PSORALENS AS PHOTOACTIVE PROBES OF NUCLEIC ACID STRUCTURE AND FUNCTION:
ORGANIC CHEMISTRY, PHOTOCHEMISTRY, AND BIOCHEMISTRY

George D. Cimino, Howard B. Gamper, Stephen T. Isaacs, and John E. Hearst

Department of Chemistry, University of California, Berkeley, California 94720

CONTENTS

PERSPECTIVES AND SUMMARY .............................................................. 1152
INTRODUCTION ..................................................................................... 1155
ORGANIC AND STRUCTURAL CHEMISTRY OF PSORALEN
AND ITS ADDUCTS ............................................................................. 1158
Naturally Occurring Psoralens .................................................. 1158
Synthesis and Radiolabeling of Psoralens .......................... 1158
Monofunctional Psoralens ...................................................... 1160
Relative Reactivity of Different Psoralens with Nucleic Acids .... 1161
Psoralen Monoadducts and Diadducts Formed With DNA and RNA 1162
PHOTOCHEMICAL PROPERTIES OF PSORALENS .................... 1165
Spectroscopic Properties of Psoralen Molecules ............... 1166
Reactive States for Monoadduct and Crosslink Formation in Polynucleotides 1170
Controlled Production of Monoadducts and Diadducts in Double-Stranded Nucleic Acids 1172
THE USE OF PSORALENS AS PROBES FOR NUCLEIC ACID
STRUCTURE AND FUNCTION .......................................................... 1172
Chromatin Structure ............................................................... 1172
Cruciforms .................................................................................. 1175
Torsional Tension .................................................................. 1176
Secondary Structure in Single-Stranded Nucleic Acid .... 1177
Stabilization of Nucleic Acid Complexes ......................... 1184
FUTURE DIRECTIONS .............................................................. 1186

1151
Psoralens comprise the most important class of photochemical reagents for the investigation of nucleic acid structure and function. They have been used for determining the structure of both DNA and RNA in viral, bacterial, and mammalian systems, and also for studying functional questions, such as the role of the small nuclear RNAs in processing heteronuclear RNA. A list of some of the major applications of these compounds during the last ten years is presented in Table 1.

Psoralens are unique in their ability to freeze helical regions of nucleic acid. Psoralens react with DNA and RNA by a two-step mechanism. First, the planar psoralen molecule intercalates within a double helical region of nucleic acid. Covalent addition of the psoralen is effected by controlled irradiation into an absorption band of the psoralen molecule. Stable, but photoreversible, covalent adducts form with pyrimidine bases at one or both ends of the psoralen molecule. By forming covalent crosslinks with base-paired structures, psoralens can probe both static and dynamic structural features. Psoralens can trap long-range interactions which are in dynamic equilibrium. This allows both the occurrence of the interaction to be established and its position within the structure to be mapped. Psoralens can also be used temporally, such as in following the fate of short-lived nucleic acid species in vivo.

The details of the interaction between psoralens and nucleic acid are well understood at the molecular level. The structure of the psoralen adducts formed with DNA have been determined, the polarity of the reaction which converts monoadduct to crosslink established, and methods for the exclusive formation of monoaddition products worked out. This advanced state of chemical control makes the psoralens extremely versatile reagents. As more information is compiled about structure-activity parameters, a fine tuning of the reaction of psoralens with nucleic acid will be realized.

Future application of psoralens for investigating nucleic acid structure and function will be aided by the following developments. The preparation of hybridization probes which carry psoralen monoadducts is currently under way. These probes will be used to form covalent hybrids for locating particular sequences and also for site-specific placement of psoralen monoadducts in nucleic acid structures via photochemical transfer of the psoralen. The transferred monoadduct will be used for fixation of “dynamic” base paired intrastructural conformations by crosslink formation. Chemical schemes for the site specific cleavage of DNA and RNA at the position of psoralen addition are also being developed. These procedures will allow for the direct mapping of secondary structure at the position of crosslink formation. Finally, many new psoralen derivatives are being synthesized for specific applications such as site-directed crosslinking of DNA and protein-nucleic acid crosslinking. Psor-
**Table 1** Publications employing psoralen as a probe for nucleic acid structure and function

<table>
<thead>
<tr>
<th>Subject Study</th>
<th>Author</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromatin Structure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em> nuclei</td>
<td>Hanson et al</td>
<td>1976</td>
<td>34</td>
</tr>
<tr>
<td>Mouse liver nuclei</td>
<td>Cech &amp; Pardue</td>
<td>1977</td>
<td>143</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em> nuclei</td>
<td>Wiesehahn et al</td>
<td>1977</td>
<td>39</td>
</tr>
<tr>
<td>Mouse liver nuclei</td>
<td>Cech et al</td>
<td>1977</td>
<td>144</td>
</tr>
<tr>
<td>Simian virus 40 minichromosomes</td>
<td>Hallick et al</td>
<td>1978</td>
<td>145</td>
</tr>
<tr>
<td>Main-band and satellite DNAs in <em>Drosophila melanogaster</em> nuclei</td>
<td>Shen &amp; Hearst</td>
<td>1978</td>
<td>146</td>
</tr>
<tr>
<td>Ribosomal RNA genes in <em>Tetrahymena thermophila</em> cells</td>
<td>Cech &amp; Karrer</td>
<td>1980</td>
<td>151</td>
</tr>
<tr>
<td>Mitochondrial DNA in <em>Drosophila melanogaster</em> cells</td>
<td>Potter et al</td>
<td>1980</td>
<td>40</td>
</tr>
<tr>
<td><em>Escherichia coli</em> DNA</td>
<td>Hallick et al</td>
<td>1980</td>
<td>153</td>
</tr>
<tr>
<td>Mitochondrial DNA in HeLa cells</td>
<td>De Francesco &amp; Attardi</td>
<td>1981</td>
<td>152</td>
</tr>
<tr>
<td>Simian virus 40 DNA in lytically infected CV-1 cells</td>
<td>Carlson et al</td>
<td>1982</td>
<td>148</td>
</tr>
<tr>
<td>Simian virus 40 DNA in lytically infected CV-1 cells</td>
<td>Robinson &amp; Hallick</td>
<td>1982</td>
<td>149</td>
</tr>
<tr>
<td>Simian virus 40 DNA in the virion</td>
<td>Kondoleon et al</td>
<td>1983</td>
<td>150</td>
</tr>
<tr>
<td>Calf thymus chromatin</td>
<td>Mitra et al</td>
<td>1984</td>
<td>191</td>
</tr>
<tr>
<td>Cruciforms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse tissue culture cells</td>
<td>Cech &amp; Pardue</td>
<td>1976</td>
<td>154</td>
</tr>
<tr>
<td>Ribosomal RNA genes in <em>Tetrahymena thermophila</em> cells</td>
<td>Cech &amp; Karrer</td>
<td>1980</td>
<td>151</td>
</tr>
<tr>
<td>Lac operator DNA in <em>Escherichia coli</em> cells</td>
<td>Sinden et al</td>
<td>1983</td>
<td>155</td>
</tr>
<tr>
<td>Torsional Tension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prokaryotic and eukaryotic cells</td>
<td>Sinden et al</td>
<td>1980</td>
<td>156</td>
</tr>
<tr>
<td>Domains of supercoiling</td>
<td>Sinden &amp; Pettijohn</td>
<td>1981</td>
<td>192</td>
</tr>
<tr>
<td>Secondary Structure in Single-Stranded DNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteriophage fd DNA in solution</td>
<td>Shen &amp; Hearst</td>
<td>1976</td>
<td>157</td>
</tr>
<tr>
<td>Denatured simian virus 40 DNA in solution</td>
<td>Shen &amp; Hearst</td>
<td>1977</td>
<td>160</td>
</tr>
<tr>
<td>Bacteriophage fd DNA in the virion</td>
<td>Shen et al</td>
<td>1979</td>
<td>158</td>
</tr>
<tr>
<td>Denatured simian virus 40 DNA in solution</td>
<td>Shen &amp; Hearst</td>
<td>1979</td>
<td>161</td>
</tr>
<tr>
<td>Chimeric phage M13Gor1 in the virion</td>
<td>Ikoku &amp; Hearst</td>
<td>1981</td>
<td>159</td>
</tr>
<tr>
<td>Secondary Structure in Ribosomal RNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em> 18S and 26S RNA in solution</td>
<td>Wollenzien et al</td>
<td>1978</td>
<td>35</td>
</tr>
<tr>
<td><em>E. coli</em> 16S RNA in solution</td>
<td>Wollenzien &amp; Hearst</td>
<td>1979</td>
<td>162</td>
</tr>
</tbody>
</table>
Table 1 (continued)

<table>
<thead>
<tr>
<th>Subject Study</th>
<th>Author</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> 16S RNA within the 30S subunit</td>
<td>Thammana et al</td>
<td>1979</td>
<td>163</td>
</tr>
<tr>
<td>Polarity of <em>E. coli</em> crosslinked 16S RNA</td>
<td>Wollenzien et al</td>
<td>1979</td>
<td>164</td>
</tr>
<tr>
<td><em>E. coli</em> 5S RNA in solution</td>
<td>Rabin &amp; Crothers</td>
<td>1979</td>
<td>33</td>
</tr>
<tr>
<td>Sequence analysis of <em>E. coli</em> crosslinked 16S RNA</td>
<td>Cantor et al</td>
<td>1980</td>
<td>165</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em> 5S RNA in solution</td>
<td>Thompson et al</td>
<td>1981</td>
<td>168</td>
</tr>
<tr>
<td>Electrophoretic separation of <em>E. coli</em> crosslinked 16S RNA</td>
<td>Wollenzien &amp; Cantor</td>
<td>1982</td>
<td>166</td>
</tr>
<tr>
<td><em>E. coli</em> 16S RNA in solution</td>
<td>Turner et al</td>
<td>1982</td>
<td>29</td>
</tr>
<tr>
<td>Comparison of <em>E. coli</em> 16S RNA in active and inactive 30S subunits</td>
<td>Chu et al</td>
<td>1983</td>
<td>167</td>
</tr>
<tr>
<td><em>E. coli</em> 16S RNA in solution</td>
<td>Thompson &amp; Hearst</td>
<td>1983</td>
<td>30</td>
</tr>
<tr>
<td>Functional significance of long range crosslinks in <em>E. coli</em> 16S RNA</td>
<td>Thompson &amp; Hearst</td>
<td>1983</td>
<td>171</td>
</tr>
<tr>
<td><em>E. coli</em> 23S RNA in solution</td>
<td>Turner &amp; Noller</td>
<td>1983</td>
<td>193</td>
</tr>
</tbody>
</table>

Secondary Structure in Transfer RNA

| *E. coli* phenylalanine tRNA in solution                                     | Bachellerie & Hearst        | 1982 | 172  |
| Yeast phenylalanine tRNA in solution                                       | Garrett-Wheeler et al       | 1984 | 37   |

Secondary Structure in Heterogeneous and Small Nuclear RNAs

| Heterogeneous nuclear RNA in HeLa cell nuclei                              | Calvet & Pederson           | 1979 | 173  |
| Heterogeneous nuclear RNA in HeLa cells                                    | Calvet & Pederson           | 1979 | 174  |
| Complex between U1 RNA and heterogeneous nuclear RNA in HeLa cells         | Calvet & Pederson           | 1981 | 176  |
| Complex between U2 RNA and heterogeneous nuclear RNA in HeLa cells         | Calvet et al                | 1982 | 177  |
| Neurospora mitochondrial 35S precursor rRNA within ribonucleoprotein particles | Wollenzien et al            | 1983 | 175  |
| Ribonucleoprotein organization of the U1 RNA complex with heterogeneous nuclear RNA in HeLa cells | Setyono & Pederson | 1984 | 178  |

Secondary Structure in Viral RNA Genomes

| Reovirus RNA genome within the virion                                      | Nakashima & Shatkin         | 1978 | 194  |
| Rous sarcoma RNA genome within the virion                                  | Swanstrom et al             | 1981 | 195  |

Fixation of Nucleic Acid Complexes

| Ternary transcription complexes                                           | Shen & Hearst               | 1978 | 180  |
| R-loops                                                                    | Kaback et al                | 1979 | 182  |
| R-loops                                                                    | Wittig & Wittig             | 1979 | 32   |
Table 1 (continued)

<table>
<thead>
<tr>
<th>Subject Study</th>
<th>Author</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-loops</td>
<td>DeFrancesco &amp; Attardi</td>
<td>1981</td>
<td>152</td>
</tr>
<tr>
<td>Adenovirus 5 replicative intermediates</td>
<td>Revet &amp; Benichou</td>
<td>1981</td>
<td>181</td>
</tr>
<tr>
<td>R-loops</td>
<td>Chatterjee &amp; Cantor</td>
<td>1982</td>
<td>183</td>
</tr>
<tr>
<td>M13 hybridization probes</td>
<td>Brown et al</td>
<td>1982</td>
<td>185</td>
</tr>
<tr>
<td>Determination of the polarity of single-stranded RNA in the electron microscope</td>
<td>Wollenzien &amp; Cantor</td>
<td>1982</td>
<td>186</td>
</tr>
</tbody>
</table>

Tertiary Interactions in Nucleoprotein Complexes

| DNA-DNA interactions in λ bacteriophage           | Haas et al              | 1982 | 87   |
| DNA-DNA and DNA-protein interactions in bacteriophage λ | Schwartz et al       | 1983 | 86   |

Psoralen analogs which will crosslink purine to pyrimidine and purine to purine are also being considered.

It is not the intent of this review to catalog the properties and applications of every psoralen derivative known. Rather we try to show how a basic understanding of the organic chemistry, photochemistry, and biochemistry of these compounds has produced a versatile molecular tool for the elucidation of nucleic acid structure and function. The use of psoralens for the determination of nucleic acid secondary structure will be emphasized here. Recent reviews include coverage of other aspects of psoralens including clinical applications (1, 2), mutagenesis, toxicity and repair (3), and photochemistry and photobiology (4–6).

INTRODUCTION

Psoralens are bifunctional photoreagents which form covalent bonds with the pyrimidine bases of nucleic acids (4, 7). Structurally, psoralens are tricyclic compounds formed by the linear fusion of a furan ring with a coumarin. An angular fusion of the two-ring systems forms an isopsoralen, which is also known as angelicin. The structures of psoralen, isopsoralen, and several psoralen derivatives are shown in Figure 1. Both psoralens and isopsoralens can intercalate into double-stranded nucleic acid and undergo covalent photocycloaddition at either the furan or coumarin and (8, 9). The photoaddition to nucleic acid occurs with incident light of wavelength of 320–400 nm, a region of the electromagnetic spectrum to which nucleic acids are transparent. Binding of both ends of a psoralen to opposite strands of a nucleic acid helix
results in the formation of a covalent interstrand crosslink (10). Isopsoralens, due to their angular geometry, cannot form crosslinks with DNA in dilute aqueous solutions; however, they are reported to crosslink certain types of folded DNA (11-14). Psoralens also react with single-stranded nucleic acid, but to a much smaller extent than with double-stranded structures (15). Psoralens have some reactivity with proteins (16-20) and lipid membranes (21); however, these reactivities are minor compared to the reaction with nucleic acid.

Psoralen reacts primarily with thymidine in DNA and uridine in RNA, although a minor reaction with cytosine also occurs. There is a single report in the literature of 8-methoxypsoralen reacting with adenosine (22), but this is a very minor reaction compared to the reaction with the pyrimidines. Aside from the alteration which occurs at the site of the chemical modification, there is no additional degradation of either RNA or DNA associated with the photochemistry provided that it is carried out under anoxygenic conditions. If singlet oxygen is produced during the photochemistry it will bring about degradation of all biological molecules with which it comes in contact (23, 24).

Isaacs et al (25) proposed a mechanism for the reaction of the psoralens with nucleic acid helices:

\[ P + S \overset{\text{PS}}{\rightleftharpoons} \]
\[ PS + h\nu \rightarrow A \]
\[ A + h\nu \rightarrow X \]
\[ P + h\nu \rightarrow B \]

Figure 1 The structure and numbering system used for psoralen, isopsoralen (angelicin), 8-methoxypsoralen (8-MOP), 4,5',8-trimethylpsoralen (TMP), 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) and 4'aminomethyl-4,5',8-trimethylpsoralen (AMT).
where $P$ is the psoralen derivative in question, $S$ is a psoralen intercalation site in a nucleic acid helix, $PS$ is the noncovalent intercalation complex between psoralen and the DNA or RNA site, $A$ refers to the covalent monoadduct of the psoralen to the nucleic acid, $X$ refers to the covalent crosslink in the nucleic acid helix, $B$ represents photobreakdown products of psoralen, and $hv$ is a photon of light.

The versatility of psoralen photochemistry allows for control of the degree of reaction by light dose, control of the ratio of monoaddition to crosslinkage by either selection of suitable wavelengths to induce the chemistry (26, 27), or by controlled timing of the delivery of the light (28). Furthermore, the monoadducts and diadducts (crosslinks) formed between psoralen and pyrimidines can be reversed with short wavelength ultraviolet light. This property is used to great advantage in the determination of nucleic acid secondary structure using psoralens (29, 30).

The monoadduct and crosslink are chemically stable, allowing for analysis under varied chemical conditions. The photochemistry can be carried out under a wide variety of conditions including broad ranges of temperature and ionic strength, in the presence or absence of divalent cations (31), and even in some organic solvents (32). The positions of the covalent crosslinks can therefore be enzymatically mapped (33). They can also be mapped by visualization in the electron microscope when the DNA or RNA is spread under denaturing conditions (34, 35). Enzymatic methods for locating monoadducts in nucleic acids sequences have been developed (36, 37), and chemical cleavage methods are sure to follow shortly (38). Since proteins that bind DNA such as nucleosomes protect the DNA from reaction with psoralen, these mapping techniques provide information about protein–nucleic acid interactions as well (39, 40). In such experiments a very high level of chemical substitution of the DNA with psoralen can be achieved.

Psoralens are used in a number of diverse ways. Clinically, psoralens are used for the treatment of psoriasis (41), vitiligo (42), and other skin disorders (43). Extensive studies on the metabolism of these compounds have been made (44). As irreversible nucleic acid specific crosslinking reagents, psoralens are used to inactivate viruses and other pathogens with little disruption of proteinaceous antigenic structures. Many viruses that cannot be otherwise inactivated for vaccine production are effectively inactivated by psoralen photochemistry (45, 46). As ligands that bind stereospecifically to nucleic acid, psoralens form a limited set of known and characterized adducts, which make them extremely useful as model compounds for mutagenesis and repair studies (47–51). The selective reaction of psoralens with nucleic acids, which is most efficient in helical regions, allows one to probe both static and dynamic nucleic acid structure both in vitro and in vivo. The psoralens are able to penetrate most biological structures and are not highly toxic to cells in the absence of actinic
light, so they are ideal probes for nucleic acid structure wherever long wave ultraviolet light can be delivered. Both DNA and RNA structure have been probed by psoralen crosslinking (see Table 1).

Control of the crosslinking reaction at each step in the psoralen-nucleic acid interaction is possible by molecular architecture. Since the properties of a given psoralen derivative are directly a function of its structure, control at the level of chemical synthesis is possible by judicious choice of substituent, with respect to both type and placement on the tricyclic ring system (52, 53). As more information is compiled about structure-activity parameters for both the dark binding and photoaddition of psoralens to nucleic acid, a fine tuning of the interaction for the study of nucleic acid will become possible.

ORGANIC AND STRUCTURAL CHEMISTRY OF PSORALEN AND ITS ADDUCTS

Naturally Occurring Psoralens

Both psoralens and isopsoralens are natural products found in plants. They are most abundant in the **Umbelliferae**, **Rutacea**, and **Leguminosae** families (54). Several dozen psoralens have been characterized from natural sources (55); some of these are thought to act as natural insecticides (56). Psoralens have also been isolated from microorganisms including fungi. The two most widely used psoralens are 8-methoxypsoralen, isolated from *Ammi Majus*, and 4,5',8-trimethylpsoralen, derived from a fungus which grows on diseased celery (57). Both of these compounds are readily prepared by chemical synthesis (58–60).

Synthesis and Radiolabeling of Psoralens

The synthetic chemistry which has been developed falls into two areas: (a) preparation of the psoralen tricyclic ring system which is built from resorcinol, and (b) exocyclic modification of the intact ring system. The combined synthetic procedures provide different psoralen derivatives with particular characteristics for different applications.

Ring synthesis of both psoralen and isopsoralen (angelicin) proceeds by converting resorcinol to either a coumarin or benzofuran then adding the furan or pyrone ring respectively. The final substituents at the 5 and 8 positions of the psoralen or isopsoralen are those present on the resorcinol used in the initial step. When proceeding from the coumarin, the presence or absence of substituents at the 6 and 8 position of the coumarin determine if a psoralen, isopsoralen, or a mixture of the two is obtained (61–63). Substituents at the 8 position of the coumarin direct furan ring closure to the 6 position giving a psoralen, while substituents at the 6 position of the coumarin have just the opposite effect and are precursors of isopsoralens. When neither the 6 or 8
PSORALENS AS PROBE OF NUCLEIC ACIDS

position of the coumarin is substituted, a mixture of psoralen and isopsoralen usually results.

Proceeding from benzofuran, the key synthetic intermediate is 6-acetoxycoumaran (64–67). Psoralens that contain a substituent on the 3 but not the 4 carbon, such as 3-carbethoxypsoralen (68) and 3-methylpsoralen (69) are prepared by this method. The more common procedure, however, is to first synthesize the coumarin and then add the furan ring. Here resorcinol is treated with a β-ketoester giving a coumarin via the von Pechmann condensation. The substituents at the 3 and 4 positions are controlled by which β-ketoester is used for the reaction. Psoralens such as 4,5′,8-trimethylpsoralen (60) and 3-n-butyl-4,5′,8-trimethylpsoralen (70) are synthesized by this procedure. A number of methods have been developed for the addition of the furan ring to the coumarin. Psoralens and isopsoralens containing alkyl groups at the 4′ or 5′ positions (60, 63, 71) or with an unsubstituted furan ring have been prepared (72, 73).

Exocyclic modification of the intact psoralen ring system is used to fine tune the molecular characteristics of the molecule. Psoralens undergo a variety of substitution and addition reactions such as nitration, halogenation, quinone formation, catalytic reduction, chlorosulfonation, and chloromethylation (25, 52, 73–81). Psoralen derivatives substituted at the 8 position with aminomethyl, aminooethyl, hydroxymethyl, and hydroxyethyl substituents have been prepared by manipulation of a family of 8-acetylpsoralens (63).

Chloromethylation has been a singularly valuable reaction for providing a synthetic handle for the preparation of a whole family of new psoralen derivatives. Among these are the highly water soluble derivatives 4′-aminomethyl-4,5′,8-trimethylpsoralen (25), 5-aminomethyl-8-methoxypsoralen (82), 5′-aminomethyl-4,4′,8-trimethylpsoralen (83), and 4′-aminomethyl-4,5′-dimethylisopsoralen (84). Specialized psoralens for site directed crosslinking of DNA (85), for protein-nucleic acid crosslinking (86), and for investigating the packaging of phage DNA (87) have also been synthesized starting with 4′-chloromethyl-4,5′,8-trimethylpsoralen.

The use of psoralens for determining nucleic acid structure often requires that the psoralen be radiolabeled. Both tritium and carbon-14 labeled psoralens have been synthesized by routine synthetic manipulation to obtain radiolabeled derivatives. Methods used include exchange reactions with tritium gas (88, 89) and tritiated water (25), catalytic reduction (80, 81) and hydrogenolysis with tritium gas (90), alkylation with tritium (80, 81) or carbon-14 labeled alkyl halides, borotritide reductions of aldehydes and ketones (80), and various ring syntheses with carbon-14 labeled precursors (47). The exchange procedures typically give low specific activities (50–300 mCi/mmol) and do not site-specifically label the molecule. Higher specific activities (1–20 Ci/mmol) are obtained with site-specific reduction or alkylation. The position of the label is
often important, such as in metabolism studies where the structure is sequentially degraded.

**Monofunctional Psoralens**

There are a variety of uses for psoralen derivatives which only form monoadducts with helical nucleic acid (91). The debate over the relative mutagenicity and/or carcinogenicity of monoadducts vs crosslinks has resulted in the clinical community looking closely at psoralens which form only monoadducts for use in photochemotherapy (92). Investigators studying DNA repair have used monofunctional vs bifunctional psoralens extensively in order to address cellular response to the two different kinds of damage (47–50). The approach has been to use compounds that due to their geometry cannot form crosslinks, such as the isopsoralens, or to use psoralens which have been modified by substituents that limit the reactivity to one end of the molecule. In this second approach, substituents that exert either steric or electronic effects have been used.

Isopsoralens are reported not to crosslink double-stranded B form nucleic acid due to their angular structure (93, 94). The geometry of the isopsoralen is such that once a monoadduct has been formed, subsequent reaction of the remaining double bond with a pyrimidine on the adjacent nucleic acid strand is not possible due to a misalignment of the two reactive double bonds. The absence of crosslinking is usually demonstrated by the inability of DNA to “snap back” rapidly after denaturation following irradiation in the presence of these compounds (95–97). Absolute lack of crosslinking by the isopsoralens has not been confirmed by total characterization of the adducts formed by HPLC and NMR analysis. Several investigators have reported that some crosslinks are formed by isopsoralen with phage lambda DNA (11–14). The explanation proposed is that the packaged DNA is amenable to crosslinking by isopsoralen due to its special folded structure within the phage head.

Monofunctional psoralens have been prepared by chemical synthesis for the most part. An exception is 3-dimethylallylpsoralen, which is a natural product. 3-dimethylallylpsoralen is reported to form only 2.5% crosslink based on the denaturation assay (98). This presumably results from the bulky dimethylallyl side chain sterically preventing the efficient conversion of the monoaddition product to crosslink. Among the synthetic derivatives, the approach has been to attach strongly electron withdrawing or donating groups to one of the two reactive double bonds. 3-Carbethoxypsoralen is reported to form little if any crosslink due to the electron withdrawing effect of the 3-carbethoxy group (47); however, this compound binds very poorly to DNA in vitro, having virtually no detectable dark association and approximately 3–5% of the photoreactivity of
8-methoxypsoralen with DNA (81). Another unreactive psoralen with an electron withdrawing group at the 3 position is 3-cyanopsoralen (99). Psoralens with electron donating substituents at the 3 position have also been prepared as monofunctional reagents, such as 3-amino and 3-methoxypsoralen (100). The 3-amino compound reportedly formed only monooadducts and reacted faster than 8-methoxypsoralen, but the total amount of the compound bound to the DNA was much less than with 8-MOP. With the 3-methoxy compound, little binding was detected and there was a significant amount of crosslink formed by the bound compound. In an attempt to inhibit the reactivity of the furan 4',5' double bond, 5-carbomethoxy-4,8-dimethylpsoralen was prepared and surprisingly was found to be highly reactive with DNA forming a significant amount of crosslink. The isomeric 3-carbomethoxy-4',8-dimethylpsoralen, however, was found to be essentially unreactive (101). There is one report of a dihydro-psoralen, reduced at the furan end, being slightly reactive with DNA (91).

Relative Reactivity of Different Psoralens with Nucleic Acids

The substituents of a psoralen effect each step in its interaction with nucleic acid. The position, the steric, and the electronic characteristics of each group on the psoralen ring system determine its ability to dark bind and photoreact with DNA and RNA, as well as the distribution of adducts formed. Many correlations as to what the effect of various substituent groups are have been made (25, 52, 53, 63, 102, 103). Some of the general trends that are known are the following. The photochemistry of methylated psoralens and isopsoralens is relatively fast. Methylation of a psoralen or isopsoralen increases the dark binding affinity, the quantum yield of photoaddition, and the quantum yield of photobreakdown of the compound (52, 102, 103). 4,5',8-trimethylpsoralen and 4,4',6-trimethylisopsoralen are therefore much more reactive than unsubstituted psoralen or isopsoralen respectively (63). A methoxy groups at the 8 position slows the photochemistry, with 8-methoxypsoralen adding much more slowly to DNA and also being much longer lived in solution than 4,5',8-trimethylpsoralen (52).

Strong electron withdrawing or donating substituents such as hydroxy, amino, and nitro groups either drastically reduce or completely eliminate the ability of the psoralen to undergo photocycloaddition with nucleic acids. 5-Nitro-8-methoxypsoralen and 5-amino-8-methoxypsoralen have no reactivity with DNA (52). Substitution at the 3 position of the psoralen nucleus with electronically active substituents is particularly unfavorable for high reactivity with nucleic acid (47, 99–101). However, relatively bulky groups which are positively charged placed at the 4' and 5 positions of the psoralen ring system form compounds which have both high dark association constants and high photoreactivity with DNA (52, 104).
The presence or absence of a methyl group at the 4 position has been shown to have a major role in controlling the amount of pyrone-side monoadduct formed with DNA (53). Psoralens that contain a 4 methyl group, such as 4,5',8-trimethylpsoralen and 4'-hydroxymethyl-4,5',8-trimethylpsoralen, form less than 2% pyrone side monoadduct. This has been attributed to steric interference between the 4 methyl group of the psoralen and the 5 methyl group of thymidine with which the psoralen predominately reacts. Psoralens that do not contain a 4 methyl group, such as 8-methoxypsoralen and psoralen, form up to 20% pyrone-side monoadduct with DNA (10). With RNA, 4'-hydroxymethyl-4,5',8-trimethylpsoralen forms up to 20% pyrone side monoadduct (105). This result is consistent with the DNA result since the reactive base in RNA is uracil which does not contain a methyl group at the 5 position.

Psoralen Monoadducts and Diadducts Formed with DNA and RNA

The photochemical addition products formed between psoralen and the pyrimidine bases of DNA and RNA have been characterized extensively. The detailed structure of both monoadducts and diadducts formed with several psoralen derivatives have been reported (8-10, 106), including the crystal structure of the furan-side 8-methoxypsoralen-thymine monoadduct (107-109). If the reaction occurs with intact, double-stranded nucleic acid, the number of adducts is limited and their stereochemistry controlled by the geometry of the intercalation complex. A wider variety of adducts are formed when psoralen is reacted with monomeric pyrimidine bases free in solution, in a frozen matrix or as a thin film.

Irradiation of 8-methoxypsoralen and double-stranded DNA starting with one 8-MOP per four base pairs will covalently bind 25–30% of the 8-MOP to the DNA. The distribution of adducts formed is approximately 25% thymidine-8-MOP-thymidine diadduct (crosslink), 45% furan-side thymidine-8-MOP monoadduct, and 20% pyrone-side thymidine-8-MOP monoadduct. A small amount of furan-side deoxyuridine-8-MOP monoadduct (ca 2%), derived from reaction with cytosine followed by hydrolytic deamination, is also formed. 8-MOP preferentially binds to 5'-TpA crosslinkable sites in the DNA (27). If the initial 8-MOP:base pair ratio is low, this preference becomes apparent, with a larger proportion of the total adduct being crosslink (60–70%), assuming a sufficient dose of light has been provided. As the amount of psoralen available for photoaddition is increased, crosslinkable sites become filled and the ratio of crosslink to monoadduct correspondingly decreases.

The structures of the monoadducts and diadducts formed between 8-MOP and DNA are shown in Figure 2. The eight possible configurations for psoralen-thymidine monoadducts are shown schematically in Figure 3.
stereochemistry of all the 8-MOP DNA adducts shown in Figure 2 is cis-syn. 
Syn for the furan-side describes the structure having the furan 1'-oxygen and 
the pyrimidine 1-nitrogen on adjacent corners of the cyclobutane ring. Syn for 
the pyrone side is defined having the 2-carbon of the psoralen and the 1-
nitrogen of the pyrimidine on adjacent corners of the cyclobutane ring. The 
formation of only one particular set of configurational isomers reflects the 
stringent restrictions on modes of psoralen intercalation imposed by the double 
helical DNA conformation. A complete discussion of the possible stereochem-
istries for psoralen-pyrimidine monoadducts and diadducts has been presented 
elsewhere (8–10).

As shown in Figure 4, 8-MOP can react from either the 3' or the 5' face of 
thymine, depending on whether the 8-MOP is positioned on top of or under-
neath the plane of the base within the binding site. Reaction within either a 
5' -TpX (3' face) or 3' -TpX (5' face) sequence results in a pair of enantiomeric 
thymine monoadducts being formed. In the case of nucleoside monoadducts, a 
pair of diasteromers is formed since the chirality of the deoxyribose is the same.

**Figure 2**  The structure of 8-methoxypsoralen (1), the two diastereomeric furan-side 8-MOP-dT-
monoadducts (2, 2'), the pyrone-side 8-MOP-dT-monoadduct (3), and the dT-8-MOP-dT diadduct 
(4).
in both cases. Two furan-side monoadducts between 8-MOP and thymidine (Figure 2, structures 2 and 2') have been isolated from DNA as nucleoside adducts while only one pyrone side nucleoside diastereomer (structure 3) has been identified (9). The thymidine-8-MOP-thymidine diadduct (structure 4) also occurs as a pair of diastereomers with cis-syn stereochemistry. The cis-syn configuration for both mono and diadducts has been found with four different psoralen derivatives and is likely to be general for all linear furcocoumarins.

The reaction of the psoralens with RNA is similar to that of DNA. The major reactive base in RNA is uridine (110), although reaction with cytosine residues in tRNA has also been reported (37). Three major monoaddition products have been characterized from the reaction between HMT and polyuridylic-polyadenylic acid (polyU-polyA) and with bulk RNA isolated from yeast (105). With both RNA substrates, the majority of the total bound HMT formed a pair of diastereometric furan-side adducts with uridine. The ratio of the two diastereomers was found to be 1:10 in polyU-polyA and 1:1 in bulk RNA. The stereochemistry of these furan side monoadducts is cis-syn, which is the same as in the HMT-thymine monoadducts formed in DNA (8, 10). The third major product was a pyrone-side HMT-uridine monoadduct which accounted for approximately 20% of the total adduct in polyU-polyA. In DNA, only 2% of the analogous pyrone-side monoadduct is formed, presumably due to steric interaction between the 5-methyl group of thymine and the 4-methyl group of HMT. The stereochemistry of the pyrone-side monoadduct with HMT, as well as the characterization of the crosslinked structure in RNA with psoralen, remain to be determined.
Irradiation of psoralen with pyrimidine bases or nucleosides in solution, as frozen matrices or as thin films, also produce psoralen-pyrimidine mono- and diadducts. There are numerous reports in the literature which describe various adducts with stereochemistries different than those formed within helical nucleic acid (111–115). A greater variety of adducts can be formed under these conditions since the constraints of intercalation are not present to limit the stereochemical outcome of the reaction.

*Figure 4*  *Top:* Schematic drawing to show the intercalation and 8-MOP-T monoadduct formation at 5'...T-A...3' and 5'...A-T...3' sequences. *Bottom:* Schematic drawing of 2 diadduct enantiomers of *cis*-syn configuration. The representations are the same as in Figure 3.
PHOTOCHEMICAL PROPERTIES OF PSORALENS

The photophysical and photochemical properties of psoralen and its derivatives have been extensively studied since the early 1970s. The excited states of psoralen have been identified. Interactions of these excited states of psoralen with bases, nucleosides, DNA, and RNA have been observed. Although much is known about psoralen excited states and their interactions, there is still some controversy about which excited state is photoactive in reactions with polynucleotides. Extensive reviews of this literature can be found by Parsons (5) and also Song & Tapeley (4). Only a brief account will be presented here.

The reactivity of any excited molecule will be influenced by several factors. The lifetime of the excited molecule must be long enough to encounter a second molecule in order to effect a reaction. During an encounter, both steric and electronic properties contribute to the overall reactivity. The relative influence of these factors on the cycloaddition of psoralen to a pyrimidine base is further determined by the local environment of the base. When psoralen is irradiated with free nucleoside or base in solution, the reaction will be diffusion controlled, and hence, the lifetime of the excited has a greater influence on the type and distribution of adducts. Since the reaction of psoralen with polynucleotides involves intercalation of psoralen into a double helical region, the lifetime of the excited state has less influence on the reactivity and the steric and electronic effects assume larger roles.

Spectroscopic Properties of Psoralen Molecules

The synthetic psoralen derivative AMT has a strong binding constant (25) for polynucleotides ($K_d$ DNA = $6.6 \times 10^{-6}$), thus permitting the characterization of spectroscopic properties of both free and intercalated psoralen (25). The absorption and fluorescence properties of AMT (116) are shown in Figures 5 and 6. These spectra are representative of most psoralen derivatives. The absorption spectrum has maxima at about 250 nm, 300 nm, and 340 nm, and extends out to 400 nm. Fluorescence emission is observed from ca. 380 nm to 600 nm. The lifetime of the fluorescence state is small and varies with the type of psoralen derivative and the solvent conditions (4, 23, 117–119). Typical values range from 1 to 5 ns (2.0 ns for free AMT in aqueous solutions). The quantum yield for fluorescence is also very small, ranging from 0.01 to 0.02 (120). Fluorescence excitation of AMT is maximized by the band centered at 340 nm. This band is also the photoactive band for photocycloaddition. Emission from the triplet state of psoralens is significantly greater than that of the singlet due to efficient intersystem crossing (23, 120, 121). Phosphorescence is observed from ca. 450 nm to 600 nm. The lifetime of the triplet state
also varies with the type of psoralen derivative, ranging from 1 μs to 1 s (100 μs for AMT) (117, 119–121).

Psoralen and its derivatives have two transitions in the 320 nm to 400 nm range: an \( n \rightarrow \pi^* \) transition resulting from the excitation of a non-bonding electron on the C-2 carbonyl group to the \( \pi^* \) orbital, and a \( \pi \rightarrow \pi^* \) transition occurring when a \( \pi \) electron in the psoralen ring system is excited to the \( \pi^* \) orbital. Mantulin & Song (120) have measured several spectroscopic parameters of coumarins and psoralens. By using substituted derivatives, they were able to assign the energy levels of the singlet and triplet states of psoralen. The relative orientation of the energy levels of these states are shown in the Jablonski diagram of Figure 7. The lowest singlet (120, 123) and triplet states
(120) are the \((\pi, \pi^*)\) states. The reactivity of either the \(1(\pi, \pi^*)\) or the \(3(\pi, \pi^*)\) toward pyrimidine cycloaddition is determined by kinetic (i.e. lifetimes), steric, and electronic factors. The electronic factors consist of the degree of local excitation and the electron density at the 3,4 and 4',5' groups (122). By analogy with coumarin, it has been established that the 3,4 double bond of the pyrone is locally excited in the triplet state. Local excitation of the 4',5' double bond has not been observed with psoralen itself. Intramolecular charge transfer from the \(\pi\)-electron system to the 3,4 or 4',5' double bonds affects the electron density in these reactive regions. Theoretical calculations of the electron density in these excited states of psoralen show that the 3,4 double bond has more charge transfer character than the 4',5' double bond (120, 122, 125–127). Based on these results, it has been suggested that the triplet photoreactivity resides with the pyrone moiety rather than with the furyl group (4, 120, 122, 124). Reactivity of the 4',5' double bond therefore appears to be determined by steric and kinetic factors rather than electronic factors. However, solvent perturbation of the triplets of 8-MOP, 5-MOP, and 3-carbethoxypsoralen has recently been explained by redistribution of charge at both the 3,4 and 4',5' positions (128).

The separation of the energy levels in Figure 7 affects the kinetics of psoralen addition by altering the lifetimes and distribution of the singlet and triplet
states. Lai et al (118) have recently measured several luminescent properties of the excited states of 8-MOP and TMP in different solvents. They conclude that the energy gap between the $1(\pi, \pi^*)$ and the $1(n, \pi^*)$ states determine the extent to which rapid, radiationless transitions occur between the $1(\pi, \pi^*)$ state and the ground state. The fluorescence quantum yield and fluorescence lifetime of 8-MOP were observed to increase upon going from a hydrophobic to a polar solvent. The quantum yield and lifetime of the triplet state changed in parallel with the fluorescence changes. It was suggested that the energy level of the $1(n, \pi^*)$ state is raised in polar solvents, thereby increasing the energy gap between the two singlet states. The increased energy gap between these two states reduces the interactions between the states and therefore reduces the radiationless transition probability. It has been suggested by Song (122) that this same effect upon the lifetime of the excited states of psoralen can also be induced by substitutents on the psoralen ring system. Electron donating groups, such as methyl and methoxy groups, would further raise the $1(n, \pi^*)$ and enhance the photocycloaddition by way of the singlet state. Lai, Lim and Lim (118) suggest that the greater fluorescence observed with TMP (30 times greater than 8-MOP) is due to a larger separation of the $1(n, \pi^*)$ and $1(\pi, \pi^*)$ states in TMP compared to 8-MOP.
Triplet-triplet absorption has been measured for several psoralen derivatives (117, 121, 128–133). Absorption occurs between 350 nm and 600 nm, with maximum extinction coefficients in the range of 3,700 to 49,500. AMT has a triplet-triplet extinction coefficient of 24,200 (117). Quantitation of the extinction coefficients of these transitions has facilitated the measurement of intersystem crossing quantum yields, $\phi_t$, for several psoralen derivatives. Values of $\phi_t$ have been observed in the range of 0.1 to 0.4 (121). 5-methoxypsoralen is an exception, with $\phi_t < 0.01$.

**Reactive States for Monoadduct and Crosslink Formation in Polynucleotides**

The identity of the reactive state that gives rise to monoadduct and crosslink formation has been a topic of debate for many years. It was predicted by Song et al (124) that the triplet excited state is more reactive than the singlet excited state. Since the excitation of the triplet state is localized in the 3,4 carbon-carbon double bond, it was expected that the 3,4 monoadduct would be predominant in the reaction with pyrimidines (120). Several 3,4 psoralen monoadducts have been isolated and characterized from reactions between psoralens and thymine (11, 112, 115) and psoralens and DNA (9, 10, 53). The distribution of adducts observed from reactions with DNA is very different than the distribution of adducts seen upon reacting free thymine with psoralen. Formation of adducts from different excited states might contribute to this observation. Involvement of the triplet state of free psoralen in solution has been confirmed by observing triplet quenching upon addition of nucleic acid bases and amino acids (121, 134, 135). $O_2$ and paramagnetic ions were found to quench the photoreaction between psoralen and thymine (136), and also quench 8-MOP inactivation of bacteriophage lambda (137). This latter observation is not in complete agreement with the work of Goyal & Grossweiner (138), who showed that intercalation of 8-MOP with calf thymus DNA inhibited the accessibility to $O_2$.

It is now clearly established that psoralen intercalation into a DNA helix determines the stereospecificity and distribution of adduct products. The 4',5' psoralen-adduct is the predominant monoadduct formed in reactions with DNA and is the precursor to crosslink formation (10). This result is not surprising, since the 4',5' monoadduct retains absorption above 320 nm. The wavelength dependence for AMT crosslinking of pBR322 DNA (139) and the action spectra for the photo-inactivation of phage with 8-MOP (140) both parallel the absorption spectrum of 4',5' monoadducts (Figure 8). Beaumont et al (134) have suggested that the 4',5' monoadducts are formed via a singlet excited psoralen upon absorption of the first photon. Crosslink formation would then be formed via the triplet state upon absorption of a second photon. Indirect evidence for the singlet excited state formation of the 4',5' adduct has been
reported by Salet et al (117) and by Beaumont et al (119). Both groups were unable to detect the triplet excited state of AMT intercalated in calf thymus DNA. The quantum yield of AMT fluorescence when bound to DNA was also reduced compared to free AMT. The fluorescence lifetime of free AMT was found to be 2.0 ns, while intercalated AMT had a lifetime of 0.03 ns (119). Involvement of the triplet excited state in the formation of a diadduct from a 4',5' monoadduct has not been established. However, the triplet excited states of 4',5' dihydropsoralen (129) and the 4',5' photoadduct of psoralen and thymine (130) are both quenched by thymine.

The major adducts formed with 8-MOP and DNA are shown in Figure 2 and the absorption properties of these adducts are presented in Figure 8. The 4',5' monoadducts have an absorption peak at about 340 nm. Irradiation into this band converts 4',5' monoadducts in DNA to crosslinks (10, 26, 27). The presence of 4',5' monoadducts and the conversion of these adducts to crosslinks in polynucleotides can be observed spectroscopically since these adducts have a distinct fluorescence (116). Both the 3,4 monoadduct and the diadduct

---

Figure 8  The absorption spectra of the two diastereomeric furan-side 8-MOP-dT monoadducts (D1, ---, 1.0 μg/ml; D2, -- -- --, 2.7 μg/ml), the pyrone-side 8-MOP-dT monoadduct (..., 2.7 μg/ml), and the dT-8-MOP-dT diadduct(- - - - 11.5 μg/ml). The spectra were determined in 30% methanol: 70% 20 mM KH₂PO₄, pH 2.2.
have little absorption above 320 nm. The capability of a 3,4 monoadduct to form a diadduct has not been established. A low lying triplet state of the 3,4 monoadduct exists (141). Ben-Hur & Song (3) propose that this state is unreactive with respect to a second photocycloaddition because of its rapid relaxation to ground state. Further characterization of the properties of 3,4 monoadducts is needed.

**Controlled Production of Monoadducts and Diadducts in Double-Stranded Nucleic Acids**

At present there are three methods of generating monoadducts in DNA that are capable of subsequent crosslinkage. The first exploits the structural properties of a psoralen adduct in DNA. The dihedral angle in a psoralen monoadduct between the plane of a psoralen molecule and the plane of a pyrimidine base requires a conformational change in the DNA backbone to accommodate a diadduct (108). Lown & Sim (11) first observed this phenomenon by irradiating PM2 DNA in the presence of different psoralen derivatives. Monoadducts capable of subsequent crosslinkage were greatly enhanced by irradiating in frozen solutions. Johnston et al (28) subsequently showed that the controlled production of monoadducts could be achieved by irradiating a solution of AMT and DNA with a single laser pulse of short duration (15 ns) sufficient for monoadduct formation but too short for a conformational change which occurs in the DNA and facilitates diadduct formation. A second method of generating monoadducts was reported by Chatterjee & Cantor (26). 4',5' monoadducts in DNA are greatly enhanced by irradiating with wavelengths above 380 nm. This selectivity is a result of the absorption properties of 4',5' monoadducts. Tessman et al (27) have recently used this procedure to show that the quantum yield for monoadduct formation (0.008) is smaller than the quantum yield for diadduct formation from a monoadduct (0.028). The last method of generating monoadducts in DNA is applicable to short oligonucleotides containing a single crosslinked site. Controlled photoreversal of the crosslinked molecule permits the isolation of monoadducted oligonucleotides (142). The quantum yield for photoreversal and the wavelength selectivity for photoreversal (furan vs pyrone monoadducted oligonucleotide) are areas that need further research.

**THE USE OF PSORALENS AS PROBES FOR NUCLEIC ACID STRUCTURE AND FUNCTION**

**Chromatin Structure**

Hanson et al (34) first proposed the use of psoralens as probes for chromatin structure based upon the results of an electron microscopic analysis of denatured DNA isolated from TMP photoreacted *Drosophila* embryo nuclei. This
PSORALENS AS PROBE OF NUCLEIC ACIDS

DNA exhibited an alternating pattern of loops (single strands) and bridges (apparent double strands). Interstrand DNA crosslinks were confined to the bridged regions. A weight averaged histogram analysis of the loops indicated that their length peaked at 160–200 base pairs and at intervals thereof. Approximately 60% of the total DNA existed as single-stranded loops. These loops were interpreted to represent regions on the DNA protected from TMP crosslinkage by associated histones. In agreement with this interpretation purified Drosophila DNA similarly photoreacted with TMP and analyzed by denaturation electron microscopy was almost completely double-stranded. The similarity in periodicity between the nucleosome subunit of chromatin and the looped pattern exhibited by in vivo crosslinked and formaldehyde denatured DNA led to the proposal that TMP crosslinking is restricted to the DNA between nucleosomes.

The selective photoreaction of TMP with internucleosomal DNA was independently confirmed by two groups (39, 143). Their studies with TMP photoreacted nuclei showed that micrococcal nuclease, itself a probe for chromatin structure which preferentially digests internucleosomal DNA, releases TMP at a rate much greater than bulk DNA when digesting the modified chromatin. In the limit digest, 92% of the TMP was released when 45% of the DNA had been rendered acid soluble. Control experiments demonstrated that TMP crosslinkage did not significantly alter the kinetics of micrococcal nuclease digestion or the size distribution of enzyme resistant DNA fragments. It was concluded that both TMP and micrococcal nuclease react with the same subset of DNA in chromatin. TMP saturation experiments conducted by Wiesehahn et al (39) support this conclusion. These experiments showed that DNA in chromatin was 90% protected from TMP photoreaction relative to naked DNA. Saturation levels of binding for the two DNAs were, respectively, 25 psoralens or 250 psoralens per 1000 base pairs.

TMP crosslinkage has been used as an in situ probe for chromatin structure in several systems. In each case the crosslinked DNA was purified and analyzed by denaturation electron microscopy. An alternating pattern of loops and bridges characteristic of chromatin structure was observed for mouse liver DNA (144), for SV40 DNA derived from both intracellular and viral nucleoprotein complexes (145), and for the main band and satellite I, II, and III DNAs from Drosophila nuclei (146). Meaningful results were not obtained for the satellite IV fraction since this DNA consists of short A-T containing segments separated by 250 base pair tracts of noncrosslinkable polypyrimidine/polypurine sequences (147).

Sequence specific positioning of nucleosomes on DNA has been addressed by Carlson et al (148) utilizing TMP photoaddition together with Hind III restriction analysis. Hind III was used in this study because preliminary experiments with TMP modified plasmid DNA showed that its recognition
sequence was 15-fold more susceptible to psoralen inactivation than other restriction sites examined. TMP mono- or diadducts at or near the *Hind* III recognition site are postulated to prevent cleavage of one or both strands by the enzyme. This phenomenon was used to probe the positioning of nucleosomes on the SV40 minichromosome. SV40 DNA from TMP photoreacted monkey cells was restricted with *Hind* III and electrophoresed on an agarose gel to resolve both partial and complete digestion fragments. Partial fragments contained one or more intact *Hind* III sites resistant to the enzyme. Since TMP photoreaction is limited to the internucleosomal DNA of chromatin, that DNA associated with the nucleosomal core will remain unmodified and restrictable by *Hind* III. A careful analysis of the restriction pattern indicated that no one *Hind* III site was either completely resistant or completely sensitive to inactivation by the psoralen probe. The evidence described here supports earlier in vitro studies which concluded that nucleosomes on the SV40 minichromosome occupy a random or nearly random set of positions on the underlying DNA sequence.

A different conclusion regarding the distribution of nucleosomes on SV40 minichromosomes has been reached by Robinson & Hallick (149). These investigators photoreacted intracellular SV40 chromatin with 3H-labeled HMT. The viral DNA was then isolated, multiply restricted, and the fragments electrophoresed on a polyacrylamide gel. Specific activity measurements of the fragments indicated that HMT had preferentially reacted with a 400 base pair region located between 0.65 and 0.73 on the SV40 physical map. This region encompasses the origin of replication and the promoters for early and late mRNA synthesis. The accessibility of this region to HMT addition is consistent with other evidence indicating that it is nucleosome-free. In contrast, however, the same group (150) could not detect preferential reaction of HMT with the origin region of viral chromatin implying that a complete randomization of nucleoprotein structure occurs upon encapsidation of the minichromosome.

Cech & Karrer (151) have applied TMP crosslinkage to the in vivo study of transcriptionally active rDNA of *tetrahymena thermophila*. Unique properties make this DNA an ideal substrate for crosslinkage analysis. Within the nucleus it exists as a separate extranucleosomal entity which can be isolated free of bulk DNA by isopycnic centrifugation. Additionally, it has a palindromic structure which means that each half of the molecule is structurally as well as functionally equivalent to the other half. Thus in the electron microscope each molecule can be systematically divided into three nontranscribable spacer regions and two coding regions. Log phase *tetrahymena* cultures were irradiated with long wave ultraviolet light in the presence of TMP to fix the transcriptionally active rDNA. This DNA was then isolated and examined by denaturation electron microscopy. Representative molecules were scored for crosslinks (double-stranded bridged regions) and bubbles (single-stranded loops). Each structure was assigned its position on the rDNA map which in turn was divided into
known coding and spacer regions. As a control, rDNA photoreacted in vitro with TMP was similarly analyzed. The results clearly showed that the frequency of crosslinkage in vivo was greater in the coding regions than in the spacer regions of the rDNA. Furthermore, the peak loop sizes measured in the coding regions were 215, 335, and 460 base pairs as compared to 200 and 400 base pairs for the peak loop sizes of main band DNA. These observations indicate that the chromatin structure of transcriptionally active rDNA possesses an altered conformation rendering it more open and accessible. One attractive model proposed by the authors is that such chromatin exists as a series of "half-nucleosomes" exhibiting a repeat unit length of approximately 110 base pairs.

The in situ structures of three non-nuclear DNAs have been probed by psoralen crosslinkage followed by denaturation electron microscopy. Uniform crosslinkage was observed over at least 90% of the contour length of mitochondrial DNA from Drosophila embryos and HeLa cell cultures (40, 152) and over the entire contour length of E. coli DNA (153) thus confirming a non-nucleosomal structure for each. However, 6–10% of the contour length was looped out at the origin of replication in each mitochondrial DNA preparation. Since these sequences were crosslinked in vitro, the in situ protection was attributed to origin specific nucleoprotein complexes.

Cruciforms

Regions in DNA with two fold symmetry, commonly referred to as inverted repeats or palindromic sequences, can exist in two alternate configurations, a linear form and a cruciform. Methods exist for the detection of cruciforms in vitro. Psoralen photochemistry, however, is unique in providing a way to probe for the existence of such structures within the cell. By crosslinking either arm of a cruciform with psoralen, its rearrangement to the linear form is prevented. Similarly, crosslinkage of the linear form precludes its conversion to cruciform. Thus DNA crosslinked in vivo can be assayed for cruciform structure in vitro without altering the original distribution. Cech & Pardue (154) have used this approach to search for cruciforms in the DNA of intact mouse tissue culture cells. After photofixing the DNA in situ with TMP crosslinks, it was purified, denatured in the presence of glyoxal, and examined by electron microscopy. Of the inverted repeats in mouse DNA long enough to be detected by electron microscopy, essentially none were found to be in the cruciform state. Using a similar protocol, Cech & Karrer (151) have probed the in situ configuration of tetrahymena thermophila rDNA. This long 20,400 base pair palindromic DNA was crosslinked in log phase cells and examined by electron microscopy under nondenaturing conditions. Again, no cruciforms were detected.

More recently Sinden et al (155) have employed psoralen crosslinkage of DNA as part of an assay for cruciform structure in a plasmid containing the
palindromic *E. coli* lac operator sequence. This DNA sequence can exist as either a 66 base pair long linear fragment or as a cruciform with two 33 base pair long hairpins. After TMP fixation and EcoR1 restriction, the conformation of the crosslinked lac operator containing restriction fragments was determined electrophoretically. Results of the assay showed that formation of the cruciform was dependent upon the presence of torsional tension in the molecule. In highly supercoiled DNA all of the lac operator sequence was in the cruciform state while in nicked circular DNA the same sequence existed in the linear state. Interconversion between the two forms was rapid at 37°C and slow at 0°C. When the assay was performed on plasmid DNA in bacterial cells, no cruciforms were detected. This indicates that in vivo the DNA lacks sufficient torsional tension to induce cruciform formation. In contrast to the S1 nuclease digestion method, this approach, which utilizes a TMP fixation step, permits an accurate determination of the relative amount of cruciform present in an in vitro or an in vivo sample without perturbing the equilibrium between the two forms.

**Torsional Tension**

Sinden et al (156) have described an assay for the determination of DNA torsional strain in both prokaryotic and eukaryotic cells based upon the rate of photobinding of labeled TMP to DNA. Like other intercalating agents, TMP has an enhanced affinity for negatively supercoiled DNA. In vitro, the rate of photobinding to covalently closed circular plasmid DNA is proportional to its superhelical density. This dependence provides an indirect assay for torsional tension. In the assay living cells were irradiated with near ultraviolet light in the presence of TMP to induce approximately one psoralen adduct per $10^3$ base pairs. The reaction conditions were chosen to minimize photochemical nicking and perturbation of chromosome structure due to drug binding. After photoreaction, DNA and rRNA were purified and the extent of modification to each determined. The photoreaction was normally expressed as the ratio of TMP binding to DNA versus rRNA. In this manner rRNA served as an internal standard to correct for experimental variations in the absolute rate of binding.

The above assay was used to monitor the effects of ionizing radiation and coumermycin treatment on the rate of TMP photoaddition to *E. coli* DNA. In prokaryotes both agents relax supercoiled DNA by either direct nicking or by inhibiting DNA gyrase. With either treatment the initial rate of TMP photobinding to DNA was reduced by 40%. Thus the *E. coli* chromosome contains unrestrained torsional tension equivalent to that expected for a purified DNA duplex with a superhelical density of $-0.05 \pm 0.01$. The same assay was extended to *Drosophila* cells and HeLa cells. However, the initial rate of TMP photoreaction in these cells was unaffected by treatment with ionizing radiation. This supports the general belief that superhelical DNA turns in the eukaryotic chromosome are restrained in nucleosome structures and that such DNA is not subject to torsional tension.
Secondary Structure in Single-Stranded Nucleic Acid

The use of psoralen derivatives to photocrosslink double-stranded regions in single-stranded nucleic acid is without parallel as a technique for determination of secondary structure. Both localized hairpins and long range interactions can be fixed in vitro or in situ by psoralen crosslinkage. In early studies, pioneered by Hearst, Cantor, and coworkers (35, 157–167), denaturation electron microscopy was used to map crosslinked secondary structures in both single-stranded DNA and RNA. With this technique structures could be mapped to a resolution of 50 base pairs. Recent advances in the analysis of crosslinked molecules now permit secondary structures to be mapped to sequence resolution (29, 33, 168). In this approach, which has been applied to 16S ribosomal RNA by Thompson & Hearst (30), small crosslinked fragments isolated from the photoreacted molecule are sequenced after photoreversal and purification of component strands, if necessary. This section will survey the ways in which psoralen crosslinkage has been used to probe structure in single-stranded fd and SV40 DNAs and in transfer, small nuclear, ribosomal, and heterogeneous nuclear RNAs.

SINGLE-STRANDED DNA The psoralen crosslinkage technique was first applied by Shen & Hearst (157) to the single-stranded DNA genome of bacteriophage fd. Both circular and full length linear DNA were photoreacted with TMP at several different ionic strengths and then examined by electron microscopy. The single-stranded linear DNA had been prepared from replicative form by restriction and transient alkali denaturation. When crosslinkage was performed at 15°C in water no secondary structures were discernable. However, when the NaCl concentration was raised to 20 mM, TMP crosslinkage stabilized a single 200±50 base pair hairpin. This hairpin was uniquely positioned near one end of the linear DNA at a site believed to correspond to the origin of replication. Crosslinkage at higher salt concentrations fixed other less stable hairpins. A comparison of the histogram map of these hairpins with known promoter sites of fd suggested a possible correlation.

In a subsequent study Shen et al (158) investigated the structure and orientation of single-stranded circular fd DNA inside its filamentous virion. For this analysis purified virions were photoreacted with TMP. The DNA was then deproteinized and analyzed by electron microscopy. Most of the DNA molecules were visualized as extended circles exhibiting a single prominent hairpin at one end together with occasional internal crosslinks. The end hairpin was mapped relative to a unique restriction endonuclease cut introduced into the crosslinked DNA after hybridizing it to a complementary linear single strand. Analysis showed that the end hairpin was identical to the stable hairpin seen in naked fd DNA. From these results it was concluded that the stable hairpin is located at one of the ends of the fd virus and that the DNA genome occupies an extended, fixed, and nonrandom orientation within the viral capsid.
Employing similar protocols, Ikoku & Hearst (159) have determined the orientation of the single-stranded DNA genome in the filamentous chimeric phase M13goril, and Shen & Hearst (160, 161) have determined the secondary structure of single-stranded SV40 DNA. In the latter studies, after linearizing and denaturing form I DNA, helical regions on single strands were fixed by TMP photoreaction and visualized by electron microscopy. Most of the short hairpins on SV40 mapped at initiation and termination sites for replication or transcription while the larger hairpin loops bracketed previously identified intervening sequences.

RIBOSOMAL RNA  Our understanding of the secondary and tertiary structure of rRNAs has been advanced by several psoralen crosslinkage studies. The initial application of this technology utilized electron microscopy to map psoralen crosslinks in Drosophila melanogaster 18S and 26S ribosomal RNA (35). HMT crosslinkage of purified rRNA stabilized a centrally located hairpin on 26S rRNA and an open loop structure near one end of the 18S rRNA.

Application of the psoralen-crosslinking technology to E. coli 16S ribosomal RNA has been extensive (162–165). Electron microscopic visualization of isolated 16S rRNA crosslinked by HMT revealed eleven large open loop structures (162). The frequency of occurrences of these structures ranged from 2% to 30%. The loop size and stem length of each secondary structure were measured and all mapped relative to one another on the 16S molecule. The polarity of the map was determined by hybridizing DNA complimentary to specific regions of the 16S rRNA and visualizing the hybrids by electron microscopy (164). 16S rRNA in the 30S subunit has also been probed by psoralen crosslinkage and analyzed by electron microscopy (163). Both active and inactive conformations of the 30S subunit were photoreacted with HMT. Analysis of 16S rRNA crosslinked in the inactive subunit showed that it contained seven of the eleven open loop structures previously characterized for the same molecule crosslinked in solution. Of those seven structures, six were present in 16S rRNA when crosslinked in the active subunit. A computer search of the 16S rRNA sequence yielded energetically stable duplex structures corresponding in positions to each of the observed psoralen crosslinking sites (165). All of the crosslinked features can occur simultaneously in the 16S molecule because the base paired sequences are non-overlapping.

Two variations on the formation, processing, and electron microscopic detection of psoralen crosslinks in E. coli 16S rRNA have been explored. A gel electrophoretic technique for the isolation of 16S rRNA fractions containing unique psoralen crosslinks has been developed by Wollenzein & Cantor (166). Purified 16S rRNA containing ca. 10 adducts per RNA molecule was further resolved by electrophoresis through a gradient polyacrylamide gel made in formamide and low salt to promote denaturation. This gel system resolved...
crosslinked 16S rRNA into several discrete bands. Each band was analyzed separately by electron microscopy and found to contain 16S rRNA molecules with long range crosslinks at different positions within the 16S molecule. Six of the previously identified long range interactions were found by this technique. Because of the reduced heterogeneity of products, it is expected that this technique will be invaluable for the analysis of crosslinks at sequence resolution.

In the second variation of the psoralen-crosslinking technique applied to 16S rRNA, Chu et al (167) utilized a two-step irradiation protocol to compare the psoralen stabilized secondary structure of 16S rRNA in active and inactive 30S ribosomal subunits. The first photoreaction, carried out with 390 nm light, was used to generate AMT monoadducts in inactive 30S particles. The modified particles were freed of unreacted AMT and, if desired, converted into active subunits. AMT monoadducts in the 30S particles were chased into crosslinks by reirradiation at 360 nm. The locations of crosslinks were determined by electron microscopy of the purified 16S RNA. Since this approach inserts AMT into the RNA prior to conformational changes induced by activation of the ribosome, any decrease in the frequency of a crosslinked feature may be attributed to a loss of base paired structure in the vicinity of an AMT monoadduct. When the secondary structure of 16S rRNA from active and inactive subunits was examined, the overall patterns were similar except for the decreased frequency of a 3' end loop in RNA isolated from active subunits. This implies that activation of the ribosome is accompanied by a change in the secondary and possibly tertiary structure of the 3' terminus of 16S rRNA.

Sequence resolution of psoralen-crosslinked fragments was first reported by Rabin & Crothers in 1979 (33). Using AMT crosslinked *E. coli* 5S rRNA, they were able to isolate, photoreverse with 250 nm light, and sequence both strands of a crosslinked fragment of 5S rRNA. The sequence analysis of each strand confirmed the existence of a stem structure in 5S RNA formed between residues 1-10 and 110-119. This work demonstrates the utility of psoralen photochemistry in the unequivocal characterization of secondary structure features in RNA.

The approach initiated by Rabin & Crothers has been extended by Thompson et al (168). These authors photoreacted uniformly ^32^P-labeled *Drosophila melanogaster* 5S RNA in solution with HMT. After photoreaction the RNA was digested to completion with RNase T1 and the fragments were separated on a 20% polyacrylamide gel. Since eukaryotic 5S RNA is only 120 nucleotides long, the straightforward digestion pattern permitted the separation of crosslinked fragments from one another and from unmodified oligonucleotides. Each crosslinked band was extracted from the gel and photoreversed, and the sequence of the fragments determined. This information permitted the authors to unequivocally identify the location of two crosslinks and to tentatively
identify the location of three others. The secondary structure model for 5S rRNA predicted by these crosslinks is very similar to ones proposed on the basis of evolutionary and enzymatic digestion data.

The first application of sequencing analysis to psoralen crosslinked 16S rRNA was reported by Turner et al (29). After photocrosslinking uniformly 32P-labeled 16S RNA with HMT, the RNA was partially digested with RNase T1. The fragments were resolved by a two-dimensional electrophoresis system. Crosslinked fragments were completely photoreversed between dimensions. In this system, crosslinked fragments migrate as a pair of off diagonal spots. A hairpin localized at positions 434-497 was identified in this study. The putative location of the HMT crosslink was deduced to be an intercalation site formed between base pairs U458-G474 and A459-U473. Based on the known reactivity of psoralens, this site should be a reaction hotspot for HMT since it contains a pair of crosslinkable uridines in a loose helix as defined by a G-U base pair directly abutting a run of three A-U base pairs.

Thompson & Hearst (30) have refined the mapping of psoralen crosslinks and in so doing have characterized 13 unique secondary structures in E. coli 16S RNA. The crosslinked fragments were isolated by modification of a two-dimensional gel electrophoretic technique originated by Zweib & Brimacombe (169). The isolation protocol is outlined in Figure 9. Off diagonal spots containing crosslinked fragments were excised and photoreversed. Component strands were purified and sequenced by standard enzymatic methods. In Figure 10 the positions of the characterized crosslinks are superimposed on the secondary structure model of Noller & Woese (170). Several long range interactions elucidated by HMT crosslinkage appear to exist in dynamic equilibrium with local secondary interactions. These equilibria may reflect conformational changes which occur in 16S RNA during functioning of the ribosome (171).

TRANSFER RNA The photoreaction of HMT with phenylalanine tRNA has been addressed by two groups. In each case the reaction pattern was determined to sequence resolution and the results were consistent with the three-dimensional structure determined by X-ray crystallography. In the earlier study, conducted by Bachellerie & Hearst (172), the sites of HMT monoaddition on E. coli tRNA^Phe were mapped by a combination of techniques including chemical and enzymatic digestions and electrophoretic separations on gel and paper. By far the most reactive site was 51U, a residue base paired to 63G in the T-stem, although every other uracil in the molecule save one exhibited some residual activity with HMT. In the more recent study, Garrett-Wheeler et al (37) photocrosslinked yeast tRNA^Phe in solution with HMT. A partial T1 digest of the modified tRNA was 32P end-labeled and fractionated by a two-dimensional electrophoresis scheme similar to that devised by Turner et al (29).
PSORALENS AS PROBE OF NUCLEIC ACIDS

This scheme, however, incorporated a photoreversal step which was reduced in both time and intensity to minimize loss of monoadducts. Oligonucleotides formerly involved in a crosslink were isolated and sequenced enzymatically. Additionally, monoadducted pyrimidines were mapped by electrophoresing purine specific U2 and cytidine specific CL3 ribonuclease digests of the oligonucleotides. These enzymes were observed to anomalously cleave RNA adjacent to psoralen monoadducts. The exact positions of five crosslinks were determined. Four of the crosslinks occurred in stems and one bridged a tertiary interaction between two trans oriented pyrimidines. Surprisingly, the crosslinks involved one C-C interaction, three C-U interactions, and only one classical U-U interaction. The ability to enzymatically map psoralen crosslinks should prove very useful in probing the tertiary structure of other RNA molecules.

HETEROGENEOUS AND SMALL NUCLEAR RNAs In both prokaryotes and eukaryotes high molecular weight primary transcripts are processed into mature tRNA, rRNA, and mRNA. An important facet of RNA processing, not clearly understood, is the enzymatic removal of extraneous intervening sequences or
introns. The coupled cleavage-ligation reaction referred to as splicing is site specific and often involves regions of RNA secondary structure in the primary transcript. In eukaryotes heterogeneous nuclear RNA (hnRNA) is the nuclear precursor to cytoplasmic mRNA. This high molecular weight RNA possesses
an ordered ribonucleoprotein structure (hnRNP) and, unlike mRNA, contains RNase resistant double-stranded regions when deproteinized. These regions, whose existence in vivo may be regulated by specific proteins, probably play a role in RNA processing. Calvet & Pederson (173, 174), utilizing AMT cross-linkage in vivo, have verified the existence of double-stranded regions in hnRNP complexes. They digested hnRNP from control or AMT photoreacted HeLa cells with proteinase K, pancreatic DNase, and the single-strand specific enzymes pancreatic RNase and RNase T1. The enzyme resistant double-stranded RNA was purified on a Cs2SO4 gradient and analyzed for rapidly renaturing or “snap back” RNA. The RNA which had been crosslinked with AMT in vivo showed a several fold enhancement in the rapidly renaturing fraction relative to the control sample. These results demonstrated the existence in vivo of double-stranded regions in hnRNA accessible to AMT intercalation and crosslinkage.

Evidence for the potential involvement of RNA secondary structure in splicing has been obtained by Wollenzein et al (175) in an electron microscopic analysis of AMT crosslinked 35S RNA from Neurospora mitochondria. Mature 25S rRNA is formed from this precursor upon removal of a 2.2 kilobase intron. The 35S RNA used in this study had been photoreacted with AMT in isolated RNP particles and then spread for electron microscopy after deproteinization and denaturation in the presence of formaldehyde. Two mutually exclusive structures were identified: a central hairpin 105 base pairs long and a large loop 2.2 kilobases in length. When crosslinkage was conducted in RNP particles at low temperature in the presence of MgCl2 and KCl, the loop structure predominated. However, when the same RNA, previously monoadducted with AMT in RNP particles, was photoreacted at 365 nm in the absence of protein, the central hairpin was more frequent. The location and size of the loop structure indicated that it was in fact the intron sequence and that the corresponding crosslink bridged the two splice sites. Taken together, the results suggested that one or more RNP proteins stabilize a psoralen accessible base paired region formed between two complementary sequences on opposite splice sites.

The small nuclear RNAs (snRNAs) are a class of abundant RNA molecules 100–300 nucleotides in length which are postulated to play a role in RNA processing. Two members of this class, U1 and U2, have potential sequence complementarity with intron-exon junctions in messenger RNA precursors. Pederson and coworkers (176, 177) have investigated the occurrence of these interactions in vivo by psoralen crosslinkage. Very simply, they isolated hnRNA under denaturing conditions from AMT photoreacted HeLa cells. The RNA was then photoreversed and electrophoresed on a 10% polyacrylamide gel made in formamide. Gel blots with U1 and U2 specific probes lit up bands for the respective snRNAs. When the photoreversed step was omitted, the probes hybridized to the high molecular weight hnRNA band at the top of the
CIMINO ET AL
gel. This electrophoresis pattern confirmed the existence of base paired complexes between U1 and U2 with hnRNA.

In a continuation of their work, Setyono & Pederson (178) have examined the ribonucleoprotein organization of the U1 RNA fraction base paired in vivo to hnRNA. Previous studies had shown that U1 RNA within the cell exists as a ribonucleoprotein complex and it was expected that this was the species bound to hnRNA. For this analysis isolated HeLa cell nuclei were pulse-labeled with 3H-uridine and photoreacted with AMT. Nuclear hnRNA-ribonucleoprotein particles were purified on a sucrose gradient containing 50% formamide. In this concentration of formamide noncovalently bound U1 RNA was lost but ribonucleoprotein structure was retained. The purified RNP, labeled with 3H in hnRNA, was passed through an anti-U1 RNP immuno-affinity column. In the absence of photoreversal, a small fraction of the hnRNA in hnRNP particles reproducibly bound to the column. These binding results support the notion that both U1 RNA and hnRNA interact with one another as ribonucleoprotein complexes.

**Stabilization of Nucleic Acid Complexes**

The psoralen derivatives TMP, HMT, and AMT have been shown to crosslink DNA-DNA, DNA-RNA, and RNA-RNA double strands in the presence of near ultraviolet light (25, 179). This photoreaction has been used to stabilize a variety of nucleic acid structures including transcriptional complexes, replicative intermediates, R-loops, and D-loops. Psoralen crosslinkage has also been used as a tool in various hybridization, electron microscopic, and cloning procedures. This section will survey these uses.

The ability of psoralen derivatives to crosslink DNA-RNA hybrids has been utilized by Shen & Hearst (180) to study the *E. coli* RNA polymerase ternary complex. 14C-labeled SV40 DNA was transcribed by the *E. coli* enzyme in the presence of 3H-labeled ribonucleoside triphosphates. Ternary complexes were photoreacted with AMT and spun through an SDS-sucrose gradient. Only in the crosslinked samples did the nascent RNA and viral DNA cosediment. These results indicate that AMT can intercalate into and crosslink the short DNA-RNA hybrid formed within the ternary complex between nascent RNA and the coding DNA strand. The authors suggest that in vivo crosslinkage of ternary complexes could facilitate the isolation of transcriptionally active DNA via isopycnic gradient ultracentrifugation.

Revet & Benichou (181) have utilized in vivo TMP crosslinkage to fix adenovirus 5 replicative form intermediates. The replicating viral DNA molecules were purified by ultracentrifugation through consecutive sucrose and CsCl gradients and examined by electron microscopy under both native and denaturing conditions. The original distribution of replication forks on each molecule was preserved during workup and analysis by the presence of TMP.
crosslinks. These crosslinks prevented branch migration. The authors concluded that in vivo TMP crosslinkage could be extended to the study of other labile double-stranded structures such as Okasaki fragments and recombination intermediates.

When an RNA strand hybridizes to a complementary section of double-stranded DNA, the RNA displaces the homologous strand in the duplex forming a DNA-RNA hybrid with a looped out single-stranded DNA. This structure, referred to as an R-loop, is unstable under physiological conditions. Three protocols, utilizing psoralen crosslinkage to stabilize the R-loop structure have been published. In a procedure developed by Davidson and coworkers (182), the DNA was lightly crosslinked with TMP prior to R-loop formation. During hybridization the presence of crosslinks stabilized the DNA double helix by preventing complete strand separation and by reducing the probability of fragmentation due to denaturation between widely spaced nicks on opposite strands. After R-loop formation the single-stranded DNA was modified at hydrogen bonding positions with glyoxal to prevent branch migration. Together, the two treatments stabilized R-loop structures for subsequent electron microscopic analysis.

Two other protocols for stabilizing R-loop structures with psoralen crosslinks have been published. Wittig & Wittig (32) have reported that R-loops in DNA can be fixed by crosslinking with TMP and near ultraviolet light directly in the R-loop formation buffer (70% formamide at 4°C) and Chatterjee & Cantor (183) have demonstrated that R-loops formed with AMT monoadducted RNA can be stabilized by subsequent irradiation at 360 nm. The ultraviolet treatment converts many of the monoadducts to crosslinks. In both of these studies the R-loop structures were successfully cloned after removal of the RNA strand by alkali treatment or the DNA loop by S1 nuclease digestion.

The approaches described above can also be applied to the stabilization of D-loops. A D-loop is typically formed when a single-stranded DNA fragment hybridizes to a complementary double-stranded DNA and displaces the homologous strand. D-loops are even less stable than R-loops and are usually formed on negatively supercoiled DNA. DeFrancesco & Attardi (152) have used HMT crosslinkage to stabilize in situ a unique D-loop present on HeLa cell mitochondrial DNA. In their analysis the mtDNA was isolated from HMT photoreacted mitochondria, linearized with a single-hit restriction endonuclease, and spread for electron microscopy under nondenaturing conditions. Approximately 25% of the molecules contained a 560±70 base pair D-loop located at the origin of replication. This loop was linked to the in vivo synthesis of a 7S DNA fragment complementary to the L-strand of the molecule. A previous study (184) had shown that this particular D-loop in the absence of crosslinkage was rapidly lost upon restriction of the mtDNA.

Two applications of psoralen crosslinkage deserve mention here. In 1982
Brown et al (185) described the preparation and use of unique M13 hybridization probes labeled to a high specific activity with a TMP crosslinked and enzymatically extended primer. The strand specific probes are ideally suited for use in Northern hybridization experiments and for use in M13 plaque hybridization assays. Also, in 1982 Wollenzein & Cantor (186) described a procedure for marking the polarity of single-stranded DNA or RNA to be analyzed by electron microscopy. In this procedure a DNA restriction fragment complementary to an asymmetric site on the molecule of interest is monoadducted with AMT and then hybridized and crosslinked to the target nucleic acid. In a denaturing spread, the marker DNA strand serves as a point of reference.

FUTURE DIRECTIONS

From our perspective, three avenues of ongoing research hold the potential for dramatically increasing the versatility of psoralens for use as photochemical probes of nucleic acid secondary structure and function. The three areas of research are, respectively, the synthesis of modified psoralens exhibiting unique crosslinking specificities or containing reporter molecules, the development of a direct mapping protocol for psoralen monoadducts and crosslinks in both RNA and DNA, and the refinement of photochemical as well as chemical procedures for the site-specific placement of crosslinkable psoralen monoadducts onto DNA oligonucleotides. In this section the advances in each area will be reviewed and the likely applications of each technology will be discussed.

Cantor and coworkers (86) have synthesized three modified psoralen derivatives capable of forming intermolecular crosslinks in nucleoprotein complexes. One derivative, a bis psoralen compound, has been photoreacted with bacteriophage lambda (87). DNA helices in close proximity should be crosslinked to one another by this compound. Nondenaturing electron microscopic analysis of a Bgl II digest of the modified lambda DNA has verified the formation of interhelical crosslinks through the detection of X-shaped molecules. The crosslinking frequencies between all six possible pairs of the four largest Bgl II fragments of lambda DNA were determined. The results ruled out purely random packaging of the DNA within the virion but at the same time were inconsistent with two previously suggested models for intraphage DNA packing. A second psoralen derivative, also applied to the structure analysis of lambda bacteriophage, was an AMT molecule linked at the 4' position to a reactive succinate group by a long alkyl chain (86). In the lambda model system, this compound crosslinked viral capsid to the DNA genome. Protein linkage occurred chemically via the succinate group while DNA linkage took place photochemically via AMT. The third psoralen derivative synthesized by the Cantor group was another AMT molecule containing a 4' linker arm terminating in a sulfhydryl group (85). The reagent was selectively dark bound
to mercurated patches in pBR322 DNA via the formation of a Hg-S linkage. After removal of excess reagent and upon irradiation with 365 nm light the site-specifically bound AMT formed intramolecular as well as intermolecular DNA crosslinks. All three compounds described above have the potential to elucidate tertiary interactions in condensed nucleoprotein complexes.

A second class of modified psoralens with obvious application to structure determination studies are furocoumarins which contain an attached reporter group. An example of such a group is the biotin moiety. We have synthesized AMT molecules with the group attached to the 4' position by various linker chains. Preliminary studies indicate that these AMT derivatives can intercalate into and crosslink double-stranded nucleic acid. The adducts can then be detected colorimetrically or fluorescently by standard methods which are based on the avidin-biotin interaction. Our expectation is that this same interaction can be used to isolate by affinity chromatography monoadducted as well as crosslinked oligonucleotides. In structure determination analyses the oligonucleotides would be derived from partial enzymatic digests of psoralen photo-reacted single-stranded RNA.

Another potential breakthrough in psoralen based secondary structure analysis would be the development of a direct chemical mapping protocol for the localization of psoralen modified bases in DNA and RNA. Two enzymatic approaches exist but each has disadvantages. In 1982 Youvan & Hearst (36) published a protocol for mapping HMT monoadducts in RNA based on the sequencing of reverse transcriptase cDNA transcripts initiated from a unique primer on a psoralen modified RNA. This study demonstrated that reverse transcriptase terminates synthesis opposite HMT uridine monoadducts in RNA. A major drawback of the approach is the need for biologically prepared or chemically synthesized DNA primers. In 1984 a second enzymatic mapping protocol was published by Garrett-Wheeler et al (37) who reported that ribonucleases U₂ and CL3 anomalously cleave RNA after HMT monoadducted pyrimidines. A disadvantage of this approach is that the psoralen specific cleavages are superimposed upon the naturally occurring cuts dictated by the nucleotide specificities of the enzymes.

The development of a chemical cleavage method for mapping psoralen monoadducts and crosslinks in nucleic acid is premised upon the reducibility of pyrimidines containing a saturated 5,6-double bond (187). Psoralen photocycloaddition to pyrimidines occurs at this site and the resultant cyclobutane ring is saturated. Reduction of modified pyrimidines with excess NaBH₄ should ring open the base (187–189) thereby permitting acid catalyzed depyrimidination followed by base catalyzed β-elimination and strand scission. This protocol should be straightforward for psoralen modified thymidine and uridine. Application to cytosine, however, must follow deamination of the 4 position on the base converting it to a carbonyl group. This reaction occurs spontaneously.
upon saturation of the 5,6-double bond (190). Becker & Wang (38) have reported that psoralen binding to DNA can be detected by chemical cleavage induced by NaBH₄ reduction of modified pyrimidines followed by acidic aniline treatment. Perfection of this technique together with the inclusion of a crosslink photoreversal step if necessary and the use of gel electrophoresis to resolve the cleavage products should permit chemical mapping of psoralen adducts in both DNA and RNA.

A third area of activity is the development of a straightforward chemistry for the synthesis of DNA oligonucleotides containing site-specifically placed crosslinkable psoralen monoadducts. If successful, this technology could not only have a significant impact on the use of psoralens as probes for nucleic acid structure and function but also extend the advantages of psoralen based crosslinking to several biochemical techniques. Gamper et al (142) have described a photochemical protocol for the generation of psoralen monoadducted oligonucleotides. They observed that the TpA sequence in double-stranded DNA oligomers is a preferred site for psoralen photoaddition. When the self-complementary Kpn I linker CGGTACCG was photoreacted with HMT, the central TpA sequence was crosslinked. Controlled photoreversal of the crosslink generated a furan side HMT monoadduct. The photochemical approach outlined above is limited to relatively short oligonucleotides which contain a single TpA sequence. We are therefore developing a direct chemical approach to the synthesis of psoralen modified DNA oligonucleotides. Preliminary studies indicate that large quantities of the furan side 8-MOP monoadduct to thymidine can be enzymatically prepared from photoreacted DNA. This adduct, after appropriate protection and activation steps, should substitute for thymidine in the standard oligonucleotide synthesis schemes.

The availability of chemically synthesized photoactive DNA oligonucleotides provides a vehicle for the sequence specific targeting of psoralen crosslinks into both single and double-stranded nucleic acid. By hybridizing and crosslinking judiciously selected psoralen monoadducted oligonucleotides to single-stranded DNA or RNA, the role of the complementary regions in secondary structure and function can be probed under reconstitution conditions. Ligation and crosslinkage of such oligonucleotides into appropriately gapped double-stranded DNA molecules will provide unique substrates with site-specific blocks to replication and transcription. Characterization of the structure and response of the stalled enzymatic complexes should be quite informative.

Advances in our understanding of psoralen photochemistry may permit actual strand transfer of the psoralen from the modified oligonucleotide to the complementary region in the receptor nucleic acid. This could be accomplished by selectively photoreversing the crosslink to give pyrone side monoadduct now attached to the complementary strand. After removal of the targeting
oligonucleotide and renaturation of the single-stranded DNA or RNA, secondary structure in the vicinity of the monoadduct could be probed by irradiating it back to crosslink. This technique will permit site-specific positioning of a psoralen monoadduct within a large DNA or RNA molecule and will therefore expand the potential reactive sites, reduce the complexity of psoralen crosslink analysis, and ultimately provide structure-function information not obtainable by conventional means. Psoralen strand transfer should also be applicable to superhelical DNA through the formation of an intermediate D-loop with the modified oligonucleotide. Success of the strand transfer concept hinges upon the development of an effective way to crosslink pyrone side monoadducts.

The use of psoralen monoadducted oligonucleotides as hybridization probes could have a significant and far reaching impact on hybridization technology and clinical diagnostics (K. Yabusaki, H. Gamper, S. Isaacs, patent pending). By crosslinking the probe-target mixture under stringent conditions, a complex is formed that is not only hydrogen bonded but also covalently linked. This in turn permits the use of rigorous, even denaturing, wash conditions thereby reducing background and increasing sensitivity. The use of short crosslinkable probes should also reduce hybridization times, favor simple solution assays, and permit the use of techniques such as gel filtration to separate probe-target complex from excess oligomer under denaturing conditions.

ACKNOWLEDGMENTS

We are grateful to the National Institutes of Health for their long term support of psoralen chemistry research in our laboratory (GM 11180, GM 25151), and to many colleagues around the world who generously provided us with current information about the psoralen activities in their laboratories.

Literature Cited

PSORALENS AS PROBE OF NUCLEIC ACIDS

1191

1192 CIMINO ET AL


PSORALENS AS PROBE OF NUCLEIC ACIDS

CONTENTS

THEN AND NOW, Harland G. Wood 1

TIME-RESOLVED FLUORESCENCE OF PROTEINS, Joseph M. Beechem and Ludwig Brand 43

ADP-RIBOSYLATION, Kunihiro Ueda and Osamu Hayaishi 73

GENETIC ANALYSIS OF PROTEIN EXPORT IN ESCHERICHIA COLI K12, S. A. Benson, M. N. Hall, and T. J. Silhavy 101

CELL ADHESION AND THE MOLECULAR PROCESSES OF MORPHOGENESIS, Gerald M. Edelman 135

MECHANISM AND CONTROL OF TRANSCRIPTION INITIATION IN PROKARYOTES, William R. McClure 171

RECEPTORS AND PHOSPHOINOSITIDE-GENERATED SECOND MESSENGERS, Lowell E. Hokin 205

THIOREDOXIN, Arne Holmgren 237

THE MEMBRANE SKELETON OF HUMAN ERYTHROCYTES AND ITS IMPLICATIONS FOR MORE COMPLEX CELLS, Vann Bennett 273

ROLE OF REVERSIBLE OXIDATION-REDUCTION OF ENZYME THIOLS-DISULFIDES IN METABOLIC REGULATION, D. M. Ziegler 305

MOLECULAR BIOLOGY AND GENETICS OF TUBULIN, Don W. Cleveland and Kevin F. Sullivan 331

NUCLEOSIDE PHOSPHOROTHIOATES, F. Eckstein 367

GROWTH HORMONE RELEASING FACTORS, Nicholas Ling, Fusun Zeytin, Peter Böhlen, Fred Esch, Paul Brazeau, William B. Wehrenberg, Andrew Baird, and Roger Guillemin 403

INDUCIBLE DNA REPAIR SYSTEMS, Graham C. Walker 425

(continued) v
CONTENTS  (continued)

VITAMIN K–DEPENDENT CARBOXYLASE, J. W. Suttie 459

PROTEIN CARBOXYL METHYLTRANSFERASES: TWO DISTINCT CLASSES OF ENZYMES, Steven Clarke 479

EVOLVING RIBOSOME STRUCTURE: DOMAINS IN ARCHAEBACTERIA, EUBACTERIA, EUCYOTES AND EUKARYOTES, James A. Lake 507

VIROIDS, Detlev Riesner and Hans J. Gross 531

THE CHEMICAL MODIFICATION OF ENZYMATIC SPECIFICITY, Emil Thomas Kaiser, David Scott Lawrence, and Steven Edward Rokita 565

EFFECTS OF SITE-SPECIFIC AMINO ACID MODIFICATION ON PROTEIN INTERACTIONS AND BIOLOGICAL FUNCTION, Gary K. Ackers and Francine R. Smith 597

ASSEMBLY OF ASPARAGINE-LINKED OLIGOSACCHARIDES, Rosalind Kornfeld and Stuart Kornfeld 631

DNA TOPOISOMERASES, James C. Wang 665

HUMAN APOLIPOPROTEIN MOLECULAR BIOLOGY AND GENETIC VARIATION, Jan L. Breslow 699

BIOSYNTHESIS AND METABOLISM OF TETRAHYDROBIOPTERIN AND MOLYBDOPTERIN, Charles A. Nichol, Gary K. Smith, and David S. Duch 729

STRUCTURAL ANALYSIS OF GLYCOCONJUGATES BY MASS SPECTROMETRY AND NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY, Charles C. Sweeley and Hernan A. Nunez 765

ORIGIN OF IMMUNE DIVERSITY: GENETIC VARIATION AND SELECTION, Tasuku Honjo and Sonoko Habu 803

THE CREATINE-CREATINE PHOSPHATE ENERGY SHUTTLE, Samuel P. Bessman and Christopher L. Carpenter 831

TRANSPOSITIONAL RECOMBINATION IN PROKARYOTES, Nigel D. F. Grindley and Randall R. Reed 863

PROTEIN-TYROSINE KINASES, Tony Hunter and Jonathan A. Cooper 897

PROTEIN KINASES IN THE BRAIN, Angus C. Nairn, Hugh C. Hemmings, Jr., and Paul Greengard 931
CONTENTS (continued) vii

CHEMICAL PROBES OF THE MITOCHONDRIAL ATP SYNTHESIS AND TRANSLOCATION, Pierre V. Vignais and Joël Lunardi 977

THE MITOCHONDRIAL ELECTRON TRANSPORT AND OXIDATIVE PHOSPHORYLATION SYSTEM, Youssef Hatefi 1015

DEVELOPMENTAL REGULATION OF HUMAN GLOBIN GENES, Stefan Karlsson and Arthur W. Nienhuis 1071

EUkARYOTIC PROTEIN SYNTHESIS, Kivie Moldave 1109

PSORALENS AS PHOTOACTIVE PROBES OF NUCLEIC ACID STRUCTURE AND FUNCTION: ORGANIC CHEMISTRY, PHOTOCHEMISTRY, AND BIOCHEMISTRY, George D. Cimino, Howard B. Gamper, Stephen T. Isaacs, and John E. Hearst 1151

THE APPLICATION OF NEUTRON CRYSTALLOGRAPHY TO THE STUDY OF DYNAMIC AND HYDRATION PROPERTIES OF PROTEINS, Anthony A. Kossiakoff 1195

INDEXES

Author Index 1229
Subject Index 1299
Cumulative Index of Contributing Authors, Volumes 50–54 1325
Cumulative Index of Chapter Titles, Volumes 50–54 1328