notides, A. C. *Mol. Endocrinol.* **1995**, *9*, 24.
67. Brooks, S. C.; Locke, E. R.; Soule, H. D. *J. Biol. Chem.* **1973**, *248*, 6251. (b) Cailleau, R.; Joug, R.; Olive, M.; Reeves, W. J. Jr. *J. Natl. Cancer Inst.* **1974**, *53*, 661.
68. von Angerer, E.; Prekajac, J.; Strohmeier, J. *J. Med. Chem.* **1984**, *27*, 1439.
69. (a) Jencks, W. P. *Prog. Phys. Org. Chem.* **1964**, *2*, 63. (b) Jencks, W. P. *J. Am. Chem. Soc.* **1959**, *81*, 475. (c) Reeves, R. L. In *The Chemistry of the Carbonyl Group*; Interscience: London, 1966; pp 600-614. (d) Ratner, S.; Clarke, H. T. *J. Am. Chem. Soc.* **1937**, *59*, 200.
70. (a) Liu, C.-F.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 307. (b) Shao, J.; Tam, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3893. (c) Canne, L. E.; Ferré-D'Amaré, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998. (d) Liu, C.-F.; Tam, J. P. *Proc. Natl. Acad. USA* **1994**, *91*, 6584. (e) Liu, C.-F.; Tam, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 4149. (f) Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30. Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138.
71. Martin, D. P.; Bibart, R. T.; Drueckhammer, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 4660.
72. (a) Singer, B.; Kuśmierk, J. T. *Ann. Rev. Biochem.* **1982**, *52*, 655. (b) Fraenkel-Conrat, H.; Singer, B. *Biochim. Biophys. Acta* **1972**, *262*, 264.
73. Lane, C. F. *Synthesis* **1975**, 135.
74. Korenman, S. G. *Endocrinology* **1970**, *87*, 1119.
75. Moggs, J. G.; Yarema, K. J.; Essigmann, J. M.; Wood, R. D. *J. Biol. Chem.* **1996**, *271*, 7177.
76. Wood, R. D.; Biggerstaff, M.; Shivji, M. K. K. *Methods*, **1995**, *7*, 163.
77. Ariza, R. R.; Keyse, S. M.; Moggs, J. G.; Wood, R. D. *Nucleic Acids Res.* **1996**, *24*, 433.
78. (a) Available from PanVera Corporation, 545 Science Dr., Madison, WI 53711. (b) Obourn, J. D.; Koszewski, N. J.; Notides, A. C. *Biochemistry* **1993**, *32*, 6229.
79. (a) Mahal, L. K.; Yareman, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125. (b) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150.

80. (a) Witkop, B. *Photochem. Photobiol.* **1968**, *7*, 813. (b) Kunieda, T.; Witkop, B. *J. Am. Chem. Soc.* **1971**, *93*, 3493.
81. Creighton, T. E. *In Proteins 2nd Ed.*; W. H. Freeman: New York, 1993; p 338.
82. Hess, M. T.; Schwitter, U.; Petretta, M.; Giese, B.; Naegeli, H. *Chemistry & Biology* **1996**, *3*, 121.
83. Moggs, J. G.; Szymkowski, D. E.; Yamada, M.; Karran, P.; Wood, R. D. *Nucleic Acids Res.* **1997**, *25*, 480.
84. (a) Dodson, M. L.; Schrock, R. D. III; Lloyd, R. S. *Biochemistry* **1993**, *32*, 8284. (b) Mazumder, A.; Gerlt, J. A.; Absalon, M. J.; Stubbe, J.; Cunningham, R. P.; Withka, J.; Bolton, P. H. *Biochemistry* **1991**, *30*, 1119.
85. Sargeant, K.; Sheridan, A.; O'Kelly, J.; Carnaghan, R. B. A. *Nature* **1961**, *192*, 1096.
86. Busby, W. F. Jr.; Wogan, G. N. *In Chemical Carcinogens*; Washington, DC: American Chemical Society, 1984; pp 945-1136.
87. Qian, G. S.; Ross, R. K.; Yu, M. C.; Yuan, J. M.; Gao, Y. T.; Henderson, B. E.; Wogan, G. N.; Groopman, J. D. *Cancer Epidemiology, Biomarkers and Prevention* **1994**, *3*, 3.
88. Croy, R. G.; Essigmann, J. M.; Reinhold, V. N.; Wogan, G. N. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 1745.
89. Wang, T. V.; Cerutti, P. *Biochemistry* **1980**, *19*, 1692.
90. Hertzog, P. J.; Smith, J. R. L.; Garner, R. C. *Carcinogenesis* **1982**, *3*, 723.
91. (a) Hertzog, P. J.; Smith, J. R. L.; Garner, R. C. *Carcinogenesis* **1980**, *1*, 787. (b) Groopman, J. D.; Croy, R. G.; Wogan, G. N. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 5445.
92. (a) Foster, P. L.; Eisenstadt, E.; Miller, J. H. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 2695. (b) Bailey, E. A.; Iyer, R. S.; Stone, M. P.; Harris, T. M.; Essigmann, J. M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1535.
93. (a) Greenblatt, M. S.; Bennett, W. P.; Hollstein, M.; Harris, C. C. *Cancer Res.* **1994**, *54*, 4855. (b) Hsu, I. C.; Metcalf, R. A.; Sun, T.; Welsh, J. A.; Wang, N. J.; Harris, C. C. *Nature* **1991**, *350*, 427. (c) Bressac, B.; Kew, M.; Wands, J.; Ozturk, M. *Nature* **1991**, *350*, 429.
94. Gopalakrishnan, S.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **1989**, *111*, 7232.

95. (a) Misra, R. P.; Muench, K. F.; Humayun, M. Z. *Biochemistry* **1983**, *22*, 3351. (b) Raney, V. M.; Harris, T. M.; Stone, M. P. *Chem. Res. Toxicol.* **1993**, *6*, 64.
96. (a) Gopalakrishnan, S.; Harris, T. M.; Stone, M. P. *Biochemistry* **1990**, *29*, 10438. (b) Iyer, R. S.; Coles, B. F.; Raney, K. D.; Thier, R.; Guengerich, F. P.; Harris, T. M. *J. Am. Chem. Soc.* **1994**, *116*, 1603.
97. Raney, K. D.; Gopalakrishnan, S.; Byrd, S.; Stone, M. P.; Harris, T. M. *Chem. Res. Toxicol.* **1990**, *3*, 254.
98. Johnson, W. W.; Guengerich, F. P. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6121.
99. D'Andrea, A. D.; Haseltine, W. A. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4120.
100. Benasutti, M.; Ejadi, S.; Whitlow, M. D.; Loechler, E. L. *Biochemistry* **1988**, *27*, 472.
101. (a) Crothers, D. M. *Biopolymers* **1968**, *6*, 575. (b) Wilson, W. D. In *Nucleic Acids in Chemistry and Biology*; Oxford Univ. Press: New York, 1990; pp 295-336.
102. Johnson, W. W.; Harris, T. M.; Guengerich, F. P. *J. Am. Chem. Soc.* **1996**, *118*, 8213.
103. Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Rice, J. A.; Broka, C.; Dallas, J.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 437
104. Denissenko, M. F.; Chen, J. X.; Tang, M. Pfeifer, G. P. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3893.
105. (a) Wakelin, L. P. G.; Romanos, M.; Chen, T. K.; Glaubiger, D.; Canellakis, E. S.; Waring, M. J. *Biochemistry* **1978**, *17*, 5057. (b) Robledo-Luiggi, C.; Wilson, W. D.; Pares, E.; Vera, M.; Martinez, C. S.; Santiago, D. *Biopolymers* **1991**, *31*, 907.
106. Singer, B.; Essigmann, J. M. *Carcinogenesis* **1991**, *12*, 949.
107. Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Lab. Press: Plainview, NY, 2nd Ed., 1989.

