INTERACTIONS BETWEEN LAC REPRESSOR PROTEIN AND BROMODEOXYURIDINE-SUBSTITUTED OPERATOR DNA: IDENTIFICATION OF A SPECIFIC AMINO ACID-NUCLEOTIDE CONTACT USING UV FOOTPRINTING AND CROSS-LINK FORMATION

by

KYLE L. WICK

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE

Kathleen S. Matthews
Dr. Kathleen S. Matthews, Chair, Wiess Professor of Biochemistry and Cell Biology

K. M. Beckingham
Dr. Kathleen M. Beckingham, Associate Professor of Biochemistry and Cell Biology

Graham Palmer
Dr. Graham Palmer, Professor of Biochemistry and Cell Biology

John S. Olson
Dr. John S. Olson, Professor of Biochemistry and Cell Biology

Michael Gustin
Dr. Michael Gustin, Assistant Professor of Biochemistry and Cell Biology

Houston, Texas
April, 1990
INTERACTIONS BETWEEN LAC REPRESSOR PROTEIN AND
BROMODEOXYURIDINE-SUBSTITUTED OPERATOR DNA:
IDENTIFICATION OF A SPECIFIC AMINO ACID-NUCLEOTIDE
CONTACT USING UV FOOTPRINTING AND CROSS-LINK
FORMATION
by
Kyle L. Wick

ABSTRACT

As the classical model for negative transcriptional control in prokaryotes and
the subject of concentrated experimental attention, the lactose operon of Escherichia
coli presents a well-defined system for studying genetic control through protein-DNA
binding interactions. Binding of repressor at its cognate operator sequence within the
regulatory region of the operon, while responsive to environmental conditions,
efficiently inhibits transcription initiation by RNA polymerase. The high binding affinity
and degree of specificity exhibited by this protein-DNA complex has encouraged
investigation of the nature of the contacts formed.

We have explored specific contacts between the lac repressor and operator
using 5-bromodeoxyuridine-substituted DNA. Substitution of BrdU for single
thymidine positions in a synthetic 40 bp operator provides an indirect means of probing
the major groove of operator DNA for critical contacts between the repressor and the
5-methyl of individual thymidines. As a photoreactive species, BrdU provides
substrate for ultraviolet irradiation. Upon irradiation, strand scission occurs at the
BrdU residues. When bound, lac repressor protein provides protection against UV-
induced breakage depending on the nature of the sites and type of interaction. We
have confirmed thirteen unique sites of inducer-sensitive protection along the operator sequence (+1, 2, 3, 4, 6, 8, 13, 15, 16, 18, 19, 20, 21) using this method compared to per-substitution with BrdU (Ogata and Gilbert, 1977). The ability of these photosensitive DNAs to form short-range cross-links to bound protein has been used to determine the efficiency with which cross-linked protein-DNA complexes are generated at each individual site of BrdU substitution. Five sites of high efficiency cross-linking to the repressor protein have been identified (+3, 4, 14, 18, 19). Comparison of the UV protection results and the cross-linking data shows that these processes provide complementary tools for identifying and analyzing individual protein-DNA contacts.

Using these same BrdU-substituted operator DNAs, we attempted to define individual protein-DNA interactions with respect to the specific amino acid(s) making contact at a selected site within the operator sequence. With the selection of the T+3 site for our initial investigation, the cross-linked complex was formed and isolated. These polypeptide-DNA species were prepared for final analysis through a series of steps including proteolysis and anion-exchange HPLC. Protein sequence analysis on the purified peptide-operator complex identified a peptide spanning Val23 through Lys33. The data suggest His29 as the specifically cross-linked amino acid.
ACKNOWLEDGEMENTS

Of life's many treasures, my family and friends are valued above all. Their loving support and constructive encouragement through the years have given me the security and confidence to move forward. Most deserving of mention are my parents, James and Marilyn Sprague, for their unwavering belief in my abilities and concern for my continued happiness.

Dr. Kathleen Matthews, as my thesis advisor and as my friend, has been instrumental in my success and satisfaction at Rice University. Her individualized attention and keen insight into my strengths and weaknesses allowed me to develop both personally and professionally. As a successful and truly beautiful individual, she will remain an inspiration throughout my life.

Special acknowledgement of my husband, Dr. Timothy Wick, is also warranted for he is and has been a continual source of strength and love. With him, I can fully appreciate life's beauty and share in the wonders of biochemistry!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>MATERIALS AND METHODS</strong></td>
<td>35</td>
</tr>
<tr>
<td>Isolation of <em>lac</em> repressor</td>
<td>35</td>
</tr>
<tr>
<td>Inducer binding</td>
<td>36</td>
</tr>
<tr>
<td>Oligonucleotide synthesis and purification</td>
<td>36</td>
</tr>
<tr>
<td>5'-End-labeling</td>
<td>38</td>
</tr>
<tr>
<td>Strand hybridization</td>
<td>39</td>
</tr>
<tr>
<td>Equilibrium repressor-operator binding</td>
<td>39</td>
</tr>
<tr>
<td>Silanization of glassware and plasticware</td>
<td>40</td>
</tr>
<tr>
<td>Repressor-operator protection determination</td>
<td>40</td>
</tr>
<tr>
<td>Repressor-operator cross-linking efficiency measurements</td>
<td>41</td>
</tr>
<tr>
<td>Large-scale repressor-T+3/B operator cross-link formation</td>
<td>42</td>
</tr>
<tr>
<td>Repressor-operator cross-link isolation</td>
<td>43</td>
</tr>
<tr>
<td>Proteolytic digestion of cross-linked complexes</td>
<td>43</td>
</tr>
<tr>
<td>HPLC purification of the peptide-operator complex</td>
<td>44</td>
</tr>
<tr>
<td>Protein sequencing</td>
<td>45</td>
</tr>
<tr>
<td><strong>RESULTS</strong></td>
<td>46</td>
</tr>
<tr>
<td>Oligonucleotide synthesis, purification and preparation</td>
<td>46</td>
</tr>
<tr>
<td>Equilibrium binding studies</td>
<td>53</td>
</tr>
<tr>
<td>Repressor-operator protection determination</td>
<td>56</td>
</tr>
<tr>
<td>Repressor-operator cross-link formation</td>
<td>93</td>
</tr>
<tr>
<td>Large-scale repressor-T+3/B operator cross-linking</td>
<td>119</td>
</tr>
<tr>
<td>Cross-link isolation</td>
<td>122</td>
</tr>
<tr>
<td>Proteolysis</td>
<td>125</td>
</tr>
<tr>
<td>HPLC purification</td>
<td>125</td>
</tr>
<tr>
<td>Protein sequencing</td>
<td>141</td>
</tr>
<tr>
<td><strong>DISCUSSION</strong></td>
<td>145</td>
</tr>
<tr>
<td><strong>BIBLIOGRAPHY</strong></td>
<td>160</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGURE 1</td>
<td>Schematic Representation of Regulatory Interactions Operative in Inducible and Repressible Negative Control Systems</td>
<td>3</td>
</tr>
<tr>
<td>FIGURE 2</td>
<td>The Helix-Turn-Helix Structural Motif Shown Within the Major Groove</td>
<td>6</td>
</tr>
<tr>
<td>FIGURE 3</td>
<td>Genetic Organization of the Regulatory Elements of the lac Operon</td>
<td>10</td>
</tr>
<tr>
<td>FIGURE 4</td>
<td>Primary Sequence of the lac Repressor Protein</td>
<td>16</td>
</tr>
<tr>
<td>FIGURE 5</td>
<td>Lac Operator Sequence with Summary of Footprint and O(^c) Mutant Information</td>
<td>22</td>
</tr>
<tr>
<td>FIGURE 6</td>
<td>Structures of Thymidine and 5-Bromodeoxyuridine</td>
<td>24</td>
</tr>
<tr>
<td>FIGURE 7</td>
<td>Initial Photochemical Consequences of Ultraviolet Irradiation of BrdU</td>
<td>27</td>
</tr>
<tr>
<td>FIGURE 8</td>
<td>Sequence of the 40 bp lac Operator Fragment</td>
<td>32</td>
</tr>
<tr>
<td>FIGURE 9</td>
<td>Purification of BrdU-Containing Oligonucleotides</td>
<td>48</td>
</tr>
<tr>
<td>FIGURE 10</td>
<td>Autoradiogram Illustrating the Efficiency of Labeling Single-Stranded BrdU-Containing Oligonucleotides at the 5'-OH</td>
<td>50</td>
</tr>
<tr>
<td>FIGURE 11</td>
<td>Autoradiogram Illustrating Strand Hybridization Between the Labeled, BrdU-Substituted T(_+3) Strand and the Complementary Bottom Strand (B)</td>
<td>52</td>
</tr>
<tr>
<td>FIGURE 12</td>
<td>Repressor-Operator Protection Determination for T(<em>{PER}/B), T/B(</em>{PER})</td>
<td>58</td>
</tr>
<tr>
<td>FIGURE 13</td>
<td>Repressor-Operator Protection Determination for T.(_5/B) and T.(_3/B)</td>
<td>61</td>
</tr>
<tr>
<td>FIGURE 14</td>
<td>Repressor-Operator Protection Determination for T.(_3/B)</td>
<td>63</td>
</tr>
<tr>
<td>FIGURE 15</td>
<td>Repressor-Operator Protection Determination for T.(_4/B)</td>
<td>65</td>
</tr>
<tr>
<td>FIGURE 16</td>
<td>Repressor-Operator Protection Determination for T.(_6/B)</td>
<td>67</td>
</tr>
<tr>
<td>FIGURE 17</td>
<td>Repressor-Operator Protection Determination for T.(<em>{14}/B) and T.(</em>{20}/B)</td>
<td>69</td>
</tr>
<tr>
<td>FIGURE 18</td>
<td>Repressor-Operator Protection Determination for T.(<em>{21}/B) and T.(</em>{22}/B)</td>
<td>71</td>
</tr>
<tr>
<td>FIGURE 19</td>
<td>Repressor-Operator Protection Determination for T/B(_{+1})</td>
<td>73</td>
</tr>
<tr>
<td>FIGURE 20</td>
<td>Repressor-Operator Protection Determination for T/B(_{+2})</td>
<td>75</td>
</tr>
<tr>
<td>FIGURE 21</td>
<td>Repressor-Operator Protection Determination for T/B(_{+8})</td>
<td>77</td>
</tr>
<tr>
<td>FIGURE</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>22</td>
<td>Repressor-Operator Protection Determination for T+B+13</td>
<td>79</td>
</tr>
<tr>
<td>23</td>
<td>Repressor-Operator Protection Determination for T+B+15 and T+B+16</td>
<td>81</td>
</tr>
<tr>
<td>24</td>
<td>Repressor-Operator Protection Determination for T+B+18 and T+B+19</td>
<td>83</td>
</tr>
<tr>
<td>25</td>
<td>Repressor-Operator Protection Determination for T+B+24</td>
<td>85</td>
</tr>
<tr>
<td>26</td>
<td>Repressor-Operator Protection Determination for T+3/B and T+B+8</td>
<td>88</td>
</tr>
<tr>
<td>27</td>
<td>UV Protection Measurements Along the Operator Sequence</td>
<td>92</td>
</tr>
<tr>
<td>28</td>
<td>Sequence and Salt Dependence of Repressor-DNA Cross-link Formation</td>
<td>95</td>
</tr>
<tr>
<td>29</td>
<td>UV Dose-Response for Repressor-Operator Cross-link Formation</td>
<td>97</td>
</tr>
<tr>
<td>30</td>
<td>Repressor-Operator Cross-link Formation with T-5/B and T-3/B</td>
<td>100</td>
</tr>
<tr>
<td>31</td>
<td>Repressor-Operator Cross-link Formation with T+B+3/B and T+B+6/B</td>
<td>102</td>
</tr>
<tr>
<td>32</td>
<td>Repressor-Operator Cross-link Formation with T+B+14/B, T+B+20/B, T+B+21/B and T+B+22/B</td>
<td>104</td>
</tr>
<tr>
<td>33</td>
<td>Repressor-Operator Cross-link Formation with T+B+1 and T+B+2</td>
<td>106</td>
</tr>
<tr>
<td>34</td>
<td>Repressor-Operator Cross-link Formation with T+B+8 and T+B+13</td>
<td>108</td>
</tr>
<tr>
<td>35</td>
<td>Repressor-Operator Cross-link Formation with T+B+15 and T+B+16</td>
<td>110</td>
</tr>
<tr>
<td>36</td>
<td>Repressor-Operator Cross-link Formation with T+B+18 and T+B+19</td>
<td>112</td>
</tr>
<tr>
<td>37</td>
<td>Repressor-Operator Cross-link Formation with T+B+24</td>
<td>114</td>
</tr>
<tr>
<td>38</td>
<td>Repressor-Operator Cross-link Formation</td>
<td>116</td>
</tr>
<tr>
<td>39</td>
<td>Efficiency of Repressor Cross-linking to Substituted Sites</td>
<td>118</td>
</tr>
<tr>
<td>40</td>
<td>Comparison of Repressor-Operator Cross-link Efficiencies</td>
<td>121</td>
</tr>
<tr>
<td>41</td>
<td>Large-Scale Isolation of Cross-linked Repressor-T+B Operator Complex</td>
<td>124</td>
</tr>
<tr>
<td>FIGURE</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>42</td>
<td>Tryptic and Chymotryptic Cleavage Sites Within the <em>lac</em> Repressor Sequence</td>
<td>127</td>
</tr>
<tr>
<td>43</td>
<td>Proteolytic Digestion of the Repressor-T+3/B Operator Complex</td>
<td>129</td>
</tr>
<tr>
<td>44</td>
<td>Anion Exchange HPLC Profile of Digested <em>lac</em> Repressor Protein</td>
<td>132</td>
</tr>
<tr>
<td>45</td>
<td>Anion Exchange HPLC Profile of Proteolysed Repressor-T+3/B Operator Complex (1)</td>
<td>134</td>
</tr>
<tr>
<td>46</td>
<td>Anion Exchange HPLC Profile of Proteolysed Repressor-T+3/B Operator Complex (2)</td>
<td>136</td>
</tr>
<tr>
<td>47</td>
<td>Protein Sequence Analysis of the Peptide-T+3/B Operator Complex</td>
<td>143</td>
</tr>
<tr>
<td>48</td>
<td>Contact Between the <em>Lac</em> Repressor NH2-terminus and Operator DNA Derived From Molecular Dynamic Simulation of NMR Data</td>
<td>157</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>Equilibrium Dissociation Constants for BrdU-Containing Operator DNA</td>
<td>55</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>UV Protection and Cross-link Efficiency Measurements</td>
<td>90</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>Recovery of the Cross-linked Protein-DNA Complex</td>
<td>139</td>
</tr>
</tbody>
</table>
INTRODUCTION

Genetic regulation is an essential function in all living organisms [some sequences in this introduction are from a review chapter prepared by Wick, K.L. and Matthews, K.S. (1989) in Molecular Biology of Chromosome Function. (Adolph, K., ed.), Springer-Verlag, New York.]. In prokaryotes genetic control provides responsivity to a constantly changing external environment, and bacterial systems have proven extremely valuable in elucidating the variety of potential mechanisms. Both positive and negative strategies exist for transcriptional control; this format for regulation at the initiation of mRNA synthesis allows maximum conservation of cellular energy by preventing generation of unnecessary proteins at an early stage in their production. Positive control over a wide range of systems in *E. coli* is exerted through catabolite repression mediated by the CAP-cAMP system. Negative control in bacteria involves a series of repressor proteins, each specific for a particular pathway. The basic regulatory scheme involves the interaction of a specific protein with its target DNA sequences in an inducible or repressible manner (Figure 1). Inducible mechanisms generally govern catabolic pathways where expression is necessary only in the presence of substrate. These genes are "turned off" through a repressor protein-operator DNA interaction which interferes with transcription initiation by RNA polymerase. Derepression or induction is accomplished through destabilization of the repressor-operator complex upon binding of substrate, or a closely related inducer molecule, to the repressor protein. In contrast, repressible mechanisms most often apply to biosynthetic pathways where gene expression is modulated by levels of the end-product. To avoid over-accumulation, repression is established when the end-product or related compound, acting as a co-repressor, binds to and activates an
Figure 1. Schematic Representation of Regulatory Interactions Operative in Inducible and Repressible Negative Control Systems. The genes or sequences for repressor protein (R), promoter (P), operator (O) and the structural genes 1 and 2 are shown. Active repressor is depicted by \( \bigcirc \), while \( \square \) represents the induced or inactive form of the protein. The inducer or co-repressor molecule is shown as \( \square \).
Figure 1.

INDUCIBLE SYSTEM

REPRESSOR

| R | PO|GENE-1|GENE-2|

INDUCER

| R | PO|GENE-1|GENE-2|

RNA TRANSCRIPT

PROTEIN 1 PROTEIN 2

REPRESSIBLE SYSTEM

REPRESSOR

| R | PO|GENE-1|GENE-2|

COREPRESSOR

| R | PO|GENE-1|GENE-2|

RNA TRANSCRIPT

PROTEIN 1 PROTEIN 2
an aporepressor protein. The resulting complex recognizes its specific DNA sequence and thereby blocks transcription initiation.

A feature common to many of these repressors is the presence of regions homologous to a secondary structural motif found in DNA binding domains of proteins with known X-ray crystallographic structures (McKay and Steitz, 1981; Anderson et al., 1981; Anderson et al., 1984). In this general model for DNA binding, a pair of helices from each of two subunits of a DNA binding protein makes contact with adjacent major grooves in right-handed B-DNA. Within each protein, the α-helices are connected by a β-turn. A schematic representation of the two-helix structure within the major groove is shown in Figure 2. Evidence for the DNA-binding function and specificity conferred by this helical structure was presented by Wharton and Ptashne (1985) when they reported the exchange of 434 repressor binding specificity for that of the P22 repressor by a corresponding exchange of critical amino acids within the 434 repressor. The sequence changes within the α-helical region of the 434 repressor resulted in a hybrid protein with the binding specificity of the P22 repressor protein. This helix-turn-helix motif has been hypothesized based on crystal structures, genetic analysis, and DNA protection and modification studies to exist in the λCRO protein, E. coli CAP protein, trp repressor and lambda CI and cro repressors (Matthews et al., 1982; Weber et al., 1982; Sauer et al., 1982; Anderson et al., 1982; Ohlendorf et al., 1982; Schevitz et al., 1985; Kelley and Yanofsky, 1985). Sequence homologies between these proteins and structurally undefined DNA binding proteins suggest that many regulatory proteins may recognize their cognate binding sites through this same structural motif. Although differences do exist among the proteins regarding the nature of individual DNA contacts and the precise orientation of the α-helices, the evidence at least suggests that certain general or common principles may apply toward specific
Figure 2. Complex Between Two *Lac* Repressor Protein Headpieces and the Symmetric Operator Based on NMR Data (derived from Lamerichs *et al.*, 1989). Helices B and C represent the proposed helix-turn-helix DNA-binding structure of the repressor protein.
protein-DNA recognition. Transcription to yield mRNA for inducible systems occurs in the presence of the substrate for the pathway and, in the case of sugar pathways, in simultaneous absence of glucose. Thus, the enzymes necessary for catabolism are synthesized only under conditions which require their activity (e.g., for generation of a carbon source or for protection from antibiotic). The operons which produce enzymes involved in sugar metabolism are also under positive control of the catabolite activator protein (CAP). The CAP protein (Mr ~ 45,000) functions as a dimer of identical subunits. The crystal structure of the CAP-cAMP complex has been solved and refined to 2.5Å resolution (McKay and Steitz, 1981; McKay et al., 1982; Weber and Steitz, 1987). Each subunit of CAP binds a cAMP molecule in its large N-terminal domain and contains the helix-turn-helix DNA binding structure within its smaller C-terminal domain (Weber and Steitz, 1987). The structure of the protein exhibits asymmetry in that one subunit possesses a large cleft between the two structural domains and is termed “open” while the second subunit is “closed” with no obvious cleft (Weber and Steitz, 1987). CAP binds to sites within the vicinity of many E. coli promoters when activated by complex formation with cAMP. The presence of this CAP-cAMP complex stimulates transcription of the associated structural genes by RNA polymerase, potentially through protein-protein interactions with the polymerase molecule. CAP-cAMP binding induces a conformational change that has properties consistent with a bend in the DNA (Kolb et al., 1983; Liu-Johnson et al., 1986; Warwicker et al., 1987; Dripps and Wartell, 1987; Koo and Crothers, 1988). cAMP levels are modulated by the influence of glucose availability on adenylate cyclase activity; thus, catabolite repression assures that alternate carbon sources are utilized only in the absence of the more efficiently metabolized glucose.
The lactose operon of *Escherichia coli* represents the classic inducible negative control system (Jacob and Monod, 1961). The operon consists of the genes encoding enzymes responsible for the transport and metabolism of lactose as well as the I gene coding for the repressor protein (Figure 3). The structural genes, z, y and a encode β-galactosidase, lactose permease and a transacetylase enzyme, respectively. Responsible for the initial reactions of lactose metabolism, lactose permease is involved in transport of the sugar into the cell, while β-galactosidase hydrolyzes lactose to glucose and galactose and converts lactose to allolactose in small amounts as a side reaction involving transglucosylation. The function of the transacetylase enzyme is poorly understood. The structural genes are subject to both positive (CAP binding) and negative (repressor binding) transcriptional control. In the absence of lactose, the *lacI* gene product, transcriptionally regulated from its own promoter, represses initiation of *lacZY* transcription by binding with high affinity to its recognition site within the control region, the operator (Müller-Hill, 1975; Bourgeois and Pfahl, 1976; Müller-Hill et al., 1976). Due to the overlap of the promoter and operator regions, the binding of repressor exerts its negative control by interfering with transcription initiation by adjacently bound RNA polymerase. Derepression of the structural genes is accomplished through interaction of the repressor protein with 1,6-allolactose, produced by β-galactosidase, or other β-galactosides (Jobe and Bourgeois, 1972; Müller-Hill et al., 1964). The protein responds to this inducer binding with a conformational change (Laiken et al., 1972; Ohshima et al., 1972; Matthews et al., 1973; Wu et al., 1976; Boschelli et al., 1981), that translates to a destabilization of the operator binding interaction (Barkley et al., 1975; O’Gorman et al., 1980; Daly and Matthews, 1986b). In its inducer-bound state, the repressor exhibits diminished affinity for operator and binds with comparable affinity to nonspecific DNA sequences.
Figure 3. Genetic Organization of the Regulatory Elements of the *lac* Operon.

Centrally positioned, the primary operator (O) represents the target sequence for *lac* repressor. The two overlapping promoters, P1 and P2, are shown along with the position of the CAP-cAMP binding site ~. Flanking the central regulatory region, the pseudooperators , which act as secondary repressor recognition sites, are located upstream within the *lacI* gene (O₁) and downstream within the *lacZ* gene (O₂). The small arrows indicate transcription polarity from P₁ and P₂.
Figure 3.
present in the genome (Lin and Riggs, 1975). With the operator site open, transcription can proceed provided RNA polymerase and the CAP-cAMP complex are bound at their neighboring recognition sites. Transcription and translation of the structural genes continues until the lactose and consequently 1,6-allolactose pools are depleted or glucose becomes available. Once free of bound inducer, the repressor resumes its high affinity binding conformation to interact at the operator site and to shut down expression from the lac operon. In addition to its role in repression, Straney and Crothers (1987) have shown that lac repressor bound at the operator site increased the binding of RNA polymerase to the promoter. Thus, the repressor also serves to facilitate the first round of transcription once inducer is bound and therefore promotes a rapid induction response.

Modulation of the negative control exerted by the lac repressor is effected not only by inducer binding but also by environmental conditions. The binding of repressor to nonspecific DNA is electrostatic in nature and therefore sensitive to ionic strength (Butler et al., 1977; Revzin and von Hippel, 1977; deHaseth et al., 1977). Repressor-DNA binding studies performed in a variety of buffers and using Na\(^{+}\) as the cation and either Cl\(^{-}\) or CH\(_3\)COO\(^{-}\) as the anion showed a linear relationship between log(\(K_{\text{obs}}\)) and log([Na\(^{+}\)]) (deHaseth et al., 1977). Both Revzin and von Hippel (1977) and deHaseth et al. (1977) found that this binding interaction required the release of 11±2 monovalent ions while resulting in the formation of 12±2 new ionic interactions. Their data suggest that it is the displacement of cations from the DNA that accounts for much of the ion release. Binding of repressor to non-operator DNA would therefore be compromised under conditions of high ionic strength where the driving force for interaction, counterion release, would be less favorable. Other factors, such as pH and temperature, influence local ion concentration and protein structure and therefore are
also critical to repressor-DNA complex formation. The influence of ionic strength on repressor-operator interaction has indicated that in addition to ionic components, apolar interactions contribute significantly to the free energy of binding (Record et al., 1977; Barkley et al., 1981; Barkley, 1981; Winter and von Hippel, 1981; Winter et al., 1981; Whitson and Matthews, 1986; Whitson et al., 1986). While counterion release and ionic interactions appear to account entirely for the thermodynamics of repressor-DNA binding, the specific repressor-operator interaction relies on these factors to a lesser degree: 8±1 ionic interactions versus 12±2 for the complex with nonspecific DNA (Record et al., 1977). Whitson et al. (1986) studied the thermodynamics of repressor association with various lac operator-containing DNAs. Their results confirmed 8 ionic interactions between the repressor protein and operator- or operator and pseudooperator-containing plasmids. Binding with a 40 bp operator fragment, however, involved only 6 ionic interactions, indicating that the protein binding site on the DNA possibly extends beyond the central 40 bp of the operator sequence (Whitson et al., 1986). Hydrophobic interactions between the repressor and the operator sequence contribute to the specificity of binding. Kinetic analyses of repressor-operator binding suggest a two-step mechanism of association with a search of decreased dimensions once the DNA is bound nonspecifically (Winter et al., 1981; Berg et al., 1982; Whitson and Matthews, 1984). According to this mechanism, which accounts for interaction with operator at a rate beyond that possible by diffusion-control, the repressor first diffuses to a nonspecific site on the DNA and then rapidly scans adjacent sequences until the operator is recognized (Winter et al., 1981).

Sequences homologous to the operator have been found both ~100 bp upstream (O₁) and ~400 bp downstream (O₂) of the regulatory region (Pfahl et al., 1979; Reznikoff et al., 1974) (Figure 3). These sequences, termed pseudo-operators, have
been shown to bind repressor, albeit with decreased affinities compared to the primary operator, and have been implicated in the overall regulatory mechanism. These sequences in the *lacI* and *lacZ* genes influence binding to operator both *in vitro* and *in vivo*. Although the sequences individually bind repressor with affinities 20-fold (O₂) and 100-fold (Oᵢ) lower than primary operator, their presence in linear operator-containing DNA results in modest stabilization of the repressor-DNA complex (Winter and von Hippel, 1981; Whitson and Matthews, 1986; Hsieh *et al.*, 1987). Ternary complex formation with occupation of both operator sites on the protein has been suggested to explain this effect of sequence context on binding affinity (Whitson and Matthews, 1986). Significant enhancement (>1000-fold) of binding affinity by supercoiling of dual operator-containing DNAs (Whitson *et al.*, 1987; Kramer *et al.*, 1988) is consistent with ternary complex and consequent DNA loop formation. Such DNA looping is influenced by the spacing between operators in linear DNA (Kramer *et al.*, 1987) and, as anticipated based on mechanism, the effect of spacing is diminished and altered by supercoiling (Kramer *et al.*, 1988; Whitson *et al.*, 1987). Direct electron microscopic examination has confirmed the presence of DNA loop formation via operator-repressor-operator association (Kramer *et al.*, 1987;1988). *In vivo* measurements of the effect of multiple operators on expression are consistent with these *in vitro* results and with DNA loop formation (Mossing and Record, 1986; Besse *et al.*, 1986; Borowiec *et al.*, 1987; Eismann *et al.*, 1987). The role of looped DNA structures in genetic regulation is evident in other prokaryotic systems; the *ara*, *gal* and *deo* operons all require occupation of multiple operators for full repression, and DNA loop formation offers a probable explanation for the observations in all of these inducible systems (Dandanell *et al.*, 1987).
The primary structure of the lactose repressor protein is known both from its gene sequence and from analysis of the protein itself (Beyreuther et al., 1975; Farabaugh, 1978) (Figure 4). In solution the repressor exists as a tetramer of identical subunits (M_r ~150,000) (Riggs and Bourgeois, 1968) and has 4 inducer sites and 2 operator DNA binding sites (Butler et al., 1977; Culard and Maurizot, 1981; O'Gorman et al., 1980; Whitson and Matthews, 1986). Cooperativity in inducer binding is observed in the presence of operator DNA or at elevated pH in the absence of operator (O'Gorman et al., 1980; Daly and Matthews, 1986a,b). The protein can be separated into two structural domains through mild proteolytic digestion: a tetrameric core consisting of amino acids 60-360 and four amino-terminal domains composed of the first 59 amino acids (Platt et al., 1973; Jovin et al., 1977). Amino acid sequence homology between proteins of known structure and lac repressor NH2-terminus strongly suggests the existence of a helix-turn-helix DNA-binding motif (Pabo and Sauer, 1984) (Figure 4). For example, the region of the lac repressor protein from amino acids 17-26 is thought to correspond to the known DNA-binding α-helix (aa 27-36) of the λCRO repressor. The sequence homology also suggests that this stretch of amino acids within the lac protein folds into an α-helical conformation (Matthews et al., 1982). In addition, homology with the CRO repressor also suggests that amino acids 5-13 of lac repressor form an α-helix and participate in DNA binding (Matthews et al., 1982). Further support has been derived from NMR analysis of the lac repressor NH2-terminal domain in which the folding patterns obtained are consistent with the postulated two-helix structure (Kaptein et al., 1985; Zuiderweg et al., 1985). The arrangement of the core and NH2-terminal domains within the tetrameric protein has been deduced from both powder X-ray diffraction analysis of microcrystals (Steitz et al., 1974) and low angle X-ray and neutron diffraction studies of the native and core
Figure 4. Primary Sequence of the *lac* Repressor Protein (Beyreuther, *et al.*, 1975). The underlined portion from Thr5 to Gln26 represents region of the repressor protein exhibiting sequence homology to the helix-turn-helix motif identified in other DNA-binding proteins.
Figure 4.
Met-Lys-Pro-Val-Thr-Leu-Tyr-Asp-Val-Ala-Glu-Tyr-Ala-Gly-Val
10
Ser-Tyr-Gln-Thr-Val-Ser-Arg-Val-Val-Asn-Gln-Ala-Ser-His-Val
20
Ser-Ala-Lys-Thr-Arg-Glu-Lys-Val-Glu-Ala-Ala-Met-Ala-Glu-Leu
40
Asn-Tyr-Ile-Pro-Asn-Arg-Val-Ala-Gln-Gln-Leu-Ala-Gly-Lys-Gln
50
Ser-Leu-Leu-Ile-Gly-Val-Ala-Thr-Ser-Ser-Leu-Ala-Leu-His-Ala
70
Pro-Ser-Gln-Ile-Val-Ala-Ala-Ile-Lys-Ser-Arg-Ala-Asp-Gln-Leu
80
Gly-Ala-Ser-Val-Val-Val-Met-Val-Glu-Arg-Ser-Gly-Val-Glu
90
Ala-Cys-Lys-Ala-Ala-Ala-His-Asn-Leu-Leu-Ala-Gln-Arg-Val-Ser
100
Gly-Leu-Ile-Ile-Tyr-Pro-Leu-Asp-Asp-Ala-Ile-Ala
110
Val-Glu-Ala-Ala-Cys-Thr-Asn-Val-Pro-Ala-Leu-Phe-Leu-Asp-Val
120
140
Ser-Asp-Glu-Thr-Pro-Ile-Asn-Ser-Ile-Ile-Phe-Ser-His-Glu-Asp
160
Gly-Thr-Arg-Leu-Gly-Val-Glu-His-Leu-Val-Ala-Leu-Gly-His-Gln
170
Gln-Ile-Ala-Leu-Leu-Ala-Gly-Pro-Leu-Ser-Ser-Val-Ser-Ala-Arg
180
Leu-Arg-Leu-Ala-Gly-Trp-His-Lys-Tyr-Leu-Thr-Arg-Asn-Gln-Ile
190
200
Gln-Pro-Ile-Ala-Glu-Arg-Glu-Gly-Asp-Trp-Ser-Ala-Met-Ser-Gly
210
Phe-Gln-Gln-Thr-Met-Gln-Met-Leu-Asn-Glu-Gly-Ile-Val-Pro-Thr
220
230
240
250
Ala-Ile-Thr-Glu-Ser-Gly-Leu-Arg-Val-Gly-Ala-Asp-Ile-Ser-Val
260
270
Val-Gly-Tyr-Asp-Asp-Thr-Glu-Asp-Ser-Ser-Cys-Tyr-Ile-Pro-Pro-
280
Leu-Thr-Thr-Ile-Lys-Gln-Asp-Phe-Arg-Leu-Leu-Gly-Gln-Thr-Ser
290
300
310
Asn-Gln-Leu-Leu-Pro-Val-Ser-Leu-Val-Lys-Arg-Lys-Thr-Thr-Leu
320
330
Ala-Pro-Asn-Thr-Gln-Thr-Ala-Ser-Pro-Arg-Ala-Leu-Ala-Asp-Ser
340
Leu-Met-Gln-Leu-Ala-Arg-Gln-Val-Ser-Arg-Leu-Glu-Ser-Gly-Gln
350
360
repressors (Charlier et al., 1981; McKay et al., 1982; Pilz et al., 1980). These studies suggest an elongated shape for the protein with the NH2-terminal domains situated at the end of the tetrameric core. The end domains, linked to the COOH-terminal core by a hinge region, exhibit differential mobility and appear motionally flexible relative to the core region (Wade-Jardetzky et al., 1979). This flexibility becomes functionally critical to the induction response where binding and dissociation of inducer within the core region results in conformational changes in the protein, translated to the NH2-terminal DNA-binding domains through the hinge region. The specificity and affinity of DNA-binding are sensitive to these conformational alterations. A definitive structural description awaits X-ray crystallographic analysis, which has been hampered by the difficulty in obtaining crystals. A recent report, however, documents the crystallization of the tetrameric lac repressor, a lac repressor-inducer complex and the ternary repressor-operator DNA complex with an anti-inducer (Pace, et al., 1990)

Mutational and physical analyses have made possible assignment of function to the two domains. Mutational studies suggested structural independence of inducer and DNA-binding activities (Adler et al., 1972; Pfahl et al., 1974; Müller-Hill et al., 1975; Miller, 1984; Miller, 1979; Gordon et al., 1988). Mutations that affect DNA-binding cluster in the NH2-terminal region with a lesser number scattered throughout the core domain. In contrast, amino acid changes that alter the inducer-binding capability of lac repressor are found entirely within the COOH-terminal 300 amino acids of the protein. Mutations affecting subunit aggregation cluster within the COOH-terminal 100 amino acids of the protein. Genetic methods have been combined with physical characterizations to demonstrate effects of specific repressor amino acid substitution and/or operator base alteration on the interaction between the two. Müller-Hill and colleagues have designed a two-plasmid system for detecting repressors with altered
DNA specificities (Lehming et al., 1987). Their studies have resulted in the isolation of several repressor mutants that repress transcription by recognizing select operator variants. Repressor mutations have been localized to two regions of the recognition helix: amino acids 17 and 18 and amino acid 22 (Tyr17, Gln18 and Arg22 in wild-type repressor). These regions of the binding helix appear to contact the TA of position 6 and GC at 7 and the GC at position 5 (see Figure 4 for repressor sequence and Figure 5 for operator sequence), respectively (Lehming et al., 1987; 1988; Sartorius et al., 1989). These results suggest that the recognition helix is oriented such that the NH2-terminal end enters the major groove of the operator near the center of sequence symmetry while the COOH-terminus exits the groove away from the center of symmetry (Lehming et al., 1988). When compared to the P22 and 434 repressors, the recognition helix of the lac repressor appears to bind in an opposite orientation (Lehming et al., 1987; 1988). In an earlier study, Ebright (1986) likewise identified a critical contact between Gln18 and GC7 of the operator by simultaneous screening of repressor and operator variants for complementary mutations. Another study involving NMR analysis of the lac repressor NH2-terminal domain-22 bp operator interaction provided further evidence for the contrasting orientation of the lac repressor recognition helix with respect to other repressors and suggested the existence of an additional specific interaction between His29 and base pair 3 (Lamerichs et al., 1989). These studies have demonstrated the importance of specific residues within the helix-turn-helix region for operator recognition and identified the bases contacted within the operator sequence. Without a confirmed structure for the lac repressor protein or the repressor-operator complex, this information is invaluable to the deduction of possible molecular forms.
Physical and chemical methods have complemented information from genetic analysis. The tetrameric core protein produced by proteolytic digestion binds inducer with wild-type affinity (Platt et al., 1973); although this species does not bind to nonspecific DNA sequences (Friedman and Matthews, 1978), it may contain determinants for operator binding (Matthews, 1979; Manly et al., 1984). Examination of operator binding of the isolated NH2-terminal domain indicates nonspecific binding as well as interaction with operator DNA sequences (Geisler and Weber, 1977; Schnarr et al., 1983; Ogata and Gilbert, 1979; Boelens et al., 1987; Nick et al., 1982; Barbier et al., 1984). Results from chemical studies are also consistent with the placement of the subunit interface at the C-terminal region of the lactose repressor monomer (Daly and Matthews, 1986a; Sams et al., 1985). Sequence homology with periplasmic sugar binding proteins (Müller-Hill, 1983) has formed the basis for postulation of residues that form an inducer binding site (Sams et al., 1984); this model is consistent with effects of variation of inducer structure on inducer binding (Chakerian et al., 1987).

As the specific target for repression control, the operator element was hypothesized by Jacob and Monod (1961) on the basis of cis-dominant mutations that resulted in operon constitutivity in the presence of active repressor. The identification and sequence analysis of operator-constitutive mutants provided information regarding critical sites and sequence requirements for specific repressor-DNA association (Smith and Sadler, 1971; Sadler and Smith, 1971; Gilbert et al., 1975). DNase I footprinting experiments identified regions of DNA contact within the operator sequence (Schmitz and Galas, 1979; Manly et al., 1984), and exonuclease III footprint analysis of the bound complex defined the extent of DNA coverage by the repressor protein (Shalloway et al., 1980; Manly et al., 1984). In addition, major and minor groove
contacts have been identified by DNA methylation experiments (Gilbert et al., 1976; Ogata and Gilbert, 1978; 1979; Manly and Matthews, 1984). These studies defined the operator locus as a specific sequence spanning up to 40 nucleotides and displaying a high degree of two-fold symmetry. A summary of results derived from the above repressor-operator footprinting experiments is illustrated in Figure 5. The repressor appears to favor a single face of the operator and binds in an asymmetric fashion with some of the most critical contacts located within the promoter-proximal half of the sequence (Gilbert et al., 1975; Ogata and Gilbert, 1977; Schmitz and Galas, 1979; Manly et al., 1984). Characteristics of the operator base sequence essential for binding have been identified by production of synthetic substituted operators (Goeddel et al., 1977; 1978; Lin and Riggs, 1972; 1974; Ogata and Gilbert, 1977).

Base substitution offers an indirect means of identifying sites and/or groups critical for contact formation. The thymidine methyl has long been a target for investigation as a major groove recognition determinant. Substitution of thymidines within the operator sequence by uridine or 5-bromodeoxyuridine allowed identification of several critical protein-DNA contacts (Goeddel et al., 1977; 1978a,b). Dissociation rate analysis of uridine-substituted DNAs confirmed AT base pair 8 (see Fig. 5 for numbering system) as a major contact point and further implicated the thymidine methyl group as the recognition element (Goeddel et al., 1977). To extend these studies, Caruthers and colleagues utilized bromodeoxyuridine as a thymidine analog. The two bases are structurally very similar (Figure 6), as the 5-substituents differ little in size, with van der Waals radii of bromine and methyl at 1.95Å and 2.00Å, respectively. Goeddel et al. (1978a,b) chemically synthesized BrdU-containing operator sequences and analyzed the effect of single and multiple substitutions on the stability of the repressor-operator complexes. The substitution of certain sites within the operator contributed in a small
Figure 5. *Lac* Operator Sequence with Summary of Footprint and O\textsuperscript{C} Mutant Information. The sequence of the synthetic 40 bp *lac* operator used in these studies is shown. T and B refer to the top and bottom strands, respectively. The sequence is numbered relative to the transcription start site (+1). A summary of repressor-operator footprinting results and O\textsuperscript{C} mutations is also presented. Exonuclease III digestion endpoints are shown by filled bars (Shalloway *et al.*, 1980; Manly *et al.*, 1984). The solid lines above and below the sequence depict the region of protection against DNase I digestion, while partial protection is indicated by a broken line (Schmitz and Galas, 1979; Manly *et al.*, 1984). The arrow points to the site of enhanced cleavage by DNase I (Schmitz and Galas, 1979). DNA methylation protection and enhancement in the presence of repressor is illustrated by (-) and (+), respectively (Gilbert *et al.*, 1976; Ogata and Gilbert, 1978; 1979; Manly and Matthews, 1984). Thymidine positions protected by repressor against UV damage are indicated by arrowheads (Ogata and Gilbert, 1977). Operator constitutive mutations are shown below the sequence (Smith and Sadler, 1971; Sadler and Smith, 1971; Gilbert *et al*, 1975).
Figure 5.

T 5' TGGTTGTTGTTGGAATTTGGAGCGGATAACAATTTCACACAGG

B 3' ACAACACACCTAACCCTACACTCGCCCTATTGTTAAAGTGTGTC

-10 +1 +10 +20 +30

A TGGTTA C T
T ACAAT G A
Figure 6. Structures of Thymidine and 5-Bromodeoxyuridine. The compounds differ structurally at the C5 position. The van der Waals radius of CH₃ is 2.00 Å, compared to 1.95 Å for the bromine atom. Electronically, substitution of Br for CH₃ introduces polarity to the molecule with the exchange of a highly electronegative group for the relatively hydrophobic methyl function.
Figure 6.

Thymidine 5-Bromodeoxyuridine
way to the stability of the protein-DNA complex (i.e. operator positions 4, 6, 13, 14) possibly due to a direct contact between a polar group within the protein and the electronegative bromine atom (Goeddel et al., 1978a). Substitution of BrdU at positions 3, 15, 16 and 19 had no apparent effect on the free energy of binding, suggesting the absence of critical contacts by repressor at these sites (Goeddel et al., 1978a). Based on the observed decrease in complex stability in response to BrdU substitution at base pairs 1, 2, 8, 20 or 21, this group postulated that direct interaction between the repressor and the 5-methyl of thymidine occurs at each of these positions. These results, based on thermodynamic calculations from equilibrium binding measurements, suggest interpretational possibilities without providing conclusive evidence for direct repressor-operator contact sites. Ogata and Gilbert (1977) took advantage of the increased photosensitivity of bromodeoxyuridine-containing DNA and identified specific contacts on the DNA by assessing the ability of bound repressor to protect the operator against UV-induced strand scission at sites of BrdU incorporation. UV irradiation of BrdU-substituted DNA results in photodissociation of bromine with the generation of a reactive free radical at the C5 of uracil (Figure 7). In B-form DNA, the substituent at this position is in close proximity to the C2' of the deoxyribose moiety of the base immediately 5' and in the same strand. In the absence of any other reactive species, the free radical presumably abstracts a hydrogen atom from the sugar leading to decomposition of that sugar (by an unknown mechanism) and a single strand break in the DNA at the position of BrdU incorporation (Hotz and Reuschl, 1967; Wacker et al., 1964; Kohnlein and Hutchinson, 1969). A properly positioned substituent within a bound protein could donate a hydrogen to the uracil or deoxyribose radical or accept the free radical and directly cross-link to the DNA, thereby avoiding strand scission. Protection of the DNA against breakage at any of the substituted sites has been interpreted as a
Figure 7. Initial Photochemical Consequences of Ultraviolet Irradiation of BrdU. UV exposure results in the photodissociation of bromine along with the generation of a free radical at the C5 position on the uracil ring. This free radical is responsible for DNA chain breakage in the absence of closely associated hydrogen donors (i.e. protein groups, buffer components) through attack on a neighboring deoxyribose.
Figure 7.

5-Bromodeoxyuridine

\[ \text{UV IRRADIATION} \]

Uracilyl radical

\[ + \text{Br}^* \]
close contact between the protein and the DNA at that position (Ogata and Gilbert, 1977). Ogata and Gilbert (1977) generated operator DNA fragments 50% substituted with BrdU by *in vivo* incorporation of the derivatized base into a thymidine auxotroph. UV irradiation of these multiply-substituted operator sequences in the absence and presence of *lac* repressor, suggested interaction by the protein at operator positions +1, 2, 3, 4, 8, 13, 15, 16, 18, 19, 20, 21, and 22 (Figure 5).

Whenever protein and/or DNA are exposed to concentrated ultraviolet radiation, the damaging consequences must be considered. Amino acids, peptide bonds, nucleotides and deoxyribose are all susceptible in varying degrees to direct or indirect UV damage. Very little recent information regarding the photochemical sensitivity of amino acids and the peptide bond is available. It can be assumed, however, that any direct damage would require absorption of the light. One would therefore predict that the aliphatic amino acids would escape damage by radiation greater than 240 nm (commonly used in photobiology), while the aromatic amino acids could prove more susceptible. A study of the photochemistry of phenylalanine, at its absorption maximum of 257.5 nm, identified several breakdown products. Analysis of these compounds showed that cleavage of the side chain bonds had occurred through energy migration from the benzene ring (Smith and Hanawalt, 1969). The potential for energy transfer from aromatic groups to other points within the protein could increase the overall sensitivity of the molecule to UV damage. The peptide bond absorbs maximally in the 180-190 nm range, but some bond breakage is possible with wavelengths as high as 254 nm. Within proteins, cysteine and the disulfide bond are most sensitive to photochemical damage at neutral and alkaline pH. When considering the overall photochemical susceptibility of a peptide or protein, factors such as wavelength, pH, temperature and the presence of aromatic residues (for energy
transfer) all influence the absorption of energy and its outcome (Smith and Hanawalt, 1969).

Within DNA, purines, pyrimidines and deoxyribose are differentially affected by UV exposure. In native DNA, the sugar backbone is insensitive to UV irradiation above 230nm (Smith and Hanawalt, 1969). As discussed previously, however, incorporation of BrdU for thymidine can lead to indirect UV effects, namely destruction of the deoxyribose moiety following free radical attack. Purine bases are relatively insensitive to UV damage and absorption of UV rarely results in any damage to the purine ring. There is some evidence, however, for the transfer of energy from adenine to thymidine in poly d(AT) sequences, indicating that UV absorption by the purines could prove significant (Smith and Hanawalt, 1969). The pyrimidines are considerably more sensitive to UV exposure. Hydration products are fairly common in irradiated single-stranded DNA. A second consequence of UV absorption by the pyrimidines is the formation of cyclobutane-type dimers. These can form between adjacent bases or can span the two strands and form a DNA-DNA cross-link (Smith and Hanawalt, 1969). Such photoproducts occur relatively infrequently unless the DNA is irradiated in the dry, frozen or highly concentrated state. Most often, very high doses of UV radiation are required for damage to occur. Alternatively, DNA-protein cross-links can form upon UV exposure. The photochemical addition of cysteine to poly-dT, poly-dC and to DNA has been demonstrated. This phenomenon is greatly enhanced by the substitution of BrdU for thymidines (for review, see Hutchinson, 1973). In addition to enhanced UV-induced DNA-protein cross-link formation, the DNA exhibits greater susceptibility to single-strand chain breakage due to the production of a free radical upon irradiation (Figure 7). The use of UV irradiation in studies involving protein and/or DNA requires an understanding of the potential photochemical outcome for
proper analysis and interpretation of the results. In many cases, interfering UV
damage can be prevented or controlled by manipulation of the experimental conditions
(i.e. sample concentrations, pH, temperature, wavelength, duration and intensity of
exposure) (Hutchinson, 1973).

We have extended the experiments of Ogata and Gilbert (1977) to probe the
repressor-operator complex for specific contacts using synthetic operator DNA
substituted with BrdU at single and multiple sites within a single strand of the
sequence. For this study, it was assumed that measurable protection of the DNA
against UV damage by repressor constituted a close contact between the protein and
DNA at the sequence position in question. Those sites chosen for substitution with
BrdU are shown in Figure 8. Our goal in performing this phase of the research was to
confirm and/or to re-assess those previously identified contacts under better defined
conditions (i.e., degree and position of BrdU incorporation). Secondly, we were
interested in quantitating the observed protection by repressor and noting any patterns
that might suggest regions or individual sites of critical contact. Each site was further
analyzed for the ability to form covalent cross-links to the repressor protein. From this
set of experiments, we hoped to identify those repressor-operator contacts most
favorable within the limits of our materials and conditions. Cross-link formation at any
site along the operator sequence was taken as direct evidence of a close protein-DNA
contact. Comparison of the results indicates distinction between the mechanisms
involved in DNA protection and cross-link formation. These differences suggest
caution in the interpretation and analysis of protection data alone. With these
precautions in mind, we would suggest that the combined analysis by UV protection
and cross-link formation provides greater confidence in identifying contacts critical to
formation and stability of the protein-DNA interaction.
Figure 8. Sequence of the 40 bp lac Operator Fragment. As in Figure 5, boldface T and B notations refer to the top and bottom strands of the operator, and the sequence is numbered with respect to the transcription start site at +1. Arrowheads indicate those thymidine positions selected for 5-BrdU substitution. The arrow at the +11 position marks the center of sequence symmetry.
Figure 8.

\textbf{\textit{lac OPERATOR SEQUENCE}}

\begin{center}
\begin{tabular}{l}
\textbf{T} 5'\text{-}TGGTGTGGAATTGTGAGCGGATAACAATTTCACACAGG \\
-10 \hspace{1cm} +1 \hspace{1cm} +10 \hspace{1cm} +20 \hspace{1cm} +30 \\
\textbf{B} 3'\text{-}ACAACACACCTTAACACTCGCCTATTGTTAAAGTGTGTCC \\
\end{tabular}
\end{center}
In addition to its usefulness in confirming and localizing specific contacts, the generation of cross-linked protein-DNA complexes also affords a means for better defining the individual protein elements involved in isolated contacts along the operator sequence. Several groups have sought to identify the peptide or amino acid involved in selected contacts from a variety of protein-nucleic acid complexes. Konigsberg and colleagues have investigated a number of different binding complexes, including gene 5 protein (bacteriophage fd)-fd DNA, *E. coli* single stranded binding protein (SSB)-oligo p(dT)₈, and bacteriophage T4 SSB (gp32)-oligo p(dT)₈ (Paradiso et al., 1979; Paradiso and Konigsberg, 1982; Merrill et al., 1984; Shamoo et al., 1988). In each case they were able to generate and isolate photochemically-induced cross-linked complexes. Following proteolysis and HPLC purification of the resulting peptide-DNA species, the peptide, as well as the linked amino acid, were identified by protein sequence analysis. The protocols designed by these groups for detailed analysis of protein-DNA contacts are applicable to many nucleic acid-binding proteins and their targets. In an alternate approach, Valenzuela and Schulman (1986) employed a lysine-reactive cross-linker, attached to *E. coli* tRNAfMet, to identify regions of contact by methionyl-tRNA synthetase. Once coupled, the complexes were digested with trypsin and purified by anion exchange HPLC. The tRNA-linked peptides were subsequently released by reduction of a disulfide bond within the cross-linker. The peptides were separated by reverse-phase HPLC and identified following amino acid analysis. Yet another group induced cross-link formation in nucleohistones via the introduction of hydroxyl radicals (Didzaroglu et al., 1989). The chemistry of the linkage was analyzed by gas chromatography-mass spectrometry and identified as a thymine-tyrosine cross-link. Substitution of 5-bromodeoxyuridine for thymidine within a target DNA sequence generates a substrate for UV-induced protein-DNA cross-link formation. Wolfes et al.
(1986) applied this technology to their analyses of EcoRI- and EcoRV-oligonucleotide binding and restriction. Using BrCN cleavage of the cross-linked species, and SDS-PAGE separation of the digestion products, they were able to identify the region of the protein involved. Working with the lactose repressor protein, Barbier et al. (1984) generated cross-links to BrdU-containing non-operator DNA. Again, analysis of proteolytic digestion products yielded information regarding the attached peptide and the specifically linked amino acid. The above examples of cross-link formation and analysis via direct protein-DNA attachment, linkage through an amino acid-specific agent, or by introduction of a photoreactive substituent into the DNA sequence suggest a variety of approaches applied toward elucidation of the individual components of a specific contact.

We attempted a similar analysis of the lac repressor protein-operator DNA complex. Using our singly BrdU-substituted operator oligonucleotides, we have demonstrated cross-link formation on a preparative scale and have designed a protocol for the preparation and analysis of the resulting complex. The T+3/B BrdU-substituted operator was initially chosen for our studies. Protein sequence analysis of the resulting covalent complex with lac repressor identified His29 as the T+3-linked amino acid. This result is in agreement with an earlier report by Lamerichs et al. (1989) and provides further evidence for the localization of critical DNA-binding contacts within or near the proposed helix-turn-helix region of the lac repressor NH2-terminal domain.
MATERIALS AND METHODS

Isolation of lac repressor protein

The lac repressor protein was purified from *E. coli* strain CSH 46 by the method of Rosenberg *et al.* (1977) as modified by O'Gorman *et al.* (1980). The cells were lysed to release their cytoplasmic contents by suspension in breaking buffer (BB: 0.20 M Tris-HCl, pH 7.6, 0.20 M potassium chloride, 0.01 M magnesium acetate, 0.30 mM dithiothreitol, 5% glucose, 50 µg/ml phenylmethylsulphonylfluoride-PMSF). DNaseI was added in solid form to digest the released DNA. The suspension was centrifuged (Model J-21, Beckman Instruments) at 7500 rpm for 50 minutes to pellet the cell debris. The supernatant was saved while the pellet was resuspended in BB and re-centrifuged as above. The supernatant fluids were combined and stored on ice. Solid ammonium sulfate was added (23.1 g/100 ml), with the pH monitored and kept above 7. After incubation on ice for 1-2 hours, the liquid was centrifuged at 7500 rpm for 40 minutes. The pelleted proteins were resuspended in 0.12 M potassium phosphate buffer, pH 7.6 (0.10 M K₂HPO₄, 0.02 M KH₂PO₄, 0.30 mM DTT, 50 g/l glucose). This solution was dialysed against several changes of 0.12 M potassium phosphate (KP), followed by centrifugation at 9000 rpm for 30 minutes. The supernatant was layered onto a phosphocellulose (Whatman) column equilibrated with 0.12 M KP. Additional column buffer was then gently layered on top of the crude protein solution. The column was washed until all of the flow-through proteins (yellowish color) had eluted. Elution of the repressor and other adherent proteins followed using a linear gradient of 0.12 M KP to 0.30 M KP. Fractions were collected and assayed for inducer-binding activity. Purified repressor protein was divided into aliquots and stored at -20°C.
Inducer binding

The assay for inducer-binding activity provides a means for following a repressor purification procedure and offers a way to monitor purified and stored protein. The required solutions are TMS (0.01 M Tris-HCl, pH 7.4, 0.001 M EDTA, 0.01 M MgCl₂, 0.1 mM DTT, 0.20 M KCl), 70% (w/v) saturated (NH₄)₂SO₄, 10% (w/v) trichloroacetic acid (TCA). The IPTG (isopropyl-D-thiogalactopyranoside)-containing solutions A and B included 10⁻⁶ M ¹⁴C-IPTG in TMS and 10⁻⁶ M ¹⁴C-IPTG with 10⁻³ M IPTG in TMS, respectively.

The set-up for inducer-binding activity measurement involved a series of tubes (in duplicate) where one contained 200 µl of A and the next contained 200 µl of B. An aliquot of a repressor solution (5-10 µg in 200 µl or less) was added to each tube with the final volume brought to 400 µl with TMS. The tubes were incubated at room temperature for 5-10 minutes. Lysozyme (30-40 µl of 150 mg/ml stock) was added as a carrier in precipitation. The proteins were then precipitated by addition of 3 ml of 70% saturated (NH₄)₂SO₄ and incubation on ice for 10 minutes. Centrifugation at 6500 rpm for 10 minutes pelleted the protein. The supernatant was removed by aspiration, and the pellet was resuspended in 0.5 ml of TMS. An addition of 0.5 ml of 10% TCA was then made. After mixing, the samples were centrifuged at 6500 rpm for 5 minutes. Finally, 0.2 ml of the supernatant was counted for 1-2 minutes in 10 ml of toluene-based scintillation cocktail. Repressor concentration was found to be proportional to the difference in cpm of A and B and could be determined from a standard curve.

Oligonucleotide synthesis and purification

Operator DNA sequences were synthesized on a BioSearch 8600 DNA synthesizer utilizing β-cyanoethyl phosphoramidite chemistry. For BrdU-containing sequences, 5'-DMT-2'-deoxy-5-bromouracil-3'-NN-diisopropyl cyanoethyl
phosphoramidite was purchased from ChemGenes Corporation, Needham, MA. Newly synthesized oligonucleotides were collected in 30% ammonium hydroxide and de-blocked at 55°C for 5 hours. The vial was opened, cooled to room temperature and placed in a hood overnight to remove ammonia. Samples were concentrated to near dryness and resuspended to 100 µl with TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). The crude synthetic yield was determined by measuring the optical density at 260 nm (OD\textsubscript{260}) of a 1000-fold dilution in water (1 OD\textsubscript{260}=37 µg/ml). In preparation for gel purification, a sample was diluted to an approximate concentration of 6 mg/ml with a denaturing dye (95% deionized formamide, 10 mM EDTA, 0.1% w/v bromphenol blue and xylene cyanol). The DNA was heat denatured at 95°C for 3-5 minutes and cooled on ice prior to gel loading. The crude 40 b DNA fragments were purified by preparative gel electrophoresis (20% polyacrylamide/7 M urea in Tris-boric acid-EDTA buffer [TBE: 0.1 M Tris-borate, pH 8.3, 2 mM EDTA]). Gels were pre-run for 30-60 minutes at 50 mA. Approximately 20 µl of the DNA/dye solution was loaded per lane (120 µg/lane). Gels were run 6-8 hours at 50 mA (800-1200 V). DNA was visualized by UV shadow (shortwave UV light, Mineralight UVGL Products) on silica gel 60 F254 TLC plates (EM Science), and the full-length band was excised. The DNA was isolated by electroelution from the acrylamide at 100 V for 6 to 12 hours in an ELUTRAP (Schleicher & Schuell) placed within a horizontal gel apparatus with 1.5 l TBE. This procedure involved placement of the acrylamide slice(s) within the elution chamber formed by proper placement of the BT1 and BT2 membranes. A “trap” for the eluted DNA sample is formed between a BT2 and BT1 membrane at the positive end of the ELUTRAP. The BT2 membrane permitted passage of the DNA molecules from the elution chamber to the “trap”, while the BT1 membrane acted as a barrier to all DNA molecules >14 b in size. Smaller components freely passed through. Following
electroelution, the poles were reversed and voltage was applied for 30-45 seconds to release weakly (BT1) membrane-associated molecules. The purified oligonucleotides were then collected in solution form from the “trap” and transferred to a 1.5 ml Eppendorf tube. These solutions were dried, resuspended in 100 μl TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and checked for purity and yield. The purification yield was determined by measuring the OD_{260} of a 100-fold dilution. The purity of each 40 b oligo was evaluated by loading 200 ng of sample on a 20% polyacrylamide/7 M urea/TBE mini gel, run for 45-60 minutes at 100 V and stained with ethidium bromide. All BrdU-containing DNA was protected from unnecessary exposure to light throughout purification and subsequent handling. DNA samples were stored in Eppendorf tubes designed for UV sensitive samples.

5'-End-labeling

Double-stranded and single-stranded 40mers were labeled at the 5'-hydroxyl using γ^{32}P-ATP (>2000 Ci/mmole, ICN) and T4 polynucleotide kinase (Promega) according to the procedure given for blunt-ended fragments by Maniatis et al. (1987). The DNA to be labeled (0.5 μg in 1.0 μl) was combined in a sterile 1.5 ml Eppendorf tube with 3 μl of a dilution buffer (0.20 M Tris-HCl, pH 9.5, 10 mM spermidine, 1 mM EDTA) and sterile water to a final volume of 27 μl. The solution was heated at 70°C for 2-3 minutes to eliminate self-annealing and cooled on ice. A 5 μl aliquot of 10X kinase buffer (0.50 M Tris-HCl, pH 9.5, 0.10 M MgCl₂, 50 mM DTT, 50% v/v glycerol) along with 50 pmoles of the γ^{32}P-ATP was added followed by gentle mixing. Upon addition of 20 units of T4 polynucleotide kinase (Promega, 10 units/μl), the tube was tapped gently to mix and placed in a 37°C water bath for 30 minutes. The reaction was stopped with the addition of 2 μl of 0.5 M EDTA, pH 8.0. The labeled oligonucleotide was purified by extraction with an equal volume of phenol/chloroform (1:1).
aqueous layer was saved and further extracted with an equal volume of chloroform. The resulting aqueous layer was saved, and the DNA was concentrated by precipitation with 0.5 volume of 7.5 M ammonium acetate and 2.5 volumes of 95% ethanol. Precipitation involved a one hour incubation at 25°C followed by centrifugation (14,000 rpm) for 30 minutes at 4°C. After the spin, the supernatant was removed, and the pellet was rinsed once with 100 µl of 70% ethanol and then dried. The labeled DNA sample was resuspended in 40-50 µl of TE prior to strand hybridization. A 0.5 µl aliquot of this solution was loaded with 2 µl of denaturing dye onto a 20% polyacrylamide/7 M urea/TBE mini gel. The gel was run 45-60 minutes at 100 V. The purity of the labeled product was visualized by autoradiography (Kodak X-OMAT film).

**Strand hybridization**

The single-stranded, base-substituted 40mers were mixed in a 1.5 ml Eppendorf tube with equivalent amounts of their non-substituted complements in the presence of 0.1 M NaCl. Samples were heated at 95°C for 10 minutes and then slowly cooled from 80°C to room temperature. Once cooled, hybridization efficiency was assessed by gel electrophoresis. An aliquot was removed and loaded with 2 µl of a dye solution (50% v/v TE, 50% v/v glycerol, 0.1% w/v bromphenol blue, 0.1% w/v xylene cyanol) onto a 10% polyacrylamide/ TBE mini gel. The gel was run for 30 minutes at 100 V. Hybridized samples were visualized by autoradiography (if radioactively labeled) or by ethidium bromide fluorescence (if unlabeled).

**Equilibrium repressor-operator binding assay**

The nitrocellulose filter binding assay (Riggs *et al.*, 1968; Hsieh and Matthews, 1981) was used to assess the affinity of *lac* repressor for the BrdU-substituted operator fragments. The assay buffer included 0.01 M Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO, 50 µg/ml bovine serum albumin. $K_d$ values were
determined by fitting the data to the equilibrium equation $K_d = ([R] - [RO]) ([O] - [RO]) / [RO]$. Repressor, diluted to $1.5 \times 10^{-8} \text{ M}$ with assay buffer, was added in the following volumes: 5, 10, 20, 50, 100, 200, 400, 10, 400 µl, to a set of 9 tubes (in duplicate) containing operator DNA at $7.6 \times 10^{-11} \text{ M}$ in assay buffer. After mixing, IPTG was added to the last two sets of tubes to a final concentration of 1 mM. The final volume in each tube was 0.5 ml. The assay tubes were incubated at 25°C for 15-20 minutes. Prior to filtration, pre-wet nitrocellulose filter discs (0.45 µm) were rinsed with 0.5 ml assay buffer (without added BSA). The solution from each tube was filtered through individual discs under pressure from an attached vacuum pump (Cole-Parmer). Assays were performed under fluorescent lighting. Test tubes were shielded from direct light exposure during incubation only. The discs were then rinsed with 0.25 ml assay buffer, placed on absorbent towels and dried for 10-15 minutes in a 70°C oven. Each disc was counted in 5 ml of a toluene-based scintillation cocktail. The cpm from the IPTG-containing tubes were subtracted from all others and the adjusted cpm were plotted against repressor concentration.

Silanization of glassware and plasticware

To avoid adhesion of protein and/or DNA to glass or plastic during subsequent experiments and to assure that all vessels were protein-free, we neutralized the surface charges on many commonly used containers. Eppendorf tubes and pipette tips were soaked in 1% (CH$_3$)$_2$Cl$_2$Si (v/v) in CHCl$_3$ in a fume hood, rinsed thoroughly with deionized water and baked at 70-100°C. Glass vessels were filled with the silane solution, rinsed and baked before use.

Repressor-operator protection determination

Singly end-labeled 40 bp operator DNA, containing one or more sites of BrdU substitution in either the top or bottom strand, was mixed with repressor protein,
irradiated, denatured and electrophoresed in a manner similar to that reported by Ogata and Gilbert (1977). For our experiments, DNA at 57 nM and repressor at 200 nM, in a total volume of 20 µl, were incubated in 0.01 M Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO in the absence and presence of 1 mM isopropyl-β-D-thiogalactoside (IPTG). Samples were spotted onto parafilm and irradiated (254 nm, 12,800 µWatts/cm²) at room temperature for 90 seconds in a Rayonet Photochemical Reactor (Southern New England Ultraviolet Company, Hamden, CT). A second UV source was also available in our laboratory. For preliminary experiments we irradiated the protein-DNA solution with short wave (~254 nm) lamp of a hand-held source (Mineralight UVGL Products) at a vertical distance of 1.5 cm for 5 minutes. The remainder of the sample work-up was identical to that detailed below. Protection results with the two sources were comparable. DNA precipitation by ethanol followed. The dried pellet was resuspended in 20-40 µl of 0.05 M NaOH, 0.5 mM EDTA, 5 M urea, 0.025% (w/v) bromophenol blue and xylene cyanol. A 1 µl aliquot of each sample was removed and counted in 6 ml of scintillation fluid for standardization of cpm/sample. Samples were then loaded onto a pre-run preparative 20% polyacrylamide, 7 M urea, TBE gel. Equal amounts of DNA as measured by cpm were applied to the gel for parallel samples. Gels were run at 50 mA for 3-5 hours. Bands were visualized by autoradiography following 3-5 hour exposure of Kodak X-OMAT film. The degree of protection afforded by repressor at individual operator sites was quantitated by densitometry (Kratos-Schoeffel Instruments).

Repressor-operator cross-linking efficiency measurements

5'-End-labeled 40 bp operator DNA, substituted with BrdU at a single site in either the top or bottom strand, was incubated with repressor protein under the conditions described above for protection determination (repressor at 200 nM and
DNA at 57 nM in 0.01 M Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO). Irradiation of the samples followed the above specifications for the Rayonet Photochemical Reactor with the exception that T+3/B operator DNA received a shorter UV exposure (30 seconds). Only those results (protection and cross-linking) obtained by irradiation within the Photochemical Reactor were used for direct comparisons. To isolate the cross-linked complexes, samples were denatured with 1/10 20% SDS v/v, 1/10 β-mercaptoethanol (βME) v/v, and non-denaturing dye (50% v/v TE, 50% v/v glycerol, 0.02% w/v bromphenol blue, 0.02% w/v xylene cyanol) and separated on Tris-glycine-SDS 10% PAGE (per 50 ml gel, 17 ml 30% acrylamide, 6.6 ml 1% bis-acrylamide, 19 ml 1 M Tris-HCl, pH 8.7, 0.25 ml 20% SDS, 7.15 ml H2O; per 2 l running buffer, 6 g Tris-base, 28.8 g glycine, 2 g SDS; gels were cast and run on a Hoefer Model SE600, Hoefer Scientific Instruments), run for 3 hours at 150-200 V. Densitometry of autoradiograms provided quantitative analysis of the cross-linking efficiency.

Large-scale repressor-T+3/B operator cross-link formation

For the initial large-scale protein-DNA cross-link experiments, operator DNA, singly-substituted with BrdU at the +3 position in the top strand, designated T+3/B, was chosen (Figure 8). Repressor and T+3/B operator were combined in 600 ml total volume under the conditions described above for cross-linking (repressor at 200 nM and DNA at 57 nM in 0.01 M Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO). The solution was irradiated within the Rayonet reactor (90 seconds, 254 nm) in 25 ml lots within a silanized Pyrex petri dish (150x15 mm, VWR). Once irradiated, the solutions were pooled, shell frozen within a silanized lyophilization flask and concentrated by lyophilization. The nearly dried material was brought to 30-50 ml with water and dialyzed (SpectraPor 12,000 MW cut-off, American
Scientific Products) against the same solution to lower the KCl concentration. Following dialysis, the sample was further concentrated to a final volume of 1.5-2.0 ml.

Repressor-operator cross-link isolation

To isolate the cross-linked material from the free protein and DNA, the samples were denatured with 1/10 20% SDS (v/v), 1/10 βME (v/v) and loaded with 1/10 (v/v) marker dye solution [0.02% bromphenol blue (w/v), 0.02% xylene cyanol (w/v), 50% glycerol in TE] onto a pair of Tris-glycine-SDS 10% polyacrylamide gels (per 50 ml gel, 17 ml 30% acrylamide, 6.6 ml 1% bis-acrylamide, 19 ml 1 M Tris-HCl, pH 8.7, 0.25 ml 20% SDS, 7.15 ml H2O; per 2 l running buffer, 6 g Tris-base, 28.8 g glycine, 2 g SDS; gels were cast and run on a Hoefer Model SE600, Hoefer Scientific Instruments). Electrophoresis at 150-200V for 4-6 hours resulted in separation of the cross-linked protein-DNA complex from each of the free species. Autoradiography allowed for visualization of the cross-linked and free DNA. Each band was excised from the gel. The acrylamide slices were placed in the elution chambers of individual ELUTRAP devices. Once placed in a horizontal gel apparatus with 1.5 l TBE, a voltage of 100 V was applied for 6 hours. Longer periods of electroelution proved unnecessary and often resulted in sample loss due to irreversible adhesion to the outer (BT1) membrane. The resulting cross-linked and free DNA samples available in solution form were dialyzed (SpectraPor 6, 3500 MWCO, American Scientific Products) against 0.1 M NH4HCO3, pH 8.0. Following dialysis, the samples were concentrated in Eppendorf tubes to 200 μl in a SpeedVac concentrator.

Proteolytic digestion of cross-linked complexes

The cross-linked protein-DNA complex was then treated in the same vessel with a 1/25 (w/w) ratio of trypsin and chymotrypsin to total repressor. A concentrated solution (20 mg/ml) of the enzymes was made in 10^{-3} M HCl and an appropriate aliquot
was added to the cross-linked sample to begin digestion. After 4 hours at room temperature, a second addition of trypsin and chymotrypsin was made from a fresh enzyme solution, and the digestion was allowed to continue for 4 hours. Proteolysis was stopped after a total of 8 hours by precipitation of the DNA-linked species in ethanol. The isolated free DNA sample was also precipitated in ethanol at this time. Following centrifugation, the supernatants were discarded, and the pellets were dried. Each pellet was resuspended in 20 μl of 20 mM potassium phosphate, pH 7.4, 4.8% ethanol, 0.04 M KCl. The extent of proteolysis was followed by loading aliquots of the cross-linked and free DNA species onto a Tris-glycine-SDS 10% polyacrylamide gel (per 50 ml gel, 17 ml 30% acrylamide, 6.6 ml 1% bis-acrylamide, 19 ml 1 M Tris-HCl, pH 8.7, 0.25 ml 20% SDS, 7.15 ml H₂O; per 21 running buffer, 6 g Tris-base, 28.8 g glycine, 2 g SDS; gels were cast and run on a Hoefer Model SE600, Hoefer Scientific Instruments), run for 3 hours at 150-200V and visualizing the corresponding bands by autoradiography.

**HPLC purification of the peptide-operator complex**

The peptide-DNA complex was further purified from any noncross-linked peptides by anion-exchange HPLC. The column (NucleoPac PA100, 12.0 μ core resin, divinyl benzene latexed with methyacrylate, 4x250 mm, Dionex Corporation, Sunnyvale, CA) was attached to a Gilson HPLC system on line with the Gilson gradient manager software package. Sample detection involved an LKB detection system (lamp, filter, box) and chromatograms were obtained with an LKB strip chart recorder. Cross-linked or free DNA samples were injected onto the column equilibrated with 20 mM potassium phosphate, pH 7.4, 4.8% ethanol, 0.04 M KCl (buffer A). The column was eluted at 1 ml/min with a linear gradient of 1 M KCl in the 20 mM potassium phosphate, pH 7.4, 4.8% ethanol (buffer B) into buffer A, 0-90 min
(0-100% B), 90-95 min (100-0% B). This scheme is based on a similar purification protocol first reported by Merrill et al. (1984).

Protein Sequencing

Following collection of the HPLC peaks, the cross-linked peptide-DNA sample was dialysed (SpectraPor 6, 3500 MW cut-off, American Scientific Products) against water to remove the salts in preparation for protein sequence analysis. Dialysed samples were concentrated (Speed Vac) to <50 µl and applied to a Biobrene-treated filter. The filter was positioned within an ABI Protein Sequencer (Model 473) and subjected to 10 cycles of Edman degradation. Chromatograms from the individual cycles were compared to a standard containing each of the derivatized amino acids to determine the peptide sequence. Dr. Richard Cook at Baylor College of Medicine, Houston, Texas performed the sequence analysis.
RESULTS

Oligonucleotide synthesis, purification and preparation for use

The thymidine residues within the -6 to +25 region of the lac operator were selected for BrdU substitution (Figure 8). Substituted oligonucleotides are referred to by the strand (top, T, or bottom, B) and the position of BrdU incorporation within the operator sequence. 'Per'-substitution refers to the incorporation of BrdU at each of the nine central thymidine positions in either the top (TPER) or bottom (BPER) strand. These BrdU-containing DNA strands were qualitatively monitored for their synthetic yield and ability to be purified relative to non-substituted oligonucleotides. Though we did note some variability in the coupling efficiency at the site of BrdU incorporation with different lots of the derivatized phosphoramidite, the overall synthetic yield of these fragments was repeatedly close to the theoretical yield and always satisfactory for our studies. Purification of the substituted oligonucleotides proved to be straightforward requiring no adjustments to the protocol used for the non-substituted DNA strands. Figure 9 compares a purified 40 b top strand (T) with that of a T+3 fragment before and after purification. These results are representative of all twenty of the BrdU-substituted oligonucleotides prepared for our studies.

Labeling of the single-stranded DNA fragments at their 5'-hydroxyls involved \( \gamma^{32}P\)-ATP, T4 polynucleotide kinase and the method for blunt-ended fragments described by Maniatis et al. (1987). This protocol offered the greatest consistency and specificity of labeling and resulted in a clean product (Figure 10). Incorporated BrdU had no effect on the labeling results.

The formation of double-stranded oligonucleotides from complementary single strands by the method described in Materials and Methods was not complicated by the presence of BrdU. The hybridization of T+3 and B is shown as an example (Figure 11).
Figure 9. Purification of BrdU-Containing Oligonucleotides. Oligonucleotides were purified as described in Materials and Methods and stored in TE (pH 8.0) at -20°C. The purity of each was checked by electrophoresis in a 20% polyacrylamide, 7 M urea, TBE mini gel and visualized by ethidium bromide fluorescence. For comparison, purified, non-substituted lac operator top strand (T, 500 ng) is shown in lane 1. An aliquot of crude T+3 operator DNA appears in lane 2. Overloading of the lane, along with insufficient denaturation of the sample, account for the appearance of material > 40 b in length. Lanes 3 and 4 illustrate the purity of a T+3 fragment where 250 and 500 ng were loaded, respectively.
Figure 9.
Figure 10. Autoradiogram Illustrating the Efficiency of Labeling Single-Stranded BrdU-Containing Oligonucleotides at the 5'-OH. The labeling protocol was derived from that published by Maniatis et al. (1987) and is detailed in Materials and Methods. Labeled non-substituted lac 40 b operator top strand (T) is shown in lane 1. Lanes 3, 5 and 6 illustrate the labeling of BrdU-substituted operator fragments TPER, T-5 and T-3, respectively. The counts loaded in each lane were equal. Note that the degree of BrdU incorporation had no effect on the efficiency of labeling. Also, purification of the labeled fragments from non-incorporated γ^{32}P-ATP was complete as evidenced by the absence of low molecular weight radioactive intensity.
Figure 11. Autoradiogram Illustrating Strand Hybridization Between the Labeled, BrdU-Substituted T+3 Strand and the Complementary Bottom Strand (B). The single-stranded T and T+3 fragments are shown in lanes 1 and 2, respectively. Lane 3 illustrates the result of heat denaturation (95°C, 3 minutes) of a double-stranded T+3/B operator. The hybridized T+3/B operator is shown in lane 4. In some cases, a small fraction of the hybridized DNA sample remained in single-stranded form (lane 4). This result was independent of BrdU incorporation and most often represented <5% of the operator DNA. When significant levels of single strands were evident, the hybridization was repeated; otherwise the operator DNA was used directly.
Equilibrium binding studies

The effect of BrdU incorporation on repressor equilibrium binding affinity for these double-stranded operators, either singly or per-substituted with BrdU, was assessed by nitrocellulose filter binding. Our assays and subsequent studies were performed in relatively high salt (0.25 M KCl) binding buffer. This condition was set for the purpose of maximizing our working repressor and operator concentrations within equilibrium boundary conditions while maintaining operator-specific (inducer-sensitive) binding. These results are presented in Table 1. Both plasmid-derived and synthetic 40 bp operator sequences were used as controls. When compared to these DNAs, the sequences per-substituted with BrdU in the top or bottom strand bound repressor with 3 to 5-fold higher affinities. These data are in agreement with earlier work reported by Lin and Riggs (1972) who observed that DNA with BrdU incorporated bound lac repressor more tightly. The presence of a single BrdU residue within the operator sequence at any of the selected positions had minimal effect on the equilibrium repressor binding affinity. When compared to the $K_d$ values for T/B and 40mer, the data in Table 1 show a maximum deviation within a factor of two, indicating that this low level of incorporation did not appreciably perturb the binding interactions. Goeddel et al. (1978a,b) noted some variation in the dissociation rates exhibited by singly BrdU-substituted operator sequences. In most cases, the $t_{1/2}$ deviated less than 2 fold from the controls, consistent with the minimal influence of a single site of BrdU incorporation on repressor binding observed in our studies of equilibrium constants. These results allowed us to perform subsequent studies with our BrdU-substituted operator fragments with the confidence that complex formation with repressor was not significantly compromised.
Table 1. Equilibrium Dissociation Constants For BrdU-Containing Operator DNA. The assay buffer contained 10 mM Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO. T/B refers to the non-substituted synthetic 40 bp operator sequence, while 40mer indicates a plasmid-derived operator control. TPER/B stands for the 40 bp operator comprised of a top strand substituted with BrdU at each of the nine central thymidine positions and hybridized to a non-substituted bottom strand. T/BPER is the fully substituted bottom strand annealed to its non-substituted complement. Each individually substituted operator sequence is referred to by the strand (T and B) and the position of BrdU incorporation and is hybridized to its complement.
<table>
<thead>
<tr>
<th>DNA</th>
<th>$K_d$ (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{5/B}$</td>
<td>$4.1 \pm 2.4 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{3/B}$</td>
<td>$6.2 \pm 2.1 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+3/B}$</td>
<td>$6.5 \pm 3.2 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+4/B}$</td>
<td>$5.0 \pm 4.3 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+6/B}$</td>
<td>$8.0 \pm 1.2 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+14/B}$</td>
<td>$7.5 \pm 1.7 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+20/B}$</td>
<td>$7.7 \pm 3.8 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+21/B}$</td>
<td>$7.6 \pm 1.0 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{+22/B}$</td>
<td>$5.6 \pm 0.0 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T_{\text{PER}/B}$</td>
<td>$1.2 \pm 0.4 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+1}$</td>
<td>$8.3 \pm 1.7 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+2}$</td>
<td>$7.0 \pm 1.7 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+8}$</td>
<td>$6.1 \pm 0.8 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+13}$</td>
<td>$7.1 \pm 3.3 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+15}$</td>
<td>$7.4 \pm 0.4 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+16}$</td>
<td>$5.6 \pm 1.4 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+18}$</td>
<td>$4.0 \pm 0.6 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+19}$</td>
<td>$5.8 \pm 1.1 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{+24}$</td>
<td>$6.1 \pm 1.0 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B_{\text{PER}}$</td>
<td>$2.0 \pm 0.9 \times 10^{-10}$</td>
</tr>
<tr>
<td>$T/B$</td>
<td>$6.2 \pm 1.6 \times 10^{-10}$</td>
</tr>
<tr>
<td>40mer</td>
<td>$4.7 \pm 0.6 \times 10^{-10}$</td>
</tr>
</tbody>
</table>
Repressor-operator protection determination.

The ability of lac repressor to protect against UV-induced strand scission for various BrdU-substituted operators was monitored via the method of Ogata and Gilbert (1977). Results from operator sequences per-substituted at the central thymidine positions in either the top or bottom strand with BrdU are shown in Figure 12. The non-uniformity of the DNA breakage at positions of BrdU incorporation in these per-substituted DNAs (T_PER/B and T/B_PER) can be seen in the absence of protein. Ogata and Gilbert (1977) suggested that strand scission was perhaps sequence dependent or influenced by the local DNA structure. Small changes in the local structure could affect the distance between the 5-position of the pyrimidine ring and the sugar-phosphate backbone. Bound repressor decreases the band intensity at positions +3, 4, and within the 20-22 cluster for T_PER/B (Figure 12A). This protection against UV-induced strand scission offered by bound repressor at specific sites throughout the operator sequence suggests either that close contacts with protein are made at these specific positions or binding of the protein alters the DNA structure to preclude the scission reaction. When repressor and operator were incubated in the presence of saturating levels of inducer, the protection afforded by repressor was reversed. Ogata and Gilbert (1977) likewise identified repressor strand scission protection at positions 3, 4, 20, 21 and 22 (Figure 5). A similar analysis was performed for the T/B_PER operator DNA, and the results are shown in Figure 12B. Under specific binding conditions, the repressor protects the operator DNA at positions +1, 2, 8, 13, 15, 16, 18, and 19 within the bottom strand. When repressor is bound to inducer, cleavage at the above sites is no longer suppressed, indicating that the protection thus identified is inducer-sensitive and specific for the operator sequence when bound to unliganded repressor. These data are in agreement with those reported for the bottom strand by
Figure 12. Repressor-Operator Protection Determination for TPER/B, T/BPER.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of TPER/B (A) and T/BPER (B). DNA was irradiated at 254 nm for 90 seconds in the absence of repressor (lane 1), with repressor (lane 2), or in the presence of repressor saturated with IPTG (lane 3). The numbers beside the UV-generated bands indicate the positions of BrdU incorporation which are identified by the length of the resulting fragment. The heavy band at the top of each lane represents uncleaved strands.
Figure 12.

A

\[ T_{\text{PER}/B} \]

1 2 3

B

\[ T/ B_{\text{PER}} \]

1 2 3
Ogata and Gilbert (1977). All of our UV protection studies involved the use of a relatively high salt binding buffer (0.01 M Tris-HCl, pH 7.4, 0.25 M KCl, 0.1 mM DTT, 0.1 mM EDTA, 5% DMSO). This buffer system was utilized to allow for sequence-specific binding of the repressor under equilibrium conditions with elevated concentrations of both protein and DNA. The increased salt concentration effectively increased the $K_d$ for specific and nonspecific binding. Nonspecific binding, however, is more dramatically influenced by ionic strength and is selected against. This choice of buffer allowed us to increase our levels of protein and DNA for better detection without compromising the specificity of binding or violating the equilibrium condition.

When utilizing highly BrdU-substituted DNA, however, the possibility of results that can be ascribed to DNA conformational anomalies or alterations in charge distribution along the binding surface must be considered. The increased affinity of these DNAs for repressor is one indication of altered binding properties. Goeddel et al. (1978a) also reported that the calculated changes in binding free energies were not additive when single and multiple BrdU-containing operators were assayed. In fact, the increased affinity observed for multiple substitutions was much greater than expected from the free energy changes at single sites. We therefore repeated the UV protection analysis with operator DNA substituted with BrdU at single sites within the sequence. All eighteen of the central thymidine positions were individually substituted and assayed for inducer-sensitive protection by bound repressor (Figures 13-25). These results illustrate the increased clarity obtained with singly-substituted operators. It is important to note, however, that with a few of the substituted operators, secondary bands appear (e.g., Figures 15, 16, 17 and 23). While a definitive explanation of their origin would require further analysis, these “extra” bands could indicate the influence of local
Figure 13. Repressor-Operator Protection Determination for T-5/B and T-3/B.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T-5/B and T-3/B. Lane 4 shows the UV cleavage pattern of TPER. The single band produced upon UV irradiation of these singly-substituted operators is shown in lanes 1 and 5. The effect of bound repressor on DNA breakage at each site can be seen in lanes 2 and 6. Lanes 3 and 7 illustrate the influence of IPTG on the protection afforded by repressor.
Figure 14. Repressor-Operator Protection Determination for T+3/B. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T+3/B. Lane 1 illustrates the UV cleavage pattern for Tper for reference. The result of UV damage at the T+3 site is shown in lane 2. The protection afforded by bound repressor is observed in lane 3. Lane 4 demonstrates the effect of IPTG on the protein’s ability to protect the DNA at this site.
Figure 14.
Figure 15. Repressor-Operator Protection Determination for T_{+4}/B. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T_{+4}/B. The UV breakage pattern of TPER is presented in lane 4. UV-induced damage at the +4 position is shown in lane 1, while protection by bound repressor can be seen in lane 2. When IPTG is bound to the repressor, the DNA cleavage pattern is as shown in lane 3.
Figure 16. Repressor-Operator Protection Determination for $T_{+6}/B$. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of $T_{+6}/B$. The major band resulting from UV damage at the +6 site is observed in lane 1. A degree of protection is afforded by bound repressor (lane 2). Fully induced repressor (lane 3) is less able to protect the DNA at this site.
Figure 17. Repressor-Operator Protection Determination for T+14/B and T+20/B.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T+14/B and T+20/B. The UV cleavage pattern for TPER is presented in lane 7 for reference. DNA breakage at the +14 and +20 sites due to UV exposure is illustrated in lanes 1 and 4, respectively. In neighboring lanes 2 and 5, the effect of bound repressor at these positions can be noted. The influence of IPTG on the protein's ability to protect the DNA at +14 and +20 is shown in lanes 3 and 6, respectively.
Figure 17.
Figure 18. Repressor-Operator Protection Determination for T+21/B and T+22/B.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T+21/B and T+22/B. The DNA cleavage pattern of TPER is presented in lane 1. Lanes 2 and 5 illustrate the result of UV damage to the DNA at the +21 and +22 positions, respectively. The effect of bound repressor on the degree of damage at the two sites can be seen in lanes 3 (+21) and 6 (+22). In the presence of IPTG, the repressor has little or no effect on UV damage at +21 or +22 as shown in lanes 4 and 7, respectively.
Figure 18.
Figure 19. Repressor-Operator Protection Determination for T/B+1. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T/B+1. For reference, the DNA cleavage pattern for B+1,2,8,13,15 appears in lane 4. The band produced upon UV irradiation is shown in lane 1. The small degree of protection afforded at the +1 positions by bound repressor can be seen in lane 2. Lane 3 illustrates the influence of IPTG on the repressor's protection of the DNA.
Figure 19.
Figure 20. Repressor-Operator Protection Determination for T/B$_{+2}$. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of T/B$_{+2}$. The cleavage pattern of B+1,2,8,13,15 appears in lane 4. Lane 1 illustrates the band generated upon UV irradiation. Lanes 2 and 3 show the effect of repressor and repressor-IPTG, respectively, on the UV-induced DNA damage.
Figure 20.
Figure 21. Repressor-Operator Protection Determination for T/B+8. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of T/B+8. The cleavage pattern for BPER is presented in lane 1. UV damage to the singly-substituted DNA is shown in lane 2. Protection afforded by bound repressor is illustrated in lane 3. The influence of IPTG on the protection by repressor can be seen in lane 4.
Figure 21.
Figure 22. Repressor-Operator Protection Determination for T/B+13. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of T/B+13. The UV cleavage pattern for B+1,2,8,13,15 appears in lane 4 for reference. UV breakage at the +13 position is illustrated in lane 1. The protection provided by repressor can be seen in lane 2. The effect of repressor in the presence of IPTG is shown in lane 3.
Figure 22.
Figure 23. Repressor-Operator Protection Determination for T/B+15 and T/B+16.

Autoradiogram of a denaturing polyacrylamide gel showing UV-induced cleavage of T/B+15 and T/B+16. Irradiated BPER is shown in lane 4 for reference. Lanes 1 and 5 illustrate UV generated cleavage at +15 and +16, respectively. Protection by bound repressor is observed with both operators (lanes 2 and 6). A reversal of protection by repressor occurs in the presence of IPTG (lanes 3 and 7).
Figure 23.
Figure 24. Repressor-Operator Protection Determination for T/B+18 and T/B+19.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of T/B+18 and T/B+19. Lane 4 presents the cleavage pattern of BPER. DNA breakage at +18 and +19 due to UV irradiation is shown in lanes 1 and 5, respectively. Protection against UV damage by bound repressor is observed at both sites as shown in lanes 2 and 6. The influence of IPTG on the repressor’s ability to protect the DNA at these two sites can be seen in lanes 3 and 7.
Figure 24.
Figure 25. Repressor-Operator Protection Determination for T/B+24. Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced cleavage of T/B+24. Lane 1 shows the cleavage pattern of BPER and is provided as reference. The single band generated by UV irradiation of the naked DNA appears in lane 2. Bound repressor has little or no affect on the degree of DNA damage (lane 3). The effect of repressor in the presence of IPTG is shown in lane 4.
Figure 25.
secondary structures within the DNA molecule. Repressor-induced conformational changes in the DNA could result in secondary strand scission. For the purpose of illustrating and explaining our cumulative UV protection results, T+3/B and T/B+8 are presented in Figure 26, while quantitative protection data for all of the DNAs studied are compiled in Table 2 and Figure 27. With T+3/B operator DNA (Figure 26A), for example, decreased band intensity relative to free DNA is noted in the presence of bound repressor, while in the presence of inducer, protection at that site is diminished. T/B+8 DNA was analyzed in the same manner (Figure 26B), and bound repressor partially protects the DNA from UV damage under specific, but not non-specific, binding conditions. Densitometric scans of autoradiograms from each of the substituted operator sequences provided a quantitative assessment of the protection studies. The degree of protection at a selected site was measured as the fractional band intensity in the presence of repressor or repressor-inducer vs the free DNA. The results are presented in Table 2 and Figure 27. The protection patterns generated with singly-substituted operator fragments allowed us to refine our earlier contact assignments. Bound repressor affords significant protection against UV-generated cleavage of operator DNA at all of the centrally substituted sites with the exception of T+14. At the outer extremes of the sequence, substituted positions T-5, T-3, T+22, and B+24 were largely unprotected by bound repressor. These results from singly-substituted DNAs are in good agreement with the operator limits and contact sites determined by exonuclease III digestion and DNase I protection (Shalloway et al., 1980; Manly et al., 1984; Schmitz and Galas, 1979), methylation protection (Gilbert et al., 1976), and the original UV protection experiments of Ogata and Gilbert (1977) (Figure 5).
Figure 26. Repressor-Operator Protection Determination for T+3/B, T/B+8.

Autoradiogram of a denaturing polyacrylamide gel showing the UV-induced breakage of T+3/B (A) and T/B+8 (B). Lanes 1 and 8 show the UV cleavage pattern of TPER and BPER, respectively. Lanes 2 and 5 illustrate the single band produced upon irradiation of these singly-substituted operators. The effect of bound repressor on DNA breakage at each site can be seen in lanes 3 and 6. The susceptibility of the DNAs to UV damage in the presence of fully induced repressor is shown in lanes 4 and 7.
Figure 26.

A

\[ T_{+} \; 3/ \; B \]

1 2 3 4

B

\[ T/ \; B_{+} \; 8 \]

5 6 7 8

-8
Table 2. UV Protection and Cross-link Efficiency Measurements. Scanning densitometry of the respective autoradiograms allowed for quantitative analysis of the results. Protection was calculated as the fractional band intensity in the repressor (-IPTG) and repressor-inducer (+IPTG) containing lanes relative to the lane of free DNA. Measurements were made from 3-7 autoradiograms from each operator sequence with the exception of T+14/B, T+20/B, T+21/B, T+22/B, T/B+15 and T/B+16, where only two data points were available. Cross-linking measurements represent the fraction of DNA present in the cross-linked complex relative to the total DNA (bound + free).
### Table 2
UV Protection and Cross-link Efficiency Measurements

<table>
<thead>
<tr>
<th>DNA</th>
<th>% PROTECTION</th>
<th>-IPTG</th>
<th>+IPTG</th>
<th>% CROSS-LINKING</th>
<th>-IPTG</th>
<th>+IPTG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T.5/B</td>
<td>20 (±6.4)</td>
<td>0.0</td>
<td></td>
<td>2.0 (±1.5)</td>
<td>0.7 (±0.4)</td>
<td></td>
</tr>
<tr>
<td>T.3/B</td>
<td>10 (±11)</td>
<td>0.0</td>
<td></td>
<td>4.0 (±0.0)</td>
<td>0.8 (±0.2)</td>
<td></td>
</tr>
<tr>
<td>T.3+3/B</td>
<td>45 (±12)</td>
<td>0.0</td>
<td></td>
<td>37 (±4.1)</td>
<td>16 (±7.4)</td>
<td></td>
</tr>
<tr>
<td>T.4+4/B</td>
<td>70 (±12)</td>
<td>28 (±5.7)</td>
<td></td>
<td>11 (±0.8)</td>
<td>4.4 (±0.5)</td>
<td></td>
</tr>
<tr>
<td>T.6/B</td>
<td>72 (±7.5)</td>
<td>0.0</td>
<td></td>
<td>3.3 (±1.5)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T.4+14/B</td>
<td>0.0 (±0.0)</td>
<td>0.0</td>
<td></td>
<td>12 (±0.0)</td>
<td>5.0 (±1.4)</td>
<td></td>
</tr>
<tr>
<td>T.4+20/B</td>
<td>27 (±21)</td>
<td>0.0</td>
<td></td>
<td>2.7 (±1.2)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T.4+21/B</td>
<td>73 (±38)</td>
<td>0.0</td>
<td></td>
<td>3.3 (±1.5)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T.4+22/B</td>
<td>0.0 (±0.0)</td>
<td>0.0</td>
<td></td>
<td>4.0 (±1.7)</td>
<td>1.6 (±0.5)</td>
<td></td>
</tr>
<tr>
<td>T.PER/B</td>
<td>--</td>
<td>--</td>
<td></td>
<td>26 (±5.4)</td>
<td>11 (±5.4)</td>
<td></td>
</tr>
<tr>
<td>T/B+1</td>
<td>36 (±0.7)</td>
<td>16 (±0.7)</td>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T/B+2</td>
<td>71 (±10)</td>
<td>0.0</td>
<td></td>
<td>2.8 (±1.1)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T/B+8</td>
<td>52 (±11)</td>
<td>0.0</td>
<td></td>
<td>2.1 (±1.7)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T/B+13</td>
<td>76 (±6.4)</td>
<td>30 (±13)</td>
<td></td>
<td>2.2 (±0.2)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T/B+15</td>
<td>64 (±19)</td>
<td>0.0</td>
<td></td>
<td>3.2 (±1.2)</td>
<td>1.3 (±0.4)</td>
<td></td>
</tr>
<tr>
<td>T/B+16</td>
<td>81 (±0.0)</td>
<td>0.0</td>
<td></td>
<td>2.5 (±3.5)</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>T/B+18</td>
<td>84 (±0.6)</td>
<td>26 (±16)</td>
<td></td>
<td>11 (±2.3)</td>
<td>5.0 (±0.1)</td>
<td></td>
</tr>
<tr>
<td>T/B+19</td>
<td>73 (±4.7)</td>
<td>21 (±18)</td>
<td></td>
<td>28 (±5.0)</td>
<td>14 (±2.1)</td>
<td></td>
</tr>
<tr>
<td>T/B+24</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td>2.0 (±2.8)</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>T/B.PER</td>
<td>--</td>
<td>--</td>
<td></td>
<td>20 (±5.4)</td>
<td>3.3 (±2.1)</td>
<td></td>
</tr>
<tr>
<td>T/B</td>
<td>--</td>
<td>--</td>
<td></td>
<td>2.4 (±1.1)</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>
Figure 27. UV Protection Measurements Along the Operator Sequence. Graphic presentation of UV protection results. The operator sequence is shown for reference. Solid bars represent protection afforded by bound repressor alone, while shaded bars indicate the level of protection offered by repressor-inducer. The values are the average of multiple experiments for each DNA, and standard deviations are shown in Table 2.
Figure 27.

% PROTECTION

OPERATOR SEQUENCE

+1

+10

+20

GT

CT

AC

AT

TT

AA

GG

CC

TG

TA

AC

AT

GT

AT

TG

AA

GT

AC
Repessor-operator cross-link formation.

Using per- and singly-substituted operator fragments, the ability of these DNAs to covalently cross-link with the lac repressor protein was assessed under specific and nonspecific conditions. An initial control experiment was included to assess the sequence-specificity and inducer-sensitivity of cross-link formation under the conditions chosen for our study. We compared the efficiency of repressor cross-linking to the TPER/B lac operator fragment in the absence and presence of IPTG at two KCl concentrations with the degree of cross-linking to a per-substituted trp operator sequence (40 bp) under like conditions (Figure 28). In all situations, cross-linking to the lac operator sequence proved more efficient than to the non-specific DNA sequence of the trp operator. Cross-linking was greater for both DNA fragments at lower salt concentrations, although inducer-sensitivity was somewhat compromised. With the potential for UV damage to the protein, the DNA, or both under the conditions used for the protection and cross-linking studies, we attempted to assess the effect of a 90 second irradiation pulse (254 nm, 12,500 μWatts/cm²) on each of the binding species. We used the nitrocellulose filter binding assay and the conditions outlined in Materials and Methods to measure the equilibrium binding affinities of individually irradiated repressor protein and non-substituted operator DNA. The resulting equilibrium dissociation constants were $7.3(\pm0.7) \times 10^{-10} \text{M}$ for the pre-irradiated T/B operator DNA and $8.3(\pm1.4) \times 10^{-10} \text{M}$ for repressor irradiated prior to binding (compare with T/B, Table 1). In each case, the binding was sensitive to inducer. We also generated UV dose-response curves in the absence and presence of inducer for T+3/B and T/B+19 operator DNA to evaluate the rate of covalent complex formation. The results, presented in Figure 29, indicate that the cross-linking event(s) is complete within 25 seconds of UV exposure.
Figure 28. Sequence and Salt Dependence of Repressor-DNA Cross-link Formation.
Cross-link efficiency experiments were performed for repressor with both the $lac$ and $trp$
40 bp operator sequences (TPE/B) at each of two salt concentrations. Cross-linking to
the $lac$ operator is shown in panel A, while panel B illustrates cross-linking to the $trp$
operator sequence. The two bars on the left in each panel represent the level of cross-
link formation observed at 0.1 M KCl in the absence (solid bar) and presence (hatched
bar) of IPTG. The results obtained at 0.25 M KCl with the two operators are shown in
the right-hand portion of each panel. Inducer (IPTG)-sensitivity under this higher salt
condition is illustrated with cross-link formation in the absence of IPTG shown in solid
and in the presence of IPTG shown by the hatched bar.
Figure 28

A

% CROSS-LINKING (LAC)

30
25
20
15
10
5
0

0.1M
0.25M
KCL

B

% CROSS-LINKING (TRP)

30
25
20
15
10
5
0

0.1M
0.25M
KCL
Figure 29. UV Dose-Response for Repressor-Operator Cross-link Formation.

Repressor, incubated with T+3/B (A) and T/B+19 (B), under conditions previously described for cross-link studies, was irradiated at 254 nm for varying lengths of time from 0-125 seconds in the absence (■) and presence (□) of IPTG. Equal counts from each reaction were loaded onto an SDS polyacrylamide gel. Densitometric scans of the resulting autoradiogram provided measurements of the cross-linking levels at each time point. The cross-link efficiencies were plotted against irradiation time.
Figure 29.
For the repressor-operator cross-linking studies, the efficiency of cross-link formation was measured and reported with the inclusion of the non-substituted operator, T/B, as a measure of background cross-linking. Figures 30-37 illustrate the results obtained with all of the sequences tested. It should be noted that just as the degree of protection varied for each position, the efficiencies of cross-linking were not equal for the different positions of substitution. In addition, protection from strand scission and cross-linking did not correlate. Looking at Figure 38 in more detail, we see that repressor exhibited a low level of cross-linking to T/B operator DNA both in the absence and presence of inducer. In contrast, TPERIB DNA cross-links with relatively high efficiency to repressor. The presence of inducer resulted in a decreased level of cross-link formation. Of particular interest, T+3/B DNA cross-links to the repressor with an efficiency greater than the per-substituted top strand (Table 2, Figure 38). Repressor also formed cross-links with T/BPER DNA (Figure 38). Covalent complex formation through bottom strand contacts involves BrdU at positions other than +8, where no cross-link was detected. In fact, operator DNA substituted at positions 1, 2, 8, 13 and 15 within the bottom strand showed only background levels of cross-linking (Figures 33 and 37), indicating that any bottom strand sites responsible for covalent complex formation exist within the promoter distal half of the sequence. Densitometric scans of the autoradiograms provided quantitative measurement of the data as shown in Table 2 and Figure 39. Cross-linking efficiency was measured as the percent of DNA present in the bound complex relative to total DNA (bound + free). In several cases, the appearance of a doublet in the region corresponding to the cross-linked complex was observed. This result was likely due to minor protein degradation. The intensities of the two bands were summed for the purpose of calculating the total percent of bound DNA. The autoradiograms also demonstrate the double band pattern present in the region of
Figure 30. Repressor-Operator Cross-link Formation with T.5/B and T.3/B.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5, and 7 show the degree of protein cross-linking to operator fragments T/B, TPER/B, T.5/B and T.3/B, respectively. The inducer-sensitivity of cross-linking to each of these operator sequences is indicated by the band intensities in corresponding lanes 2, 4, 6 and 8.
Figure 30.

CROSS-LINKED

FREE DNA
Figure 31. Repressor-Operator Cross-link Formation with T+3/B and T+6/B.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5 and 7 show the degree of protein cross-linking to TPER/B, T/BPER, T+3/B and T+6/B, respectively. The effect of IPTG on the level of cross-linking to each of these sequences is shown in corresponding lanes 2, 4, 6 and 8.
Figure 31.

1 2 3 4 5 6 7 8

CROSS-LINKED—

FREE DNA—
Figure 32. Repressor-Operator Cross-link Formation with $T_{+4}/B$, $T_{+14}/B$, $T_{+20}/B$, $T_{+21}/B$ and $T_{+22}/B$. Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5, 7, 9 and 11 show the degree of protein cross-linking observed to $T/B$, $T_{+4}/B$, $T_{+14}/B$, $T_{+20}/B$, $T_{+21}/B$ and $T_{+22}/B$, respectively. The inducer-sensitivity of cross-linking to each of these operators can be seen in corresponding lanes 2, 4, 6, 8, 10 and 12.
Figure 32.

CROSS-LINKED

FREE DNA
Figure 33. Repressor-Operator Cross-link Formation with T/B+1 and T/B+2.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5 and 7 illustrate the level of protein cross-linking to T/B, T/B+1,2,8,13,15, T/B+1 and T/B+2, respectively. The inducer-sensitivity of cross-linking to these sequences is shown in corresponding lanes 2, 4, 6 and 8.
Figure 33.

CROSS-LINKED —

FREE DNA —
Figure 34. Repressor-Operator Cross-link Formation with T/B+8 and T/B+13.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3 and 5 show the level of cross-linking to T/B, T/B+8 and T/B+13, respectively. The effect of IPTG on cross-linking efficiency to these sequences is shown in corresponding lanes 2, 4 and 6.
Figure 34.

CROSS-LINKED —

FREE DNA —
Figure 35. Repressor-Operator Cross-link Formation with T/B+15 and T/B+16.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5 and 7 illustrate the level of cross-linking to T/B, T/BpER, T/B+15 and T/B+16, respectively. The inducer-sensitivity of cross-linking to these sequences is shown in corresponding lanes 2, 4, 6 and 8.
Figure 35.

CROSS-LINKED

FREE DNA
Figure 36. Repressor-Operator Cross-link Formation with T/B+18 and T/B+19.

Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5 and 7 show the degree of protein cross-linking to T/B, T/BPER, T/B+18 and T/B+19, respectively. The inducer-sensitivity of cross-linking to these sequences is illustrated in corresponding lanes 2, 4, 6 and 8.
Figure 36.

CROSS-LINKED

FREE DNA
Figure 37. Repressor-Operator Cross-link Formation with T/B+24. Autoradiogram of an SDS polyacrylamide gel used to resolve free and cross-linked DNA species. Lanes 1, 3, 5 and 7 show the degree of protein cross-linking to T/B, T/B+1,2,8,13,15, T/BPER and T/B+24, respectively. The inducer-sensitivity of cross-linking to each of these sequences can be seen in corresponding lanes 2, 4, 6 and 8.
Figure 37.

CROSS-LINKED—

FREE DNA—
Figure 38. Repressor-Operator Cross-link Formation. Autoradiogram of an SDS polyacrylamide gel used to isolate the cross-linked complexes from noncovalently bound protein and DNA. Lanes 1, 4, 6, 9, and 11 show the degree of protein cross-linking to non-substituted operator (T/B), TpER/B, T_{+3}/B, T/BpER, and T/B_{+8} DNA, respectively. The inducer-sensitivity of cross-linking to each of these operator sequences is indicated by the band intensities in corresponding lanes 2, 5, 7, 10 and 12.
Figure 38.

CROSS-LINKED-

FREE DNA—
Figure 39. Efficiency of Repressor Cross-linking to Substituted Sites Along the Operator Sequence. Quantitative summary of the cross-linking efficiencies exhibited by the BrdU-substituted operators. The operator sequence is shown for reference. Solid bars represent the degree of specific cross-link formation while the shaded bars illustrate the level of protein-DNA cross-linking in the presence of inducer. The numerical values are the average of 2-3 experiments, with standard deviations presented in Table 2.
the free DNA. To investigate the origin of the two bands, each was excised and
electroeluted from an SDS-polyacrylamide gel. These separate samples, corresponding
to the upper and lower free DNA bands, were then electrophoresed on a non-
denaturing polyacrylamide gel along with double and single-stranded lac operator
controls (data not shown). An autoradiogram of the gel revealed that the upper free
DNA band migrated with the double-stranded control. The faster migrating species,
however, appeared as two bands with one running just below the double-stranded
control and the second migrating just beyond the single-stranded control. We believe
that these last two bands represent UV-induced DNA cleavage products of a small
fraction of the noncross-linked operator DNA.

Five positions of thymidine substitution exhibited high efficiencies of cross-
linking to the repressor protein under specific binding conditions: +3, 4, 14, 18 and 19.
Lower levels of cross-linking to these operator sites measured in the presence of IPTG
indicate a significant degree of inducer-sensitivity in each case.

Large-scale repressor-T+3/B operator cross-linking

With the cross-linking profiles for each of the BrdU-substituted operators
compiled, we then directed our attention toward defining the repressor's contribution to
particular contacts at the level of the amino acid or peptide directly involved in the cross-
link. Encouraged by its high cross-linking efficiency, we chose to analyze the protein
interaction at the T+3 position within the operator sequence. To accommodate potential
losses throughout sample preparation, it was necessary to increase the scale of the
experiment 3 x 10⁴-fold. As a result of the greater initial working volume, we irradiated
the solution of protein and DNA in 25 ml lots which were later combined and
concentrated. The autoradiogram shown in Figure 40 compares the cross-linking
efficiencies obtained under the small- and large-scale irradiation conditions. Our results
Figure 40. Comparison of Repressor-Operator Cross-link Efficiencies Under Small- and Large-Scale Irradiation Conditions. Repressor, incubated with T₄3/B lac operator, was irradiated in 20 µl and 25 ml lots. Equal counts from each were loaded onto an SDS polyacrylamide gel and analyzed by autoradiography. Lane 1 illustrates the small-scale results with the cross-linking under large-scale conditions presented in lane 2.
Figure 40.

CROSS-LINKED—

FREE DNA—
suggest that the efficiency of cross-linking was not compromised by the increase in scale. Following irradiation, the solution required concentration for the purpose of convenient sample handling throughout subsequent steps. Volumes of 150-200 ml were shell frozen in a dry ice-isopropanol bath within a silanized 600 ml lyophilization flask. In this manner, the 600 ml of irradiated repressor/operator solution was concentrated by lyophilization to 30-50 ml. This solution was dialysed against several changes of 80 volumes of water until the conductivity of the outer solution dropped to that of water indicating that the KCl concentration had been sufficiently reduced. The dialysate was returned to the original lyophilization flask and concentrated to <10 ml. This solution was removed, and the flask was rinsed with water to collect residual radioactivity. Concentration of the solution to 1.5 to 2.0 ml in 2-1.5 ml Eppendorf tubes followed in the Speed Vac concentrator.

Cross-link isolation

The cross-linked protein-DNA complex was readily separated from free repressor and operator by gel electrophoresis. The contents of each of the above Eppendorf tubes was loaded onto an SDS-10% polyacrylamide gel. Under the denaturing conditions used, the free DNA migrated with the dye front while migration of the covalent complex was significantly retarded (Figure 41). Non-cross-linked repressor ran just below the complex (data not shown). The autoradiogram allowed for isolation of the cross-linked species from free protein and DNA by band excision. The band corresponding to free DNA was also excised and carried through the work-up as a control. The protein-DNA and DNA species were readily purified from the acrylamide by electroelution. Now available in solution form, each sample was dialysed against 0.1 M NH₄HCO₃ to remove any TBE and SDS in preparation for proteolytic digestion (of the cross-linked complex) and subsequent treatments.
Figure 41. Large-Scale Isolation of Cross-linked Repressor-T+3/B Operator Complex. Following irradiation, concentration and dialysis of the reaction mix, the sample was loaded onto an SDS polyacrylamide gel to resolve the free and cross-linked DNA. This autoradiogram illustrates the differential electrophoretic mobility of each. The regions of the two DNA species are shown. Band excision from the gel allowed for complex isolation.
Figure 41.

CROSS-LINKED

FREE DNA
Proteolysis

In order to remove the bulk of the repressor molecule and to expose the specific region of the protein involved in the cross-link, the isolated complex was subjected to extensive tryptic and chymotryptic digestion. Figure 42 illustrates the potential cleavage points as determined by Beyreuther et al. (1975). The reaction was terminated by precipitation of the DNA-linked species with ethanol. This step was also included as a means of separating much of the digested protein material from the sample of interest (i.e., the peptides remain in solution). The extent of proteolysis was determined by SDS gel electrophoresis of aliquots of the complex, removed just prior to and after digestion, alongside a lane of free DNA (Figure 43). These results confirm that the 8 hour digestion of the protein-DNA complex was sufficient to remove the bulk of the repressor protein. Lanes 1 and 3 show the migration of the repressor-operator complex and the free DNA, respectively. In this denaturing gel system, the proteolysed sample (lane 2) appears to migrate with the free DNA. While this system confirms complete digestion, it does not confirm the presence of a linked peptide. Indeed, one would not expect the contribution of a relatively short (<10 aa) peptide to the mobility of a 40 bp DNA fragment in an SDS-polyacrylamide gel to be great; however, the possible lability of the covalent linkage to the protease treatment cannot be ignored. Other groups have noted peptide release from the DNA during proteolysis (personal communication). Often digestion with an alternate enzyme eliminated the problem.

HPLC purification

Prior to HPLC purification, the proposed peptide-DNA complex was partially isolated from contaminating digestion products by precipitation with ethanol. In order to determine the elution pattern of any residual peptides and also to evaluate the degree
Figure 42. Tryptic and Chymotryptic Cleavage Sites Within the \textit{lac} Repressor Sequence (Beyreuther et al., 1975). The primary sequence of the \textit{lac} repressor is shown with the trypsin-sensitive sites indicated by arrowheads above the "bond" and chymotrypsin cleavage sites marked below the sequence.
Figure 42.

Met-Lys-Pro-Val-Thr-Leu-Tyr-Asp-Val-Ala-Glu-Tyr-Ala-Gly-Val-
Ser-Tyr-Gln-Thr-Val-Ser-Arg-Val-Val-Asn-Gln-Ala-Ser-His-Val-
Ser-Ala-Lys-Thr-Arg-Glu-Lys-Val-Glu-Ala-Ala-Met-Ala-Glu-Leu-
Asn-Tyr-Ile-Pro-Asn-Arg-Val-Ala-Gln-Gln-Leu-Ala-Gly-Lys-Gln-
Ser-Leu-Leu-Ile-Gly-Val-Ala-Thr-Ser-Ser-Leu-Ala-Leu-His-Ala-
Pro-Ser-Gln-Ile-Val-Ala-Ala-Ile-Lys-Ser-Arg-Ala-Asp-Gln-Leu-
Gly-Ala-Ser-Val-Val-Val-Ser-Met-Val-Glu-Arg-Ser-Gly-Val-Glu-
 Ala-Cys-Lys-Ala-Val-His-Asn-Leu-LeuAla-Gln-Arg-Val-Ser-
Gly-Leu-Ile-Ile-Asn-Tyr-Pro-Leu-Asp-Asp-Gln-Asp-Ala-Ile-Ala-
Val-Glu-Ala-Ala-Cys-Thr-Asn-Val-Pro-Ala-Leu-Phe-Leu-Asp-
Ser-Asp-Glu-Thr-Pro-Ile-Asn-Ser-Ile-Ile-Phe-Ser-His-Glu-Asp-
Gly-Thr-Arg-Leu-Gly-Val-Glu-His-Leu-Val-Ala-Leu-Gly-His-Gln-
Gln-Ile-Ala-Leu-Leu-Ala-Gly-Pro-Leu-Ser-Ser-Val-Ser-Ala-Arg-
Leu-Arg-Leu-Ala-Gly-Trp-His-Lys-Tyr-Leu-Thr-Arg-Asn-Gln-Ile-
Gln-Pro-Ile-Ala-Glu-Arg-Glu-Gly-Asp-Trp-Ser-Ala-Met-Ser-Gly-
Phe-Gln-Gln-Thr-Met-Gln-Met-Leu-Asn-Glu-Gly-Ile-Val-Pro-Thr-
 Ala-Met-Leu-Val-Ala-Asn-Asp-Gln-Met-Ala-Leu-Gly-Ala-Met-Arg-
 Ala-Ile-Thr-Glu-Ser-Gly-Leu-Arg-Val-Gly-Ala-Asp-Ile-Ser-Val-
Val-Gly-Tyr-Asp-Asp-Thr-Glu-Asp-Ser-Ser-Cys-Tyr-Ile-Pro-Pro-
Leu-Thr-Ile-Lys-Gln-Asp-Phe-Arg-Leu-Leu-Gly-Gln-Thr-Ser-
Val-Asp-Arg-Leu-Leu-Val-Asp-Gln-Val-Gly-Val-Thr-Leu-
Asn-Gln-Leu-Leu-Pro-Val-Ser-Leu-Val-Lys-Arg-Lys-Thr-Thr-Leu-
 Ala-Pro-Asn-Thr-Gln-Thr-Ala-Asp-Pro-Ala-Leu-Ala-Asp-Ser-
Leu-Met-Gln-Leu-Ala-Arg-Gln-Val-Ser-Arg-Leu-Glu-Ser-Gly-Gln
Figure 43. Proteolytic Digestion of the Repressor-T₃₋₃/B Operator Complex.

Autoradiogram of an SDS polyacrylamide gel. The extent of proteolysis after an eight hour double digestion was determined by measurement of the electrophoretic mobility of the complex (lane 2) relative to undigested repressor-operator (lane 1) and free DNA (lane 3).
Figure 43.

REPRESSOR-OPERATOR COMPLEX

FREE DNA—
of purification afforded by the precipitation step, *lac* repressor protein was digested under identical conditions and injected onto the anion-exchange column before and after treatment with ethanol (Figure 44). Comparison of chromatograms of proteolyzed repressor alone, with and without the recommended ethanol precipitation step, illustrated not only the early elution of the peptides (5-30 min, Figure 44B), but also the efficient removal of most of the digested protein material with the supernatant following DNA precipitation by ethanol (Figure 44C). This simple control indicates that many of the peptides generated by proteolytic digestion are not precipitated by ethanol and therefore are readily removed from the DNA-linked complex.

Anion exchange HPLC proved to be a valuable purification tool for the peptide-DNA sample in preparation for protein sequence analysis. The early elution of contaminating peptides demonstrated in Figure 44 allowed for isolation of the more highly charged free and complexed DNA species from residual unwanted digestion products. Figures 45A-C and 46A-C present a series of chromatograms illustrating the elution profiles of both free DNA and the peptide-DNA cross-linked species from two separate experimental attempts. These two large-scale preparations represent the culmination of at least five complete small-scale trials with the T+3/B operator DNA conducted to evaluate the protocol. The free DNA, shown in Figure 45B, elutes at 51.0-53.5 minutes into the gradient. In Figure 45C, the presumed peptide-DNA complex appears to elute in three distinct peaks, with two minor peaks at 43.0-45.0 minutes (P1) and 47.0-48.2 minutes (P2) and a major product at 50.0-52.0 minutes (P3) into the gradient. The results of the second HPLC run, shown in Figure 46, are similar with a major free DNA peak appearing at 53-55.5 minutes (Figure 46B) and the cross-linked sample eluting in three distinct peaks: P1, 45.8-46.9 minutes; P2, 48.8-50.0 minutes; P3, 51.0-54.0 minutes (Figure 46C). Elution of these DNA-containing species was
Figure 44. Anion Exchange HPLC Profile of Digested *lac* Repressor Protein. Repressor (0.4 mg) was digested with trypsin and chymotrypsin under the conditions detailed for the protein-DNA complex in Materials and Methods. The proteolysed sample was injected onto an anion exchange HPLC column and eluted with a salt gradient (panel B). Panel C illustrates the elution profile of a similarly digested protein sample that was subsequently treated with ethanol and centrifuged. The pelleted material was resuspended and injected onto the column. A buffer profile is shown in panel A. The peak appearing at 32-33 minutes can be attributed to an air bubble.
Figure 44.

A

B

C

TIME (MINUTES)
Figure 45. Anion Exchange HPLC Profile of Proteolysed Repressor-T₄₃/B Operator Complex (1). The cross-linked protein-DNA complex was digested with trypsin and chymotrypsin and precipitated with ethanol. The pelleted material was resuspended in 20 µl of column buffer and injected onto the column. Elution with a salt gradient yielded the profile shown in Panel C. Elution was monitored spectrophotometrically using a 224 nm cut off filter. Peaks were also identified by radioactive monitoring and were collected for further analysis. The three peaks noted at 43-52 minutes in Panel C were collected and labeled P1, P2 and P3 in order of their elution. The elution profile of a noncross-linked DNA sample appears in Panel B. The peak at 51-53.5 minutes was collected. Buffer alone exhibits the profile shown in Panel A.
Figure 45.
Figure 46. Anion Exchange HPLC Profile of Proteolysed Repressor-T₃/B Operator Complex (2). A second repressor-operator complex was treated with trypsin and chymotrypsin, precipitated with ethanol and injected onto the anion exchange column. The gradient elution is presented in Panel C. Again, three peaks were observed and collected for further analysis (P₁, P₂ and P₃). The elution profile shown in Panel B represents the free DNA sample. Only the peak at 53-55.5 minutes was collected. Panel A illustrates the background profile of buffer alone.
Figure 46.
subsequently monitored and confirmed by following the radioactivity of the effluent. In addition to the degree of purification afforded by HPLC, a comparison of the elution profiles of free and complexed DNA from Figures 45 and 46 provides evidence for the presence of a covalently-linked peptide within the complexed sample. The earlier elution of the cross-linked species with respect to free DNA is in agreement with results of a similar experiment reported by Merrill et al. (1984). The limited resolution of free and complexed DNA observed under our gradient conditions would suggest minimal influence of the attached peptide on the overall elution properties of the peptide-DNA complex relative to the free DNA. If both size and charge determine the retention of a sample on the column, a peptide of 500 to 1000 MW would not be expected to exert much influence on the elution of a highly charged DNA fragment of 26,400 MW.

Further analysis of each sample by protein sequencing and/or amino acid analysis required removal of the salts present in solution. A dialysis step, against water, was included to lower the ionic strength. Table 3 chronicles the experimental yield from irradiation of the protein-DNA solution through HPLC purification and subsequent dialysis. The data illustrate the considerable losses encountered due to extensive, yet necessary, sample handling. The use of silanized glassware and exhaustive rinsing of vessels aided our efforts, but could not eliminate all losses due to handling and transfer of the sample. Clearly, the most dramatic losses occurred at the early stages of the procedure. In the first experiment, up to 85% of the total radioactive material was lost to the lyophilization flask and dialysis membrane. This was improved to 50% in the second experiment through more extensive rinsing of the silanized flask. Electrophoresis and subsequent electroelution of the free and cross-linked DNA samples also
Table 3. Recovery of the Cross-linked Protein-DNA Complex From Experiments 1 (Figure 45) and 2 (Figure 46). In each instance, a maximum yield of 12.6 nmole of protein-linked DNA would be expected based on a cross-linking efficiency of 37% for the T4+3/B operator fragment. The overall yields for experiments 1 and 2 were approximately 220 pmole and 120 pmole, respectively. Within the table, 100% represents the total cpm available prior to irradiation and not the maximum attainable yield. Sample loss was monitored throughout the work-up procedure by following the recovered cpm at selected stages within the protocol. The stages are defined below.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Sample Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Incubation of repressor and operator (total cpm)</td>
</tr>
<tr>
<td>2</td>
<td>Irradiation, initial concentration, dialysis and secondary concentration of sample prior to gel electrophoresis</td>
</tr>
<tr>
<td>3</td>
<td>Electrophoretic separation of free and cross-linked DNA, electroelution from acrylamide and dialysis</td>
</tr>
<tr>
<td>4</td>
<td>Proteolytic digestion of the repressor-operator complex followed by sample precipitation with ethanol prior to HPLC isolation</td>
</tr>
<tr>
<td>5</td>
<td>HPLC purification, dialysis and concentration prior to final analysis</td>
</tr>
</tbody>
</table>

P1, P2 and P3 represent the three peaks observed in the elution of the cross-linked sample. They are listed in order of elution.
Table 3
Recovery of the Cross-linked Repressor-T+3/B Operator Complex at Various Stages in Sample Preparation

<table>
<thead>
<tr>
<th>Stage</th>
<th>(1) cpm</th>
<th>%</th>
<th>(2) cpm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9.7 x 10^8</td>
<td>100</td>
<td>3.6 x 10^8</td>
</tr>
<tr>
<td>2</td>
<td>1.4 x 10^8</td>
<td>14.4</td>
<td>1.8 x 10^8</td>
</tr>
<tr>
<td>3 cross-link</td>
<td>2.8 x 10^7</td>
<td>2.9</td>
<td>2.2 x 10^7</td>
</tr>
<tr>
<td>free DNA</td>
<td>----</td>
<td>----</td>
<td>2.2 x 10^7</td>
</tr>
<tr>
<td>4 cross-link</td>
<td>9.0 x 10^6</td>
<td>0.9</td>
<td>7.1 x 10^6</td>
</tr>
<tr>
<td>free DNA</td>
<td>7.6 x 10^6</td>
<td>0.8</td>
<td>6.5 x 10^6</td>
</tr>
<tr>
<td>5 cross-link (P1)</td>
<td>2.4 x 10^5</td>
<td>0.0</td>
<td>8.6 x 10^4</td>
</tr>
<tr>
<td>cross-link (P2)</td>
<td>3.5 x 10^4</td>
<td>0.0</td>
<td>2.8 x 10^5</td>
</tr>
<tr>
<td>cross-link (P3)</td>
<td>6.3 x 10^6</td>
<td>0.6</td>
<td>1.5 x 10^6</td>
</tr>
<tr>
<td>free DNA</td>
<td>5.0 x 10^6</td>
<td>0.5</td>
<td>2.2 x 10^6</td>
</tr>
</tbody>
</table>
accounted for significant loss of material. The sample handling in experiment 1 appeared to be more efficient at this stage. With an expected cross-linking efficiency of the T+3/B operator DNA of 37%, a recovery of 2.9% of cross-linked material from the gel was satisfactory. A considerably less efficient recovery of the repressor-operator complex from the acrylamide was noted in the second experiment. The factors responsible for this outcome are unknown, although irreversible attachment of the protein-linked material to the BT1 membrane following electroelution often proved troublesome. The extremely low yield of free DNA noted in the second experiment was most likely due to cross-linking of the intact fragments to the acrylamide. Proteolysis and ethanol precipitation of the samples also led to losses due to adhesion to the walls of the Eppendorf tubes and to incomplete resuspension of the sample following precipitation. It should be noted that overall recovery of the free DNA fraction was low primarily due to less careful handling of these samples at the various stages. Already present in excess over the cross-linked species and necessary only as a control, we were concerned only with obtaining at least an equivalent quantity of the uncomplexed DNA. Given the small amount of material available at the HPLC stage, our recoveries from the column and subsequent dialysis were acceptable. When calculated relative to the gel-purified repressor-DNA complex (Table 3, stage 3), the overall yields for the peptide-DNA species from experiments 1 and 2 were 21% and 6.6%, respectively.
Protein sequencing

Protein sequence analysis of the peptide-DNA complex (P3) was performed in an effort to define the specific region of the repressor protein involved in the contact at the T+3 operator site. Our second experiment yielded 120 pmole of cross-linked material from P3 (Figure 46, Table 3). This sample was analysed for peptide sequence on an ABI Protein Sequencer, Model 473. An equivalent amount of free DNA was similarly analysed to control for chemical interference by the DNA portion of the complex. The protein sequencing results from 10 cycles revealed no protein contamination. The sequencing chemistry appeared to react with the DNA such that non-protein degradation products consistently appeared. These peaks gradually decreased in size with successive sequencing cycles. The protein sequencing results from the peptide-T+3/B operator complex are depicted in Figure 47. Ten cycles of Edman degradation were performed by the instrument and the resultant chromatograms are shown (Figure 47A-K). Figure 47L represents the free DNA elution pattern from the 8th sequencing cycle. The double injection front, visible for every sequencing cycle, is noted along with the recurring peaks for DMPTU and DPTU (diphenylthiourea-based reaction by-products). Peaks appearing in panels B-K and apparently due to DNA degradation are indicated. The sequence derived from these data is: Val-Val-Asn-Gln-Ala-Ser-?-Val-Ser-Ala. When reading the sequence, one notes, with each successive cycle, an additional peak corresponding to the previous amino acid. This is a common lag response. In panel E, the major peak corresponding to glutamic acid appearing just after the glutamine peak results from the elimination of ammonia from the parent glutamine (Fairwell, 1983). The identified sequence directly corresponds to a peptide within the lac repressor protein spanning positions 23-32. Within this peptide, the missing amino acid at position 29 is histidine.
Figure 47. Protein Sequence Analysis on the Peptide-T_{+3}/B Operator Complex. The sequence analysis was performed on an ABI Protein Sequencer (Model 473) by Dr. Richard Cook at Baylor College of Medicine. The HPLC elution chromatograms from the ten cycles of Edman degradation are shown in panels B-K. Panel A depicts the elution of the amino acid standards for reference. Panel L illustrates the chromatogram generated from the 8th sequencing cycle performed on the free DNA sample. The major contaminating peaks noted in B-K derived from DNA interference are indicated by arrows. The double injection front (IF) along with the DMPTU and DPTU peaks are likewise marked. The peptide sequence derived from this analysis is: Val-Val-Asn-Gln-Ala-Ser-?-Val-Ser-Ala. This corresponds to the lac peptide spanning amino acids 23 through 32.
Figure 47.
Figure 47 (continued).
DISCUSSION

Direct substitution of DNA with BrdU, a thymidine analog, at single sites was chosen as a means of probing the binding surface within the major groove of operator DNA for individual repressor contacts. Our experimental goals required us to focus initially on the synthetic substitution of the lac operator sequence and subsequent purification of the strands. Figure 9 illustrates the comparable degrees of purification attained for both the non-substituted (lane 1) and the BrdU-substituted (lanes 3 and 4) oligonucleotides. The higher molecular weight bands appearing in the non-purified sample (lane 2) reflect incomplete denaturation and overloading of the lane. The remaining preparatory steps, 5'-end labeling and strand hybridization, were likewise uncomplicated by incorporated BrdU (Figs 10 and 11, respectively).

The ability of the substituted operators to bind lac repressor was measured under equilibrium conditions (Table 1). When compared to control DNAs, the per-substituted operators (T\textsubscript{PER}/B and T/B\textsubscript{PER}) bound lac repressor 3-5 fold more tightly. This overall increase in binding affinity, first noted by Lin and Riggs (1972), has been hypothesized to derive from the increased dipole moment of BrdU relative to thymidine; this alteration could yield a more stable complex through direct interaction of protein substituents with the bromine atom. This explanation is plausible for those contacts involving ionic or polar repressor groups (Goeddel \textit{et al.}, 1978a,b). BrdU incorporation might also influence complex stability through its effect on base stacking, charge distribution along the helix and/or overall DNA conformation (Goeddel \textit{et al.}, 1978a). Lin and Riggs (1971) observed a comparable increase in repressor binding affinity for poly[\textit{dA},BrdU)], while Ogata and Gilbert (1977) were unable to detect UV protection of the same DNA by repressor, suggesting that the increased stability involves minor groove contacts and is a nonspecific effect. Our results indicate that incorporation of
BrdU at a single site within the operator minimally affects the equilibrium binding affinity for repressor, as the difference in equilibrium binding constants obtained for the singly-substituted operators is within the error observed for this assay.

The BrdU-substituted operators were examined for effects of bound repressor protein by observing protection against UV-induced strand breakage at each specific site of BrdU incorporation. The bound protein can offer protection through shielding of the DNA at the site, donation of a hydrogen to the uracil or deoxyribose radical, or direct cross-linking to the DNA through an amino acid. Additionally, the protein could protect the DNA via indirect contacts through solvent molecules (i.e. hydrogen donors such as Tris, EDTA, DTT) at a given BrdU position or by eliciting conformation changes in the DNA that preclude strand scission. Initial determinations involved the per-substituted operator sequences, with operator-specific contacts identified at positions +1, 2, 3, 4, 8, 13, 15, 16, 18, 19, and within 20-22. These results were confirmed and extended by identical assays performed with the individually BrdU-substituted sequences. The singly-substituted operator sequences afforded improved clarity and resolution of the protection pattern, especially where neighboring sites were involved. Densitometric scanning of the autoradiograms allowed quantitation of the degree of protection afforded at each position. Significant protection is found at positions +1, 2, 3, 4, 6, 8, 13, 15, 16, 18, 19, 20, and 21, while lesser protection is observed at -5 and -3 (Table 2, Figure 27). The bar graph shown in Figure 27 illustrates the pattern of protection afforded by bound repressor at sites across the operator sequence. With the understanding that our probe for protein-DNA contacts was limited to thymidine positions within the sequence, protection of 70-85%, potentially corresponding to the more favorable contacts, was noted throughout the central 20 bp of the operator. When analyzed in conjunction with DNA methylation
results (Gilbert et al., 1976, Figure 5), the observed protection at selected sites is better understood. Methylation protection within the major groove was noted at positions 5, 7, 11, 12 and 17. Each of these sites is adjacent to or very near a repressor-operator contact site defined by UV protection results. Protection against operator methylation in the minor groove was observed at -3, 3, 4, 13 and 19. While we measured significant protection at each of these positions, except at -3, it is particularly interesting to note that at 4, 13 and 19 we additionally observed a significant degree of inducer-insensitive protection. It is unlikely that simultaneous major and minor groove contacts are made at these positions; however, it is plausible that induced alterations in the DNA conformation could account for methylation protection at -3, 3, 4, 13 and 19 and allow for the significant degree of nonspecific binding observed in our study at 1, 4, 13, 18 and 19. In summary, the protection pattern generated by bound repressor protein suggests that strong contacts are made within the major groove throughout the central portion of the operator binding site and that nonspecific interactions may also function in repressor-operator binding at select positions within both halves of the sequence.

Analysis of the UV-irradiated samples for covalently cross-linked protein-DNA complexes yielded somewhat contrasting results. A direct cross-link to the DNA would imply a close interaction between the repressor and the C5 position of thymidine within the major groove. In addition to distance constraints, cross-link formation is also dependent upon orientation of the interacting groups and particularly on the relative reactivity of the amino acid(s) making the contact at a specific site. The hydrogen-donating capacity of a particular amino acid side chain could influence its ability to cross-link. Smith (1969) measured the photochemical addition of 22 amino acids to uracil and found that the eleven reactive amino acids were (in order of de-
creasing reactivity) cysteine, tyrosine, phenylalanine, cystine, histidine, arginine, lysine, methionine, tryptophan, serine, and glycine. These and other factors are critical in the cross-linking and protection processes.

Covalent complexes were easily identified in an SDS-PAGE system by their lower mobility with respect to free DNA. Densitometry of autoradiograms provided a means of quantitating the efficiency of cross-linking. With a clear means of visualizing and quantitating the cross-linkability of repressor to variously BrdU-substituted DNA sequences, we first assessed the sequence-specificity of covalent complex formation. The bar graph in Figure 28 illustrates the efficiencies of repressor cross-linking to both its cognate operator sequence (TPER/B, panel A) and to a representative non-specific sequence, the trp operator (trp TPER/B, panel B). When comparing the levels of cross-linking measured for the two sequences at either 0.10 M or 0.25 M KCl, one notes that the repressor protein preferentially forms covalent links to the lac operator sequence. These results confirm the sequence specificity of the binding interaction and support our selection of experimental conditions for this study. The results shown in Figure 28 also illustrate the inducer-sensitivity of repressor-DNA binding. In panel A, a significant level of inducer-sensitivity is measured under both the low and high salt conditions, but the sensitivity to IPTG is greater at 0.25 M KCl. This observation is consistent with the competitive effect of increasing ionic strength on nonspecific binding interactions. For a direct illustration of this competition, note the absence of measurable cross-link formation to the trp operator, a nonspecific DNA for lac repressor, at 0.25 M KCl (panel B). Binding of the trp sequence (panel B) was not responsive to the presence of IPTG as would be expected for any nonspecific sequence.
As an added control for the UV protection and cross-linking studies, we separately assessed the effect of UV irradiation on the repressor protein and the operator DNA by measuring the equilibrium binding affinities before and after UV treatment. The $K_d$ obtained for pre-irradiated repressor binding to T/B operator DNA was $8.3(\pm1.4) \times 10^{-10} \text{M}$ while that for the irradiated T/B DNA binding to repressor was $7.3(\pm0.7) \times 10^{-10} \text{M}$. When compared to the $K_d$ for untreated repressor-T/B binding [$6.2(\pm1.6) \times 10^{-10} \text{M}$, Table 1], these results suggest that 90 seconds of UV irradiation did not appreciably damage either the protein or the DNA. The UV dose-response curves generated for $T_{+3}/B$ and $T/B_{+19}$ (Figure 29) indicate that inducer-sensitive cross-link formation maximizes within 20-25 seconds of irradiation. This observation would suggest that photoreaction occurs and is completed before permanent damage to either species results and that the data obtained under the conditions chosen for our study reflect native binding phenomena.

The cross-linking results obtained from the BrdU-substituted $lac$ operator sequences are summarized in Table 2 and Figure 39. In the absence of inducer, five positions of BrdU incorporation yielded cross-linking with significantly high efficiencies. Substitution for thymidine at symmetrically related sites, $T_{+3}$ and $B_{+19}$, resulted in the greatest degree of UV-induced cross-linking: 37 (±4.1)% and 28 (±5.0)%, respectively. It is interesting to note that these values surpass those measured for the corresponding per-substituted operators. This result suggests that once an initial protein-DNA cross-link forms, subsequent sites in the per-substituted strand were prevented from cross-linking, potentially due to resultant conformational alterations. In an earlier study, Goeddel et al., (1978a,b) found that BrdU substitution at either of these two positions had no effect on the binding free energies of the complexes. Based on these results, they concluded that no favorable major groove contact was made by
the repressor at these sites. In contrast, our data provide evidence for close protein-DNA interaction at both positions. In addition, the very high level of IPTG-sensitive cross-linking observed at position T+3 required only 1/3 the irradiation time compared to all other positions, indicating that all aspects of this protein-DNA contact were favorable. A recent NMR analysis of complexes between the isolated NH₂-terminus fragment from repressor and a 22 bp symmetric operator found evidence for the proximity of His29 to the T+3 position (Lamerichs et al., 1989). Given the potential photoreactivity of histidine, our cross-linking results are consistent with this contact determination, and direct sequencing of the cross-linked site confirms this contact (see below). It is interesting to note that an earlier study identified His29 as the amino acid involved in cross-linking of repressor to BrdU-containing non-operator DNA (Barbier et al., 1984). While overall inducer sensitivity is observed, any nonspecific contribution to the binding could account for the significant degree of cross-linking measured in the presence of IPTG. If one assumes a symmetric association of the repressor protein, the high level of cross-linking at B+19, as well as the degree of inducer-insensitive cross-linking, could be explained similarly. A second set of symmetrically oriented positions, T+4 and B+18, show a lower, but significant, level of cross-linking. Combining the data from our UV protection results and methylation patterns observed previously (Gilbert et al., 1976), bound repressor appears to cover or affect the protection at the +4 position in both the major and minor grooves, while the +18 position is affected only in the major groove. Taking into account the degree of cross-linking and the juxtaposition of +4 and +18 relative to +3 and +19, it is possible that the cross-links formed at +4 and +18 involve the same groups as those at +3 and +19, respectively. One additional substitution, at +14, resulted in significant cross-link formation with the repressor protein. While no protection against strand scission by
repressor was noted at this position, UV protection results indirectly suggest major
groove coverage at this site since the surrounding bases (13, 15, 16) are protected
against UV damage in the major groove.

The remaining thirteen substituted positions exhibited cross-linking efficiencies
at or near the level measured for the non-substituted operator. The UV protection
assay failed to detect significant protein effects at the outer extremes of the operator
sequence (positions -5, -3, +22, +24). Exonuclease III and DNase I footprinting
results (Schmitz and Galas, 1979; Shalloway et al., 1980; Manly et al., 1984) likewise
indicate that these sites are at the edge or just beyond the contacted region of the
operator. Positions +1, 2, 6, 8, 13, 15, 16, 20 and 21, though unable to cross-link with
repressor, were protected against UV damage by bound protein. DNA structural
changes that accompany repressor binding (Zweib et al., 1989) may account in part for
these differences; changes induced in the orientation and distance between the C5 on
the pyrimidine and the sugar-phosphate backbone may preclude the scission reaction.
Other explanations for the difference between protection and cross-linking at a specific
site are the absence of a photoreactive amino acid at the contact site and indirect con­
tact made by the protein through a solvent molecule. Along this vein, Goeddel et al.
(1978a,b) found that BrdU substitution at positions 1, 2, 8, 20 or 21 destabilized the
bound complex. They hypothesized that the primary contact in each case was made
through the thymine methyl and likely involved a hydrophobic side chain. This
hypothesis is consistent with our cross-linking results, since such an amino acid would
not be photoreactive based on the findings of Smith (1969). It is also possible that the
observed UV protection resulted from an indirect contact made by the protein through a
solvent molecule. If so, this situation could account for the observed results at
positions 15 and 16 where BrdU substitution had no effect on the binding free energies
of the protein-DNA complexes (Goeddel et al., 1978a,b). Our protection and cross-linking results would also indicate the participation of an indirect interaction which could preclude repressor-operator cross-link formation at these otherwise important contact sites.

Although more specific descriptions of the various contacts is hampered by the limited information regarding specific amino acid-nucleotide interactions currently available, a few projections can be made. The contact at position 6, for example, could be influenced by the neighboring GC5-Arg22 interaction predicted by Lehming et al. (1988), Sartorius et al. (1989), and Lamerichs et al. (1989). Given a contact at the +5 position within the major groove involving Arg22, interaction with neighboring Ser21 at +6 could account for the observed protection and complex stability. Though demonstrated to be active in photoaddition to uracil, the reactivity of serine was lowest (along with glycine) among the eleven reactive amino acids identified (Smith, 1969). The limited photoreactivity of serine (and its neighboring amino acids within the lac repressor primary sequence) could explain the absence of detectable cross-link formation at the +6 position. A similar situation likely governs the interaction of repressor at the +8 position of the operator. Early work by Goeddel et al. (1977) demonstrated the importance of the 5-methyl group at this position to repressor-operator complex stability. Favorable interaction through the hydrophobic methyl group suggests the participation of a nonpolar or hydrophobic side chain. The proximity of this amino acid to the C5 position of the base would allow for protection of the DNA against UV damage, but the photoreactivity of such an amino acid would be low. Very little information is available regarding specific amino acid involvement with bases in the promoter distal half of the operator sequence. For this reason, we can only
postulate that similar interactions might influence contacts made at symmetrically related positions.

These interpretations are consistent with the observed absence of a direct relationship between protection against UV damage and protein-DNA cross-linking. We have noted several instances where bound repressor was able to protect specific sites within the operator against UV-induced breakage, but was unable to form cross-links at those sites (i.e., +1, 2, 6, 8, 13, 15, 16, 20, 21). In several cases cited, amino acid reactivity and/or solvent participation at the contact site can account for the results. However, it is also likely that protein-induced changes in DNA conformation influence the protection and cross-linking patterns. Our results obtained with lac operator substituted with BrdU at the +14 position emphasize this point. While bound repressor was unable to protect the operator against UV-induced strand scission at position 14, cross-linking to this site was observed (Table 2, Figures 27, 39). This result would suggest an as yet undescribed reaction mechanism whereby protein-DNA cross-link formation and DNA chain breakage both occur. It is possible that the nature and orientation of the interacting molecules at this particular site are unique. In particular, the local DNA conformation may contribute to the factors involved in the unusual behavior at this site. In fact, DNA bending in response to repressor binding has been noted (Culard and Maurizot, 1981; 1982). Zwieb et al. (1989) found that the repressor-induced bend was located at or near the dyad symmetry axis of the operator sequence at +11. It is possible that the conformational change in the operator results in a strained structure in the vicinity of the +14 position which alters the usual pattern of DNA breakage and provides simultaneous protein-DNA cross-linking.

In this study we have applied two distinct methods of probing the major groove of operator DNA for repressor interaction at selected thymidine positions. Our
approach enabled us to quantify and to compare the UV protection and cross-linking assays using the same substrate. Assessing the ability of repressor to protect the operator against UV-induced strand scission, apparent contacts are identified at positions +1, 2, 3, 4, 6, 8, 13, 15, 16, 18, 19, 20, and 21 with the degree of protection varying from 27% to 84%. Under identical experimental conditions, protein-DNA cross-links were detected only at +3, 4, 14, 18, and 19. This more limited set of cross-linkable sites suggests that protection and cross-linking are not directly related, and individual sites are influenced by local structural and environmental factors. Positions within the operator both protected against strand scission and able to participate in cross-linking to bound protein can be confidently identified as sites of close contact by repressor. Protection without cross-linking at a given operator position points to environmental factors influencing strand scission (e.g., DNA conformational changes, absence of reactive amino acid, etc.). Similarly, cross-link formation without measurable protection at a site within the DNA may indicate localized regions of altered DNA conformation that contribute to an unusual reaction mechanism.

Combined measurement of UV protection and cross-linking efficiencies to singly BrdU-substituted target DNA provides dual assessment using the same substrate and thereby increases confidence in interpreting the occurrence of protein contact at a specific residue.

With the goal of more specifically characterizing and defining individual repressor-operator contacts, we devised a protocol for generating and purifying large amounts of a singly cross-linked species. For our initial attempts we chose to work with the T+3/B operator fragment, given its high efficiency of cross-linking with repressor (Table 2, Figure 39). It is important to note at this point that the population of cross-linked protein-DNA complexes following sample irradiation need not be
homogeneous with respect to the destination of the cross-link within the repressor molecule. While linkage to a particular amino acid group might be favored within the resultant population, protein and DNA conformational flexibilities and/or constraints along with the local environment of the particular contact could influence the possibility of multiple free radical destinations. This consideration is critical to the interpretation of any final protein sequence or amino acid composition analysis.

The covalent repressor-operator complexes were isolated from free protein and DNA by gel electrophoresis, band excision and electroelution from the acrylamide. Figure 41 illustrates the separation of the cross-linked complex from free DNA and depicts the regions excised. With the autoradiogram of the gel as a guide, excision of the well-resolved free and cross-linked DNA bands allowed for complete isolation of the two species. Purification of the two samples from the acrylamide by electroelution resulted in high recovery of each in solution form. Extensive digestion of the repressor molecule was necessary for isolation of the peptide involved in contacting the operator at the T+3 site. The expected tryptic/chymotryptic digestion pattern of the lac repressor is shown in Figure 42 (Beyreuther et al., 1975).

Given the potential for multiple cross-link destinations from a single site within the operator sequence, the generation of a family of singly DNA-linked peptides, either closely nested within the primary sequence or widely separated due to repressor folding patterns, is a possible result. This point is of particular importance during final analysis of the peptide-linked DNA. Previously reported repressor-operator contact determinations suggest, however, that the critical interactions with operator DNA occur within the amino-terminal domain of the protein (Ebright, 1986; Lehming et al., 1988; Sartorius et al., 1989; Lamerichs et al., 1989). More specifically, Lamerichs et al. (1989) reported that His29 is involved in a contact at the +3 position within the
operator sequence. Given complete digestion with both enzymes (Figure 42), this result suggests the possibility for DNA linkage to a peptide spanning amino acids 27-29 or 27-33 (if linkage at His29 interfered with chymotryptic cleavage). A sequence was identified upon analysis of the P3 sample from the second cross-link preparation. The data shown in Figure 47 indicate the linkage of the Val23-Ala32 peptide. The terminal Lys, predicted from the enzyme cleavage patterns (Figure 42), would likely have been identified following an 11th cycle of Edman degradation. The absence of an amino acid signal in the 7th cycle (Figure 47H) strongly suggests that covalent attachment to the DNA at the +3 position occurs through the predicted His29. Such an attachment would likely interfere with the sequencing chemistry and effectively prevent detection of the linked residue. This determination supports our earlier prediction and confirms the results of Lamerichs et al. (1989). The model for the complex between lac repressor NH2-terminus and operator DNA reported by de Vlieg et al. (1989) also upholds our His29-T+3 contact determination. Derived from molecular dynamics simulation and based on NMR data, their sketch (Figure 48) reveals the stable positioning of the His29 side chain near the AT3 base pair. They observed hydrogen bonds bridged by water between His29 and the 5' oxygen, along with the phosphate, of adenosine at the +1 position in the operator. Nonpolar contacts were likewise noted between His29 and adenines at +1 and +2 (de Vlieg et al., 1989). While no direct His29-T+3 contact was implicated in their model, the predicted positioning of the His29 side chain is consistent with covalent cross-link formation between this amino acid and the +3 position within the operator sequence. It is interesting to note that His29 was also shown to participate in cross-link formation with BrdU-containing non-operator DNA (Barbier et al., 1984). These results suggest that this region of the lac repressor and specifically His29, located just outside the putative helix-turn-helix
Figure 48. Contact Between the *Lac* Repressor NH$_2$-terminus and Operator DNA Derived From Molecular Dynamic Simulation of NMR Data (de Vlieg *et al.*, 1989). Contacts are indicated by circles: D, hydrogen bond; W, hydrogen bond bridged by water; N, nonpolar contact.
Figure 48.
region (aa5-26), is of critical importance to the binding interaction with both specific
and non-specific DNA sequences.

The elution of 3 distinct peaks (P1, P2 and P3) from the HPLC column might
indicate that sub-populations of cross-linked peptides were generated. As discussed
earlier, the possibility for distinct cross-link destinations within the protein cannot be
ignored. Incomplete proteolysis could also generate cross-linked peptides of varying
length and charge. Either of these possibilities (or a combination of the two) could
account for the multiple HPLC peaks observed for the cross-linked sample. While we
were able to isolate the three peaks, we were only able to analyze the major
component (P3) given the instrumentation available to us. Analysis of these lesser
components could provide further insight into the nature of the contact made by
repressor at this site.

This study has been extended to include identical cross-linking experiments
between the lac repressor protein and the T/B+19 BrdU-substituted operator sequence.
The symmetric opposition of the +3 and +19 positions with respect to the center of
sequence symmetry at +11 affords a means of addressing the symmetry of repressor
binding to the operator. This analysis is currently underway in the laboratory. Future
plans for this project also include analysis of the repressor contact at the +14 position.
Our earlier UV protection and cross-linking efficiency experiments with the BrdU-
substituted T+14/B operator suggested that contact at this site might be influenced by
strain within the local DNA conformation. While our approach could not directly
address this possibility, identification of the peptide and/or the amino acid involved in
this interesting contact would add to our knowledge of the protein-DNA binding
interface.
BIBLIOGRAPHY


Daly, T.J., and Matthews, K.S. (1986b) *Biochemistry* 25, 5479-5484.


