Molecular Studies on the Regulation of the Adaptive Response to Alkylating Agents in *Escherichia coli*

by

Stephen J. Hughes

A thesis submitted for the degree of Ph.D
at
The University of London
November, 1989

Imperial Cancer Research Fund
and
University College London
ACKNOWLEDGEMENTS

The work described in this thesis was supported by a bursarship from the Imperial Cancer Research Fund, to whom I am indebted. I would especially like to thank Dr. Tomas Lindahl and Dr. Barbara Sedgwick, my supervisors at the I.C.R.F, for their guidance and encouragement throughout the course of this work. I am also grateful to my external supervisor at University College, Dr. Peter Swann, for his kind assistance. Many thanks must go to my colleagues in the Mutagenesis Group for their advice and good humour, and to John Nicholson for preparing the photographs presented in this thesis. Last, but not least, I would like to acknowledge the continuous support given by my family and friends.
When *Escherichia coli* cells are exposed to low, non-lethal doses of simple alkylating agents, such as N-methyl-N'-nitro-N-nitrosoguanidine and N-methyl-N-nitrosourea, they acquire an increased resistance to the mutagenic and cytotoxic effects of a subsequent challenge with a high dose of these agents. This cellular response, referred to as the adaptive response to alkylating agents, results from the increased expression of at least four genes, *ada, alkA, alkB* and *aidB*. The products of the *ada* and *alkA* genes are DNA repair enzymes that remove potentially mutagenic and cytotoxic alkylation lesions from cellular DNA. The Ada protein also acts as a positive regulator of expression of the inducible genes of the adaptive response.

In this study the regulatory *ada* gene has been cloned from four independently-isolated *E. coli* mutants which express the adaptive response constitutively in the absence of alkylation damage. Nucleotide sequence analysis has shown that each gene contains two G-C to A-T transition mutations within the coding region which result in the synthesis of a mutant Ada protein containing two amino acid substitutions in the N-terminal region. Three of the mutant strains have the same two *ada* mutations. *E. coli* transformed with recombinant plasmids carrying the mutated *ada* genes overexpressed both the mutated *ada* gene and the chromosomal *alkA* gene, indicating that the mutant Ada proteins are strong inducers of *ada* and *alkA* gene expression *in vivo*. The Ada protein purified from one of the constitutive mutants was also shown to be a strong inducer of expression of the wild-type *ada* gene in an *in vitro* coupled transcription-translation system. One amino acid substitution, methionine-126 substituted by isoleucine, occurred in the Ada protein synthesized by all four mutant strains, and this substitution alone has been shown to be sufficient to convert the Ada protein into a strong activator of *ada* and *alkA* gene expression *in vivo*. Models for the mechanism by which this substitution exerts its effect on gene expression are presented.
## CONTENTS

### CHAPTER 1 - GENERAL INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. DNA Repair and the Maintenance of the Genetic Material</td>
<td>11</td>
</tr>
<tr>
<td>B. DNA Repair Pathways of <em>Escherichia coli</em></td>
<td>12</td>
</tr>
<tr>
<td>1. Mismatch correction</td>
<td>12</td>
</tr>
<tr>
<td>2. Enzymatic photoreversal</td>
<td>18</td>
</tr>
<tr>
<td>3. Excision repair</td>
<td>19</td>
</tr>
<tr>
<td>C. Regulation of the DNA Repair Pathways of <em>E. coli</em></td>
<td>29</td>
</tr>
<tr>
<td>1. SOS response</td>
<td>29</td>
</tr>
<tr>
<td>2. Adaptive response to oxidation damage</td>
<td>33</td>
</tr>
<tr>
<td>D. DNA Damage Induced by Simple Alkylating Agents</td>
<td>37</td>
</tr>
<tr>
<td>E. The Adaptive Response to Alkylating Agents</td>
<td>42</td>
</tr>
<tr>
<td>F. Aims and Rationale of this Study</td>
<td>58</td>
</tr>
</tbody>
</table>

### CHAPTER 2 - MATERIALS AND METHODS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Materials</td>
<td>61</td>
</tr>
<tr>
<td>1. Bacterial strains</td>
<td>61</td>
</tr>
<tr>
<td>2. Plasmids</td>
<td>61</td>
</tr>
<tr>
<td>3. Bacteriophage strains</td>
<td>61</td>
</tr>
<tr>
<td>4. Media</td>
<td>61</td>
</tr>
<tr>
<td>5. Radiochemicals</td>
<td>63</td>
</tr>
<tr>
<td>6. Molecular biology reagents</td>
<td>63</td>
</tr>
<tr>
<td>7. Oligodeoxyribonucleotides</td>
<td>63</td>
</tr>
</tbody>
</table>
8. Electrophoresis reagents
9. General chemical/biochemical reagents

B. Methods

1. Bacterial growth conditions
2. Preparation of bacteriophage f1 suspension
3. Survival curves-sensitivity to killing by methylmethane-sulphonate
4. Preparation of crude cell extracts
5. Assay for O\textsuperscript{6}-methylguanine-DNA methyltransferase activity
6. Assay for methylphosphotriester-DNA methyltransferase
7. Assay for N\textsuperscript{3}-methyladenine-DNA glycosylase activity
8. Immunological screening method for the detection of transformants synthesizing high cellular levels of Ada protein
9. Transformation of \textit{E. coli}
10. Centrifugation
11. Recombinant DNA techniques
12. Quantitation of DNA
   a. Measurement of OD\textsubscript{260} and OD\textsubscript{280}
   b. Ethidium bromide-mediated fluorescence
13. Purification of DNA
   a. Extraction with phenol
   b. Centrifugation through Sephadex G50 columns
14. Concentration of DNA by ethanol precipitation
15. Isolation of chromosomal DNA from \textit{E. coli} BS strains
16. Isolation of plasmid DNA
   a. Large scale preparations
   b. Small scale preparations
17. Restriction endonuclease digestion of DNA
18. Use of calf intestinal alkaline phosphatase
19. Ligation of DNA
20. End-labelling of DNA fragments
21. Agarose gel electrophoresis
22. Polyacrylamide gel electrophoresis
   a. Non-denaturing gels
   b. Denaturing gels for DNA sequencing
23. Autoradiography
24. Purification of DNA fragments from agarose
and polyacrylamide
25. DNA sequencing by the Maxam-Gilbert chemical cleavage method 7 8
26. DNA sequencing by the dideoxy chain termination method 7 8
   a. Isolation of single-stranded pEMBL DNA 7 9
   b. Denaturation of plasmid DNA using NaOH 7 9
27. In vitro DNA-directed protein synthesis 8 0
28. SDS-polyacrylamide gel electrophoresis 8 0

CHAPTER 3 - EXPERIMENTAL RESULTS

A. Four E. coli mutants constitutively express the adaptive response in the absence of alkylation damage 8 3

B. The ada genes of the constitutive mutants could not be cloned using a high copy number vector 8 4

C. Molecular cloning of the mutant ada genes using a low copy number vector 8 7

D. Determination of the nucleotide sequences of the cloned ada genes 9 2

E. The mutant Ada proteins are strong activators of ada gene expression 9 9

F. The mutant Ada proteins repair methylphosphotriesters in DNA 1 0 9

G. A single amino acid substitution converts the Ada protein into a strong activator of expression of the ada and alkA genes 1 1 3

H. E. coli containing plasmids pSHR13 and pSHR41 do not exhibit an increased resistance to killing by an alkylating agent 1 1 9

CHAPTER 4 - DISCUSSION

Discussion of experimental results 1 2 3
The Adaptive Response to Alkylation Damage: constitutive expression through a mutation in the coding region of the ada gene.
J. Biol. Chem., in press.
<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1: Schematic model for methyl-directed mismatch repair</td>
<td>16</td>
</tr>
<tr>
<td>1.2: Schematic model for base excision repair</td>
<td>22</td>
</tr>
<tr>
<td>1.3: Schematic model for postreplication repair</td>
<td>26</td>
</tr>
<tr>
<td>1.4: Schematic representation of the model for the regulation of the SOS response</td>
<td>31</td>
</tr>
<tr>
<td>1.5: Structures of the DNA adducts induced by simple alkylating agents</td>
<td>39</td>
</tr>
<tr>
<td>1.6: Nucleotide sequence of the promoter region of the <em>E. coli</em> ada and alkA genes</td>
<td>53</td>
</tr>
<tr>
<td>1.7: Schematic model for the induction of the adaptive response</td>
<td>55</td>
</tr>
<tr>
<td>3.1: Schematic map of pEMBL18[+]</td>
<td>85</td>
</tr>
<tr>
<td>3.2: Schematic map of pHSG415</td>
<td>88</td>
</tr>
<tr>
<td>3.3: Immunological screening of ApR HB101 transformants in the experiment to isolate the <em>ada</em>-21 gene</td>
<td>90</td>
</tr>
<tr>
<td>3.4: Schematic map of pSH11</td>
<td>91</td>
</tr>
<tr>
<td>3.5: Nucleotide sequence of the wild-type <em>ada</em> gene from <em>E. coli</em> F26</td>
<td>94</td>
</tr>
<tr>
<td>3.6: Autoradiograph of a DNA sequencing gel showing the mutation at nucleotide +17 of the <em>ada</em>-11, <em>ada</em>-31 and <em>ada</em>-41 genes</td>
<td>95</td>
</tr>
<tr>
<td>3.7: Autoradiographs of DNA sequencing gels showing the G to A transition mutation at nucleotide +378 of the <em>ada</em>-11, <em>ada</em>-21 and <em>ada</em>-41 genes</td>
<td>97</td>
</tr>
<tr>
<td>3.8: Autoradiograph of a DNA sequencing gel showing the G to A transition mutation at nucleotide +289 of the <em>ada</em>-21 gene</td>
<td>98</td>
</tr>
<tr>
<td>3.9: Summary of the amino acid substitutions in the Ada proteins synthesized by the constitutive mutants</td>
<td>100</td>
</tr>
<tr>
<td>3.10: Schematic map of pSHR41</td>
<td>102</td>
</tr>
<tr>
<td>3.11: Assays of O6-methylguanine-DNA methyltransferase activity in cell extracts of GW7101 strains containing the plasmids pSHR1, pSHR21 and pSHR41 and of the constitutive mutants BS21 and BS41</td>
<td>103</td>
</tr>
<tr>
<td>3.12: Stimulation of plasmid directed Ada protein synthesis by the unmethylated Ada-11 protein</td>
<td>108</td>
</tr>
<tr>
<td>3.13: Assays of methylphosphotriester-DNA methyltransferase activity in the BS constitutive mutants</td>
<td>110</td>
</tr>
<tr>
<td>3.14: Sensitivity of GW7101 strains containing plasmids pSHR1, pSHR12, pSHR13 and pSHR41 to methylmethanesulphonate</td>
<td>120</td>
</tr>
</tbody>
</table>
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td><em>E. coli</em> genes controlled by the SOS regulatory network</td>
<td>34-35</td>
</tr>
<tr>
<td>2.1</td>
<td>Bacterial strains</td>
<td>62</td>
</tr>
<tr>
<td>3.1</td>
<td>O(^6)-methylguanine-DNA methyltransferase activity in cell extracts of the GW7101 transformants containing plasmids pSHR1, pSHR21 and pSHR41 and the original mutants BS21 and BS41</td>
<td>105</td>
</tr>
<tr>
<td>3.2</td>
<td>Methylphosphotriester-DNA methyltransferase activity in cell extracts of the BS constitutive mutants</td>
<td>112</td>
</tr>
<tr>
<td>3.3</td>
<td>O(^6)-methylguanine-DNA methyltransferase activity in cell extracts of the GW7101 transformants containing plasmids pSHR1, pSHR41, pSHR12 and pSHR13</td>
<td>115</td>
</tr>
<tr>
<td>3.4</td>
<td>N3-methyladenine-DNA glycosylase II activity in cell extracts of the BK2012 transformants containing plasmids pSHR1, pSHR12, pSHR13 and pSHR41</td>
<td>117</td>
</tr>
</tbody>
</table>
CHAPTER 1 - GENERAL INTRODUCTION
GENERAL INTRODUCTION

A. DNA Repair and the Maintenance of the Genetic Material

The DNA forming the genetic material of a cell is subject to continuous modification or damage. DNA damage can arise from the spontaneous chemical degradation of the purine and pyrimidine DNA bases, and through the DNA reacting with intermediates of cellular metabolism or with agents of external origin, such as the ultraviolet component of sunlight and chemicals present in the environment. Damage to the DNA poses a serious threat to the survival of the cell since it may result in the induction of mutations which have deleterious effects on cellular metabolism by causing the partial or complete inactivation of an essential gene product. Furthermore, certain DNA lesions themselves are directly cytotoxic as a result of a capacity for inhibiting DNA replication. In order to maintain the integrity of the genetic material, and thereby promote cell survival, cells possess DNA repair pathways that serve to remove potentially mutagenic and cytotoxic DNA lesions.

The organism in which DNA repair processes have been most extensively studied, and are therefore best understood, is the bacterium Escherichia coli. The E. coli cell contains a number of distinct DNA repair pathways that enable it to respond efficiently to the wide range of different lesions that can arise in DNA. The study of the DNA repair pathways of E. coli has proven to be of considerable benefit to researchers analysing DNA repair mechanisms in the cells of higher organisms as it has recently become clear that the basic biochemical mechanisms of DNA repair have been well-conserved during evolution.

One of the major classes of DNA damaging agent present in the environment is the alkylating agents, which transfer methyl, ethyl or larger alkyl groups to sites in the DNA. When E. coli are exposed to simple alkylating agents a DNA repair pathway which acts specifically on alkylation damage is induced. The induction of this DNA repair
pathway results in the bacteria acquiring an increased resistance to the mutagenic and cytotoxic effects of alkylating agents. This cellular response to alkylation damage in \textit{E. coli} is termed the adaptive response to alkylating agents. This thesis describes a study on the molecular mechanism through which the expression of the adaptive response is regulated.

In the next section of this General Introduction, section B, the biochemistry of the major DNA repair pathways of \textit{E. coli} (excluding the adaptive response to alkylating agents) is reviewed in order to highlight the diverse nature of the processes involved in the maintenance of the genetic material. The cellular levels of some of the DNA repair enzymes involved in these pathways are increased under conditions of genotoxic stress. In section C of the Introduction the regulatory mechanisms that control the synthesis of the inducible enzymes are briefly summarized. This is followed in section D with a review of the nature of the DNA damage induced by simple alkylating agents, and in section E with a detailed introduction to the adaptive response to alkylating agents. The General Introduction is completed with a brief summary of the aims and rationale of my own studies on the regulation of the adaptive response.

\section*{B. DNA Repair Pathways of \textit{Escherichia coli}}

The DNA repair processes considered here are mismatch correction, enzymatic photoreversal, excision repair, postreplication repair and error-prone repair (SOS processing).

\subsection*{1. Mismatch correction}

Base pair mismatches occur when non-complementary bases are sited opposite one another in double-stranded DNA. One source of mismatches is erroneous DNA replication. In \textit{E. coli} there is good biochemical and genetic evidence that the high fidelity of DNA replication is maintained through the 3' to 5' exonuclease activity
associated with DNA polymerases I, II and III performing a "proofreading" role and removing incorrect deoxyribonucleotides immediately following their incorporation into DNA (Loeb and Kunkel, 1982). However, misincorporated deoxyribonucleotides do escape this proofreading activity at a finite frequency and a mismatch is generated. A second source of mismatches is the process of genetic recombination between homologous, but non-identical DNA sequences. Furthermore, in *E. coli* K12 strains the internal cytosine in the sequences 5'-CCAGG and 5'-CCTGG is methylated at the C5 position by the enzyme DNA cytosine methylase (May and Hattman, 1975). These sequences are hotspots for mutation due to the spontaneous deamination of 5-methylcytosine to form thymine, which results in the formation of T:G mismatches (Duncan and Miller, 1980). The mechanisms through which mismatches may be generated have been reviewed by Claverys and Lacks (1986).

The concept that mismatches within heteroduplex DNA generated by genetic recombination might be processed or repaired was first proposed by Holliday (1964) to account for the phenomenon of gene conversion in fungi, and by Ephrussi-Taylor and Gray (1966) to explain the low transformation efficiency of certain markers in *Streptococcus pneumoniae* transformation. Subsequently, Wagner and Meselson (1976) suggested that mismatch correction limited to the newly-synthesized strand of replicated DNA could serve to eliminate errors and enhance the fidelity of DNA replication. This hypothesis included the proposal that strand discrimination in *E. coli* might be dependent on the state of methylation of the DNA. It is now known that *E. coli* possesses at least three distinct mismatch repair pathways, one of which is dependent on methylation to direct repair to the newly-synthesized strand of replicated DNA. These pathways have been reviewed by Radman and Wagner (1986), Claverys and Lacks (1986) and Modrich (1989), and are discussed briefly below.

a. Methyl-directed mismatch repair

*E. coli* DNA is methylated at the N6 position of adenine in the sequence d(GATC) by the activity of the *dam* methylase (Marinus and Morris, 1973; Geier and Modrich, 1979). The methylation of newly-synthesized DNA lags behind replication (Lyons and
Schendel, 1984), such that the new strand is transiently unmethylated. It is this phenomenon that permits a specific repair pathway to identify the newly-synthesized DNA strand and repair replication errors. The involvement of adenine methylation in mismatch repair was demonstrated by studies of the effects of dam methylation on repair of mismatches in bacteriophage lambda and f1 DNA heteroduplexes upon transfection into E. coli (Pukkila et al, 1983; Lu et al, 1983). The repair of mismatches in heteroduplexes which were methylated on one strand (hemimethylated DNA heteroduplexes) was found to be biased towards correction of the unmethylated strand, whereas little strand bias was observed in heteroduplexes in which neither strand was methylated. Repair of mismatches in DNA heteroduplexes in which both strands were methylated was reduced compared to that observed in unmethylated and hemimethylated heteroduplexes.

Transfection of E. coli with artificially-constructed DNA heteroduplexes established that methyl-directed mismatch repair requires the products of the mutH, mutL, mutS and mutU (uvrD) genes (Nevers and Spatz, 1975; Rydberg, 1978; Bauer et al, 1981; Schaaper, 1988). Mutations within these genes confer a high spontaneous mutability (Cox, 1976). A mutation in the mutD (dnaQ) gene, which encodes the epsilon subunit of DNA polymerase III containing the 3' to 5' exonuclease activity (Scheuerman and Echols, 1984), also results in a defect in methyl-directed mismatch repair (Schaaper, 1988), indicating that DNA polymerase III has a role in the repair mechanism. All types of mismatch in transfected heteroduplex DNA are not repaired with equal efficiency (Dohet et al, 1985). The G:T, A:C, G:G and A:A mismatches are repaired more efficiently than T:T, C:T and A:G mismatches.

The development of an in vitro system for detecting repair of mismatches in f1 heteroduplexes by E. coli cell extracts (Lu et al, 1983; Lu et al, 1984) proved to be of considerable benefit in the elucidation of the mechanism of repair. In the in vitro system the repair of mismatches was biased towards correction of the unmethylated strand (Lu et al, 1983; Lu et al, 1984), and it was observed that mismatch repair could occur with the d(GATC) sequence sited over one thousand nucleotides away from
the mismatch (Lahue et al., 1987). The cell-free system demonstrated the same
genetic requirements as mismatch repair in vivo, in that extracts prepared from
mutH, mutL, mutS and mutU mutant strains were defective in methyl-directed repair,
but complementation was observed when the extracts were mixed (Lu et al., 1983). In
addition, extracts of ssb mutant strains were defective in in vitro repair, indicating
that single-strand binding protein may play a role in the repair mechanism (Lu et al.,
1984). The complementation between extracts of mut" strains has permitted the
purification of the MutH (25kDa), MutL (70kDa) and MutS (97kDa) proteins (Su and
Modrich, 1986; Welsh et al., 1987; Grilley et al., 1988). In an in vitro system
containing these three proteins, DNA helicase II (the mutU/uvrD gene product),
single-strand binding protein, DNA polymerase III, a 55kDa stimulatory protein, the
four dNTPs, NAD+ cofactors and ATP, methyl-directed mismatch repair is mediated
with the same efficiency as observed with cell extracts (Modrich, 1989). Attempts
have been made to ascribe specific functions to the Mut proteins. Using the technique of
DNase I footprinting Su et al. (1988) have shown that the MutS protein binds to
mismatches in DNA, and found a clear correlation between the affinity of binding and
the efficiency of repair in vitro. The MutH protein appears to have an endonuclease
activity that cleaves the unmethylated strand of hemimethylated DNA heteroduplexes 5'
to the dG of the d(GATC) sequences (Modrich, 1989). No distinct activity has yet been
assigned to the MutL protein, but Modrich (1989) has postulated that it may function
to direct MutH to incise the unmethylated strand at the d(GATC) sequence following
recognition of a mismatch by MutS protein.

Several models for the mechanism of removal of the mismatch in methyl-directed
mismatch repair have been described. Figure 1.1 illustrates the model proposed by
Modrich and co-workers (Modrich, 1989), which is based on observations of in vitro
repair. The incision of the unmethylated strand at the d(GATC) site by MutH is followed
by excision of a region of the strand incorporating the mismatch and the d(GATC)
sequence. The gap is repaired by 5' to 3' DNA synthesis, and the strand would be closed
by DNA ligase. The essential features of this model are that a single d(GATC) site is
Mismatch Me

Mismatch generated in DNA

Mismatch recognized by MutS protein

MutS

MutH

MutL

Methyl-directed strand cleavage at d(GATC) site by MutH protein

Excision process initiated at d(GATC) site removes lesion

DNA synthesis repairs gap

Figure 1.1: Schematic model for methyl-directed mismatch repair
(Adapted from Modrich, 1989)
involved, and that excision is initiated at the site of incision at the d(GATC) site and proceeds towards the mismatch. For such a mechanism to operate there is a requirement for the repair system to be able to recognize on which side of the incision the mismatch is sited. A related model has been proposed by Radman and co-workers (Langle-Rouault et al, 1987), who showed that a persistent strand interruption introduced into bacteriophage ØX174 DNA at a d(GATC) site can alleviate the requirement for the mutH gene product in mismatch repair.

b. Methyl-independent mismatch repair

Three mismatch repair processes which are independent of methylation at d(GATC) sites have been identified in E. coli. From studying the in vivo and in vitro repair of fully-methylated pBR322 heteroduplexes containing two mismatches, Fishel and Kolodner (1983,1984) have identified a short patch repair process (repair tracts generally <300 nucleotides), which requires the products of the recF, recJ and ssb genes. A second repair process that specifically repairs T:G mismatches to restore C:G base pairs has been identified by Lieb (1983, 1985) from the analysis of crosses between mutations in the cl gene of bacteriophage lambda. Mutations resulting from a C to T transition at the internal C in the sequences 5'-CC(A/T)GG were found to generate excess numbers of recombinants in crosses with other mutations in the same gene. This was attributed to the 5'-CC(A/T)GG sequences being restored by the action of a very short patch mismatch correction process involving repair tracts of <20 nucleotides (Lieb, 1985). Very short patch repair requires the products of the mutS and mutU genes, and a gene located close to dcm (Lieb, 1987). Its main role may be to repair T:G mismatches generated by spontaneous deamination of 5-methylcytosine at dcm sites. Recently, a third methyl-independent mismatch correction process that specifically repairs A:G mismatches to yield C:G base pairs has been identified (Lu and Chang, 1988; Su et al, 1988). This process is dependent on the mutY gene product (Au et al, 1988).
2. Enzymatic Photoreversal

Ultraviolet light (UV) induces potentially lethal and mutagenic DNA damage. The two major classes of UV-induced DNA adducts (photoproducts) which have been identified are the cyclobutane pyrimidine dimer, which arises through the formation of a cyclobutane ring across the 5-6 double bond of adjacent pyrimidines, and the 6-4'-(pyrimidin-2'-one)-pyrimidine photoproduct, often termed the (6-4)-photoproduct, which results from the formation of a covalent bond between the 6-position of one pyrimidine and the 4-position of an adjacent pyrimidine (see Wang, 1976, for a detailed discussion of the photochemistry of DNA). Chan et al (1985) have employed UV-irradiated phage M13 DNA as a template in polymerization reactions using the Klenov fragment of E. coli DNA polymerase I and found that cyclobutane pyrimidine dimers and (6-4)-photoproducts quantitatively block DNA polymerization. This indicates that both of these lesions are potentially lethal in E. coli. UV induces a wide spectrum of mutations in E. coli including base substitutions, deletions, and frame-shifts (Miller, 1983). For mutagenesis to be induced, the UV-damaged DNA must be processed by an error-prone repair pathway induced during the SOS response, which is discussed later in this Introduction. There is good evidence that the (6-4)-photoproduct is an important premutagenic lesion that can result in mutagenesis when the DNA is subject to error-prone repair (Brash and Haseltine, 1982; Wood, 1985; Franklin and Haseltine, 1986).

In E. coli there are four DNA repair pathways involved in the repair of pyrimidine-pyrimidine photoproducts. The most direct of these is enzymatic photoreversal mediated by DNA photolyase (photoreactivating enzyme). This enzyme catalyses the direct monomerization of cyclobutane pyrimidine dimers in situ in a reaction driven by near-UV or visible light. The DNA photolyase is the product of the phr gene, which has been cloned (Sancar, A. and Rupert, 1978) and sequenced (Sancar, G.B. et al, 1984). The protein has a molecular weight of 54kDa (Sancar, G.B. et al, 1984), and contains two chromophores, a flavin adenine dinucleotide and 5,10-methenyltetrahydrofolate (Sancar, A. and Sancar, G.B., 1984; Johnson et al, 1988).
In vitro, the purified protein binds to UV-irradiated plasmid DNA in the dark and on exposure to light of 300-500nm catalyses the repair of the damage (Sancar, G.B. et al, 1983). It has been postulated that the reaction mechanism involves the light-driven transfer of an electron between the cyclobutane dimer and the chromophores (Sancar, G.B. et al, 1987). The photolyase is unable to repair (6-4)-photoproducts (Brash et al, 1985).

The three other repair pathways involved in processing UV damage are also active in repairing DNA adducts induced by a range of different DNA damaging agents, and these pathways will be discussed in following sections of this Introduction.

3. Excision Repair

A large spectrum of spontaneous and induced DNA lesions are removed from DNA by excision repair. There are two distinct excision repair pathways in E. coli. One pathway, often termed base-excision repair, functions to remove modified bases and is initiated by the action of a DNA glycosylase. The second pathway serves to remove more bulky adducts which may cause some distortion of the DNA helix, and is mediated by a multienzyme complex, the UvrABC excinuclease. This has been termed the nucleotide excision repair pathway.

a. Base excision repair

E. coli contains a class of DNA repair enzymes termed DNA glycosylases which remove damaged or unconventional bases by hydrolytic cleavage of the sugar-base (glycosylic) bond, liberating the base in free form and leaving an apurinic or apyrimidinic (AP) site in the DNA. The enzymes are of low molecular weight, do not require co-factors, and have a narrow substrate specificity, each enzyme repairing a single lesion or a small number of structurally-similar lesions (see Lindahl, 1979; Lindahl, 1982; Friedberg, 1985, for reviews).

The first DNA glycosylase discovered was uracil-DNA glycosylase (Lindahl, 1974). Uracil can arise in DNA through the spontaneous hydrolytic deamination of cytosine, the incorporation of dUMP during DNA replication, or through the
deamination of cytosine by treatment with bisulphite or nitrous acid (see Lindahl, 1979). Uracil mispairs with adenine, and the biological importance of removing this unconventional base is demonstrated by the increased frequency of spontaneous G:C to A:T mutations observed in *E. coli* ung mutants lacking the uracil-DNA glycosylase (Duncan and Weiss, 1982). Adenine in DNA undergoes spontaneous deamination to generate hypoxanthine (Karran and Lindahl, 1980), and *E. coli* contains a hypoxanthine-DNA glycosylase to remove this lesion (Karran and Lindahl, 1978). Oxidative base damage is induced by reactive hydroxyl and oxygen radicals generated in cells either as by-products of aerobic respiration or by exposure to agents such as hydrogen peroxide, bleomycin and ionizing radiation. This damage involves the ring-saturation, ring-opening and ring-fragmentation of pyrimidines, and imidazole ring-opening of purine residues (Cerutti, 1976; Hutchinson, 1985). These lesions are potentially mutagenic (Glickman *et al*, 1980; Levin *et al*, 1982). Two DNA glycosylases of *E. coli* are active in the repair of oxidative base damage. Thymine glycol-DNA glycosylase removes the damaged thymine residues and formamidopyrimidine-DNA glycosylase removes the damaged purine residues from DNA (see Breimer and Lindahl, 1985). Formamidopyrimidine-DNA glycosylase is also active in the repair of an imidazole ring-opened derivative of N7-methylguanine which is generated in DNA exposed to alkylating agents (Chetsanga and Lindahl, 1979). Two further DNA glycosylases required for repair of alkylation damage have been characterized in *E. coli*. N3-methyladenine-DNA glycosylase I (Riazuddin and Lindahl, 1978) removes N3-methyladenine residues, and N3-methyladenine-DNA glycosylase II (Karran *et al*, 1982; Evensen and Seeberg, 1982) removes N3-methyladenine, N3-methylguanine, O2-methylcytosine and O2-methylthymine residues, from alkylated DNA. These two glycosylases will be considered in detail elsewhere.

*E. coli* repairs AP sites generated through the activity of a DNA glycosylase by an excision mechanism initiated by an AP endonuclease, which specifically recognizes an AP site and incises the damaged strand immediately 5' of the lesion (see Lindahl, 1979; Lindahl, 1982; Friedberg, 1985). Endonuclease IV and the endonucleolytic
activity associated with exonuclease III are the 5'-acting AP endonucleases of *E. coli* (Friedberg, 1985; Cunningham *et al*, 1986). Following incision by the AP endonuclease, the deoxyribose-phosphate residue is then excised and the resulting gap in the DNA is filled by DNA polymerase I and the strand is sealed by DNA ligase (Friedberg, 1985). Franklin and Lindahl (1988) have discovered an enzyme that catalyzes the release of 2-deoxyribose-5-phosphate from single-strand interruptions in DNA which have a base-free residue on the 5' side. This DNA deoxyribophosphodiesterase activity may be responsible for the removal of the deoxyribose-phosphate residue at an AP site following incision by an AP endonuclease (Franklin and Lindahl, 1988). Figure 1.2 summarizes the model for the removal of a damaged/modified base by the base excision repair pathway.

In addition to being generated by the action of a DNA glycosylase, AP sites can arise in DNA through spontaneous depurination (Lindahl and Nyberg, 1972), or through specific chemical modifications of the DNA bases resulting in depurination or depyrimidination by the labilization of the glycosylic bond (Singer *et al*, 1978; Drinkwater *et al*, 1980). These sites are also repaired by the mechanism involving an AP endonuclease described above. The efficient repair of AP sites is important since they are potentially mutagenic lesions (Loeb, 1985).

b. Nucleotide excision repair

The repair of cyclobutane pyrimidine dimers by an excision repair process was first demonstrated by Setlow and Carrier (1964) and by Boyce and Howard-Flanders (1964). The genetic analysis of *E. coli* mutants defective in excision repair of pyrimidine dimers revealed that the process involved the products of three genes, *uvrA*, *uvrB* and *uvrC* (Howard-Flanders *et al*, 1966). The cloning of these genes (Pannekoek *et al*, 1978; Sancar and Rupp, 1979; Auerbach and Howard-Flanders, 1979; Yoakum *et al*, 1980), and the identification and purification of the gene products (Sancar, A. *et al*, 1981a; Sancar, A. *et al*, 1981b; Sancar, A. *et al*, 1981c; Yoakum and Grossman, 1981; Thomas *et al*, 1986) has enabled the mechanism of excision repair to be elucidated. Indeed, it is now known that the UvrA, UvrB and UvrC
Uracil residue in DNA

Uracil base removed by Uracil-DNA glycosylase

5'-AP endonuclease incises DNA strand

Enzymatic release of deoxyribose-phosphate

DNA synthesis and ligation restore integrity of DNA

Figure 1.2: Schematic model for base-excision repair
(Adapted from Lindahl, 1979)
proteins associate to form a complex with ATP-dependent endonuclease activity (often termed the UvrABC excinuclease) that hydrolyzes the eighth phosphodiester bond 5’ and the fourth or fifth phosphodiester bond 3’ of the DNA lesion. The lesion is released as part of a 12 or 13 nucleotide-long oligomer, and the gap is filled-in by DNA polymerase I and the DNA strand is closed by DNA ligase (Sancar, A. and Rupp, 1983; Yeung et al, 1983; Friedberg, 1985, Sancar, A. and Sancar, G.B., 1988). The repair process mediated by the excinuclease not only functions to remove cyclobutane pyrimidine dimers, but also repairs (6-4)photoproducts (Sancar and Rupp, 1983; Yeung et al, 1983) and adducts generated by a diverse range of chemical agents including cis-platinum (Husain et al, 1985a), Benzo(a)pyrene diolepoxide (Seeberg et al, 1983), psoralen and acetylaminofluorene (Sancar et al, 1985), and alkylating agents (van Houten and Sancar, 1987; Voigt et al, 1989). To account for this wide-ranging substrate specificity it is believed that the UvrABC excinuclease recognizes a distortion of the DNA helix generated by these “bulky” lesions (Thomas et al, 1986).

The functions of the individual Uvr proteins in the repair process have been extensively studied using in vitro techniques (see Sancar and Sancar, 1988, for a review). The important point that became apparent from these experiments is that the UvrABC excinuclease does not exist free in solution under physiological conditions, but the constituent proteins associate in a sequential manner on the damaged DNA. A number of models have been proposed to describe the events leading to incision of the damaged DNA on the basis of these in vitro data (Seeberg and Steinum, 1983; Husain et al, 1985b; Grossman et al, 1988). In these models the initiating event is the binding of a UvrA protein dimer to the damaged DNA. Interaction of UvrB protein with the UvrA-DNA complex leads to the formation of a stable UvrAB-DNA complex at the site of the DNA lesion. Binding of the UvrC protein to this UvrAB-DNA complex results in the formation of the functional UvrABC excinuclease, which cleaves the eighth phosphodiester bond 5’ and the fourth or fifth phosphodiester bond 3’ of the adduct.

When the repair of UV-irradiated DNA was followed in an in vitro system containing purified UvrA, UvrB and UvrC proteins it was observed that the excinuclease incised
the DNA, but an oligonucleotide did not appear to be released and the excinuclease remained bound to the DNA at the site of incision, indicating that the system lacked important components which function to complete the repair process (Yeung et al, 1983). Kumura et al (1985), Caron et al (1985) and Husain et al (1985b) showed that the simultaneous addition of purified DNA helicase II (uvrD gene product) and DNA polymerase I to such in vitro repair systems results in the turn-over of the UvrABC excinuclease (i.e. its dissociation from DNA) and the release of the oligonucleotide containing the pyrimidine dimer. Furthermore, addition of DNA ligase creates an in vitro excision repair system capable of repairing DNA lesions at a rate approaching that calculated for repair in vivo, leading to the proposal that these six proteins constitute the activities that perform the entire excision repair process in vivo (Husain et al, 1985b). Thus, models describing post-incision events propose that the combined action of DNA helicase II and DNA polymerase I displaces the excinuclease and oligonucleotide, the gap is repaired by DNA polymerase, and the nicked strand sealed by DNA ligase (Husain et al, 1985b; Sancar and Sancar, 1988).

4. Postreplication (Recombinational) Repair

Recombination-deficient E. coli recA mutants are sensitive to UV radiation (Clark and Margulies, 1965), and recA uvrA double mutants are much more sensitive to UV than uvrA or recA mutants (Howard-Flanders and Boyce, 1966). This indicates that recombination processes are involved in the repair of UV-induced DNA damage. To elucidate the nature of the recombination repair pathway Rupp and Howard-Flanders (1968) undertook a study of DNA synthesis in UV-irradiated E. coli. Excision-deficient (uvrA) cells were irradiated with UV and then pulse-labelled with [³H]-thymidine to label the newly-synthesized daughter DNA strands. The size of the daughter DNA was measured immediately after labelling, or after cells had been incubated for 70 minutes, using sedimentation on alkaline sucrose gradients. The daughter DNA analysed immediately after labelling was found to be significantly smaller in size than daughter DNA from control non-irradiated cells. However, after
70 minutes incubation the daughter DNA from irradiated cells was comparable in size to that of control cells. They proposed that the smaller size of the daughter DNA in irradiated cells results from gaps being generated in the daughter strand by DNA synthesis arresting at the site of a pyrimidine dimer and resuming at some distance past the lesion (at the next site of initiation of an Okasaki fragment), and the increase in size of the daughter DNA during incubation after pulse-labelling indicates that these gaps are subsequently repaired. Rupp and Howard-Flanders (1968) proposed a postreplication repair mechanism in which the daughter strand gaps are filled with the corresponding region of DNA derived from the sister DNA duplex by a recombinational exchange. The gaps generated in the sister duplex by this process would be repaired by DNA synthesis. Experimental evidence which demonstrates that daughter strand gaps occur opposite pyrimidine dimers in newly-replicated DNA, and sister-strand exchanges accompany the repair of these gaps (see Hanawalt et al, 1979; Friedberg, 1985), has indicated that this model of postreplication repair is fundamentally correct.

Genetic analysis has determined that a number of gene products, in addition to the RecA protein, are involved in postreplication repair. These include the recB, recC, recF, lexA, uvrD, and ssb gene products (see Hanawalt et al, 1979; Whittier and Chase, 1981). The requirement for these gene products is consistent with the roles that the proteins are known to play in recombination and its regulation (see Smith, 1988). In vitro studies have indicated that the RecA protein participates directly in mediating strand exchange during postreplication repair (West et al, 1981; West et al, 1982; Zivneh and Lehman, 1982). A model for postreplication repair of daughter strand gaps by RecA-mediated strand exchange is illustrated in figure 1.3.

It is important to note that postreplication repair does not remove the primary adduct, the pyrimidine dimer (either cyclobutane pyrimidine dimer or [6-4]-photoproduct), from the DNA. Rather, it repairs the secondary lesions, the daughter strand gaps, generated as a consequence of replication of UV-damaged DNA. However, in repairing the gaps and restoring double-stranded DNA duplexes, the primary adducts
Figure 1.3: Schematic Model of Postreplication Repair  
(From West et al, 1981)

(a) Helical filament of RecA protein, shown as a cylinder, binds to DNA at a daughter strand gap
(b) RecA protein promotes homologous pairing with sister duplex. The intact sister DNA is nicked by a "cutting in trans" activity
(c) RecA protein mediates the transfer of the 3' terminus of the nicked strand into the gap
(d) Reciprocal strand transfer generates a crossed strand exchange (Holliday structure)
(e) Resolution of crossover completes postreplication repair
may subsequently be repaired by nucleotide excision repair. Postreplication repair can therefore be considered to be a mechanism by which cells can temporarily tolerate UV-induced DNA adducts.

5. Error-prone Repair (SOS Processing)

In *E. coli*, mutagenesis by UV light and a wide range of chemical agents including 4-nitroquinoline-1-oxide and methylmethanesulphonate (MMS) results from the activity of a cellular pathway which processes the damaged DNA in such a manner that mutations are induced. This pathway is induced during the SOS response and has been termed error-prone repair or SOS processing (Radman, 1975; Witkin, 1976; Walker, 1984; Walker, 1985). The apparent function of the pathway is to permit DNA replication to proceed past a non-coding DNA adduct, such as a cyclobutane pyrimidine dimer. The SOS response will be discussed in section C1 of this Introduction. However, the genetics and biochemistry of the error-prone repair pathway will be considered here. Error-prone repair requires the products of at least three genes, *umuC*, *umuD* and *recA* (Walker, 1984; Ennis et al, 1985). The *umuC* locus was identified by screening for mutations which render *E. coli* non-mutable by UV or 4-NQO (Kato and Shinoura, 1977; Steinborn, 1978). Mutations at the *umuC* locus were found to lie in two adjacent genes, *umuD* and *umuC*, which form an operon and encode proteins of molecular weight 16kDa and 45kDa, respectively (Elledge and Walker, 1983; Shinagawa et al, 1983). The *umuDC* operon is induced by UV and agents including 4-NQO, MMS and mitomycin C, and is repressed by the LexA protein (Bagg et al, 1981; Elledge and Walker, 1983; Shinagawa et al, 1983). The biochemical functions of the UmuD and UmuC proteins have not been determined. Some naturally-occurring plasmids carry analogues of *umuD* and *umuC* (Walker, 1984). The most studied of these is pKM101, which is able to suppress the non-mutability of *umuDC* mutants (Walker and Dobson, 1979). The plasmid carries two genes, *mucA* and *mucB*, which are organized into a LexA-repressible operon, and encode proteins of molecular
weight 16 and 45kDa, respectively (Perry and Walker, 1982; Elledge and Walker, 1983).

Genetic evidence indicates that activated RecA protein plays a second role in error-prone repair besides mediating the cleavage of the LexA protein that is required for the induction of the umuDC operon (Witkin and Kogama, 1984; Ennis et al., 1985). Perry et al. (1985) reported that the UmuD protein shows a 30% amino acid sequence homology to the carboxy-terminal region of the LexA protein, which contains the site at which RecA-mediated cleavage occurs. On the basis of this homology they suggested that the second requirement for RecA protein might be to interact with, or cleave, the UmuD protein. It has since been shown that activated RecA protein does mediate the cleavage of the UmuD protein in vitro (Burckhardt et al., 1988) and in vivo (Shinagawa et al., 1988), and this cleavage activates the UmuD protein for its role in error-prone repair (Nohmi et al., 1988). These experiments demonstrate that the second role of the RecA protein is, at least in part, the activation of UmuD protein.

Very recently, Woodgate et al. (1989) have found that the UmuC and RecA-activated UmuD (UmuD') proteins associate strongly in vivo and in vitro, and have suggested that the UmuC-UmuD' protein complex is the agent that mediates error-prone repair.

The biochemical mechanism of error-prone repair has not been defined. However, on the basis of the recent observations on the activation of UmuD and its interaction with UmuC, Woodgate et al. (1989) have presented a model in which the UmuC-UmuD' and RecA proteins form a complex with the DNA polymerase III holoenzyme which has stalled at a non-coding DNA lesion. The interaction of these proteins with the DNA polymerase is proposed to allow the enzyme to incorporate a nucleotide opposite the non-coding DNA lesion, thereby allowing replication to continue. It is suggested that the interaction of UmuC-UmuD' and RecA proteins with the DNA polymerase may inhibit the editing exonuclease activity of the enzyme, modify the enzyme such that the requirements for precise base-pairing are relaxed, or induce a change in the conformation of the DNA template which compensates for the distortion caused by the lesion. Mutagenesis would result from the incorporation of a non-complementary
nucleotide opposite the DNA lesion. Like postreplication repair, error-prone repair
does not remove the primary DNA lesion, but it allows the replication apparatus to
bypass what would otherwise represent a replication-blocking lesion, albeit at the
possible expense of the induction of a mutation.

C. Regulation of the DNA Repair Pathways of *E. coli*

In *E. coli* the expression of a number of the genes encoding proteins involved in
DNA repair processes is increased when the cells are exposed to DNA-damaging agents.
The resulting elevation in the cellular levels of these repair proteins confers an
increased capacity to deal with the DNA damage. Three different regulatory networks of
genes which are induced upon exposure to DNA-damaging agents have been
characterized in *E. coli*. These three regulatory networks are those of the SOS
response, the adaptive response to oxidation damage, and the adaptive response to
alkylating agents.

1. SOS Response

The exposure of *E. coli* to DNA-damaging agents, such as UV radiation, mitomycin
C, MMS, and aflatoxin B1, results in the induction of a diverse set of physiological
phenomena which include an increased capacity for DNA repair, enhanced mutagenesis,
cessation of respiration, induction of prophages, and filamentation of cells. Together,
these phenomena serve to promote cell survival, and they collectively constitute what
is termed the SOS response of *E. coli* (see Little and Mount, 1982; Walker, 1984;
Walker, 1985; Ossanna et al, 1986, for detailed reviews).

The SOS response results from the induced expression of a coordinately regulated
set of at least 18 chromosomal genes. The mechanism by which the expression of these
genes is regulated has been defined from extensive genetic analysis of the two
regulatory genes, *recA* and *lexA*, and biochemical analysis of the RecA and LexA
proteins (refer to reviews listed above). The current model for the regulation of the response is summarized diagrammatically in figure 1.4. In an uninduced cell the product of the \textit{lexA} gene acts as a repressor of the SOS-inducible genes by binding to nucleotide sequences in the promoter regions of these genes. However, a number of the SOS genes are expressed at significant levels even in the repressed state. When the DNA is damaged, or DNA replication is inhibited, an inducing signal is generated which activates the RecA protein. The interaction of the activated RecA protein with the LexA protein promotes the autodigestion of LexA to generate two polypeptides which are inactive as repressors. As the cellular levels of intact LexA protein decrease, the SOS genes are derepressed and the SOS response is induced. The LexA protein shows a varying affinity for binding to the promoter regions of the SOS genes, so that it binds more tightly to some gene promoters than others. The SOS genes which bind LexA protein relatively weakly are the first to be expressed at high levels, whereas those that bind LexA tightly require the inducing signal to be maintained and the cellular levels of intact LexA protein to decline to very low levels in order to be expressed fully. As the SOS response enables the cell to recover from the DNA-damaging treatment the inducing signal stops being generated, and the RecA protein is no longer activated. The cellular levels of intact LexA protein consequently increase, and the SOS genes become repressed once again.

One of the most important concepts of the SOS regulatory model is that RecA protein must be activated in order for it to promote the cleavage of the LexA protein and induce the SOS response. Evidence that supports this proposal includes the observation that a number of SOS functions, such as prophage induction, are not induced merely by the presence of high cellular levels of RecA protein (Uhlin and Clark, 1981). The nature of the inducing signal that leads to the activation of the RecA protein has not been defined, although a number of proposals have been made on the basis of \textit{in vitro} and \textit{in vivo} observations. RecA protein is activated for repressor cleavage \textit{in vitro} when it forms a complex with single-stranded DNA and a nucleoside triphosphate (Craig and Roberts, 1981). Roberts and Devoret (1983) have suggested
Figure 1.4: Schematic representation of the model for the regulation of the SOS response.
that binding of RecA protein to single-stranded DNA, generated where a replication fork is blocked at the site of a DNA lesion, activates the protein in vivo. The observations that the SOS response is constitutively induced at the non-permissive growth temperature in E. coli mutants carrying thermosensitive mutations in genes encoding proteins involved in DNA replication, such as dnaB, dnaE, lig and polA (Casaregola et al, 1982; Roberts and Devoret, 1983), support the hypothesis that the interruption of DNA replication generates the inducing signal. Other alternatives for inducing signals have been proposed. One possibility is that oligonucleotides generated by the enzymatic degradation of damaged DNA may act as an inducing signal (Smith and Oishi, 1978). Alterations in the cellular nucleotide pools generated by DNA damaging treatment may also be involved in the activation of RecA protein (Das and Loeb, 1984). Mutants carrying a recA441(Ts) mutation constitutively express the SOS response at 42°C in the absence of DNA damage (Castellazzi et al, 1972). The RecA protein from these strains is more efficient at interacting with single-stranded DNA and a nucleoside triphosphate in vitro (Phizicky and Roberts, 1981), which has led to the proposal that constitutive expression in recA441 mutants in vivo may be the result of the ability of the mutant protein to be activated by single-stranded regions of DNA generated ahead of replication forks or by oligonucleotides generated during normal DNA metabolism (Phizicky and Roberts, 1981). The situation is further complicated by some evidence that single-strand DNA-binding (SSB) protein may be involved in activation of RecA. Strains carrying ssb mutations exhibit defects in the induction of some SOS functions (Vales et al, 1980; Lieberman and Witkin, 1981), and SSB appears to stimulate the rate of RecA-mediated cleavage of lambda phage repressor in vitro (Resnick and Sussman, 1982).

The SOS response is induced when LexA repressor is proteolytically cleaved to generate two polypeptides that are non-functional as repressors. The induction of lambda prophages during the SOS response results from the cleavage of the phage repressor. Both of these reactions are mediated by activated RecA protein, however, the mechanism of cleavage has not been defined. Little (1984) reported that the LexA
and bacteriophage lambda repressors can undergo specific proteolytic cleavage in vitro in the absence of activated RecA. Incubation of purified LexA or lambda repressor under mild alkaline conditions in the presence of divalent cations resulted in the cleavage of the repressors at the same site that is cleaved by activated RecA protein under neutral conditions. This indicates that the activated RecA protein may not be a protease, but its interaction with the repressor molecule may stimulate a protease activity inherent to LexA so that it autodigests (Little, 1984; Siliaty and Little, 1987).

The chromosomal genes that are known to be controlled by the SOS regulatory network are shown in table 1.1. The table indicates the function(s) of the gene products, and lists the appropriate references that demonstrate that the genes are under SOS regulation. The majority of these genes were identified as being SOS-inducible from gene fusion studies using the Mu d(ApR lac) bacteriophage constructed by Casadaban and Cohen (1979). Kenyon and Walker (1980) screened a set of random Mu d(ApR lac) fusions in the E. coli chromosome for fusions where the expression of β-galactosidase was increased by exposing the cells to DNA-damaging agents, and the increase in expression was inhibited by recA(Def) and lexA(Ind') mutations. Using this procedure a set of five damage-inducible (din) loci were identified, dinA, dinB, dinD, dinE and dinF. The dinE insertion has been mapped to the uvrA gene, but the functions of remaining loci have not been identified (Kenyon and Walker, 1980). The Mu d(ApR lac) phage has subsequently been used to generate insertion mutations in genes suspected of being SOS-inducible. Using this approach it has been demonstrated that the uvrA, uvrB, uvrD, himA, recA, recN, ruv and sulA genes are regulated by the SOS regulatory network (see table 1.1 for references). Gene fusion studies with the lacZ gene have also shown that the expression of the lexA and recQ genes are controlled by the SOS network (Brent and Ptashne, 1980; Irino et al, 1986).

2. Adaptive Response to Oxidation Damage

Single-strand breaks are induced in DNA by reactive hydroxyl or oxygen radicals generated by agents such as hydrogen peroxide (Lesko et al, 1980; Repine et al,
Table 1.1: *E. coli* genes controlled by the SOS regulatory network

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function of gene product</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>dinA</td>
<td>Unknown</td>
<td>Kenyon and Walker, 1980</td>
</tr>
<tr>
<td>dinB</td>
<td>Unknown</td>
<td>As above</td>
</tr>
<tr>
<td>dinD</td>
<td>Unknown</td>
<td>As above</td>
</tr>
<tr>
<td>dinF</td>
<td>Unknown</td>
<td>As above</td>
</tr>
<tr>
<td>himA</td>
<td>Site-specific recombination</td>
<td>Miller <em>et al.</em>, 1981</td>
</tr>
<tr>
<td>lexA</td>
<td>Repressor of SOS genes</td>
<td>Brent and Ptashne, 1980</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Little <em>et al.</em>, 1981</td>
</tr>
<tr>
<td>recA</td>
<td>Cleavage of LexA repressor</td>
<td>Little <em>et al.</em>, 1981</td>
</tr>
<tr>
<td></td>
<td>Role in recombination/recombinational repair</td>
<td>Casaregola <em>et al.</em>, 1982</td>
</tr>
<tr>
<td>recN</td>
<td>Role in recombination/recombinational repair</td>
<td>Lloyd <em>et al.</em>, 1983</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Picksley <em>et al.</em>, 1984</td>
</tr>
<tr>
<td>recQ</td>
<td>Role in recombination/recombinational repair</td>
<td>Irino <em>et al.</em>, 1986</td>
</tr>
<tr>
<td>rpsU-dnaG-rpoD</td>
<td>Macromolecular synthesis</td>
<td>Lupski <em>et al.</em>, 1984</td>
</tr>
<tr>
<td>ruv</td>
<td>Role in recombination/recombinational repair; inhibition of cell division?</td>
<td>Shurvinton and Lloyd, 1982</td>
</tr>
<tr>
<td>sulA (sfa)</td>
<td>Inhibition of cell division</td>
<td>Huisman and D'Ari, 1981</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cole, 1983</td>
</tr>
<tr>
<td>(ssb)</td>
<td>Single-strand DNA binding protein; roles in postreplication and mismatch repair</td>
<td>Brandsma <em>et al.</em>, 1983</td>
</tr>
<tr>
<td></td>
<td>Activation of RecA protein?</td>
<td>Salles <em>et al.</em>, 1983</td>
</tr>
<tr>
<td>Gene</td>
<td>Function</td>
<td>References</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td><em>uvrD</em></td>
<td>DNA helicase II; role in excision, mismatch and postreplication repair</td>
<td>Arthur and Eastlake, 1983, Siegel, 1983</td>
</tr>
</tbody>
</table>
1981), bleomycin (Gilioni et al, 1981) and ionizing radiation (Repine et al, 1981; Hutchinson, 1985). The breaks induced by exposure to bleomycin and ionizing radiation have phosphate or phosphoglycolate residues at the 3' DNA termini (Gilioni et al, 1981; Henner et al, 1983). Such termini cannot support DNA synthesis (Hutchinson, 1985) and so the strand breaks are potentially lethal lesions. Likewise, the strand breaks induced by hydrogen peroxide (H$_2$O$_2$) cannot support in vitro DNA synthesis, indicating that the 3' termini have similar or identical blocking groups (Demple et al, 1986). In E. coli, these blocking groups are removed by exonuclease III and endonuclease IV (Demple et al, 1986; Levin et al, 1988), and mutant strains lacking these enzymes show an increased sensitivity to killing by oxidizing agents (Demple et al, 1983; Cunningham et al, 1986). E. coli recA mutants are more sensitive to killing by H$_2$O$_2$ than wild-type E. coli under anaerobic conditions (Carlsson and Carpenter, 1980), indicating that recombinational repair involving RecA protein may be involved in the repair of oxidative DNA damage.

E. coli cells pretreated with a low, non-toxic concentration of H$_2$O$_2$ prior to a challenge with a high concentration of H$_2$O$_2$ show an increased resistance to killing by the challenge when compared to cells that have not been pretreated (Demple and Halbrook, 1983). The increased resistance to H$_2$O$_2$ is accompanied by an increased resistance to killing by a number of other agents, including ionizing radiation, organic peroxides, heat, and the sulphydryl reagent N-ethylmaleimide (Demple and Halbrook, 1983; Greenberg and Demple, 1986). This phenomenon of inducible resistance to the cytotoxicity of agents which cause oxidative damage to cells has been termed the adaptive response to oxidative damage and appears to result, in part, from an enhanced capacity to repair oxidative DNA damage (Demple and Halbrook, 1983).

The adaptive response induced by H$_2$O$_2$ involves the increased synthesis of at least 30 proteins (VanBogelen et al, 1987). The cellular levels of some of these proteins is also increased in response to other stresses, including heat-shock (VanBogelen et al, 1987). The synthesis of nine of the H$_2$O$_2$-inducible proteins is under the control of the oxyR gene (VanBogelen et al, 1987), which was first identified and studied in
Salmonella typhimurium (Christman et al, 1985). Four of the proteins whose synthesis is regulated by oxyR have been identified as catalase/peroxidase, manganese superoxide dismutase, glutathione reductase and NAD(P)H-dependent alkyl hydroperoxide reductase (Christman et al, 1985), which may increase resistance to killing by destroying peroxides and oxygen radicals before they can damage vital cellular components. None of the enzymes known to be involved in the repair of oxidative DNA damage appear to be significantly induced in E. coli upon exposure to H\textsubscript{2}O\textsubscript{2}. Halbrook and Demple (1983) have found that the levels of exonuclease III and endonuclease III (thymine glycol-DNA glycosylase) are unchanged in response to H\textsubscript{2}O\textsubscript{2}. The cellular level of endonuclease IV is increased up to two-fold in response to exposure to H\textsubscript{2}O\textsubscript{2}, but this increase is dependent on the growth medium employed (Chan and Weiss, 1987). However, a 10 to 20-fold increase in endonuclease IV levels is induced by exposure to paraquat, which generates superoxide radicals (Chan and Weiss, 1987). It is possible that the proteins induced by H\textsubscript{2}O\textsubscript{2} may include DNA repair activities that have not been previously characterized.

The signal(s) that result in the induced synthesis of the proteins involved in the adaptive response to oxidative damage are not well understood. However, adenylated nucleotides have been observed to accumulate in E. coli and S. typhimurium following oxidative stress, and it has been suggested that these may play a role in the induction process (Bochner et al, 1984; VanBogelen et al, 1987).

### D. DNA Damage Induced by Simple Alkylation Agents

The simple monofunctional alkylating agents form a major class of environmental mutagens and carcinogens. These agents include the N-alkyl-N-nitrosoureas, the N-alkyl-N'-nitro-N-nitrosoguanidines, the alkyl alkane sulphonates, and the dialkylsulphates. The alkylation of DNA by these agents has been extensively studied in vitro and in vivo (see Singer, 1975; Singer and Grunberger, 1983, for reviews). For
a specific agent, the spectrum of alkylation products generated in cellular DNA in vivo is generally similar to that formed in double-stranded DNA exposed to the agent in vitro (Singer and Grunberger, 1983). The sites in DNA which are prone to alkylation are the ring nitrogen and exocyclic oxygen atoms of the DNA bases and the oxygen atoms of the phosphate linkages in the DNA backbone. Alkylation at the oxygens of the phosphate linkages generates alkylphosphotriesters of two diastereoisomeric conformations, the R and S stereoisomers. Figure 1.5a illustrates the sites in the DNA bases which may be alkylated in double-stranded DNA, and figure 1.5b shows the structures of the R and S stereoisomeric conformations of the alkylphosphotriester adduct.

The extent of alkylation at a particular site is dependent on the chemical characteristics of the alkylating agent itself (Singer and Grunberger, 1983; Saffhill et al, 1985). The alkylating agents with a greater tendency to alkylate the oxygen atoms of DNA are generally more mutagenic and carcinogenic than agents which preferentially alkylate nitrogen atoms (Singer, 1979; Singer and Grunberger, 1983). With the methylating agents N-methyl-N-nitrosourea (MNU) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) the principal site of alkylation is at the N7 position of guanine, N7-methylguanine constituting 60-70% of the alkylation lesions induced in double-stranded DNA treated with these agents in vitro (Singer, 1979; Singer and Grunberger, 1983). Methylphosphotriesters represent 15-20% of the alkylation lesions in MNU and MNNG-treated DNA, and methylation at the O6 position of guanine and the N3 position of adenine each contribute 5-10% of the total number of lesions. The N3 position of guanine, the N1 and N7 positions of adenine, the O4, O2 and N3 positions of thymine, and the O2 and N3 positions of cytosine are minor sites of methylation with these two agents (Singer, 1975; Singer and Grunberger, 1983). The carcinogenicity of MNU correlates with the formation and persistence of O6-methylguanine in cellular DNA (Kleihues and Margison, 1974; Frei et al, 1978; Cox and Irving, 1979).
a: Sites of the DNA bases which are alkylated

b: Structures of the two alkylphosphotriester isomers
(R represents the alkyl group)

Figure 1.5: Structures of the DNA adducts induced by simple monofunctional alkylating agents.
The O-alkylated bases are potentially mutagenic lesions. O\(^6\)-methylguanine (O\(^6\)-\text{MeGua}) residues in DNA preferentially pair with thymine during DNA replication \textit{in vitro} (Abbott and Saffhill, 1979; Snow \textit{et al}, 1984). Such mispairing can result in the induction of a GC to AT transition mutation at the site of an O\(^6\)-MeGua adduct \textit{in vivo} (Loechler \textit{et al}, 1984). Similarly, O\(^4\)-methylthymine (O\(^4\)-\text{MeThy}) residues in DNA can mispair with guanine during DNA replication (Singer \textit{et al}, 1983). AT to GC transition mutations are generated at the sites of O\(^4\)-MeThy lesions as a consequence of this mispairing (Preston \textit{et al}, 1986). O\(^2\)-methylthymine (O\(^2\)-\text{MeThy}) and O\(^2\)-methylcytosine (O\(^2\)-\text{MeCyt}) are also thought to contribute to alkylation-induced mutagenesis, although they exhibit weaker miscoding properties than O\(^6\)-MeGua and O\(^4\)-MeThy (Singer \textit{et al}, 1983; Singer, 1986).

Direct miscoding by O-alkylated base residues appears to be the principal cause of mutagenesis by alkylating agents. However, some mutations may also result indirectly from the generation of AP sites in DNA (Loeb, 1985; Foster and Davis, 1987). AP sites arise in DNA damaged by methylating agents through the activity of DNA glycosylases (as described previously), or through spontaneous depurination at sites of N\(^7\)-methylguanine (N\(^7\)-MeGua) and N\(^3\)-methyladenine (N\(^3\)-\text{MeAde}) residues, which are labile in DNA (Lawley and Warren, 1976). The mutations induced in \textit{E. coli} by MNNG and MNU are predominantly GC to AT transitions (Burns \textit{et al}, 1987; Richardson \textit{et al}, 1987). This is consistent with mutagenesis occurring at sites of O\(^6\)-MeGua, the promutagenic lesion generated at the highest frequency by these agents.

N\(^3\)-\text{MeAde} is considered to be the major cytotoxic DNA lesion induced by simple methylating agents. It has been proposed that the methyl group protrudes into the minor groove of the DNA helix and will prevent base-pairing (see McCarthy \textit{et al}, 1984). Indeed, this adduct forms a block to DNA replication \textit{in vitro} (Larson \textit{et al}, 1985). Furthermore, \textit{E. coli} mutants lacking one or both of the N\(^3\)-methyladenine-DNA glycosylase activities are sensitized to killing by alkylating agents (Karran \textit{et al}, 1982), consistent with N\(^3\)-\text{MeAde} being a potential cell-killing lesion. In common with other replication-blocking lesions, such as cyclobutane pyrimidine dimers, the
persistence of N3-MeAde adducts in *E. coli* DNA may result in the induction of the SOS response (Costa de Oliveira *et al*, 1986). N3-methylguanine (N3-MeGua) adducts in DNA are structurally analagous to N3-MeAde adducts, with the methyl group protruding into the minor groove of the helix (see McCarthy *et al*, 1984). Although N3-MeGua is only a minor alkylation product, it may nonetheless contribute to the cytotoxic effects of agents such as MNU and MNNG. Another minor alkylation product, N3-methylthymine, inhibits DNA replication *in vitro* (Huff and Topal, 1987), and may therefore contribute to cell killing *in vivo*.

The biological effects of alkylphosphotriesters have not been well defined. These lesions are unlikely to cause misincorporation during DNA replication because the alkylation is not at a site directly required for base-pairing (see Singer, 1986). However, Miller *et al* (1982) have found that the rate of *in vitro* replication of a deoxyribonucleotide template containing a single ethylphosphotriester adduct is slower than that of a control template lacking the adduct. In these experiments template DNA containing an ethylphosphotriester of the R stereoisomeric conformation was replicated by *E. coli* DNA polymerase I at a rate 25% slower than the control template, and template containing the adduct in the S stereoisomeric conformation was replicated 50% slower than the control (Miller *et al*, 1982; Weinfeld *et al*, 1985). This suggests that alkylphosphotriesters may reduce the rate of DNA replication *in vivo*, but not function as complete blocks in the manner of N3-MeAde. Alkylphosphotriesters could reduce the rate of replication through inducing a change in the conformation of the template which affects the binding of DNA polymerase or, alternatively, the neutralization of the negative charge associated with the phosphodiester group by its esterification could affect the binding of polymerase (Miller *et al*, 1982).

In summary, the spectrum of mutations and level of cell killing arising from the exposure of cells to an alkylating agent will depend on a combination of factors, including the relative frequency of alkylation at the various sites in DNA, and the capacity of the cell to repair the potentially mutagenic and cytotoxic adducts. As a final
consideration regarding DNA alkylation damage, it should be noted that the exposure of cells to alkylating agents is not the sole mechanism by which DNA alkylation adducts can be generated in vivo. The DNA may also be nonenzymatically methylated by the intracellular methyl group donor S-adenosyl-L-methionine (Ryderberg and Lindahl, 1982; Barrows and Magee, 1982). It has been proposed that such methylation could result in a background level of mutagenesis in cells (Ryderberg and Lindahl, 1982).

E. The Adaptive Response to Alkylating Agents

During the course of a study on the accumulation of mutations in E. coli exposed to low concentrations of mutagenic agents, Samson and Cairns (1977) observed that cells grown in medium containing a low concentration of MNNG became increasingly resistant to the mutagenic and cytotoxic effects of a subsequent challenge with a high dose of the same agent. This physiological response to MNNG was shown to require de novo protein synthesis and it was suggested that it might result from the induction of an uncharacterized pathway for DNA repair (Samson and Cairns, 1977).

Further studies on the phenomenon, initially referred to as the MNNG-adaptive response (Jeggo et al, 1977), revealed that E. coli which had been pretreated with a low concentration of MNNG were more resistant than untreated cells to a challenge with high concentrations of a number of other alkylating agents, including MNU, methyl methanethiosulphonate (MMS), N-ethyl-N-nitrosourea (ENU), and N-ethyl-N'-nitro-N-nitroso-N-nitrosoguanidine (ENNG). However, the MNNG-pretreated ("MNNG-adapted") cells did not show an enhanced resistance to mutagenesis and killing by UV radiation or 4-NQO (Jeggo et al, 1977). The exposure of E. coli to low concentrations of a number of methylating agents other than MNNG, including MNU and MMS, also resulted in the cells acquiring an increased resistance to mutagenesis and killing by high concentrations of these agents (Jeggo et al, 1977). The observation that recA and lexA mutants exhibited an increased resistance to mutagenesis and killing by a toxic
dose of MNNG when they were pretreated with a low concentration of this agent demonstrated that this adaptive response was not a part of the SOS response (Jeggo et al, 1977). Overall, these cross-reactivity studies by Jeggo and co-workers indicated that the adaptive response is not a phenomenon which only occurs in cells exposed to MNNG. Rather, it is a response induced in *E. coli* upon exposure to a number of simple alkylating agents, and which results in the cells acquiring an increased resistance to the mutagenic and cytotoxic effects of these agents (Jeggo et al, 1977).

It is now known that the adaptive response to alkylating agents involves the increased expression of at least four genes, *ada*, *alkA*, *alkB* and *aidB*. The products of two of these genes, *ada* and *alkA*, have been identified as DNA repair enzymes which repair potentially mutagenic and cytotoxic alkylation adducts. The *ada* gene is also the regulatory gene of the adaptive response. An inducing signal generated when the cell is exposed to an alkylating agent converts the *ada* gene product into a positive regulator of transcription of the inducible genes.

The induction of the synthesis of a DNA repair enzyme in the adaptive response was first indicated from an analysis of the levels of different base adducts in the DNA of MNNG-adapted and non-adapted *E. coli* following a brief challenge with [3H]-MNNG (Schendel and Robins, 1978). The levels of N7-MeGua and N3-MeAde adducts were found to be comparable in the adapted and non-adapted cells. However, the adapted cells had a much lower level of O6-MeGua in the DNA. The reduced level of this adduct in the DNA of adapted cells resulted from it being removed from the DNA at a higher rate than in non-adapted cells, and not due to a lower frequency of formation, indicating that the adapted cells have an increased capacity to repair O6-MeGua (Schendel and Robins, 1978). It had previously been observed that adapted *E. coli* are initially refractory to mutagenesis when exposed to a challenge dose of MNNG, but on prolonged exposure to the agent they begin to accumulate mutations (Schendel et al, 1978). Schendel and Robins (1978) observed that adapted cells began to accumulate O6-MeGua during a prolonged challenge with MNNG, and the time of onset of O6-MeGua accumulation corresponded with the time at which mutations began to accumulate. Two important
concepts were derived from this observation. First, adapted cells are able to protect themselves from the mutagenic effects of MNNG by specifically repairing O\(^6\)-MeGua adducts. Secondly, this repair capacity of the adapted cells becomes exhausted if cells are exposed to a high dose of the DNA damaging agent for a prolonged period of time (Schendel and Robins, 1978). In order to account for this apparent limit in the capacity of adapted cells to repair O\(^6\)-MeGua, Robins and Cairns (1979) proposed that exposure of *E. coli* to alkylating agents induces the synthesis of molecules that can repair O\(^6\)-MeGua, but these molecules are only able to act once. An adapted cell would initially contain a large number of repair molecules that had been synthesized during adaptation, and not consumed in the repair of the small number of O\(^6\)-MeGua adducts generated by the low adapting concentration of MNNG. These molecules would be immediately capable of repairing the O\(^6\)-MeGua adducts generated when the cell is challenged with a high concentration of alkylating agent. However, prolonged exposure to a high concentration of alkylating agent would consume all of these repair molecules. Thereafter, the rate of removal of O\(^6\)-MeGua would be limited by the actual rate of synthesis of the repair molecules in the cell (and will be the same as the rate of repair in non-adapted cells that have been challenged with alkylating agent).

The biochemical mechanism of inducible O\(^6\)-MeGua repair associated with the adaptive response was elucidated following the development of an *in vitro* repair system in which crude cell extracts prepared from adapted *E. coli* could remove radioactively-labelled O\(^6\)-MeGua adducts from a substrate prepared by treating DNA with N-[methyl-\(^3\)H]-nitrosourea (Karran *et al.*, 1979). An unusual aspect of the *in vitro* repair of O\(^6\)-MeGua was that radioactive material was not released in acid-soluble or ethanol-soluble form. This led to the suggestion that O\(^6\)-MeGua adducts are not repaired by a DNA glycosylase or a nuclease, or by demethylation to form methanol (Karran *et al.*, 1979). It was subsequently found that the methyl groups from the O\(^6\)-MeGua in the DNA substrate were enzymatically transferred to cysteine residues in protein, thereby regenerating normal guanine residues in the DNA (Olsson and Lindahl, 1980; Foote *et al.*, 1980). The enzyme that mediated this repair process,
termed the O^6^-methylguanine-DNA methyltransferase, was purified to homogeneity (Demple et al, 1982). This enzyme, molecular weight 19kDa, was shown to be the methyl group acceptor protein itself (Demple et al, 1982). Thus, the O^6^-MeGua-DNA methyltransferase repairs O^6^-MeGua adducts by transferring the methyl group from the methylated base to one of its own cysteine residues. Lindahl et al (1982) demonstrated that the enzyme is suicide-inactivated by its repair of an O^6^-MeGua adduct in vitro, in agreement with the hypothesis of Robins and Cairns (1979) which was based on observations on the repair of O^6^-MeGua in vivo.

The substrate specificity of the methyltransferase has been analysed in vitro. The enzyme is capable of repairing O^6^-ethylguanine, although at a slower rate than O^6^-MeGua (Sedgwick and Lindahl, 1982). O^6^-MeGua in single-stranded DNA is repaired at a rate approximately 1000-fold slower than rate of repair for double-stranded DNA (Lindahl et al, 1982). This may render the single-stranded regions of DNA situated close to replication forks especially susceptible to mutagenesis, and account for the ability of MNNG to induce clusters of mutations at replication forks in E. coli (Guerola et al, 1971). Cell extracts of adapted E. coli were found to be capable of repairing mutagenic O^4^-MeThy adducts in DNA by a methyl transfer reaction accompanied by the appearance of methyl groups in protein (McCarthy et al, 1983; Ahmmed and Laval, 1984). The 19kDa O^6^-MeGua-DNA methyltransferase was shown to be the activity that mediates this reaction, the methyl group being transferred to one of its own cysteine residues (McCarthy et al, 1984). E. coli are therefore able to protect themselves against the mutagenic effects of alkylating agents by the induction of a DNA repair enzyme that repairs both O^6^-MeGua and O^4^-MeThy.

The observation that a functional DNA polymerase I was required for adaptation to the lethal effects of MNNG, but was not needed for adaptation to mutagenesis, first led to the suggestion that these might be separate components of the adaptive response (Jeggo et al, 1977). This hypothesis was proven to be correct when a DNA glycosylase which repairs replication-blocking DNA adducts was found to be induced during the adaptive response. In 1978, Riazuddin and Lindahl characterized a DNA glycosylase in
E. coli that repairs N³-MeAde adducts. Karran et al (1980) later discovered that E. coli tag mutants, which are deficient in this enzyme, still contain a small residual N³-MeAde-DNA glycosylase activity. This residual glycosylase activity was shown to be biochemically distinct from the major N³-MeAde-DNA glycosylase activity described by Riazuddin and Lindahl, and was designated as N³-methyladenine-DNA glycosylase II (Karran et al, 1980; Thomas et al, 1982). In wild-type cells N³-MeAde-DNA glycosylase II constitutes 5-10% of the total N³-MeAde-DNA glycosylase activity, the other 90-95% of activity being contributed by N³-methyladenine-DNA glycosylase I, the designation given to the enzyme identified by Riazuddin and Lindahl (Karran et al, 1980; Thomas et al, 1982). The biological significance of N³-MeAde-DNA glycosylase II became apparent when it was discovered that exposure of E. coli to adapting concentrations of MNNG induced a 10-20 fold increase in the cellular levels of this enzyme (Karran et al, 1980; Evensen and Seeberg, 1982). N³-MeAde-DNA glycosylase II has a wider substrate specificity than glycosylase I. Whereas N³-MeAde-DNA glycosylase I is only active in the repair of N³-MeAde, N³-MeAde-DNA glycosylase II can repair N³-MeAde, N³-MeGua, O²-MeCyt, O²-MeThy, and, at a much lower rate, N⁷-MeGua (Karran et al, 1980; Thomas et al, 1982; McCarthy et al, 1984). The increase in the cellular levels of N³-MeAde-DNA glycosylase II in E. coli exposed to an alkylating agent will result in the bacteria acquiring an increased resistance to killing as a result of their enhanced capacity for repair of N³-MeAde and N³-MeGua.

Yamamoto et al (1978) isolated an E. coli mutant that was specifically sensitive to killing by alkylating agents and unable to reactivate MMS-treated lambda phage. The mutation responsible for this phenotype was mapped to a previously unidentified locus, termed alkA, located between 43' and 44' on the genetic map (Yamamoto et al, 1978). This alkA mutant was shown to be defective in the induction of N³-MeAde-DNA glycosylase II (Evensen and Seeberg, 1982). The alkA gene has been cloned and its nucleotide sequence has been determined (Clarke et al, 1984; Nakabeppu et al, 1984a; Nakabeppu et al, 1984b). From an analysis of the nucleotide sequence it was deduced that the alkA gene product is composed of 282 amino acid residues and has a molecular
weight of 31,400 Da (Nakabeppu et al, 1984b). The deduced amino acid composition and amino-terminal amino acid sequence of the alkA product were found to be identical to that of the purified N³-MeAde-DNA glycosylase II, demonstrating that the glycosylase is the product of the alkA gene (Nakabeppu et al, 1984b).

The identity of the gene encoding the O⁶-MeGua-DNA methyltransferase remained unknown until the gene that regulates the adaptive response was cloned and its product characterized. E. coli mutants unable to induce resistance to both mutagenesis and killing when exposed to an adapting concentration of MNNG were first isolated by Jeggo (1979). The mutations responsible for the phenotype of these ada mutants mapped to 47' on the E. coli genetic map (Sedgwick, 1982). Both the O⁶-MeGua-DNA methyltransferase and N³-MeAde-DNA glycosylase II were found to be non-inducible in these mutants, although these enzymes are still present in the cells at levels comparable to non-adapted wild-type E. coli (Mitra et al, 1982; Evensen and Seeberg, 1982). The phenotype of these mutants indicated that the ada locus is involved in the regulation of the adaptive response. Sedgwick (1983) cloned the ada+ gene from E. coli F26, an E. coli B strain, by ligating fragments of chromosomal DNA into a cosmid vector, transforming ada mutants with the hybrid cosmids, and screening for transformants that exhibited resistance to mutagenesis and killing by MNNG. E. coli ada mutants or wild-type strains containing the cloned ada gene on a multicopy plasmid were found to synthesize high cellular levels of O⁶-MeGua-DNA methyltransferase and N³-MeAde-DNA glycosylase II in the absence of alkylation damage (Sedgwick, 1983). On the basis of this observation, Sedgwick (1983) proposed that the ada gene encoded a positive regulator of the adaptive response.

The product of the ada gene was identified as a 39kDa protein (Teo et al, 1984). Unexpectedly, antibodies raised against homogeneous 19kDa O⁶-MeGua-DNA methyltransferase were found to cross-react with the 39kDa Ada protein. In addition, synthetic oligonucleotides corresponding to known amino acid sequences of the methyltransferase specifically hybridized with plasmid DNA containing the ada gene (Teo et al, 1984). These results indicated that the ada gene encoded the O⁶-MeGua-DNA
methyltransferase. Indeed, Teo et al (1984) showed that the 39kDa Ada protein is susceptible to proteolysis in cell extracts, the protein being cleaved to generate a 19kDa polypeptide which corresponds to the O^6-MeGua-DNA methyltransferase originally purified by Demple et al (1982). The 39kDa and the 19kDa forms of the Ada protein repair O^6-MeGua in DNA in an indistinguishable manner (Teo et al, 1984).

Subsequent to the cloning of the ada gene from E. coli B, and the discovery that the inducible O^6-MeGua-DNA methyltransferase is a function of the Ada protein, the ada gene was also cloned from E. coli K12 strains (LeMotte and Walker, 1985; Margison et al, 1985; Nakabeppu et al, 1985a). The complete nucleotide sequences of the ada genes of E. coli B and E. coli K12 have been determined (Demple et al, 1985; Nakabeppu et al, 1985b). The gene from both strains encodes a protein composed of 354 amino acids. However, there are 25 nucleotide differences between the ada gene of the two strains, which will result in the Ada proteins differing at 6 amino acid residues. The physiological significance of these differences is unknown, although it has been reported that E. coli B show a greater increase in resistance to alkylation damage when the adaptive response is induced (Sedgwick and Robins, 1980). By comparing the known amino acid sequences of fragments of the 19kDa O^6-MeGua-DNA methyltransferase with the deduced amino acid sequence of the Ada protein, Demple et al (1985) showed that the O^6-MeGua-DNA methyltransferase activity comprises the C-terminal half of the Ada protein and identified the site of proteolytic cleavage which generates the 19kDa polypeptide as the bond between lysine\textsuperscript{178} and glutamine\textsuperscript{179}. Furthermore, the active cysteine residue that accepts a methyl group from O^6-MeGua or O^4-MeThy was identified as the cysteine at amino acid 321 of the Ada protein (Demple et al 1985).

A further DNA repair activity is associated with the 39kDa Ada protein. McCarthy et al (1983) discovered that a methyltransferase activity which repairs methylphosphotriesters is induced during the adaptive response. This repair activity was shown to be a function of the Ada protein, the methyl group of a
methylphosphotriester adduct being irreversibly transferred to a cysteine residue in
the N-terminal region of the Ada protein (McCarthy and Lindahl, 1985; Margison et
al, 1985). Amino acid sequence analysis of peptides derived from methylated Ada
protein indicated that the active residue is Cys\textsuperscript{69} (Teo et al, 1986; Sedgwick et al,
1988). Interestingly, only methylphosphotriesters of the S-stereoisomeric
conformation are repaired by this mechanism (Hamblin and Potter, 1985; McCarthy

The products of two other genes, \textit{alkB} and \textit{aidB} are also involved in the adaptive
response. Kataoka \textit{et al} (1983) isolated an \textit{E. coli} mutant which exhibited increased
sensitivity to mutagenesis and killing by MMS, but showed wild-type sensitivity to
MNNG. The mutation responsible for this phenotype mapped to a locus, termed \textit{alkB},
situated at 47' on the chromosome (Kataoka \textit{et al}, 1983). The observation that the
\textit{alkB} mutant showed a similar degree of sensitivity to MNNG as wild-type \textit{E. coli}
indicated that the mutation did not lie within the \textit{ada} gene which also maps at 47'. The
\textit{alkB} gene was subsequently cloned (Kataoka and Sekiguchi, 1985), and the nucleotide
sequence has been determined (Kondo \textit{et al}, 1986). From an analysis of the nucleotide
sequence it has been deduced that the \textit{alkB} gene product contains 216 amino acid
residues and has a molecular weight of 23.9kDa (Kondo \textit{et al}, 1986). Comparison of
the \textit{alkB} nucleotide sequence with that of the \textit{ada} gene revealed that the two genes are
immediately adjacent on the chromosome, the final nucleotide of the termination codon
of the \textit{ada} gene also functioning as the first nucleotide of the initiation codon of \textit{alkB}
(Kondo \textit{et al}, 1986). The expression of the \textit{alkB} gene is regulated by the promoter
region of the \textit{ada} gene (Kondo \textit{et al}, 1986), demonstrating that the two genes constitute
an operon. The AlkB protein has been purified (Kondo \textit{et al}, 1986), but its function
has not been characterized. \textit{E. coli} \textit{alkB} mutants exhibit a decreased capacity for
reactivating MMS-treated lambda phage (Kataoka \textit{et al}, 1983), which suggests that
the protein is involved in a DNA repair process.

Volkert and Nguyen (1984) undertook a search for \textit{E. coli} genes which are induced
specifically by alkylating agents by generating a set of random Mu d(Ap\textsuperscript{R} lac) insertion
mutations in the chromosome and screening for mutants in which β-galactosidase activity was induced by exposure to sub-lethal concentrations of MMS but not by UV radiation. Genetic analysis of the positive isolates identified alkylation-inducible genes situated at two loci, which were termed aidA and aidB. The introduction of an ada mutation into aidA or aidB insertion mutants rendered β-galactosidase activity non-inducible by MMS, indicating that the inducible genes at these loci are regulated by the ada gene. The observations that the Mu d(ApR lac) insertion mutations in the aidA mutants showed similar linkage to the his operon as alkA mutations, and that aidA mutants were unable to exhibit adaptation to killing by alkylating agents or induce N³-MeAde-DNA glycosylase II, demonstrated that the gene at the aidA locus is alkA (Volkert and Nguyen, 1984). In contrast, aidB mutants were found to exhibit a normal adaptive response to alkylating agents and, in some cases, were more resistant to MNNG than wild-type E. coli (Volkert and Nguyen, 1984). The phenotype of the mutants, and the position of the aidB locus at 95' on the chromosome, indicated that the inducible gene at this locus had not been previously characterized. Indeed, the specific function of the aidB gene in the adaptive response has not yet been determined.

The molecular mechanism through which the ada gene controls the induction of the adaptive response has only recently been elucidated. As described previously, Sedgwick (1983) proposed that the ada gene encoded a positive regulator of the adaptive response after observing that E. coli containing a multicopy plasmid carrying the ada gene produce considerable amounts of O⁶-MeGua-DNA methyltransferase and N³-MeAde-DNA glycosylase II in the absence of alkylation damage. A series of experiments by Nakabeppu et al (1984b) and LeMotte and Walker (1985) using alkA-lacZ and ada-lacZ gene fusions established that this hypothesis was correct.

Nakabeppu et al (1984b) constructed a plasmid carrying an alkA-lacZ fusion. E. coli ada⁺ cells containing this plasmid expressed a low level of β-galactosidase in the absence of alkylating agents. Exposure of these cells to a low concentration of MNNG resulted in a large increase in the β-galactosidase levels. However, when E. coli ada mutants harbouring the same plasmid were exposed to MNNG the induction of β-
Galactosidase was inhibited. LeMotte and Walker (1985) constructed a plasmid containing an ada-lacZ fusion and performed a similar series of experiments. *E. coli* ada+ strains carrying this plasmid expressed a low basal level of β-galactosidase in the absence of alkylating agents, and a large induction in activity was observed upon exposure to MNNG. However, no induction of β-galactosidase activity was detected in an ada10::Tn10 insertion mutant carrying the same plasmid when the cells were exposed to MNNG. Thus, the induction of the *ada* and *alkA* genes by MNNG is Ada-dependent. In a further experiment LeMotte and Walker (1985) constructed an *E. coli* strain in which the chromosomal *ada* gene was replaced by the ada-lacZ fusion. This strain expressed a low basal level of β-galactosidase which was not increased by the introduction of a single copy of the *ada* gene. However, when the cells were transformed with a multicopy plasmid carrying the *ada* gene the basal level of β-galactosidase activity increased approximately 10-fold, and this level of expression was increased even further when the cells were exposed to MNNG. This data is consistent with the observation of Sedgwick (1983) that the high levels of O6-MeGua-DNA methyltransferase and N3-MeAde-DNA glycosylase in *E. coli* containing *ada* on a multicopy plasmid could be further increased by exposing the cells to MNNG. Overall, the studies of LeMotte and Walker and Nakabeppu *et al* indicated that the Ada protein is a weak positive regulator of *ada* and *alkA* gene expression in the absence of alkylating agent, but this activity is greatly enhanced when the cell is exposed to an alkylating agent (Walker, 1985). The regulation of expression by Ada protein was proposed to occur at the level of transcription (LeMotte and Walker, 1985).

In its repair of DNA damage generated by agents such as MNNG the Ada protein is irreversibly methylated. Methylation of the Ada protein was suggested as a possible mechanism through which it might be activated as a positive regulator (LeMotte and Walker, 1985; Nakabeppu *et al*, 1985a). Using the techniques of *in vitro* DNA-directed protein synthesis and *in vitro* run off transcription Teo *et al* (1986) demonstrated that the methylation of Ada protein by its incubation with a methylated DNA substrate converts the protein from a weak to a strong activator of transcription.
of the *ada* and *alkA* genes. Similarly, Nakabeppu and Sekiguchi (1986) observed that methylated Ada protein acted as a strong activator of run-off transcription of the *ada* gene. However, they found that unmethylated and methylated protein were equally effective at promoting *alkA* transcription. The Ada protein can be methylated at two cysteine residues, Cys^{321} in the C-terminal region of the protein by repair of an O^{6}-MeGua or O^{4}-MeThy adduct, and Cys^{69} in the N-terminal region by repair of a methylphosphotriester (Demple *et al*, 1985; McCarthy and Lindahl, 1985; Margison *et al*, 1985; Teo *et al*, 1986). In order to determine whether one or both of these residues must be methylated for the protein to be converted to a transcriptional activator, Teo *et al* (1986) prepared DNA substrates containing either O^{6}-MeGua or methylphosphotriesters as the alkylation adduct and incubated Ada protein with these substrates to generate singly-methylated protein. The singly-methylated forms of Ada protein were examined for their ability to activate transcription of the *ada* gene. Ada protein methylated at Cys^{321} did not act as an activator, whereas protein methylated by repair of a methylphosphotriester was found to be a strong transcriptional activator (Teo *et al*, 1986).

Demple *et al* (1985) and LeMotte and Walker (1985) suggested that Ada protein might activate transcription through its interaction with the promoter regions of the inducible genes and facilitating the binding of RNA polymerase. Teo *et al* (1986) used the technique of DNase I footprinting to examine the ability of the unmethylated and methylated Ada proteins to bind to the promoter regions of the *ada* and *alkA* genes. These studies demonstrated that Ada protein methylated at Cys^{69} binds specifically to the *ada* and *alkA* promoter regions. The Ada protein binding sites in the *ada* and *alkA* promoters contain a common nucleotide sequence, AAANNAAGCGCA. In the *ada* promoter this sequence lies immediately upstream of the putative RNA polymerase binding site, and the sequence overlaps with the putative RNA polymerase binding site in the *alkA* promoter. The nucleotide sequences of the promoter regions of the *ada* and *alkA* genes are shown in figure 1.6. The Ada protein binding site in each promoter is indicated.
**ada** promoter

```
-101                     -35 seq                     -10 seq                     +1
AAGCTTTCCTTGTCAGCGAAAAAAATTAAGCGCAAGATTGTTTTGCTTGATGGTGACCAGCGCCTAAAGCTATCCTTAAGC
```  
Ada binding site

```
start of transcription
```

```
Met
```

**alkA** promoter

```
-81                     -35_sequence                     -10_seq                     +1
CGTCGCGACAACCGGAATATGAAAGCgAAAGCGCAAGCgGATGAAATAACGTTATGCTGAAAGCGGATGAAATAAGGAGATCG
```  
Ada binding site

```
start of transcription
```

```
Met
```

Figure 1.6: Nucleotide sequence of the promoter region of the *E. coli* ada and alkA genes

The sequence of the *ada* promoter is from Demple *et al* (1985) and that of the *alkA* promoter is from Nakabeppu *et al* (1984b). The nucleotides protected against DNase I digestion by methylated Ada protein are shown bold and underlined (data from Teo *et al*, 1986) The nucleotides common to the two Ada binding sites are shown in large lettering. The putative RNA polymerase binding site of the *ada* gene is that proposed by Teo *et al* (1986) and for the *alkA* gene is that proposed by Nakabeppu and Sekiguchi (1986).
Teo et al (1986) and Nakabeppu and Sekiguchi (1986) have proposed essentially identical models for the regulation of the adaptive response based on the results of the in vivo and in vitro studies of the expression of the ada and alkA genes. The model described by Teo et al is summarized schematically in figure 1.7. In the absence of alkylating agents the inducible genes of the response are expressed at a low basal level. Upon exposure to an agent such as MNNG the cellular DNA becomes methylated. The repair of a methylphosphotriester is proposed to induce a change in the conformation of the Ada protein which permits it to bind to the promoter regions of the inducible genes and activate transcription by RNA polymerase. The resulting increase in the cellular levels of the products of these genes confers an increased resistance to mutagenesis and killing. The model described by Nakabeppu and Sekiguchi differs only in one respect. They propose that the methylated Ada protein activates transcription of the ada gene, and the increased cellular level of the unmethylated Ada protein may be sufficient to activate transcription of the alkA gene.

The vast majority of the studies on the biochemistry and regulation of the adaptive response described here have involved the use of methylating agents as the inducing and challenging agents. In vivo evidence indicates that the adaptive response induced by a methylating agent is also able to confer protection against ethylating and propylating agents (Jeggo et al, 1977; Todd and Schendel, 1983). However, ethylating and propylating agents themselves are poor inducers of the adaptive response (Sedgwick and Lindahl, 1982; Nakabeppu et al, 1984b; Otsuka et al, 1985). It seems likely that the repair of ethylphosphotriesters and propylphosphotriesters by Ada protein is either inefficient, or ethylation or propylation of the protein does not convert the protein into a strong activator of transcription. The adaptive response can therefore be considered to be primarily concerned with the repair of DNA damage generated by methylating agents.

A second O6-MeGua-DNA methyltransferase activity distinct from the Ada protein has recently been discovered in E. coli (Potter et al, 1987; Rebeck et al, 1988; Shevell et al, 1988). This activity, associated with a protein of molecular weight
Pro | aid B | Pro | ada | alk B | Pro | alk A | Low level of expression of repair genes in non-induced cells.

Ada protein

N - C

3-methyladenine-DNA glycosylase II

DNA damaged by alkylating agent.
Repair of a methylphosphotriester induces a change in the conformation of Ada protein which converts it into an activator of transcription.

Methylated Ada protein binds to the promoter regions of the inducible genes and stimulates transcription by RNA polymerase.

Elevated levels of the DNA repair enzymes allow the cell to deal with the potentially mutagenic and toxic DNA lesions.

Figure 1.7: Schematic model for the induction of the adaptive response
19kDa, is present in cells at a low constitutive level and is non-inducible by alkylating agents, demonstrating that it is not involved in the adaptive response (Rebeck et al, 1988; Shevell et al, 1988). It has been proposed that this second methyltransferase may function to protect cells from low levels of alkylation damage produced by endogenous agents such as S-adenosyl-methionine (Shevell et al, 1988).

A number of other microorganisms have been screened for an ability to induce an adaptive response to MNNG. Those species which exhibit an adaptive response include Bacillus subtilis (Hadden et al, 1983; Morohoshi and Munakata, 1983), Micrococcus luteus (Ather et al, 1984) and Streptomyces fradiae (Baltz and Stonesifer, 1985). Furthermore, B. subtilis and M. luteus contain an O\(^6\)-MeGua-DNA methyltransferase activity which is induced during the adaptive response in these organisms. In contrast, Haemophilus influenzae (Kimball, 1980), Neisseria gonorrhoeae (Campbell and Yasbin, 1984), Saccharomyces cerevisiae (Maga and McEntee, 1985) and Salmonella typhimurium (Guttenplan and Milstein, 1982) do not exhibit an adaptive response to MNNG. However, S. typhimurium contains a constitutive repair activity that removes O\(^6\)-MeGua from DNA (Guttenplan and Milstein, 1982).

Mammalian cells contain an O\(^6\)-MeGua-DNA methyltransferase which, like the E. coli Ada protein, removes the methyl group from the alkylated base and transfers it to one of its own cysteine residues (Bogden et al, 1981; Mehta et al, 1981; Pegg et al, 1982; Harris et al, 1983; Hora et al, 1983; Pegg et al, 1983; Yarosh et al, 1984). The mammalian methyltransferase is smaller in size than the Ada protein, estimates of the molecular weight of the mammalian enzyme purified from mouse, rat and human tissues range from 20-24kDa. Unlike the Ada protein, the mammalian enzyme is unable to repair methylphosphotriesters in DNA (Pegg et al, 1983). An issue that has been the subject of considerable debate is whether the mammalian enzyme is specifically inducible by alkylating agents. Early animal studies demonstrated that the chronic treatment of rats with the alkylating agent dimethylnitrosamine (Montesano et al, 1979) or diethylnitrosamine (Margison et al, 1979) induces a 2-3 fold increase in the capacity of the liver cells to remove O\(^6\)-MeGua from their DNA. However,
partial hepatectomy (Pegg and Perry, 1981) and chronic treatment of rats with a wide variety of non-alkylating carcinogens and other agents, such as acetylaminofluorene (Cooper et al, 1982) and phenobarbital (Den Engelse et al, 1986), are now known to induce a similar increase in the capacity of the hepatic cells to repair O\textsuperscript{6}-MeGua (see Saffhill et al, 1985, for a review). Many of these agents, including dimethylnitrosamine and diethylnitrosamine, are hepatotoxic and the increased repair capacity may be related to the restorative proliferation of the liver cells. Significantly, the inducibility of O\textsuperscript{6}-MeGua repair in hepatic cells is limited to the rat. The treatment of mice and gerbils with alkylating or non-alkylating carcinogens, or with partial hepatectomy, does not result in the liver cells acquiring an increased capacity to remove O\textsuperscript{6}-MeGua from DNA (Maru et al, 1982; Bamborschke et al, 1983).

The inducibility of the O\textsuperscript{6}-MeGua-DNA methyltransferase in cultured mammalian cells treated with low concentrations of alkylating agents has been examined (see Frosina and Abbandandolo, 1984, for a review). Waldstein et al (1982a, 1982b) have reported that the treatment of some human tumour cell lines and HeLa cells with multiple doses of MNNG induces a 2.5 and 3-fold increase in the levels of the methyltransferase, respectively. In addition, Frosina and Laval (1987) have found that a single dose of MNNG induces a 3-fold increase in the levels of the enzyme in rat hepatoma cells. However, other laboratories have not observed increases in the levels of O\textsuperscript{6}-MeGua-DNA methyltransferase in MNNG-treated human fibroblasts (Karran et al, 1982a) or human tumour cells (Yarosh et al, 1984). Whether the small increase in the cellular levels of the methyltransferase observed in some studies is the result of increased gene transcription, or simply a phenomenon resulting from a change in the steady state equilibrium between enzyme synthesis and degradation, for instance, is unknown. However, what is clear is that mammalian cells do not exhibit an increase in the levels of the O\textsuperscript{6}-MeGua-DNA methyltransferase comparable to the >100 fold induction observed in adapted E. coli (see Lindahl et al, 1983).
F. Aims and Rationale of this Study on the Adaptive Response to Alkylating Agents in *Escherichia coli*

Three important points regarding the structure of the Ada protein became apparent when Teo *et al* (1986) and Nakabeppu and Sekiguchi (1986) elucidated the mechanism through which the Ada protein mediates the induction of the adaptive response. First, the Ada protein will contain a DNA binding region which interacts with the promoter regions of the inducible genes. Second, assuming that the protein acts through stimulating RNA polymerase to initiate transcription, the protein is likely to contain a region specifically involved in interacting with RNA polymerase at the promoter. Finally, the self-methylation of the Ada protein by repair of a methylphosphotriester must induce some change in the structure of the protein that permits the protein to bind to the promoter region and interact with RNA polymerase. However, since detailed structural studies of the Ada protein had never been performed, the regions of the protein likely to be involved in its role as a regulator of gene expression could not be predicted. Furthermore, although the DNase I footprinting studies of Teo *et al* (1986) indicated the region of the *ada* and *alkA* promoters with which the methylated Ada protein interacts, the critical nucleotide residues within these regions that make contact with the protein were unknown. Similarly, the RNA polymerase binding sequences within the *ada* and *alkA* promoters had not been absolutely defined, although the likely sequences had been indicated (Teo *et al*, 1986; Nakabeppu and Sekiguchi, 1986).

In our laboratory Sedgwick and Robins (1980) had previously isolated four mutant strains of *E. coli* which express the adaptive response constitutively in the absence of alkylating agents. The mutations responsible for the phenotype of these mutants were mapped to the region of the chromosome containing the *ada* gene (Sedgwick, 1982; Sedgwick, personal communication). It was considered that the characterization of the mutations responsible for the constitutive expression of the
adaptive response in these strains might provide an excellent means of defining the
critical nucleotide residues within the ada promoter required for the binding of Ada
protein or RNA polymerase, or equally, might result in the identification of the
regions of the Ada protein involved in its role as a positive regulator of gene
expression.

The initial aim of my study on the regulation of the adaptive response was to
isolate the ada gene from each of the four mutants and determine the nucleotide
sequence changes that result in the constitutive expression of the response. The results
of nucleotide sequence analysis would then provide the basis for a detailed analysis of
the ada promoter region using techniques such as site-directed mutagenesis, or
alternatively, might lead to structural studies of the Ada protein to determine the
functions of the regions implicated as being important in its regulatory role by the
presence of amino acid substitutions in the Ada proteins synthesized by the constitutive
mutants. Either of these approaches would enable us to gain a new insight into the
regulation of the adaptive response by the Ada protein.
CHAPTER 2 - MATERIALS AND METHODS
MATERIALS AND METHODS

A. Materials

1. Bacterial strains

The *E. coli* strains used in this study are described in table 2.1

2. Plasmids

The high copy number cloning vectors pEMBL8[-] and pEMBL18[+] (Dente *et al*, 1983) were laboratory stocks. The low copy number vector pHSG415 (Hashimoto-Gotoh *et al*, 1981) was obtained from Dr. S. Picksley. Plasmids pCS42 (Sedgwick, 1983) and pCS70 (Teo *et al*, 1984), which contain the *ada* gene, were obtained from Dr. B. Sedgwick.

3. Bacteriophage strains

Bacteriophage f1 (Denhardt *et al*, 1975) was used in the procedure to isolate single-stranded pEMBL DNA template for DNA sequencing by the dideoxy method (Dente *et al*, 1983).

4. Media

In the following descriptions of media the contents are given as the amounts required to prepare 1 litre of medium with distilled H2O. LB (Luria-Bertani) broth contained 10g of Difco bacto-tryptone, 5g of Difco bacto-yeast extract and 10g of NaCl. L agar and L top agar contained, in addition to the contents described for LB broth, 15g and 7g of Difco bacto-agar, respectively. LB broth and L agar were supplemented with thymine at a final concentration of 20μg/ml for growth of *E. coli* F26 strains. 2 x TY medium contained 16g of bacto-tryptone, 10g of bacto-yeast extract and 5g of NaCl. The minimal media used were based on M9 minimal salts (Miller, 1972), which contain 6g of NaH2PO4, 3g of KH2PO4, 0.5g of NaCl and 1g of NH4Cl, pH 6.7-7.2. The salts were supplemented with 0.2%(w/v) glucose, 0.1mM CaCl2, 1mM MgSO4,
Table 2.1: Bacterial Strains

<table>
<thead>
<tr>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli K12 Strains</strong></td>
<td></td>
</tr>
<tr>
<td>HB101</td>
<td>F^- hsdS20 (rB m^-B) supE44 ara-14 galK2</td>
</tr>
<tr>
<td></td>
<td>lacY1 proA2 rpsL20 xyl-5 leu myl1 recA13</td>
</tr>
<tr>
<td>DIH101</td>
<td>HB101/F' pro+ lac::Tn5</td>
</tr>
<tr>
<td>GW7101</td>
<td>F^- thr-1 leu-6 proA2 his-4 argE3 lacY1 galK2</td>
</tr>
<tr>
<td></td>
<td>ara-14 xyl5 mtl-1 tsx-33 rpsL31 supE37 Δada25</td>
</tr>
<tr>
<td></td>
<td>Evensen and Seeberg (1982)</td>
</tr>
<tr>
<td><strong>E. coli B Strains</strong></td>
<td></td>
</tr>
<tr>
<td>F26</td>
<td>his thy sulA</td>
</tr>
<tr>
<td>BS11</td>
<td>As F26 but ada-11</td>
</tr>
<tr>
<td>BS21</td>
<td>As F26 but ada-21</td>
</tr>
<tr>
<td>BS31</td>
<td>As F26 but ada-31</td>
</tr>
<tr>
<td>BS41</td>
<td>As F26 but ada-41</td>
</tr>
</tbody>
</table>
1μg/ml thiamine and the required L-amino acids at 10μg/ml. Where necessary, antibiotics were added to media to give the following final concentrations: ampicillin at 50μg/ml for selection of strains carrying high copy-number plasmids, and 20μg/ml to select for strains carrying plasmids derived from the low copy number vector pHSG415; kanamycin at 20μg/ml; chloramphenicol at 10μg/ml for selection of *E. coli* GW7101 strains. Antibiotics were obtained from Sigma Chemical Co.

5. **Radiochemicals**

\[ \alpha^{-32P}\text{-dATP} \text{ (3000Ci/mmol)}, \alpha^{-32P}\text{-dCTP} \text{ (3000Ci/mmol)}, \alpha^{-35S}\text{-dATP} \text{ (>600Ci/mmol)}, \text{ L-}[^{35S}]\text{-methionine} \text{ (>800Ci/mmol)}, \text{ N-[methyl-}^{3}\text{H}\text{-N-nitrosourea} \text{ (23Ci/mmol) and Rainbow coloured [}^{14}\text{C}\text{-methylated protein molecular weight markers were obtained from Amersham International PLC. [methyl-}^{3}\text{H}\text{-dimethyl sulphate} \text{ (1.6Ci/mmol) was obtained from NEN-Du Pont.}}

6. **Molecular biology reagents**

Restriction endonucleases, Klenow fragment of DNA polymerase I and T4 DNA ligase were purchased from New England Biolabs and Gibco BRL. Calf intestinal alkaline phosphatase was obtained from Boehringer Mannheim. RNase A was from Millipore Corp., and DNase I was from Worthington Chemicals. Deoxyribonucleotides were obtained from Sigma Chemical Co. The prokaryotic DNA-directed translation kit and the dideoxy DNA sequencing kit employing the Klenow fragment of DNA polymerase I were from Amersham International PLC. The “Sequenase” DNA sequencing kit was purchased from United States Biochemical Corp. Reagents were stored under the conditions recommended by the manufacturer.

7. **Oligodeoxyribonucleotides**

Oligodeoxyribonucleotide primers for DNA sequencing were synthesized by Mr. I. Goldsmith on an Applied Biosystems 380B DNA synthesizer.
8. Electrophoresis reagents

Agarose and urea were purchased from Gibco BRL. Acrylamide, N,N'-methylene-bis-acrylamide (bis-acrylamide), ammonium persulphate, N,N,N',N'-tetramethyl-ethylenediamine (TEMED) and sodium dodecyl sulphate (SDS) were obtained from Bio-Rad Laboratories.

9. General chemical/biochemical reagents

All other reagents were obtained from BDH Chemicals Ltd., Fisons, Gibco BRL and Sigma Chemical Co., unless otherwise stated. Chemicals were stored as recommended by the manufacturer.

B. Methods

1. Bacterial growth conditions

*E. coli* were routinely grown as cultures in LB broth in environmental incubator shakers (New Brunswick Scientific Co.), or as colonies on L agar plates in laboratory incubators (Gallenkamp). For strains carrying plasmids derived from the temperature-sensitive vector pHSG415 the incubation temperature was 30°C. All other strains were grown at 37°C, unless otherwise stated. Growth of liquid cultures was followed by measuring the optical density at 600nm (OD$_{600}$) in a Cecil Instruments CE292 digital ultraviolet spectrophotometer.

Strains were maintained as frozen stocks of liquid cultures made 15% (v/v) with sterile glycerol and stored at -20°C and -80°C.

2. Preparation of bacteriophage f1 stock suspension

A 50ml culture of *E. coli* DIH101 grown to an OD$_{600}$ of 0.2 in LB broth containing kanamycin was inoculated with bacteriophage f1 at a multiplicity of infection (m.o.i.) of 20. The culture was grown for a further 5 hours, and then the bacterial cells pelleted by centrifugation at 10,000 revolutions per minute (rpm) for 10 minutes.
The supernatant was collected and incubated at 60°C for 30 minutes. 50ml of supernatant was mixed with 12.5ml of 20%(w/v) polyethylene glycol 8000 (PEG), 2.5M NaCl and the suspension placed on ice for 60 minutes. The phage particles were pelleted by centrifugation at 10,000 rpm for 10 minutes, and resuspended in 0.4ml of 10mM Tris.HCl pH 8.0, 1mM EDTA, 10mM NaCl. The stock suspension was titrated and then stored at 4°C.

For titrating the f1 suspension 0.1ml volumes of DIH101 culture grown to OD600 of 0.8 were added to 2.5ml aliquots of molten L top agar (45°C). To these suspensions were added 0.1 ml samples of serial 10-fold dilutions of the f1 suspension prepared in M9 salts solution. Each mixture was poured onto a L agar plate, and the phage titre calculated from the numbers of plaques that developed on the plates after overnight incubation at 37°C.

3. Survival curves - sensitivity to killing by methylmethanesulphonate

Cultures were grown in supplemented M9 media to an OD600 of 0.3. 0.5ml volumes of culture were mixed with 0.5ml samples of M9 minimal salts containing methylmethanesulphonate (MMS; Aldrich Chemicals) at 2 x the required final concentration (prepared immediately before use). The cell suspensions were incubated for the required length of exposure at the permissible growth temperature, and then serial 10-fold dilutions of the suspensions prepared using M9 salts solution. 0.1ml samples of the diluted suspensions were spread on L agar plates and incubated to allow surviving cells to grow up into colonies. Percent survival at a specific MMS concentration was calculated from the ratio of the number of survivors relative to the total number of cells plated (determined by spreading samples of control cell suspensions not exposed to MMS).

4. Preparation of crude cell extracts

25ml cultures were grown to mid-exponential growth phase in LB broth. Bacteria were harvested by centrifugation at 5,000 rpm for 5 minutes and washed in 5ml
phosphate buffered saline (PBS; 1% NaCl, 0.025% KCl, 0.14% Na$_2$HPO$_4$, 0.025%
KH$_2$PO$_4$, pH 7.2). The cells were re-pelleted by centrifugation and resuspended in
0.15ml of ice-cold extraction buffer (50mM Tris.HCl pH 8.0, 10mM dithiothreitol,
1mM EDTA). The cells were disrupted by two 2 second ultrasound treatments in an
ultrasonic disintegrator (MSE Soniprep 150), and the extract cleared by
centrifugation for 10 minutes at 12,000 rpm and 4°C. Aliquots of extract were snap-
frozen and stored at -80°C. The protein concentration of extracts was determined by
the method of Bradford (1976), employing a commercial protein assay reagent
(Pierce).

5. Assay for O$_6$-methylguanine-DNA methyltransferase activity

The O$_6$-methylguanine-DNA methyltransferase activity in cell extracts was
assayed using the method described by Demple et al (1983), with minor modifications
to the volume of reaction mixture and incubation time.

An alkylated DNA substrate containing O$_6$-methylguanine as the principal adduct
was prepared by treating Micrococcus luteus DNA with N-[methyl-$^3$H]-N-nitrosourea
using the method of Karran et al (1979). The alkylated substrate was resuspended in
assay buffer (70mM HEPES.KOH pH 7.8, 10mM DTT, 1mM EDTA) to give a reaction
mixture containing 10,000-12,000 cpm/ml. 0.1ml samples of the reaction mixture
were incubated with varying amounts of cell extract, usually 0.25-2 µg of protein,
for 20 minutes at 37°C. 10µl of 10mg/ml carrier DNA solution (herring sperm DNA,
Boehringer Mannheim) and 0.12ml of ice-cold 0.8M trichloroacetic acid (TCA) were
added and the samples chilled on ice for 10 minutes. The samples were centrifuged for
10 minutes at 12,000 rpm and 4°C and the supernatants removed. 0.1ml of 0.1M HCl
was added to the pellets and the substrate hydrolyzed at 70°C for 30 minutes. The
samples were placed back on ice for 5 minutes and then centrifuged for 10 minutes at
12000 rpm and 4°C. Under these conditions the self-methylated Ada protein remains
precipitated on the walls of the reaction tube, and the products of hydrolysis of the DNA
substrate are released into the supernatant. The supernatants were removed and their
radioactivity measured by mixing with 5ml of liquid scintillation solution (Picofluor 15, Packard), and counting in a liquid scintillation counter (Packard Tri-Carb 4000 Series). The amount of radioactivity that has been removed from the DNA substrate and remains precipitated on the walls of the reaction tube, calculated by subtracting the radioactivity of the supernatant from that of control supernatants (from assay mixtures that contained no cell extract), is a measure of the amount of O6-MeGua repaired by O6-MeGua-DNA methyltransferase.

6. Assay for methylphosphotriester-DNA methyltransferase activity

The methylphosphotriester-DNA methyltransferase activity of Ada protein in cell extracts was measured using the assay described by McCarthy and Lindahl (1985), but with modifications to the volume of reaction mixture and length of incubation.

A substrate containing methylphosphotriesters as the major alkylation lesion was prepared by treatment of poly[dT] with N-[methyl-3H]-N-nitrosourea, and then annealing with an equimolar amount of poly[dA] (McCarthy and Lindahl, 1985). The substrate was resuspended in assay buffer (70mM HEPES.KOH pH 7.8, 10mM DTT, 1mM EDTA) to give a reaction mixture containing 27,000 cpm/ml. 0.1ml samples of reaction mixture were incubated with cell extract for 20 minutes at 37°C. 0.1µl of carrier DNA and 0.15ml of ice-cold 0.8M TCA were added and the samples chilled on ice for 5 minutes. Samples were centrifuged for 15 minutes at 12,000 rpm and 4°C and the supernatants removed. 0.1ml of 0.1M Tris.HCl pH 8.0, 10mM EDTA and 20µl of proteinase K solution (BDH Ltd., 25mg/ml in H2O) was added to the pellets, and the samples incubated for 60 minutes at 37°C. 0.12ml of ice-cold 0.8M TCA was added and the samples placed on ice for 5 minutes. The samples were centrifuged for 15 minutes at 12,000 rpm and 4°C, and the supernatants removed and their radioactivity determined. The radioactivity of the supernatant is a measure of the number of methyl groups transferred from the alkylated substrate to a protease-sensitive form, and represents the repair of methylphosphotriesters by the Ada protein.
7. Assay for N³-methyladenine-DNA glycosylase activity

The assay method of Riazuddin and Lindahl (1978) was used to measure the N³-methyladenine-DNA glycosylase activity in cell extracts of *E. coli* BK2012.

A DNA substrate containing N³-[methyl-³H]-adenine was prepared by treating *M. luteus* DNA with [methyl-³H]-dimethyl sulphate (Riazuddin and Lindahl, 1978). The alkylated DNA was resuspended in assay buffer (70mM HEPES.KOH pH 7.8, 1mM DTT, 1mM EDTA) to generate a reaction mixture containing $2.5 \times 10^5$ cpm/ml. 0.1 ml aliquots of reaction mixture were incubated with samples of cell extract containing 1-10μg of protein for 60 minutes at 37°C. Samples were placed on ice and 20μl of a solution containing 1mg/ml carrier DNA, 1M NaCl and 0.3ml of absolute ethanol were added. The samples were vortexed and placed at -20°C for 30 minutes. Following this incubation the samples were centrifuged for 15 minutes at 12,000 rpm and 4°C, and the radioactivity of 0.3ml of supernatant was determined. The radioactivity of the supernatant is a measure of the amount of N³-methyladenine removed from the substrate by the action of N³-MeAde-DNA glycosylase.

8. Immunological screening method for detection of transformants synthesizing high cellular levels of Ada protein

An immunological screening method based on that of Helfman *et al* (1983) was used to identify colonies synthesizing high cellular levels of Ada protein.

The colonies to be screened were replica-plated. One set of plates were kept as master plates. The duplicates were placed at 4°C for 2 hours and then the colonies were lifted onto 82mm nitrocellulose filters (BA85, Schleicher and Schuell). The filters were suspended in a chloroform vapour tank for 60 minutes. Each filter was then placed in 15ml of 50mM Tris.HCl pH 7.5, 150mM NaCl, 5mM MgCl₂ containing 3%(w/v) bovine serum albumin, 1μg/ml DNase I and 40μg/ml lysozyme, and agitated gently on an orbital shaker for 5 hours at room temperature. Filters were washed three times in 50mM Tris.HCl pH 7.4, 200mM NaCl, 0.05% Nonidet P40 (TSN buffer), and then agitated gently for 5 minutes in 20ml of the same buffer containing 10%(v/v) horse serum. Rabbit polyclonal antiserum raised against homogenous
39kDa Ada protein was added at 1000-fold dilution, and the filters gently agitated on a rocker for 12-16 hours at room temperature. The filters were washed three times in TSN buffer and incubated for 2 hours in 20ml of TSN containing 10% horse serum and GAR/IgG(H+L)/PO secondary antibody (horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin, Nordic Immunology). Filters were washed three times in PBS and then incubated for 5-15 minutes in 20ml of peroxidase substrate solution, composed of 0.5mg/ml 4-chloro-1-naphthol, 0.01%(v/v) H2O2 in PBS. A colony synthesizing high levels of Ada protein was indicated by the development of an intense purple colouration on the surface of the filter.

9. Transformation of *E. coli*

*E. coli* were made competent for transformation with DNA using treatment with CaCl2 as described by Maniatis *et al* (1982). Frozen stocks of competent cells were prepared by adding sterile glycerol to the cell suspensions to 15%(v/v) and storing at -80°C in 0.1ml aliquots. For transformation, cells were thawed on ice and DNA added (1-50ng in maximum volume of 10µl TE or ligation buffer). The cells were maintained on ice for 30 minutes and then heat-shocked at 42°C for 1 minute. Cells were placed back on ice for 2 minutes and 0.4ml of LB broth added. The cells were incubated for 90 minutes at 30°C or 37°C with shaking to allow expression of antibiotic resistance, and then spread on selective L agar plates.

10. Centrifugation

Samples of <1.5ml volume were centrifuged in microcentrifuges (Eppendorf or Hettich). Samples of larger volume were routinely centrifuged in a Sorvall RC-5B refrigerated centrifuge (Du Pont Instruments), using an HB4 rotor with 15ml and 30ml Corex tubes, and GSA and GS3 rotors with 250ml and 450ml polypropylene centrifuge buckets.
11. Recombinant DNA techniques

Gloves were worn at all times during experimental procedures. Reactions were routinely performed in 1.5ml Eppendorf tubes, and tubes and micropipette tips were sterilized by autoclaving prior to use. DNA solutions of large volume were handled in 15ml and 30ml Corex tubes which had been oven-sterilized and treated with dimethyldichlorosilane (Repelcote, BDH Chemicals). Sterile distilled H₂O was used in the preparation of all buffers and solutions. Where possible, buffers and solutions were sterilized by autoclaving or filtration.

12. Quantitation of DNA

a. Measurement of OD₂₆₀ and OD₂₈₀

The optical density of the DNA solution at 260nm and 280nm was measured (Perkin Elmer Lambda UV5 spectrophotometer). The reading at 260nm was used to calculate the DNA concentration. An OD₂₆₀ of 1 corresponds to a concentration of 50µg/ml for double-stranded DNA. The purity of DNA was determined from the ratio OD₂₆₀/OD₂₈₀. A value of 1.8 corresponds to pure DNA, and significantly smaller values indicate possible contamination by protein or phenol. Purified DNA used in this work had OD₂₆₀/OD₂₈₀ of 1.7-1.9.

b. Ethidium bromide-mediated fluorescence

When the concentration of DNA was too low to determine using OD₂₆₀, it was estimated by spotting equal volumes of the solution and reference solutions of known DNA concentration onto a 1%(w/v) agarose plate containing 1µg/ml ethidium bromide (EtBr), and comparing the intensity of fluorescence of the unknown solution with the reference samples when the plate was placed on a UV transilluminator.

13. Purification of DNA

a. Extraction with phenol

Proteins were removed from DNA solutions by extracting once with an equal volume of buffered phenol, once with an equal volume of a 1:1 mixture of
phenol/chloroform, and once with chloroform. The phenol used was re-distilled analytical or chromatography grade, buffered by extraction three times with an equal volume of 1M Tris.HCl pH8.0 and once with 0.1M Tris.HCl pH8.0, 0.2%(w/v) β-mercaptoethanol. The chloroform used was a 24:1 mixture of chloroform-isooamyl alcohol. Following extraction the DNA was ethanol-precipitated to remove any traces of organic solvent.

b. Centrifugation through Sephadex G50 spun columns

End-labelled DNA fragments were purified from unincorporated dNTPs by centrifugation through 1ml spun columns of Sephadex G50 (Maniatis et al, 1982). This technique was also used to change the buffer in which DNA was dissolved during procedures involving the sequential modification of DNA by enzymes with different buffer requirements. The maximum sample volume loaded on the columns was 0.1ml.

14. Concentration of DNA by ethanol precipitation

The DNA solution was made to either 0.25M sodium acetate by addition of 3M sodium acetate pH 5.2, or 0.1M NaCl by addition of 5M NaCl. 2 volumes of absolute ethanol were added, mixed, and the sample incubated at -20°C for >60 minutes or -70°C for 5-10 minutes. The solution was centrifuged for 10-15 minutes at 10,000-12,000 rpm and 4°C to pellet the precipitated DNA. The supernatant was removed and the pellet washed in 70% ethanol. Following re-centrifugation the 70% ethanol was removed and the DNA pellet dried briefly under vacuum to remove all traces of ethanol. The DNA was resuspended in the required buffer.

15. Isolation of chromosomal DNA from E. coli BS strains

Cells from a 50ml overnight culture were harvested by centrifugation and washed twice in 10ml of STE buffer (20mM Tris.HCl pH 7.5, 5mM EDTA, 100mM NaCl). The cells were resuspended in 5ml of STE containing 0.5mg/ml of lysozyme and 12.5μg/ml of RNase A, and incubated for 30 minutes at 37°C (stock solutions of 10mg/ml RNase A in 10mM Tris.HCl pH 7.5, 15mM NaCl were prepared and stored
according to Maniatis et al, 1982). The cells were lysed by addition of 5ml of STE containing 1%(w/v) SDS and 0.2mg/ml of proteinase k (BDH Chemicals) and incubating for 60 minutes at 60°C. 10ml of phenol was added and the aqueous and organic phases mixed for 30 minutes by gently inverting the tube containing the solution. The phases were separated by centrifugation at 10,000 rpm for 60 minutes and the aqueous phase collected. The solution was extracted once again with phenol, once with phenol/chloroform, and once with chloroform, but with 5 minutes mixing of phases and 5 minutes centrifugation at 5,000 rpm in each case. The DNA was ethanol-precipitated (overnight incubation at -20°C), and collected by spooling onto the tip of a sterile Pasteur pipette. After washing in 70% ethanol, the DNA was dried in air for 30 minutes and dissolved in 4ml of TE pH 8.0. The DNA was stored at 4°C.

16. Isolation of plasmid DNA

Plasmid DNA was extracted from E. coli on both a small and large scale using the alkaline-lysis procedure (Birnboim and Doly, 1979, Ish-Horowicz and Burke, 1981). In large scale preparations the DNA was further purified by caesium chloride-ethidium bromide density gradient ultracentrifugation (Radloff et al, 1967).

a. Large scale preparations.

For isolation of plasmids with a high copy-number, a 500ml culture was grown overnight in LB broth containing a selective antibiotic. Cells were harvested by centrifugation at 5,000 rpm for 10 minutes, resuspended in 20ml of 50mM glucose, 25mM Tris.HCl pH 8.0, 10mM EDTA containing 2mg/ml of lysozyme (solution I), and incubated at room temperature for 10 minutes. 40ml of 0.2N NaOH, 1%(w/v) SDS (solution II) was added, mixed by swirling, and the solution placed on ice for 10 minutes. 30ml of ice-cold 5M potassium acetate pH 4.8 (solution III) was added, mixed by inverting the tubes, and the solution incubated on ice for 10 minutes. The solution was centrifuged for 30 minutes at 11,000 rpm and 4°C to pellet the flocculent precipitate, and the supernatant filtered through four layers of muslin cloth to remove remaining traces of precipitate. The DNA was precipitated by adding 0.6
volumes of propan-2-ol to the supernatant and leaving to stand on ice for 60 minutes. The DNA was recovered by centrifugation at 10,000 rpm for 30 minutes, washed in 70% ethanol, and dried briefly in a vacuum desiccator. The DNA pellets were dissolved in 8ml of TE pH 8.0, and the volume of the solution measured using a graduated plastic pipette. The volume was made up to 10ml with TE, and 11g of CsCl added and allowed to dissolve. 1ml of 10mg/ml ethidium bromide was added and the solution loaded into a single ultracentrifuge tube (Beckman Quickseal, 16x76mm). The gradient was centrifuged at 38,000 rpm and 20°C for 36-48 hours (50Ti fixed-angle rotor on a Beckman L8-55 ultracentrifuge). The covalently closed circular plasmid band was collected into a plastic syringe through a 19 gauge hypodermic needle inserted into the side of the tube (the DNA band was usually visible in ordinary light, but if this was not the case the tube was illuminated with long wavelength UV light). EtBr was removed by extraction with TE-saturated propan-2-ol, and the solution dialysed against three changes of 1 litre of TE to remove the CsCl. The DNA was recovered by ethanol precipitation, redissolved in TE pH 8.0, and stored in aliquots at -20°C.

To isolate pHSG415 and plasmids derived from this low copy-number vector, 5 litres of overnight culture were lysed using a 10-fold scaling-up of the lysis procedure described above. Following propan-2-ol precipitation the DNA was dissolved in 40ml of TE pH 8.0 and extracted once with phenol, phenol/chloroform, and chloroform. The plasmid DNA was recovered by ethanol precipitation and resuspended in 10ml TE. 11ml of solution was mixed with 12g of CsCl, and 1ml of 10mg/ml EtBr added. This solution was centrifuged at 12,000 rpm for 60 minutes at room temperature to pellet proteins. A further 1g of CsCl and 1ml of EtBr was added to the cleared supernatant and the solution loaded into a single ultracentrifuge tube. Centrifugation and recovery of plasmid DNA were performed as described previously.

b. Small scale preparations

For the rapid isolation of plasmid DNA from a number of transformants small scale alkaline-lysis "minipreps" were performed on 1.5ml overnight cultures using
the method described by Maniatis et al (1982). The procedure was routinely used to prepare DNA suitable for restriction analysis from putative recombinants.

17. Restriction endonuclease digestion of DNA

DNA was digested in the reaction buffer recommended by the supplier of the restriction enzyme. Reaction mixtures routinely contained 2-4 units of enzyme per μg of DNA, and were incubated at 37°C (or 30°C for SmaI) for 1-2 hours. Following digestion, a sample of the reaction mixture was removed and the reaction products analysed by agarose gel electrophoresis. The DNA samples were loaded onto agarose gels in gel loading buffer (5 x loading buffer contained 17.5% Ficoll 400, 0.1% bromophenol blue, 100mM Tris.HCl pH 8.0, 20mM EDTA).

18. Use of calf intestinal alkaline phosphatase

To prevent re-circularization of vector DNA during ligation reactions the DNA fragments were pre-treated with calf intestinal alkaline phosphatase (CIP) to remove the terminal 5'-phosphate groups.

DNA, containing up to 20 picomoles of 5' ends, was incubated in CIP buffer (50mM Tris.HCl pH 9.0, 1mM MgCl₂, 0.1mM ZnCl₂, 1mM spermidine) containing 1.0 unit of CIP for 30 minutes at 37°C. The reaction mixture was heated to 70°C for 10 minutes, extracted once with phenol/chloroform, once with chloroform, and passed through a Sephadex G50 column.

19. Ligation of DNA

Vector DNA, up to 100ng, was incubated with a 1-10 fold molar excess of insert DNA in ligation buffer (66mM Tris.HCl pH 7.6, 10mM MgCl₂, 10mM DTT, 1mM ATP, 50μg/ml BSA) containing 100-400 units of T4 DNA ligase. Incubation was for 5-6 hours at 16°C or 12-16 hours at 4°C. The volume of ligation mixtures was kept to a minimum (10-15μl) to maintain a high concentration of ligatable ends.
20. End-labelling of DNA fragments

Klenow fragment of DNA polymerase I was used to radioactively-label recessed 3'-termini of DNA fragments by an end-filling reaction. The reaction could be performed in the restriction endonuclease buffer in which the DNA had been previously digested.

To 25-30μl of DNA solution (containing up to 10μg of DNA fragments) was added 50μCi of a suitable [α-32P]dNTP, the three remaining dNTPs to a final concentration of 0.1mM, and 10 units of Klenow fragment. The labelling reaction mixture was incubated at 22°C for 30 minutes, and the reaction terminated by extraction with phenol/chloroform. The DNA solution was centrifuged through a Sephadex G50 column equilibrated in TE pH 8.0 to remove the unincorporated dNTPs.

21. Agarose gel electrophoresis

This technique was used for the separation and analysis of DNAs >1000bp in size. Electrophoresis was performed in a horizontal apparatus (Pharmacia GNA100; Bio-Rad mini sub cell; BRL model H5). Electrophoresis buffer was Tris-borate (TBE; 89mM Tris base, 89mM boric acid, 2.5mM EDTA).

Agarose was dissolved in TBE at the required concentration, 0.7-1.0% (w/v), by heating. After cooling to below 70°C, EtBr was added to a final concentration of 1μg/ml (if required) and the gel was cast in the mould. Samples were loaded in gel-loading buffer and electrophoresed at 70-150 volts for 1-3 hours or 15-20v overnight (the voltage and time period employed was dependent on the apparatus used and the DNA band separation required in each experiment). Following electrophoresis the DNA bands in gels containing EtBr were visualized by placing the gel on a UV transilluminator. Gels were photographed through a red filter using a Polaroid 545 Land Camera Back and Polaroid type 57 film. The DNA marker fragments used were BRL 1kb ladder, size range 210-12,216bp.
Polyacrylamide gel electrophoresis was used for the separation and analysis of DNA fragments <1000bp in size and the analysis of the products of DNA sequencing reactions. Sequencing products were electrophoresed on denaturing gels containing urea to separate the DNA into single strands. Acrylamide was stored as a 40%(w/v) stock solution, prepared by dissolving 38g of acrylamide and 2g of bis-acrylamide in distilled H₂O to a final volume of 100ml. Electrophoresis buffer was TBE.

a. Non-denaturing gels

A vertical gel of dimensions 20 x 18 x 0.15cm (BRL type V16 apparatus) was used for analysis of double-stranded DNA fragments. 60ml of acrylamide gel solution of the required concentration was prepared by mixing appropriate volumes of acrylamide stock solution, distilled H₂O, and 6ml of 10 x TBE. The solution was chilled on ice and 0.12ml of 25%(w/v) ammonium persulphate and 30μl of TEMED were added. The gel was cast immediately after the addition of these polymerization catalysts. Samples were loaded onto the gel using gel-loading buffer, and electrophoresed at 150-200v. To visualize the DNA the gel was soaked in TBE containing 1μg/ml EtBr for 30 minutes and placed on a UV transilluminator.

b. Denaturing gels for DNA sequence analysis

Sequencing reaction products were electrophoresed on a vertical gel of dimensions 40 x 20 x 0.04cm (Raven Scientific apparatus). The [³²P]-labelled products of Maxam-Gilbert sequencing were analysed on 8% and 20% gels with sample wells of 5mm width, and the [³⁵S]-labelled products of the dideoxy sequencing method were analysed on 6% gels with wells of 3mm width. 40ml of acrylamide gel solution was prepared by mixing appropriate volumes of acrylamide stock solution and H₂O with 16.8g of urea and 4ml of 10 x TBE. Urea was dissolved by heating and the solution was then chilled on ice. 80μl of 25% ammonium persulphate and 60μl of TEMED were added and the gel cast in the mould immediately. Prior to the loading of samples the gel was pre-run for 30 minutes at 15mA (20% gels) or 25mA (6% and 8% gels) to bring the temperature of the gel up to the optimum running temperature of 50-60°C.
The DNA samples, in formamide dye mix (95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF), were heated to 90°C for 2 minutes and then loaded into the sample wells using a microsyringe. The loading volumes were 4μl/well for Maxam-Gilbert products and 2μl/well for dideoxy products. Gels were run at the currents described above. To maximize the length of sequence that could be determined from the electrophoresis of dideoxy sequencing products on 6% gels it was usual to load each set of samples twice on the same gel. The first loading was electrophoresed until the xylene cyanol dye had migrated to 5-6cm from the bottom of the gel; the second loading was then made and electrophoresis continued until the bromophenol blue dye in the second loading had migrated to 2-3cm from the bottom of the gel. Following electrophoresis the gels were prepared for autoradiography. 20% gels were left attached to one of the glass plates, covered with a layer of Saranwrap, and autoradiographed at -70°C. 6% and 8% gels were fixed in 10%(v/v) methanol, 10%(v/v) acetic acid for 20 minutes, transferred onto a sheet of Whatman 3MM paper, and dried at 80°C for 60-90 minutes using a Bio-Rad slab gel dryer. The dried gels were autoradiographed at room temperature.

23. Autoradiography

The film used was Kodak X-OMAT or Fuji RX. The film cassettes were obtained from Genetic Research Instrumentation Ltd., and contained integral intensifying screens. Exposed film was developed in a Fuji RG II x-ray film processor.

24. Purification of DNA fragments from agarose and polyacrylamide

DNA was extracted from agarose and polyacrylamide by electroelution into a dialysis bag. The gel slice was placed in a dialysis bag containing 0.5 x TBE buffer and submerged under 0.5 x TBE in a horizontal electrophoresis box. The DNA was electroeluted into the dialysis bag at 100v for 120 minutes (agarose) or 150v for 90 minutes (polyacrylamide). The current was reversed for 1 minute and the buffer removed from the bag. The solution was passed through a column of siliconized plastic
wool packed in a 1 ml plastic pipette tip to remove pieces of gel, and the DNA recovered by ethanol precipitation. DNA eluted from agarose was phenol extracted before precipitation.

25. DNA sequencing by the Maxam-Gilbert chemical cleavage method

The sequence that needed to be determined was isolated on a DNA fragment that was radioactively-labelled at one of its termini. This was achieved by digesting plasmid DNA (purified on a CsCl gradient) with a suitable restriction endonuclease to generate a fragment containing the unknown sequence adjacent to one of the termini. The fragments were end-labelled using Klenow fragment and further digested with a second restriction endonuclease to generate a fragment labelled at one end, and with the unknown sequence adjacent to this labelled end. The DNA fragment was purified by electrophoresis and electroelution. The DNA was sequenced using the chemical cleavage method of Maxam and Gilbert (1980). The cleavage reactions performed were those specific for G, G+A, C+T, and C, using the standard buffers and reaction conditions (Maxam and Gilbert, 1980). The products of the reactions were analysed by polyacrylamide gel electrophoresis followed by autoradiography.

26. DNA sequencing by the dideoxy chain termination method

DNA was sequenced by the dideoxy chain termination method of Sanger et al (1977) using the Klenow fragment of DNA polymerase I or the modified bacteriophage T7 DNA polymerase ("Sequenase") described by Tabor and Richardson (1987). The enzyme and necessary reagents were purchased as commercial kits (Amersham International PLC, United States Biochemical Corp.), and the sequencing reactions were performed using the protocol supplied with the kit. The radioactive deoxyribonucleotide used was [α-35S]dATP. The sequencing products were analysed by polyacrylamide gel electrophoresis. Single-stranded DNA template was obtained either by purifying single-stranded forms of pEMBL-derived plasmids by infecting E. coli
DIH101 clones with bacteriophage f1, or by denaturing double-stranded plasmid DNA using NaOH, as described below.

a. Isolation of single-stranded pEMBL DNA

An overnight culture of the DIH101 strain harbouring the plasmid was grown in LB broth containing ampicillin and kanamycin. 30μl of this culture was inoculated into 1.4ml of 2 x TY medium containing ampicillin and 0.2% glucose and grown at 37°C for 2 hours. Bacteriophage f1 suspension was added at a m.o.i. of 20 and the culture incubated for a further 5 hours. Cells were pelleted by centrifugation and 1ml of supernatant removed and mixed with 0.25ml of 20% PEG 8000, 2.5M NaCl. The sample was left to stand at room temperature for 15 minutes and then centrifuged at 12,000 rpm for 5 minutes. The supernatant was removed and the phage particles resuspended in 0.1ml of TE pH 8.0. The suspension was phenol extracted and the DNA recovered by ethanol precipitation. The mixture of single-stranded f1 and pEMBL DNA was resuspended in 30μl of TE. 1-2μg of pEMBL DNA was used in each annealing reaction.

b. Denaturation of plasmid DNA using NaOH

This technique was used for direct sequencing of DNA cloned in the vector pHSG415. 4-5μg of plasmid DNA, purified by CsCl-EtBr gradient ultracentrifugation, was denatured in 10μl of 0.2M NaOH for 5 minutes at room temperature. The solution was neutralized by adding 0.4 volumes of 5M ammonium acetate pH 7.5, and the DNA precipitated by addition of 4 volumes of ethanol and placing at -70°C for 5 minutes. After centrifugation and washing in 70% ethanol, the DNA was dissolved in 7μl of distilled H₂O and used immediately as the template for a single annealing reaction.

For rapid determination of short lengths of sequence in pHSG415-derived plasmids (without the necessity of performing large scale plasmid preparations), plasmid DNA of sufficient purity was prepared using the method of Kraft et al (1988). In this method plasmid DNA is isolated by a small scale alkaline lysis procedure and purified by precipitation with PEG. Plasmid DNA isolated from 10ml of overnight
culture by this protocol was resuspended in 20 µl H₂O and denatured by addition of NaOH to 0.2M as described above. The denatured DNA was resuspended in 7 µl H₂O and used as the template for a single annealing reaction.

27. *In vitro* DNA-directed protein synthesis

The expression of the *ada* gene cloned in plasmid pCS70 (Teo *et al*, 1984) was studied *in vitro* using an *E. coli* cell-free coupled transcription-translation system (DNA-directed translation kit, Amersham International PLC). Using this system, based on those described by Zubay (1973) and Collins (1979), expression of *ada* was examined by following the incorporation of [³⁵S]-methionine into newly-synthesized Ada protein.

The reactions were performed according to the recommended protocol. Reaction mixtures (30 µl) containing an *E. coli* S-30 extract, amino acids, nucleotides, tRNA, an energy regenerating system and necessary co-factors were supplemented with 2 µg plasmid DNA, 20 µCi L-[³⁵S]-methionine and 1 µg of unmethylated or methylated Ada protein and incubated at 37°C for 30 minutes. After a 5 minute chase with unlabelled methionine, the newly-synthesized proteins were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), followed by autoradiography.

The wild-type Ada protein used in these experiments was purified by M. Olsson using the method of McCarthy and Lindahl (1985). The Ada-11 protein was isolated from the BS11 mutant strain by M. Olsson using the same procedure, except that a higher salt concentration (0.5M NaCl) was required to elute the protein from the DNA cellulose column. Methylated Ada protein was prepared by incubating 3 µg of Ada protein with 7.5 µg MNU-treated poly[dA].poly[dT] in 6 µl 70 mM HEPES.KOH pH 7.8, 10 mM DTT, 1 mM EDTA and 5% glycerol at 22°C for 20 minutes.

28. SDS-polyacrylamide gel electrophoresis

Electrophoresis was performed on vertical discontinuous gels of dimensions 8 x 7 x 0.075 cm (Bio-Rad Mini Protean II apparatus), using the methods of Laemmli
The stacking gel was 5% polyacrylamide and the resolving gel 12% polyacrylamide. Rainbow coloured [\(^{14}\text{C}\)]-labelled protein molecular weight markers (Amersham) were run on all gels. The markers were lysozyme (14.3kDa), trypsin inhibitor (21.5kDa), carbonic anhydrase (30kDa), ovalbumin (46kDa), bovine serum albumin (69kDa), phosphorylase b (92.5kDa) and myosin (200kDa). Samples were prepared in loading buffer (50mM Tris.HCl pH 6.8, 2% SDS, 2mM EDTA, 10% glycerol, 5% β-mercaptoethanol, 0.05% bromophenol blue) and denatured by heating at 100°C for 10 minutes immediately before loading. Samples were electrophoresed through the stacking gel at 50v until the markers had stacked at the interface of the two gels, and the voltage increased to 150v to run the proteins through the resolving gel. The gels were fixed by soaking for 4 hours in two changes of 5% methanol, 7% acetic acid, 5% glycerol, and then dried. \(^{35}\text{S}\)]-methionine-labelled proteins synthesized in DNA-directed protein synthesis reaction mixtures were analyzed by autoradiography.
EXPERIMENTAL RESULTS

A. Four *E. coli* mutants constitutively express the adaptive response in the absence of alkylation damage

Four mutant strains of *E. coli* F26 (an *E. coli* B/r derivative) which express the adaptive response to alkylating agents constitutively have been isolated in our laboratory (Sedgwick and Robins, 1980). These strains, designated BS11, BS21, BS31 and BS41, were obtained independently by mutagenizing four F26 cultures with MNNG and repeatedly challenging these cultures with a toxic concentration of MNU to select for mutants hyper-resistant to alkylation. All four of the mutants are more resistant to mutagenesis and killing by MNNG and MNU than wild-type *E. coli* in which the adaptive response has been induced (Sedgwick and Robins, 1980).

The BS11, BS31 and BS41 strains synthesize higher levels of Ada protein than adapted wild-type *E. coli*, and contain levels of N³-MeAda-DNA glycosylase II which are comparable to those synthesized by adapted wild-type cells (Lindahl *et al.*, 1983). In contrast, BS21 synthesizes higher levels of Ada protein than adapted wild-type *E. coli*, but contains only 2-3 fold higher levels of N³-MeAda-DNA glycosylase II than nonadapted wild-type cells (Karran *et al.*, 1982a; Lindahl *et al.*, 1983). The mutations responsible for the phenotype of these mutants have been mapped to the region of the ada gene at 47 minutes on the *E. coli* genetic map (Sedgwick, 1982; B. Sedgwick, personal communication). The initial aim of this study was to clone the ada gene from each of the four mutants and determine the nucleotide sequence changes that result in the constitutive expression of the adaptive response in the absence of alkylation damage.
B. The *ada* genes of the constitutive mutants could not be cloned using a high copy number vector

Sedgwick (1983) isolated the *ada* gene from wild-type *E. coli* F26 on a 3.1 kb *Hind*III fragment of chromosomal DNA. Digestion of this 3.1kb DNA fragment with *Sma*I generates a 1.3kb *Hind*III/*Sma*I chromosomal fragment containing the entire *ada* gene (Teo et al, 1984; Demple et al, 1985). Making use of these observations, a strategy was devised to clone the *ada* gene from each of the four constitutive mutants using the plasmid vector pEMBL18[+] (Dente et al 1983).

The pEMBL plasmids are derived from the pUC series of vectors and contain the *bla* gene as a selectable marker, a polylinker with multiple cloning sites inserted within the gene encoding the α-peptide of β-galactosidase, and a segment of the genome of the f1 single-strand DNA bacteriophage carrying the phage origin of replication (Dente et al, 1983). Infection of *E. coli* carrying a pEMBL-derived plasmid with bacteriophage f1 causes the plasmid to switch to the phage mode of replication, resulting in the synthesis of single-stranded plasmid DNA. This DNA is packaged into viral coat particles, which are secreted from the infected cell. The single-stranded plasmid DNA isolated from these phage particles can be used as template for DNA sequencing by the dideoxy chain termination method (Dente et al, 1983). Indeed, the nucleotide sequence of the wild-type *ada* gene of *E. coli* F26 was in part determined by this method after subcloning the gene into a pEMBL vector on the 1.3kb *Hind*III/*Sma*I fragment of chromosomal DNA (Demple et al, 1985; B. Sedgwick, personal communication).

Plasmid pEMBL18[+], illustrated schematically in figure 3.1, contains a polylinker with an *Hind*III and a *Sma*I site. The aim of the strategy to isolate the *ada* genes from the constitutive mutants was to insert the 1.3kb *Hind*III/*Sma*I fragment of chromosomal DNA containing the *ada* gene into the corresponding sites in the polylinker of pEMBL18[+]. The insert DNA would be in such an orientation that the single-stranded plasmid DNA packaged into phage coat particles in an f1-infected culture of
Figure 3.1: Schematic map of pEMBL18(+)
the recombinant clone would contain the non-coding strand of the *ada* gene. This DNA could be isolated and sequenced to generate the coding nucleotide sequence of the *ada* gene.

Chromosomal DNA isolated from the constitutive mutant was digested with *Hind*III and the reaction products resolved by agarose gel electrophoresis. DNA fragments of 2.9 to 3.3kb were cut out of the gel in a slice of agarose and purified by electroelution. These purified DNA fragments were further digested with *Sma*I, and the products ligated to pEMBL18[+] DNA which had been digested with *Hind*III and *Sma*I, precipitated with propan-2-ol to remove the small fragment of polylinker DNA, and treated with calf intestinal alkaline phosphatase (CIP) to prevent the vector DNA molecules from ligating to each other. *E. coli* DIH101 were transformed with the ligation products, and the transformed cells plated on ampicillin. To identify recombinant clones containing the *ada* gene, the ampicillin-resistant (ApR) transformants were analysed using an immunological screening test that specifically detects colonies of cells synthesizing high levels of Ada protein.

The cloning protocol described was employed for all four mutant strains. In each case, 800-1000 ApR colonies were screened for high cellular levels of the Ada protein. However, none of the colonies gave a positive reaction. Colonies of the mutant strains were tested as a control during the screening procedures and gave a strong positive reaction, demonstrating that the immunological screening method was capable of detecting the levels of Ada protein synthesized from a single copy of a constitutively-overexpressed *ada* gene. The complete absence of positive transformants was unexpected since the initial purification of the *Hind*III chromosomal fragments of size 2.9 to 3.3kb should have ensured that at least some of the ApR transformants would contain a recombinant plasmid with a 1.3kb *Hind*III/*Sma*I DNA insert carrying the *ada* gene.

When the four mutants were originally isolated it was observed that alkylation-sensitive revertants unable to express the adaptive response accumulated on stored slope cultures of the strains (Sedgwick and Robins, 1980). This suggested that
overproduction of the Ada protein, or one of the other inducible gene products, might be deleterious to the cell. The possibility that such a phenomenon could account for the failure to clone the \textit{ada} genes from these mutants was considered. The pEMBL plasmids have a high copy number (Dente \textit{et al}, 1983), and the presence of a large number of copies of the constitutively-expressed \textit{ada} gene in a cell might possibly result in the synthesis of lethal amounts of the Ada protein. A new cloning strategy using a low copy number vector was therefore designed in order that any such problems of toxicity might be reduced.

\textbf{C. Molecular cloning of the mutant \textit{ada} genes using a low copy number vector}

Plasmid pHSG415, shown schematically in figure 3.2, is a low copy number vector maintained at 4-6 copies per chromosome (Hashimoto-Gotoh \textit{et al}, 1981). The 7.1kb plasmid contains the \textit{bla}, \textit{cat}, and \textit{kan} genes, which confer resistance to ampicillin, chloramphenicol and kanamycin, respectively. There are unique cutting sites for a number of restriction endonucleases, including \textit{HindIII}. The replication of pHSG415 is temperature-sensitive. Replication proceeds normally at 30°C, but ceases at 37°C and above. The aim of the second strategy to clone the \textit{ada} genes from the constitutive mutants was to ligate the 3.1kb \textit{HindIII} fragment of chromosomal DNA containing the \textit{ada} gene into the \textit{HindIII} site of pHSG415. From hereon the \textit{ada} genes of BS11, BS21, BS31 and BS41 will be referred to as \textit{ada-11}, \textit{ada-21}, \textit{ada-31} and \textit{ada-41}, respectively.

Vector DNA was cleaved with \textit{HindIII} and then dephosphorylated using CIP. \textit{HindIII} fragments of chromosomal DNA in the size range 2.9kb to 3.3kb were ligated to pHSG415 DNA and \textit{E. coli} DIH101 or HB101 transformed with the ligation products. Transformed cells were plated on ampicillin and the Ap\textsuperscript{R} colonies screened immunologically for high cellular levels of Ada protein. In the experiment to clone the \textit{ada-11} gene, 3 out of a total of 178 colonies screened gave a strong positive reaction.
Figure 3.2: Schematic map of pHSG415
Colonies synthesizing high levels of Ada protein were also obtained in the experiments to clone the *ada-21, ada-31 and ada-41* genes, the numbers of positive transformants relative to the total number of Ap\(^R\) transformants screened being 9/242, 4/110 and 1/98, respectively. Figure 3.3 shows a photograph of a nitrocellulose filter used to screen for Ap\(^R\) HB101 transformants carrying the cloned *ada-21* gene. Five positive transformants synthesizing high levels of Ada protein are visible within a total of 87 Ap\(^R\) colonies.

Transformants that gave a positive reaction in the immunological screen were purified from the master plates and re-screened to ensure that the correct colonies had been selected. Plasmid DNA was isolated from these recombinant clones and analyzed by restriction endonuclease digestion. In the experiment to clone the *ada-11* gene, two out of the three recombinants contained a 10.2kb plasmid composed of the 7.1kb vector DNA with a single 3.1kb *HindIII* insert. Further restriction analysis demonstrated that the 3.1kb insert contained the *ada-11* gene. The insert DNA was in the same orientation in both plasmids, and this construct was named pSH11. A schematic map of plasmid pSH11 is shown in figure 3.4. The third recombinant generated in this experiment contained a plasmid larger in size than 10.2kb, but which nevertheless produced 3.1kb and 7.1kb fragments among the products when digested with *HindIII*. In the experiments to isolate the *ada-31* and *ada-41* genes, a single clone containing the 10.2kb plasmid carrying the *ada* gene was obtained in each case. The orientation of the chromosomal DNA inserts were identical to that in pSH11, and these constructs containing the *ada-31* and *ada-41* genes were designated pSH31 and pSH41. In the experiment to clone the *ada-21* gene, four out of the nine positive recombinants carried a 10.2kb plasmid containing the *ada-21* gene. Two of these plasmids had the insert DNA in the same orientation as pSH11, pSH31, and pSH41, and were named pSH21. The other two plasmids had the insert DNA in the reverse orientation and were termed pSHR21. The five other recombinants generated in this experiment contained plasmids larger than 10.2kb, and yielded additional products to a 3.1 and 7.1kb fragment upon digestion with *HindIII*.
Figure 3.3: Immunological screening of ApR HB101 transformants in the experiment to isolate the *ada*-21 gene.

ApR transformants were lifted onto nitrocellulose filters and screened immunologically for high cellular levels of Ada protein (see Materials and Methods). This photograph shows a filter after incubation with the peroxidase substrate solution. Five positive transformants giving an intense colour reaction are visible in a background of negative colonies.
Figure 3.4: Schematic map of plasmid pSH11

Vector DNA is indicated by the thin line and the 3.1 kb chromosomal DNA insert by the thick line. The direction of transcription of the *ada-11* gene (dark shading) from its own promoter (Pr, hatched shading) is indicated. The DNA downstream of the *ada* gene (lighter shading) contains the *alkB* gene. Plasmids pSH21, pSH31 and pSH41 are the corresponding constructs containing the *ada-21*, *ada-31* and *ada-41* genes, respectively.
The success in cloning the *ada* genes of the constitutive mutants in pHSG415 following initial unsuccessful attempts with a similar cloning strategy using pEMBL18[+] suggested that the presence of these genes on a high copy number plasmid is lethal to the cell. To provide support for this suggestion, and to confirm that the failure to clone the *ada* genes in pEMBL18[+] had not been the result of an unforeseen technical problem, an attempt was made to subclone the *ada-11* gene from pSH11 into pEMBL18[+]. Plasmid pSH11 was digested with *Hind*III and *Sma*I and the 1.3kb *Hind*III/*Sma*I DNA fragments containing the *ada-11* gene were gel-purified. These fragments were then ligated to pEMBL18[+] DNA which had been digested with *Hind*III and *Sma*I, precipitated with propan-2-ol and then treated with CIP. Analysis of the ligation products by agarose gel electrophoresis confirmed that the 1.3kb *Hind*III/*Sma*I DNA fragments had ligated to the vector DNA. Competent *E. coli* DH101 were transformed with the ligation products and the cells plated on ampicillin to select for transformants containing recombinant plasmids. No ApR colonies developed on the selective plates, even though control transformations of the competent DH101 with intact pEMBL18[+] yielded large numbers of ApR colonies. The lack of ApR transformants clearly indicated that *E. coli* is unable to tolerate the presence of a high copy number plasmid carrying the *ada-11* gene.

D. Determination of the nucleotide sequence of the cloned *ada* genes

The nucleotide sequence of the promoter region of the *ada-11, ada-21, ada-31* and *ada-41* genes was determined by the Maxam-Gilbert chemical cleavage method of DNA sequencing. Plasmids pSH11, pSH21, pSH31 and pSH41 were isolated by a large scale alkaline lysis procedure and purified by caesium chloride-ethidium bromide density gradient ultracentrifugation. The plasmid DNA was digested with *Hind*III and the DNA fragments end-labelled with [α-32P]dCTP using the Klenow fragment of DNA polymerase I. The fragments were further digested with *Sal*I to generate a 372bp fragment of the *ada* gene consisting of the entire promoter region and the first 272
nucleotides of the structural gene. This DNA fragment, radioactively-labelled in the non-coding DNA strand, was purified and subjected to the Maxam-Gilbert sequencing reactions. Electrophoresis of the reaction products on 8% and 20% sequencing gels permitted the nucleotide sequence of the non-coding DNA strand from the start of the promoter through to the EcoRI site at nucleotides +76 to +81 within the structural gene to be determined.

The nucleotide sequence of the wild-type ada gene isolated from E. coli B strain F26 was determined by Demple et al (1985) and is shown in figure 3.5. The nucleotide sequences of the ada-11, ada-21, ada-31 and ada-41 genes determined by Maxam-Gilbert sequencing were compared with that of the wild-type ada gene and it was discovered that none of these genes have mutations in the promoter region. However, the ada-11, ada-31 and ada-41 genes were found to have a C to T transition mutation in the non-coding DNA strand at nucleotide +17 of the structural gene (see Figure 3.6), and must therefore have a G to A transition mutation in the coding gene at this site. To confirm this observation the nucleotide sequence of this region of the coding DNA strand was determined. Plasmids pSH11, pSH31 and pSH41 were digested with EcoRI and the DNA fragments end-labelled with \[\alpha^{32}\text{P}]dATP. The DNA was then digested with HindIII to generate a 176 bp HindIII/EcoRI fragment of the ada gene, consisting of the promoter region and the first 76 nucleotides of the structural gene, labelled at the 3'-terminus of the coding strand. Sequencing of this DNA by the Maxam-Gilbert method enabled the sequence of the coding strand from the start of the promoter through to nucleotide +70 within the structural gene to be determined. The presence of a G to A transition mutation at nucleotide +17 in all three genes was confirmed. This mutation results in the synthesis of an altered Ada protein in which the cysteine (TGC) residue at amino acid 6 is replaced by tyrosine (TAC).

To complete the nucleotide sequence of the structural genes, a fragment of the ada gene lacking the promoter region and the first part of the structural gene was subcloned into a pEMBL vector and the ada DNA sequenced by the dideoxy chain termination method. Plasmids pSH11, pSH21 and pSH41 were cleaved with EcoRI and
Figure 3.5: Nucleotide sequence of the wild-type ada gene from E. coli F26 (Demple et al, 1985)
Figure 3.6: Autoradiograph of a DNA sequencing gel showing the mutation at nucleotide +17 of the *ada-11, ada-31* and *ada-41* genes

The 372bp *HindIII/SmaI* fragment of the *ada-11, ada-31* and *ada-41* genes, end-labelled at the *HindIII* terminus, were sequenced by the Maxam-Gilbert method. A purified 176bp *HindIII/EcoRI* fragment of the *ada*\(+\) gene, obtained from Dr. B. Sedgwick, was also end-labelled at the *HindIII* terminus and sequenced by the Maxam-Gilbert method. The sequencing products of the four genes were electrophoresed on an 8% polyacrylamide sequencing gel. This autoradiograph shows the nucleotide sequence of the non-coding DNA strand of the *ada* genes from nucleotide -44 in the promoter region through to nucleotide +40 within the structural gene. The intact arrow indicates the position of the C to T transition mutation at nucleotide +17 of the *ada-11, ada-31* and *ada-41* genes. These genes must have a G to A transition in the coding strand at this position. The broken arrow indicates a site where there is a band missing in the T+C and C lanes of the *ada-11, ada-31* and *ada-41* sequences. This nucleotide is the internal C in the sequence CCTGG, and is methylated by the *dcm* methylase in *E. coli* K12 strains (the HB101 and DH101 hosts for plasmids pSH11, pSH31 and pSH41 are K12 strains). The Maxam-Gilbert cleavage reaction specific for cytosine does not cleave DNA at methylcytosine residues. The *ada*\(+\) DNA was cleaved at this nucleotide residue because it had been isolated from a plasmid propagated in an *E. coli* B strain, which lack the *dcm* methylase.
SmaI to generate a 1.2kb EcoRI/SmaI fragment containing the ada gene without the promoter and first 76 nucleotides of the structural gene. This fragment was ligated into pEMBL8[-] (Dente et al., 1983) between the EcoRI and HincII sites of the polylinker region. The 5.3kb constructs so formed do not permit the synthesis of intact Ada protein, and potential problems of cytotoxicity due to the overproduction of the Ada protein were therefore avoided. Cultures of DIH101 transformants carrying these constructs were infected with bacteriophage f1. Phage particles were harvested from the culture medium and the single-stranded plasmid DNA which contained the non-coding strand of the ada gene was isolated. Using this DNA as template, and employing the M13 reverse universal primer and a further five oligonucleotide primers complementary to the non-coding strand of the ada gene, the nucleotide sequences of the coding strands of the ada genes were generated by the chain termination method. The patterns of bands on the autoradiographs of the sequencing gels were very clear, and the sequences generated by adjacent primers overlapped substantially. Thus, determination of the complete sequence of the opposite DNA strand was considered to be unnecessary. However, there were two specific nucleotide residues within the coding sequence where bands persistently occurred in all four lanes of the sequencing gel. The nucleotide residue at these two sites was determined by generating the sequence of the non-coding DNA strand in these regions by double-stranded DNA sequencing using pSH11, pSH21 and pSH41 DNA as template.

Chain termination sequencing identified a second mutation in the ada-11 and ada-41 genes, and two mutations in the ada-21 gene. The ada-11, ada-21 and ada-41 genes all have a G to A transition mutation at nucleotide +378 (Figure 3.7), which will result in the synthesis of an Ada protein containing isoleucine (ATA) in place of methionine (ATG) at amino acid residue 126. In addition, the ada-21 gene was found to have a G to A transition mutation at nucleotide +289 (see Figure 3.8), which will result in the glutamic acid (GAA) residue at amino acid 97 of the Ada protein being replaced by lysine (AAA). Although the ada-31 gene was not sub-cloned into pEMBL8[-] and completely sequenced, partial double-stranded sequencing of pSH31.
Figure 3.7: Autoradiographs of DNA sequencing gels showing the G to A transition mutation at nucleotide +378 of the *ada-11*, *ada-21* and *ada-41* genes.

The 1.2 kb *EcoRI/Smal* fragment of the *ada-11*, *ada-21* and *ada-41* genes was subcloned into pEMBL8[-]. Single-stranded plasmid DNA template was isolated and sequenced by the dideoxy chain termination method to generate the nucleotide sequence of the coding DNA strand of the *ada* genes. The sequencing products were electrophoresed on 6% polyacrylamide sequencing gels and analyzed by autoradiography. The arrows show the position of the G to A transition mutation at nucleotide +378 of these genes.
Figure 3.8: Autoradiograph of a DNA sequencing gel showing the G to A transition mutation at nucleotide +289 of the ada-21 gene.

Single-stranded ada-21 DNA template was sequenced by the dideoxy chain termination method to generate the nucleotide sequence of the coding DNA strand. The sequencing products were electrophoresed on a 6% polyacrylamide sequencing gel and analyzed by autoradiography. The arrow indicates the position of the G to A transition mutation at nucleotide +289 of the ada-21 gene.
showed that this gene also has a G to A transition mutation at nucleotide +378. Thus, BS11, BS31 and BS41 have two identical mutations in the *ada* gene. BS21, which differs phenotypically from these three strains, has one *ada* mutation identical to BS11, BS31 and BS41, but contains a different second *ada* mutation. Figure 3.9 summarizes the amino acid substitutions which occur in the Ada proteins synthesized by the four constitutive mutants. The proteins synthesized by BS11, BS21, BS31 and BS41 have been designated as Ada-11, Ada-21, Ada-31 and Ada-41, respectively.

A point of note is that all of the mutations in the *ada* genes are GC to AT transitions. This is consistent with them having been induced by MNNG or MNU during the selection procedure employed to isolate the mutants, since both of these alkylating agents primarily induce GC to AT transition mutations in *E. coli* (Burns *et al.*, 1987; Richardson *et al.*, 1987).

**E. The mutant Ada proteins are strong activators of *ada* gene expression**

The adaptive response is induced through the self-methylation of Cys\(^{69}\) in the Ada protein by its repair of a methylphosphotriester lesion activating the protein as a strong positive regulator of transcription of the inducible repair genes (Teo *et al.*, 1986; Nakabeppu and Sekiguchi, 1986). The finding that the four constitutive strains contained mutations in the *ada* structural gene, and not in the *ada* promoter region, suggested that the adaptive response might be expressed constitutively because the resulting amino acid substitutions convert the Ada proteins into strong activators of expression of *ada* and the other inducible genes of the response. Two approaches were taken to establish whether the altered Ada proteins synthesized by the constitutive mutants were strong activators of gene expression in the absence of cellular alkylation damage.

In the first approach the ability of the mutant Ada proteins to act as inducers of their own synthesis was examined by assaying the levels of Ada protein in cells
Figure 3.9: Summary of the amino acid substitutions in the Ada proteins synthesized by the constitutive mutants
carrying the mutated ada genes on the low copy number plasmids. The ada-21 and ada-41 genes were taken to represent the two types of mutated ada gene. The 3.1 kb chromosomal DNA inserts in plasmids pSH21 and pSH41 are within the kan gene of pHSG415, and in such an orientation that the kan and ada genes are transcribed in the same direction (see figure 3.4). To avoid the synthesis of Ada protein from mRNA transcripts initiated at the promoter of the kan gene, the inserts were inverted to form plasmids pSHR21 and pSHR41. Figure 3.10 illustrates the structure of pSHR41. The 3.1 kb HindIII fragment of F26 chromosomal DNA containing the ada+ gene was subcloned from pCS42 (Sedgwick, 1983) into the HindIII site of pHSG415 so that the insert DNA was in the same orientation as the chromosomal inserts containing the ada-21 and ada-41 genes in pSHR21 and pSHR41. This ada+ homologue of pSHR21 and pSHR41 has been termed pSHR1. Plasmids pSHR1, pSHR21 and pSHR41 were transformed into E. coli GW7101 (Shevell et al., 1988), a derivative of AB1157 in which the chromosomal ada gene has been deleted, and the cellular levels of Ada protein determined by assaying the O6-MeGua-DNA methyltransferase activity in cell extracts.

Two cell extracts of the strains GW7101, GW7101/pSHR1, GW7101/pSHR21, GW7101/pSHR41, BS21 and BS41 were prepared from independent cultures grown at 30°C. The O6-MeGua-DNA methyltransferase activity in samples of extract containing 0.25, 0.5, 1.0, 1.5, and 2.0 μg of protein was assayed. Figure 3.11 shows a graph of the radioactivity (counts per minute, cpm) removed from the [3H]MNU-treated DNA substrate versus amount of protein for one set of extracts assayed under identical conditions (the second set of extracts gave very similar results). The BS21 and BS41 extracts, employed as the positive controls containing high levels of Ada protein, exhibited a linear increase in radioactivity with increasing protein concentration. In the case of the untransformed GW7101 host strain, the radioactivity varied between 0 and 70 cpm in a manner independent of the amount of protein in the assay reaction mixture. Although GW7101 is deleted for the chromosomal ada gene, it does contain the non-inducible second O6-MeGua-DNA methyltransferase activity encoded by the ogt
Figure 3.10: Schematic map of plasmid pSHR41

Vector DNA is indicated by the thin line and the 3.1kb chromosomal DNA insert by the thick line. The direction of transcription of the *ada-41* gene (dark shading) from its own promoter (Pr, hatched shading) is indicated. The DNA downstream of the *ada* gene (lighter shading) contains the *alkB* gene. Plasmids pSHR1 and pSHR21 are the corresponding constructs containing the wild-type *ada* gene and the *ada-21* gene, respectively.
Radioactivity (cpm) removed from $[^3H]$MNU-treated DNA

O$^6$-[methyl-$^3$H] guanine constitutes approximately 75% of the radioactive counts in the DNA substrate (Karran et al, 1979)

Figure 3.11: Assays of O$^6$-methylguanine-DNA methyltransferase activity in cell extracts of GW7101 strains containing the plasmids pSHR1, pSHR21 and pSHR41 and of the constitutive mutants BS21 and BS41.
gene (Shevell et al, 1988). However, the cellular level of this enzyme is very low (Shevell et al, 1988), and appears to have been below the minimum limit of detection of the assay for the range of protein concentrations which were employed. GW7101 transformed with pSHR1 exhibited a similar variation in radioactivity about the baseline, indicating that the level of Ada protein synthesized from the *ada*+ gene carried on the low copy number vector was below the limit of detection of the assay. In contrast, the cell extracts of GW7101/pSHR21 and GW7101/pSHR41 exhibited an even greater capacity to remove radioactivity from the DNA substrate than the BS21 and BS41 extracts, indicating that these transformants synthesize higher levels of Ada protein than the original mutant strains.

One unit of O6-MeGua-DNA methyltransferase activity removes one picomole of methyl groups from O6-MeGua in DNA (Demple et al, 1982). To allow a more quantitative comparison of the levels of Ada protein synthesized by the GW7101 transformants and the BS strains, the O6-MeGua-DNA methyltransferase activities in the cell extracts (expressed as units/mg of protein) were calculated from the results of the assays. The initial gradients of the plots of radioactivity versus amount of protein were determined and used to calculate the radioactivity that would be removed from the DNA substrate by the methyltransferase present in 1mg of protein. Knowing the specific activity of the N-[methyl-3H]-N-nitrosourea used to prepare the assay substrate (23Ci/mm, equivalent to 5.06 x 10^{13} dpm/mmol), and the percent counting efficiency of the scintillation counter, these radioactivity determinations were converted to their equivalent in picomoles of methyl groups removed from O6-MeGua. The O6-MeGua-DNA methyltransferase levels in the two cell extracts prepared for each strain are shown in table 3.1. With the untransformed GW7101 host strain and GW7101/pSHR1 the enzyme levels could not be calculated since they had been below the lower limits of detection of the assay at the particular range of protein concentrations used. In the table they have been assigned as being <1.0 unit/mg of protein, this representing the minimum level of methyltransferase at which the enzyme-catalyzed release of radioactivity from the substrate would have been detected.
Table 3.1: O\textsuperscript{6}-MeGua-DNA methyltransferase activity in cell extracts of the GW7101 transformants containing plasmids pSHR1, pSHR21 and pSHR41 and the original mutants BS21 and BS41

<table>
<thead>
<tr>
<th>Strain</th>
<th>units/mg of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS21</td>
<td>16.5, 21.3</td>
</tr>
<tr>
<td>BS41</td>
<td>13.4, 8.9</td>
</tr>
<tr>
<td>GW7101</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>GW7101/pSHR1</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>GW7101/pSHR21</td>
<td>29.1, 28.3</td>
</tr>
<tr>
<td>GW7101/pSHR41</td>
<td>32.3, 45.5</td>
</tr>
</tbody>
</table>

One unit of methyltransferase removes 1 picomole of methyl groups from O\textsuperscript{6}-MeGua in DNA under the standard assay conditions.
over the background variation in radioactivity along the baseline which is inherent to
the assay. For the four strains in which O⁶-MeGua-DNA methyltransferase activity
was detected, the two determinations of the cellular level of the enzyme were in close
agreement. GW7101/pSHR21 contained approximately 1.5-fold higher levels of O⁶-
MeGua-DNA methyltransferase than BS21, and the levels in GW7101/pSHR41 were
2.5 to 5-fold greater than in BS41. The observation that GW7101/pSHR21 and
GW7101/pSHR41 contained very high cellular levels of methyltransferase, but no
such activity could be detected in GW7101/pSHR1 (ada⁺), strongly suggested that the
Ada-21 and Ada-41 proteins act as constitutive inducers of their own synthesis.

The 3.1 kb chromosomal DNA inserts in plasmids pSHR1, pSHR21 and pSHR41
contained the alkB gene, which is located immediately downstream of ada. These two
genes form an operon, the expression of alkB being regulated by the promoter of the
ada gene (Kondo et al, 1986). The function of the 24kDa AlkB protein has not been
defined. However, the observation that E. coli alkB mutants are able to induce the ada
and alkA genes in response to treatment with MNNG (Kataoka and Sekiguchi, 1985)
suggested that AlkB is unlikely to be involved in the regulation of the adaptive
response. Nevertheless, to exclude the possibility that mutations in the alkB genes of
BS21 and BS41 might have been responsible for the synthesis of the high levels of Ada
protein observed in the GW7101/pSHR21 and GW7101/pSHR41, a derivative of
pSHR41 lacking the alkB gene was constructed and the cellular levels of O⁶-MeGua-
DNA methyltransferase in GW7101 transformants containing this new construct were
determined.

Plasmid pSHR41 was digested with SmaI to generate two products of sizes 2.1kb
and 8.1kb. The 8.1kb DNA fragment was circularized using T4 DNA ligase to form a
new plasmid, designated pSHR14, which contained the entire ada-41 gene and only the
first 156 nucleotides of the alkB gene of BS41. The corresponding construct containing
the ada⁺ gene, termed pSHR15, was derived from pSHR1 using the same procedure. No
O⁶-MeGua-DNA methyltransferase activity was detected in two different cell extracts
of GW7101/pSHR15. However, constitutively-high levels of activity were observed
in GW7101/pSHR14, confirming that the mutations within the ada-41 gene were responsible for the synthesis of the high levels of Ada protein that had been measured in the GW7101/pSHR41.

The expression of genes contained on a plasmid or bacteriophage genome can be studied in vitro using an E. coli cell-free coupled transcription-translation system (Zubay, 1973; Collins, 1979). Employing such a system, Teo et al. (1986) demonstrated that the methylated wild-type Ada protein acts as a strong inducer of ada gene expression. In the second approach taken to determine whether the Ada proteins synthesized by the constitutive mutants are strong activators of gene expression, an in vitro coupled transcription-translation system was used to examine the ability of the purified Ada-11 protein to induce expression of the wild-type ada gene carried on a plasmid vector.

Plasmid pCS70 contains the ada operon, composed of the ada and alkB genes, cloned in the vector pAT153 (Teo et al., 1984). Different forms of purified wild-type Ada protein and Ada-11 protein were added to in vitro transcription-translation reaction mixtures containing pCS70 and L-[35S]-methionine. The proteins synthesized from the genes carried by pCS70, radioactively-labelled as a result of the incorporation of [35S]-methionine, were separated by SDS-PAGE and analyzed by autoradiography. The levels of ada+ gene expression in the various reaction mixtures were compared by measuring the intensity of the 39kDa Ada protein bands on the autoradiographs using an LKB Ultrosan XL laser densitometer.

Figure 3.12 shows a representative autoradiograph from these studies. The major product of protein synthesis directed by the parental vector pAT153 (lane 1) was the 31kDa β-lactamase. There was no visible synthesis of the 36kDa tet gene product. In the absence of added Ada protein the principal product of pCS70-directed protein synthesis (lane 2) was β-lactamase, and there was also a low basal level of synthesis of the Ada protein. Addition of unmethylated wild-type Ada protein to a reaction mixture containing pCS70 (lane 3) resulted in a 2 to 3-fold increase in the level of Ada protein synthesis, consistent with previous observations that it is a weak activator.
Different forms of purified Ada-11 and wild-type Ada protein were added to an *in vitro* coupled transcription-translation system containing plasmid pCS70 as the DNA template. This plasmid is derived from pAT153 and contains the *ada*+ operon. The proteins synthesized in the reaction mixtures were resolved on a 12% SDS-polyacrylamide gel and visualized by autoradiography. The reaction mixtures contained: lane 1, pAT153 as template; lane 2, pCS70 as template; lane 3, pCS70 and unmethylated wild-type Ada protein; lane 4, pCS70 and unmethylated Ada-11 protein; lane 5, pCS70 and methylated wild-type Ada protein. Lane M contained [14C]-labelled protein molecular weight markers, which were trypsin inhibitor (21.5kDa), carbonic anhydrase (30kDa), ovalbumin (46kDa), bovine serum albumin (68kDa), phosphorylase b (92.5kDa) and myosin (200kDa).
of transcription of the \textit{ada} gene (Teo \textit{et al}, 1986). However, addition of unmethylated Ada-11 protein (lane 4) caused a 15 to 20-fold stimulation of Ada protein synthesis, which was comparable to that induced by the methylated wild-type Ada protein (lane 5). Thus, the unmethylated Ada-11 protein containing its two amino acid substitutions acted as a strong inducer of \textit{ada} gene expression \textit{in vitro}.

The O$^6$-MeGua-DNA methyltransferase assays and the \textit{in vitro} translation studies together demonstrated that the Ada proteins synthesized by the constitutive mutants are strong activators of \textit{ada} gene expression. This indicated that the constitutive expression of the adaptive response in the mutant strains results from the synthesis of structurally-modified Ada proteins which induce the expression of the \textit{ada} gene and the other inducible genes of the response in the absence of cellular alkylation damage.

\textbf{F. The mutant Ada proteins repair methylphosphotriesters in DNA}

Self-methylation of Cys$^{69}$ by the repair of a methylphosphotriester converts the wild-type Ada protein into a strong activator of transcription of the \textit{ada} and \textit{alkA} genes (Teo \textit{et al}, 1986). To determine if the amino acid substitutions in the N-terminal region of the Ada proteins synthesized by the constitutive mutants affect the ability of the proteins to repair methylphosphotriesters, the levels of methylphosphotriester-DNA methyltransferase activity in cell extracts of the four mutants and the wild-type F26 strain were determined.

Two cell extracts of strains BS11, BS21, BS31, BS41 and F26 were prepared from independent cultures. The methylphosphotriester-DNA methyltransferase activity in samples of extract containing 1, 2.5, 5 and 10\(\mu\)g of protein was assayed. Figure 3.13 shows graphs of the radioactivity (cpm) transferred to a protease-sensitive form versus \(\mu\)g of protein for one set of extracts assayed under identical conditions. The extracts of the four constitutive mutants had a far greater capacity for repairing methylphosphotriesters than the F26 extract. Indeed, the plots for the BS
Radioactivity (cpm) transferred from the \[^{3}H\]MNU-treated poly(dT).poly(dA) substrate to a protease-sensitive form.

Figure 3.13: Assays of methylphosphotriester-DNA methyltransferase activity in the BS constitutive mutants
strains start to plateau at protein concentrations above 2.5μg, indicating that the availability of methylphosphotriester lesions in the substrate became a limiting factor in the assay mixtures containing 5 and 10μg of protein.

Like the levels of O\(^6\)-MeGua-DNA methyltransferase in cell extracts, the levels of methylphosphotriester-DNA methyltransferase are routinely expressed as the number of units of activity/mg of protein, one unit of activity being that which removes 1 picomole of methyl groups from methylphosphotriesters under the standard assay conditions. The levels of methylphosphotriester-DNA methyltransferase in the two extracts prepared for each strain were calculated from the results of the assays and are shown in table 3.2. All four mutants contained substantially higher levels of activity than the wild-type strain. Of the four mutants, BS31 contained the lowest level of activity, and this was 30 to 33-fold greater than that of F26. BS21 contained the highest level of activity of the four mutants, which was 115 to 120-fold greater than the level of activity in F26. The high cellular levels of methylphosphotriester-DNA methyltransferase in the constitutive mutants indicated that the amino acid substitutions in the Ada proteins synthesized by these strains do not cause an alteration to the structure of the active site containing Cys\(^{69}\) which prevents the protein from repairing methylphosphotriesters. It is of note that the methylphosphotriester-DNA methyltransferase levels in BS21 and BS41 were very similar to the O\(^6\)-MeGua-DNA methyltransferase levels determined for these two strains (see table 3.1). This is what one would have expected for a strain synthesizing Ada protein in which both active sites are functional, since each Ada protein molecule can only repair a single methylphosphotriester and a single O\(^6\)-MeGua or O\(^4\)-MeThy adduct (Lindahl et al, 1982; McCarthy and Lindahl, 1985).
Table 3.2: Methylphosphotriester-DNA methyltransferase activity in cell extracts of the BS constitutive mutants

<table>
<thead>
<tr>
<th>Strain</th>
<th>Methylphosphotriester-DNA methyltransferase (units/mg protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F26</td>
<td>0.2, 0.2</td>
</tr>
<tr>
<td>BS11</td>
<td>8.9, 14.7</td>
</tr>
<tr>
<td>BS21</td>
<td>21.3, 22.7</td>
</tr>
<tr>
<td>BS31</td>
<td>6.0, 6.5</td>
</tr>
<tr>
<td>BS41</td>
<td>6.8, 11.7</td>
</tr>
</tbody>
</table>

One unit of methylphosphotriester-DNA methyltransferase removes 1 picomole of methyl groups from methylphosphotriesters in DNA under standard assay conditions.
G. A single amino acid substitution converts the Ada protein into a strong activator of expression of the \textit{ada} and \textit{alkA} genes

The \textit{ada} genes of the four constitutive mutants each contain two mutations. To establish whether both of the mutations are required for constitutive expression of the adaptive response, or if one of them is sufficient to cause constitutive expression, derivatives of \textit{ada} containing single mutations were constructed and the ability of the encoded Ada proteins to induce expression of the \textit{ada} gene was examined.

The \textit{ada-11}, \textit{ada-31} and \textit{ada-41} genes contain G to A transition mutations at nucleotides +17 and +378. An EcoRI restriction site is located within the \textit{ada} gene between the sites of these two mutations. Mutated \textit{ada} genes containing a single G to A transition at nucleotide +17 or nucleotide +378 were constructed by cleaving \textit{ada}^+ and \textit{ada-41} DNA with EcoRI, and ligating the \textit{ada}^+ and \textit{ada-41} gene fragments together.

The 3.1 kb \textit{HindIII} fragment of \textit{E. coli} F26 chromosomal DNA carrying the \textit{ada}^+ gene was subcloned from pCS42 (Sedgwick, 1983) into the \textit{HindIII} site of pHSG415 to generate plasmid pSH1. The chromosomal DNA insert in pSH1 was in the same orientation as the insert containing the \textit{ada-41} gene in pSH41 (see figure 3.4). Plasmids pSH1 and pSH41 were digested with EcoRI and the reaction products, a 1.9kb and an 8.3kb DNA fragment, were separated by agarose gel electrophoresis and purified. The 1.9kb EcoRI fragment of pSH41 was ligated to the 8.3kb EcoRI fragment of pSH1 to form plasmid pSH12 containing an intact \textit{ada} gene, designated as \textit{ada-12}, with a single G to A transition mutation at nucleotide +17. Likewise, the 1.9kb fragment of pSH1 was ligated to the 8.3kb fragment of pSH41 to generate plasmid pSH13 containing an intact \textit{ada} gene, termed \textit{ada-13}, with a single G to A transition mutation at nucleotide +378 (DNA sequencing confirmed that the \textit{ada} genes in pSH12 and pSH13 had only one mutation). To avoid the synthesis of Ada protein from mRNA transcripts initiated at the promoter of the \textit{kan} gene, the orientation of the 3.1kb \textit{HindIII} fragments of pSH12 and pSH13 were reversed, forming plasmids pSHR12 and
pSHR13. To determine whether the Ada-12 and Ada-13 proteins could induce expression of the ada gene, the O⁶-MeGua-DNA methyltransferase levels in E. coli GW7101 transformants containing pSHR12 and pSHR13 were determined.

Two cell extracts of BS41, GW7101, GW7101/pSHR1, GW7101/pSHR41, GW7101/pSHR12 and GW7101/pSHR13 were prepared from independent cultures grown at 30°C. The O⁶-MeGua-DNA methyltransferase activity in samples of extract containing 0.25, 0.5, 1.0, 1.5 and 2.0µg of protein was assayed. The units of activity per mg of protein are shown in table 3.3. GW7101/pSHR12 contained levels of methyltransferase that were below the lower limits of detection of the assay, indicating that the Ada-12 protein, in which cysteine-6 is replaced by tyrosine, does not act as a constitutive inducer of ada gene expression. However, the levels of O⁶-MeGua-DNA methyltransferase in GW7101/pSHR13 were 2-3 fold higher than those in BS41, demonstrating that the Ada-13 protein, which contains isoleucine in place of methionine at amino acid 126, is a strong activator of ada expression. The levels of methyltransferase detected in GW7101/pSHR13 were slightly lower than those in GW7101/pSHR41. Additional experimental evidence suggested that this difference was not significant. In an experiment using HB101 as the host strain for the pSHR series of plasmids, the levels of O⁶-MeGua-DNA methyltransferase in HB101/pSHR13 were high and directly comparable with those detected in HB101/pSHR41. This observation indicated that the Ada-13 protein is as effective an inducer of ada expression as the Ada-41 protein.

The BS11, BS31 and BS41 mutants synthesize high cellular levels of both the Ada protein and N³-MeAde-DNA glycosylase II (the product of the alkA gene) in the absence of alkylating agents (Lindahl et al, 1983). BS21 constitutively synthesizes high levels of Ada protein, but contains only slightly larger amounts of N³-MeAde-DNA glycosylase II than uninduced wild-type E. coli (Karran et al, 1982a; Lindahl et al, 1983). Like the Ada-11, Ada-31 and Ada-41 proteins, the Ada-21 protein has the Met¹²⁶ to Ile substitution which was shown in the experiments described above to be sufficient to convert the Ada protein into a constitutive activator of ada gene expression. The Ada-
Table 3.3: O₆-MeGua-DNA methyltransferase activity in cell extracts of the GW7101 transformants containing plasmids pSHR1, pSHR41, pSHR12 and pSHR13.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Amino acid substitutions in plasmid-encoded Ada protein</th>
<th>O₆-MeGua-DNA methyltransferase. units/mg of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS41</td>
<td>None</td>
<td>13.4, 8.9</td>
</tr>
<tr>
<td>GW7101</td>
<td>Cys₆ to Tyr, Met¹²₆ to Ile</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>GW7101/pSHR1</td>
<td>None</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>GW7101/pSHR41</td>
<td>Cys₆ to Tyr, Met¹²₆ to Ile</td>
<td>32.3, 45.5</td>
</tr>
<tr>
<td>GW7101/pSHR12</td>
<td>Cys₆ to Tyr</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>GW7101/pSHR13</td>
<td>Met¹²₆ to Ile</td>
<td>23.5, 28.3</td>
</tr>
</tbody>
</table>

One unit of methyltransferase removes 1 picomole of methyl groups from O₆-MeGua in DNA under the standard assay conditions.
21 protein differs from the Ada-11, Ada-31, and Ada-41 proteins in its second amino acid substitution. Ada-21 protein has the glutamic acid residue at amino acid 97 replaced by lysine and the Ada-11, Ada-31 and Ada-41 proteins have the Cys\(^6\) to Tyr substitution as the second amino acid change. Thus, it was considered possible that the constitutive expression of the \(alkA\) gene observed in BS11, BS31 and BS41 might require the Cys\(^6\) to Tyr substitution in the Ada protein.

To determine whether the Cys\(^6\) to Tyr substitution was necessary to convert the Ada protein into an activator of \(alkA\) gene expression, *E. coli* BK2012 (tag\(^{-}\) ada\(^{+}\)) were transformed with plasmids pSHR1, pSHR12, pSHR13 and pSHR41, and the N\(^3\)-MeAde-DNA glycosylase activity in the transformants was assayed. Since BK2012 is deficient in N\(^3\)-MeAde-DNA glycosylase I, the product of the tag gene, the activity assayed was that of N\(^3\)-MeAde-DNA glycosylase II synthesized from the chromosomal \(alkA\) gene. The levels of the enzyme in the untransformed host strain and the four different transformants are shown in table 3.4. Two estimates are given for each strain, which represent the enzyme activity measured in two cell extracts prepared from independent cultures grown at 30°C. For the range of protein concentrations used in the assays, which was 1-10\(\mu\)g of protein per assay mixture, the minimum level of glycosylase that could be detected was approximately 1 unit/mg of protein. Those extracts in which there was no apparent glycosylase activity have been assigned as having a N\(^3\)-MeAde-DNA glycosylase II level of <1.0 unit/mg of protein.

No N\(^3\)-MeAde-DNA glycosylase II activity was detected in the untransformed BK2012 cells. Karran *et al* (1980) had previously observed low levels of N\(^3\)-MeAde-DNA glycosylase II in ammonium sulphate fractionated extracts of BK2012 resulting from a basal level of expression of the \(alkA\) gene. The basal level of expression of \(alkA\) was therefore below the lower limits of detection of the assay at the range of protein concentrations which were employed. As expected, BK2012/pSHR41 contained high levels of glycosylase activity as a result of the Ada-41 protein acting as a constitutive inducer of \(alkA\) expression. The BK2012/pSHR12 strain contained no detectable N\(^3\)-MeAde-DNA glycosylase II activity, clearly showing that the Cys\(^6\) to Tyr amino acid
Table 3.4: N³-methyladenine-DNA glycosylase II activity in cell extracts of the BK2012 transformants containing plasmids pSHR1, pSHR12, pSHR13 and pSHR41.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Amino acid substitutions in plasmid-encoded Ada protein</th>
<th>N³-MeAde-DNA glycosylase II units/mg of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>BK2012</td>
<td>None</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>BK2012/pSHR1</td>
<td>Cys⁶ to Tyr, Met¹²⁶ to Ile</td>
<td>37.8, 23.4</td>
</tr>
<tr>
<td>BK2012/pSHR41</td>
<td>None</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>BK2012/pSHR12</td>
<td>Cys⁶ to Tyr</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>BK2012/pSHR13</td>
<td>Met¹²⁶ to Ile</td>
<td>16.9, 24.2</td>
</tr>
</tbody>
</table>

One unit of glycosylase removes 1 picomole of methylated base from alkylated DNA under the standard assay conditions.
substitution alone does not convert the Ada protein into a strong activator of \textit{alkA} expression. However, BK2012/pSHR13 contained levels of the enzyme which were comparable with those detected in BK2012/pSHR41, demonstrating that the Met$^{126}$ to Ile amino acid substitution, on its own, is sufficient to convert the Ada protein into a constitutive activator of \textit{alkA} expression.

In summary, the analysis of the levels of O$^6$-MeGua-DNA methyltransferase and N$^3$-MeAde-DNA glycosylase II in \textit{E. coli} containing plasmids pSHR12 and pSHR13 demonstrated that the Met$^{126}$ to Ile substitution converts the Ada protein into a constitutive activator of expression of the \textit{ada} and \textit{alkA} genes. Ada protein containing Cys$^6$ to Tyr as the only amino acid substitution does not act as an inducer of \textit{ada} or \textit{alkA} expression. Furthermore, the presence of the Cys$^6$ to Tyr substitution in Ada protein containing the Met$^{126}$ to Ile substitution does not appear to have a significant effect on the capacity of the protein for inducing expression of the \textit{ada} and \textit{alkA} genes.

The second mutation in the \textit{ada-21} gene is at nucleotide +289 and results in the Ada-21 protein having a lysine residue instead of glutamic acid at amino acid 97. Unfortunately, an \textit{ada} gene containing this mutation alone could not be constructed in the low copy number vector due to the lack of a suitable restriction site between the two mutations in the \textit{ada-21} gene. However, the finding that the Met$^{126}$ to Ile substitution converted the Ada protein into a strong activator of \textit{alkA} gene expression suggested a possible effect of the Glu$^{97}$ to Lys substitution. The Ada-21 protein contains the Met$^{126}$ to Ile substitution, but BS21 does not synthesize high levels of N$^3$-MeAde-DNA glycosylase II. It therefore appeared that the Glu$^{97}$ to Lys amino acid substitution in the Ada-21 protein suppressed its ability to constitutively activate expression of the \textit{alkA} gene. Unexpectedly, when plasmid pSHR21 was transformed into BK2012 and the cellular levels of N$^3$-MeAde-DNA glycosylase II were assayed to confirm that the Ada-21 protein cannot induce expression of the \textit{alkA} gene, high levels of the enzyme were detected in the transformants. Indeed, the glycosylase levels in two different cell extracts of BK2012/pSHR21 were 14.1 and 22.3 units/mg of protein, within the
range detected in BK2012/pSHR13 and BK2012/pSHR41. The significance of this observation will be considered in the Discussion.

**H. E. coli containing plasmids pSHR13 and pSHR41 do not exhibit an increased resistance to killing by an alkylating agent**

The studies on the ability of the Ada-12, Ada-13 and Ada-41 proteins to induce expression of the *ada* and *alkA* genes failed to reveal a role for the Cys⁶ to Tyr amino acid substitution in Ada protein. However, the occurrence of this substitution in the Ada proteins synthesized by three of the four constitutive mutants suggested that it was likely to have some physiological effect. The constitutive mutants were isolated on the basis of their increased resistance to killing by an alkylating agent (Sedgwick and Robins, 1980). To determine whether the synthesis of Ada protein containing both the Cys⁶ to Tyr and Met¹²⁶ to Ile amino acid substitutions conferred a greater resistance to cell killing than the synthesis of Ada protein with the single Met¹²⁶ to Ile substitution, the methylmethanesulphonate (MMS) sensitivity of the GW7101 strains containing plasmids pSHR1, pSHR12, pSHR13 and pSHR41 was examined.

Samples of a mid-exponential culture of each strain were treated with 0, 0.25 or 0.5%(v/v) MMS for 30 minutes at 30°C. The ratio of the numbers of surviving cells in the MMS-treated cell suspensions relative to the total number of viable cells in the control cell suspension (no MMS) was determined. Figure 3.14 shows graphs of percent (%) cell survival versus MMS concentration for the four different strains. GW7101/pSHR1, GW7101/pSHR12 and GW7101/pSHR13 showed a similar degree of sensitivity to MMS, and GW7101/pSHR41 was slightly more sensitive.

The similarity in the sensitivity of strains GW7101/pSHR13 (Met¹²⁶ to Ile) and GW7101/pSHR1 (wild-type Ada protein) was unexpected. *E. coli* alkA mutants are highly sensitive to killing by MMS (Yamamoto *et al.*, 1978), demonstrating that N³-MeAde-DNA glycosylase II confers protection against the cytotoxic effects of this alkylating agent. GW7101/pSHR13 was shown above to have constitutively-high

119
Figure 3.14: Sensitivity of GW7101 strains containing plasmids pSHR1, pSHR12, pSHR13 and pSHR41 to methylmethanesulphonate.
cellular levels of this enzyme. This strain was therefore expected to be less sensitive to MMS than GW7101/pSHR1 in which the \textit{alkA} gene would not be induced by the low cellular levels of wild-type Ada protein prior to exposure to the alkylating agent. GW7101/pSHR41 (Cys\textsuperscript{6} to Tyr, Met\textsuperscript{126} to Ile) exhibited a greater sensitivity to killing by MMS than GW7101/pSHR13, suggesting that the synthesis of Ada protein containing the Cys\textsuperscript{6} to Tyr and Met\textsuperscript{126} to Ile substitutions does not confer a greater resistance to killing than the synthesis of Ada protein containing only the Met\textsuperscript{126} to Ile substitution. However, the fact that GW7101/pSHR41 did show a greater sensitivity than GW7101/pSHR13 and GW7101/pSHR12 (Cys\textsuperscript{6} to Tyr) indicated that the Cys\textsuperscript{6} to Tyr substitution does have some physiological effect when it occurs in combination with the Met\textsuperscript{126} to Ile substitution.

A further observation concerning the growth of \textit{E. coli} containing the pSHR plasmids has suggested that both the Cys\textsuperscript{6} to Tyr and Glu\textsuperscript{97} to Lys amino acid substitutions have a physiological effect when present with the Met\textsuperscript{126} to Ile substitution. Colonies of strains containing plasmids pSHR21 or pSHR41 grow more slowly on L-agar plates than strains containing pSHR1, pSHR12 or pSHR13. The possible significance of these observations on the MMS sensitivity and growth of strains carrying the pSHR plasmids will be considered further in the discussion.
DISCUSSION

The ada gene has been cloned from each of four independently derived E. coli mutants which constitutively express the adaptive response to alkylating agents. Nucleotide sequence analysis has shown that each gene contains two GC to AT transition mutations within the coding region which will result in the synthesis of a mutant Ada protein containing two amino acid substitutions within the N-terminal region of the protein. Three of the mutant strains have the same two ada mutations. E. coli carrying the mutated ada genes on recombinant plasmids overexpressed both the mutated ada gene and the chromosomal alkA gene, indicating that the mutant Ada proteins act as strong inducers of ada and alkA gene expression in vivo in the absence of DNA alkylation damage. The Ada protein purified from one of the mutant strains has been shown to be a strong inducer of expression of the wild-type ada gene in an in vitro coupled transcription-translation system. One amino acid substitution, Met₁²⁶ to Ile, occurred in the Ada protein synthesized by all four mutants. This substitution alone has been found to be sufficient to convert the Ada protein into a strong inducer of ada and alkA gene expression in vivo.

A number of recent observations on the structure of the wild-type Ada protein provide some insight into the possible mechanisms by which the Met₁²⁶ to Ile substitution might convert the Ada protein into a strong activator of ada and alkA gene expression. Thus, prior to discussing the observations on the mutant Ada proteins, I will briefly review the recent findings on the wild-type protein.

A property of the Ada protein that is now better understood as a result of recent studies is its susceptibility to cleavage by an endogenous proteolytic activity. Teo et al (1984) first observed that the Ada protein undergoes proteolytic cleavage when E. coli are lysed to generate cell extracts. It is due to this proteolytic processing that the inducible O⁶-MeGua-DNA methyltransferase was originally purified as a 19kDa polypeptide (Demple et al, 1982) and it only became possible to purify the 39 kDa Ada
protein to homogeneity after the ada gene had been cloned into an expression vector (McCarthy and Lindahl, 1985; Nakabeppu et al, 1985b). The site at which the Ada protein is cleaved to yield the 19kDa O^6-MeGua-DNA methyltransferase and a 20 kDa polypeptide corresponding to the N-terminal half of the protein was identified as the peptide bond between Lys^{178} and Gln^{179} (Demple et al, 1985). In their models for the regulation of the adaptive response, Teo et al (1986) and Nakabeppu and Sekiguchi (1986) suggested that the proteolytic cleavage of the Ada protein might be responsible for the down-regulation of the adaptive response following the completion of DNA repair. Teo (1987) has since discovered that the Ada protein is cleaved at another site to generate a 15kDa N-terminal fragment and a 24kDa C-terminal fragment of the protein. This second site of cleavage appears to be the peptide bond between Lys^{129} and Ala^{130} (Teo, 1987). Very recently, Sedgwick (1989) has identified the activity responsible for the proteolytic processing of Ada protein in cell extracts as the OmpT protease, which is located within the outer membrane of E. coli. In intact cells the Ada protein would not be exposed to this enzyme. Indeed, following the induction of the adaptive response the levels of Ada protein in wild-type E. coli and in cells in which the ompT gene has been deleted decay at a similar rate (Sedgwick, 1989), showing that the OmpT protease is not involved in the down-regulation of the adaptive response in vivo. Nevertheless, this observation does not exclude the possibility that the adaptive response might be down-regulated by proteolytic processing of the methylated Ada protein by another protease. However, cleavage of the Ada protein has not been observed in vivo.

One of the sites at which the OmpT protease cleaves the Ada protein lies within a region of the protein molecule which is highly sensitive to attack by a number of proteolytic enzymes. Sedgwick et al (1988) partially digested the purified Ada protein with four different reagent proteases (trypsin, chymotrypsin, subtilisin and V8 protease) to determine if any regions of the Ada protein are particularly sensitive to proteolytic cleavage. All four proteases preferentially cleaved the Ada protein at a site between amino acid residues 171-181 to yield two polypeptides, of similar size,
corresponding to the N and C-terminal halves of the protein molecule. The two fragments of the Ada protein retained their respective DNA repair activities. This data suggests that the intact Ada protein is comprised of two functional domains linked by a protease-sensitive hinge region of approximately ten amino acid residues (Sedgwick et al, 1988).

Using the technique of DNase I footprinting, Teo et al (1986) showed that the Ada protein methylated at Cys⁶⁹ binds to specific nucleotide sequences within the promoter regions of the ada and alkA genes. Employing this same technique, Sedgwick has obtained evidence that the N-terminal domain of the Ada protein contains the DNA binding site which interacts with the promoter regions (see Sedgwick and Hughes, 1988). A mixture of 20kDa N-terminal and 19kDa C-terminal proteolytic fragments of the Ada protein, self-methylated by incubation with MNU-treated DNA, was observed to protect the ada and alkA promoters from digestion by DNase I. However, the methylated purified 19kDa fragment of Ada protein was unable to protect ada or alkA promoter DNA from DNase I digestion. These observations clearly suggest that the DNA binding site of the Ada protein is within the N-terminal domain of the molecule.

Detailed structural studies on the trp repressor and cyclic AMP receptor protein (CRP) of E. coli, and the Cro protein and repressor of bacteriophage λ, have revealed that these regulatory proteins bind to operator DNA as dimers using α-helices to contact adjacent major grooves on one face of the DNA (see Pabo and Sauer, 1984; Schleif, 1988; Brennan and Matthews, 1989a; Brennan and Matthews, 1989b, for reviews). All four of the proteins show amino acid sequence homology within the α-helical regions that interact with DNA. The amino acid sequence motif common to these proteins, referred to as the helix-turn-helix motif, has also been observed in a large number of other regulatory proteins from prokaryotes (see Pabo and Sauer, 1984; Brennan and Matthews, 1989b). The mode of DNA binding used by the trp repressor, CRP, Cro and λ repressor may therefore be widespread. However, a computer search of the amino acid sequence of the Ada protein performed by M. Ginsburg in our laboratory has failed to identify sequences homologous to the set of ten helix-turn-helix motifs
described by Brennan and Matthews (1989b). Furthermore, there is no experimental
evidence for dimerization of the Ada protein. Thus, the actual mode of DNA binding
employed by the Ada protein is unclear.

Perhaps the most relevant of the recent observations on the structure of the Ada
protein are those that have given an indication of how methylation of Cys69 permits the
Ada protein to bind to the promoter regions of the ada and alkA genes and induce their
expression. Teo et al (1986) proposed that self-methylation of the Ada protein by its
repair of a methylphosphotriester lesion might induce a change in the conformation of
the protein molecule which allows it to bind to the promoter regions of the ada and alkA
genes and activate transcription through facilitating the binding of RNA polymerase to
the gene promoters. Sedgwick et al (1988) have found that the N-terminal region of
Ada protein methylated by incubation with MNU-treated DNA is less sensitive to
trypsin digestion than the N-terminal region of the unmethylated Ada protein,
suggesting that methylation induces a change in the conformation of the N-terminal
domain of the protein. Furthermore, it has also been observed that the methylated Ada
protein is less sensitive to cleavage by the endogenous proteolytic activity in cell
extracts than the unmethylated protein (Yoshikai et al, 1988), suggestive of a
methylation-induced conformational change. An alteration in the conformation of the
N-terminal domain of the Ada protein could possibly allow the DNA binding site located
in this region to interact with the specific regulatory sequences within the promoter
regions of the Ada-inducible genes. Using DNase I footprinting, Sakumi and Sekiguchi
(1989) have studied the interaction of RNA polymerase with the promoter region of
the ada gene. When incubated with RNA polymerase alone, the ada promoter was not
protected from digestion by DNase I. Similarly, when ada promoter DNA was incubated
with unmethylated wild-type Ada protein and RNA polymerase together, no protection
from DNase I digestion was observed. However, when incubated with the methylated Ada
protein and RNA polymerase, a large section of the promoter containing the Ada protein
binding site and sequences downstream of this site, including the putative -35 and -10
boxes, was protected from digestion. These results strongly suggest that the binding of
methylated Ada protein to the promoter region of the ada gene facilitates the binding of RNA polymerase to the promoter.

The latter observations on the protease-sensitivity of the unmethylated and methylated Ada proteins, and the binding of RNA polymerase to the promoter region of the ada gene, have indicated that the mechanism by which the Ada protein induces the expression of the genes of the adaptive response bears some similarity to the way in which cyclic AMP receptor protein activates the expression of a number of catabolic operons in E. coli (see de Crombrugghe et al, 1984, for a review on CRP). Upon binding cAMP, CRP undergoes a conformational change and shows an increased affinity to bind to specific nucleotide sequences within the promoter regions of the catabolite-sensitive genes. The binding of the cAMP-CRP complex to the gene promoters stimulates the binding of RNA polymerase, and thereby results in the activation of transcription of the catabolite-sensitive genes.

The work described in this thesis has shown that a single amino acid substitution in the N-terminal domain converts the Ada protein into a strong activator of expression of the ada and alkA genes. The critical substitution, Met$^{126}$ to Ile, is a conservative change in the sense that both amino acids have a hydrophobic R group. If, as evidence now suggests, the methylation of Cys$^{69}$ converts the wild-type Ada protein into an activator of gene expression through inducing a change in the conformation of the protein molecule, then the most plausible mechanism by which the Met$^{126}$ to Ile substitution might cause the constitutive expression of the adaptive response is by inducing a similar conformational change which permits the mutant protein to bind to the promoter regions of the ada and alkA genes and stimulate transcription by RNA polymerase. Mutant forms of CRP which act as inducers of gene expression in the absence of endogenous cAMP have been isolated and shown to have a conformation which is similar to that of the wild-type CRP complexed with cAMP (Blazy and Ullmann, 1986; Harman et al, 1986).

In our laboratory, B. Sedgwick has analysed the protease sensitivity of the Ada-11 protein isolated from the constitutive mutant BS11 and has obtained evidence that it
may have an altered conformation (see Hughes and Sedgwick, 1989). When the proteolytic fragments generated by partial trypsin digestion of the Ada-11 and unmethylated wild-type Ada proteins were compared, it was observed that the two proteins differ in the pattern of appearance and further degradation of fragments arising from tryptic cleavage in the N-terminal domain of the Ada protein. This difference in the sensitivity to trypsin digestion suggests that the N-terminal domain of the Ada-11 protein has an altered conformation. Whether the Ada-11 protein has a similar sensitivity to trypsin as the methylated wild-type Ada protein, and therefore a similar conformation, is presently unknown. The Ada-11 protein has two amino acid substitutions, Cys6 to Tyr and Met126 to Ile, and both may have contributed to the conformational change indicated by the altered sensitivity to proteolysis. It would therefore be highly informative to isolate the Ada-13 protein from GW7101/pSHR13 and compare the trypsin sensitivity of this protein with the wild-type Ada protein to ascertain whether the Met126 to Ile substitution alone alters the conformation of the protein.

Although I favour the Met126 to Ile substitution causing the constitutive expression of the adaptive response through inducing a change in the conformation of the Ada protein, there is an alternative mechanism by which this substitution could convert the Ada protein into a strong activator of ada and alkA gene expression. In this study the unmethylated wild-type Ada protein has been shown to be a weak activator of expression of the ada gene in an in vitro DNA-directed translation system. Teo et al (1986) also observed that the unmethylated Ada protein acted as a weak inducer of in vitro run-off transcription of the ada gene. It therefore appears that the unmethylated Ada protein can bind with a low affinity to the promoter region of the ada gene (and the conformational change induced upon methylation of Cys6 would increase the affinity for binding to promoter DNA through allowing a more favourable interaction of the DNA binding site with the promoter region). Since the DNA binding site of the Ada protein can apparently interact with the promoter region of the ada gene when the protein is in its unmethylated form, albeit very weakly, it is possible that an amino
acid substitution located within the DNA binding site itself could result in the Ada protein binding to promoter DNA with an increased affinity if it permitted improved contacts between the binding site and the DNA. Specific amino acid substitutions within the DNA-binding α-helices of the E. coli trp repressor and the repressor of bacteriophage λ have been found to increase their affinity for binding to operator DNA as a result of allowing the DNA binding site to form tighter contacts with the operator DNA sequences (Kelley and Yanofsky, 1985; Nelson and Sauer, 1985). One of the sites at which the Ada protein is susceptible to proteolytic cleavage in cell extracts is the peptide bond between Lys^{129} and Ala^{130}. It is therefore likely that these two amino acid residues, and those in the close vicinity, are exposed on the exterior of the protein molecule. Thus, Met^{126} might be located on the surface of the Ada protein within the DNA binding site which interacts with the promoter DNA, and its substitution by isoleucine could possibly permit the DNA binding site to form a tighter contact with the promoter DNA.

The two different proposals for the way in which the Met^{126} to Ile amino acid substitution could convert the Ada protein into a strong activator of ada and alkA gene expression both involve this amino acid substitution enhancing the capacity of the Ada protein to bind to the promoter regions of these genes, the tight binding of the unmethylated form of the mutant protein to the promoters stimulating the binding of RNA polymerase to its own regulatory sequences within the gene promoters and leading to an increased level of transcription of the ada and alkA genes. At our current level of understanding of the regulation of the adaptive response by the Ada protein, there does not appear to be an obvious alternative mechanism by which an amino acid substitution within the N-terminal domain might convert the Ada protein into a strong activator of expression of the ada and alkA genes. In order to show that the mutant Ada proteins have an increased affinity to bind to ada and alkA promoter DNA, the DNA binding affinity of the wild-type Ada protein could be compared with the binding affinities of the purified mutant proteins using a filter binding assay.
The Ada proteins synthesized by the four constitutive mutants contain two amino acid substitutions. Since the Met$^{126}$ to Ile substitution common to all four mutant Ada proteins converts the protein into a strong activator of expression of the *ada* and *alkA* genes, the reason for the selection of a second amino acid substitution is unclear. The Cys$^6$ to Tyr substitution occurred in the Ada protein synthesized by three of the mutants, BS11, BS31 and BS41, and is therefore unlikely to have arisen by chance. The engineered Ada-12 protein containing only the Cys$^6$ to Tyr substitution was no more effective an activator of *ada* and *alkA* expression in vivo than the unmethylated wild-type Ada protein. Furthermore, the Ada-13 protein with the Met$^{126}$ to Ile substitution and the Ada-41 protein containing both the Cys$^6$ to Tyr and Met$^{126}$ to Ile substitutions were equally effective at inducing *ada* and *alkA* expression. Thus, the Cys$^6$ to Tyr substitution does not enhance the effect of the Met$^{126}$ to Ile substitution in converting the Ada protein into a strong inducer of expression of the *ada* and *alkA* genes. However, the observations that GW7101/pSHR41 transformants were more sensitive to killing by MMS than GW7101/pSHR12 and GW7101/pSHR13, and *E. coli* transformed with plasmid pSHR41 grew less well on solid media than cells transformed with pSHR12 or pSHR13, suggests that the Cys$^6$ to Tyr substitution does have some physiological effect when present with the Met$^{126}$ to Ile substitution. The Cys$^6$ to Tyr substitution may possibly have been selected if it stabilizes a conformational change induced by the Met$^{126}$ to Ile substitution. Alternatively, it may enhance the ability of the Ada protein to induce the expression of genes other than *ada* and *alkA* whose products might confer an increased resistance to alkylating agents. An increase in the stability of the mutant protein conferred by the Cys$^6$ to Tyr substitution could account for the observations on MMS survival and cell growth, as will be discussed later.

The second amino acid substitution in the Ada protein synthesized by BS21, the mutant strain which does not contain high levels of N$^3$-methyladenine-DNA glycosylase II (Karran *et al.*, 1982a; Lindahl *et al.*, 1983), is Glu$^{97}$ to Lys. When it was discovered that the Met$^{126}$ to Ile substitution converts the Ada protein into a strong activator of
expression of both the *ada* and *alkA* genes, it seemed possible that the Glu97 to Lys substitution might modulate the effect of the Met126 to Ile substitution such that the Ada-21 protein only induces the expression of the *ada* gene. However, BK2012/pSHR21 was subsequently found to contain high levels of N3-MeAde-DNA glycosylase II similar to those detected in BK2012/pSHR13 and BK2012/pSHR41. Thus, the Ada-21 protein appears to be a strong activator of expression of the *alkA* gene. The BS strains were isolated from *E. coli* cultures which had been highly mutagenized and may therefore contain additional mutations outside the *ada* gene. The low level of N3-MeAde-DNA glycosylase II activity in BS21 could be the result of a mutation within the promoter region of the *alkA* gene which prevents induction by the mutant Ada protein, or a mutation within the *alkA* structural gene which results in the synthesis of a mutant glycosylase with reduced repair activity. As proposed for the Cys6 to Tyr substitution in the Ada-11, Ada-31 and Ada-41 proteins, the Glu97 to Lys substitution in the Ada-21 protein may have been selected because it increases the stability of the mutant protein or allows the Ada protein to induce the expression of genes other than *ada* and *alkA*.

The finding that the *ada* genes of the four constitutive strains could not be cloned in the high copy number vector pEMBL18[+] suggests that excessive overproduction of the mutant Ada proteins is lethal to *E. coli*. Some observations on other DNA-binding regulatory proteins indicate a possible reason why the mutant Ada proteins might be cytotoxic. Mutant forms of the *λ* repressor and the *lac* repressor of *E. coli* which have an increased affinity to bind to operator DNA have also been shown to have an increased non-specific DNA binding affinity (Nelson and Sauer, 1985; Pfahl, 1981). When expressed at high levels the tight-binding mutant *λ* repressors are toxic to *E. coli* and their cytotoxicity correlates with the increased non-specific DNA binding affinity of the protein (Nelson and Sauer, 1985). I have postulated that the Ada proteins synthesized by the constitutive mutants are able to act as strong inducers of *ada* and *alkA* gene expression as a result of their having an increased ability to bind to the promoter regions of these genes. It is possible that these mutant Ada proteins may also
have a propensity to bind non-specifically to DNA in a manner analogous to the mutant \( \lambda \) and \( lac \) repressors. The synthesis of very high levels of such a protein from one of the mutated \( ada \) genes carried on a high copy number vector could severely disrupt normal DNA metabolism (since the protein will bind all over the chromosomal DNA), consequently leading to cell death.

The proposed cytotoxicity of the mutant Ada proteins could explain a number of unexpected observations on the MMS sensitivity and growth of \( E. coli \) carrying members of the pSHR series of plasmids. When the MMS sensitivity of strains GW7101/pSHR1 (Ada\(^+\)), GW7101/pSHR13 (Met\(^{126}\) to Ile) and GW7101/pSHR41 (Cys\(^6\) to Tyr, Met\(^{126}\) to Ile) was compared, strains GW7101/pSHR1 and GW7101/pSHR13 showed a similar degree of sensitivity to killing and GW7101/pSHR41 was slightly more sensitive. Since the Met\(^{126}\) to Ile substitution converts the Ada protein into a strong activator of expression of the \( alkA \) gene, one would have expected the strains carrying plasmids pSHR13 and pSHR41 to be less sensitive to killing by MMS than the strain carrying pSHR1 because the former two strains contain high cellular levels of N\(^3\)-MeAde-DNA glycosylase II prior to the treatment with the alkylating agent. The unexpected MMS sensitivity of GW7101/pSHR13 and GW7101/pSHR41 may have resulted from the mutant Ada proteins binding non-specifically to the cellular DNA and blocking access of the glycosylase to the cytotoxic N\(^3\)-MeAde and N\(^3\)-MeGua lesions. Alternatively, the fact that the mutant proteins are deleterious to the cell might render GW7101/pSHR13 and GW7101/pSHR41 generally less capable of tolerating damage to the DNA, and the benefit of having high cellular levels of N\(^3\)-MeAde-DNA glycosylase II is therefore counteracted. The Cys\(^6\) to Tyr substitution in the Ada-41 protein may stabilize a conformational change caused by the Met\(^{126}\) to Ile substitution. The increased stability of the toxic mutant Ada protein could account for the greater MMS sensitivity of GW7101/pSHR41 compared with GW7101/pSHR13 and the observation that \( E. coli \) containing pSHR41 grow less well on agar plates than cells containing pSHR13.

A number of studies in other laboratories have identified mutations within the \( ada \)
gene which result in an increased level of expression of this gene in the absence of alkylation damage. Nakamura et al (1988) used synthetic oligonucleotides to construct a library of random base substitution mutations within the promoter region of the ada gene. Their effect on ada gene expression was examined by placing the lacZ structural gene under the control of the mutated ada promoters in a plasmid vector and measuring the levels of β-galactosidase synthesized in E. coli containing these constructs after growth in medium with or without MMS. The β-galactosidase levels were compared with those detected in E. coli containing the same plasmid, but with the lacZ structural gene placed under the control of the wild-type ada promoter. In the absence of MMS, two single mutations within the ada promoter resulted in the synthesis of levels of β-galactosidase which were considerably higher than the level of enzyme synthesized from the lacZ gene under the control of the wild-type ada promoter. One of these mutations, an AT to TA transversion, occurred in the region of the promoter that is protected from DNase I digestion (in footprinting experiments) by the methylated Ada protein. This mutation could possibly increase the expression of the ada gene by permitting the unmethylated Ada protein to bind to the ada promoter and stimulate transcription by RNA polymerase. The second mutation, a GC to TA transversion, was sited between the -35 and -10 boxes. This mutation may increase the expression of the ada gene by allowing RNA polymerase to bind to the promoter without the binding of methylated Ada protein. An important observation is that the levels of enzyme synthesized from lacZ under the control of the wild-type ada promoter in MMS-treated E. coli were 3-4 fold higher than those synthesized from lacZ under the control of the two mutated promoters in cells grown in the absence of MMS. Thus, although the two promoter mutations result in an elevated level of expression of the ada gene in the absence of alkylation damage, the degree of overexpression is not as considerable as that observed in the four BS mutants, which in the absence of alkylating agents synthesize even higher levels of Ada protein than adapted wild-type E. coli.

Using site-directed mutagenesis, Takano et al (1988) constructed mutated ada genes which encoded Ada proteins in which the methyl group acceptors, Cys69 and
Cys\textsuperscript{321}, were replaced by alanine residues. The ability of the unmethylated and methylated forms of the mutant Ada proteins to induce \textit{in vitro} transcription of the \textit{ada} and \textit{alkA} genes was examined. When the unmethylated Ada protein containing the Cys\textsuperscript{321} to Ala amino acid substitution was added to an \textit{in vitro} transcription system containing \textit{ada} or \textit{alkA} promoter DNA, a small increase in the level of transcription of the \textit{ada} and \textit{alkA} templates was observed. However, addition of methylated Ada protein containing the Cys\textsuperscript{321} to Ala substitution induced a large increase in the level of \textit{ada} and \textit{alkA} transcription. Indeed, this mutant Ada protein exhibited transcription-promoting properties which were essentially identical to the wild-type Ada protein. With the mutant Ada protein containing the Cys\textsuperscript{69} to Ala amino acid substitution, neither the unmethylated or methylated forms of the protein were able to stimulate transcription of the \textit{ada} and \textit{alkA} DNA. This data confirms the observation of Teo \textit{et al} (1986) that the methylation of Cys\textsuperscript{69} converts the Ada protein into a strong activator of transcription of the \textit{ada} and \textit{alkA} genes. An interesting observation was made when the ability of the mutant proteins to induce expression of the \textit{ada} gene was examined \textit{in vivo}. The wild-type and mutated \textit{ada} structural genes placed under the control of the \textit{lac} promoter were inserted into a plasmid vector which contained the gene encoding the enzyme chloramphenicol acetyltransferase (CAT) under the control of the wild-type \textit{ada} promoter. The three plasmids were transformed into an \textit{E. coli} strain deleted for the chromosomal \textit{ada} operon. By assaying the levels of CAT synthesized in \textit{E. coli} grown in medium lacking an alkylating agent, the capacity of the unmethylated wild-type and mutant Ada proteins for inducing the expression of the \textit{ada} gene were compared. Similarly, by measuring the levels of CAT synthesized in cells exposed to MMS, the ability of the methylated Ada proteins to induce \textit{ada} gene expression was examined. In \textit{E. coli} synthesizing wild-type Ada protein, a low level of CAT activity was detected in cells grown in the absence of MMS. Treatment of these bacteria with MMS resulted in a 10-fold increase in the cellular levels of CAT due to the methylated Ada protein inducing transcription of the \textit{cat} gene from the \textit{ada} promoter. With the \textit{E. coli} synthesizing the mutant Ada protein containing the Cys\textsuperscript{69} to Ala amino acid
substitution, a very low level of CAT was detected in cells in the absence of MMS, and no induction of CAT synthesis occurred when the bacteria were exposed to the alkylating agent. However, without being exposed to MMS, the cells synthesizing the mutant Ada protein containing the Cys\textsuperscript{321} to Ala substitution produced levels of CAT which were comparable to those detected in MMS-treated \textit{E. coli} synthesizing wild-type Ada protein. Thus, the unmethylated form of the mutant Ada protein containing the Cys\textsuperscript{321} to Ala amino acid substitution acted as a strong inducer of \textit{ada} gene expression \textit{in vivo}. It should be noted that \textit{E. coli} having this amino acid substitution in the Ada protein would not exhibit a greatly increased resistance to alkylating agents because the mutant protein is unable to repair O\textsuperscript{6}-MeGua and O\textsuperscript{4}-MeThy.

Shevell \textit{et al} (1988) constructed a series of plasmids carrying ordered deletions from the 3' end of the \textit{ada} gene which encoded fusion proteins deriving their N-terminal region from the \textit{ada} gene and their C-terminal region from transcribed vector (pBR322) DNA sequences. Fusion proteins which contained 278, 308, 309 and 315 \textit{ada} specified amino acids acted as strong inducers of expression of the \textit{ada} gene \textit{in vivo} in the absence of alkylation damage. The nature of the vector-specified amino acid sequence at the C-terminus of the fusion proteins strongly affected the degree of gene activation. It is of note that these fusion proteins were considered to be cytotoxic when synthesized at high levels because the plasmids which encoded them mutated to reduce the plasmid copy number.

The observations of Takano \textit{et al} (1988) and Shevell \textit{et al} (1988) that alterations to the C-terminal domain of the Ada protein can convert the protein into a strong inducer of expression of the \textit{ada} gene have suggested that this domain not only functions to repair O\textsuperscript{6}-MeGua and O\textsuperscript{4}-MeThy lesions, but it may also play a mechanistic role in the regulation of gene expression by the Ada protein. The concept that the C-terminal region of the Ada protein might be involved in gene activation was first proposed by Demple (1986) following the observation that two \textit{ada}\textsuperscript{−} mutants that synthesize Ada proteins which repair O\textsuperscript{6}-MeGua in DNA at a much reduced rate, but are not defective in the repair of methylphosphotriesters, exhibit delayed induction of \textit{ada} and \textit{alkA} gene
expression upon exposure to MNNG. The precise function of the C-terminal domain in the regulation of gene expression is unknown, although Demple (1986) has suggested that it may be involved in interacting with RNA polymerase at the gene promoter.

Undoubtedly the most interesting aspect of the regulation of the adaptive response that has yet to be determined is the way in which the Ada protein specifically interacts with the promoter regions of the inducible genes. Since the Ada protein does not contain an apparent helix-turn-helix amino acid sequence motif, it may utilize a mode of DNA binding which has not been previously characterized. Ultimately, the amino acid sequences and the structural features of the protein which are involved in its interaction with promoter DNA are likely to be identified by a combination of physical, biochemical and genetic analyses. The mutant Ada-13 protein which acts as a constitutively strong inducer of *ada* and *alkA* gene expression may prove to be a useful tool in these studies. X-ray crystallographic analysis of the Ada protein is likely to provide the most detailed information on the overall structure of the Ada protein. Moody and Demple (1988) have recently crystallized the purified 19kDa C-terminal fragment of the Ada protein containing the O6-MeGua-DNA methyltransferase activity. However, X-ray diffraction studies on this 19kDa fragment are unlikely to reveal the structural features of the Ada protein involved in its specific interaction with the promoter DNA of the *ada* and *alkA* genes. To gain insight into the way in which the Ada protein interacts with promoter DNA, the form of the Ada protein which should ideally be analysed by X-ray crystallography is the 39kDa wild-type protein which has been methylated at Cys69 and then co-crystallized with promoter DNA. Unfortunately, at the present time the methylated form of the Ada protein cannot be purified in sufficient quantities to allow physical studies to be undertaken. The Ada-13 protein acts as a strong inducer of *ada* and *alkA* gene expression, possibly as a result of the Met126 to Ile substitution inducing a change in the conformation of the protein so that mutant protein has a similar conformation to the wild-type Ada protein methylated at Cys69. This mutant protein could easily be purified in large quantities from *E. coli* GW7101/pSHR13. X-ray crystallographic analysis of the Ada-13 protein co-
crystallized with ada promoter DNA might therefore provide an excellent means of studying the Ada protein-promoter DNA interaction. A genetic approach which has been taken to identify amino acid residues in the λ and trp repressors which are involved in interacting with operator DNA has been to characterize mutations which reduce the ability of the repressor to bind to operator DNA (Hecht et al, 1983; Kelley and Yanofsky, 1985). The amino acid substitutions which reduced the operator DNA binding affinity were observed to cluster within the helix-turn-helix regions of the repressor molecules. Assuming that the Ada-13 protein is shown to have an increased affinity for binding to promoter DNA, then a similar approach using this mutant Ada protein could theoretically be taken to identify amino acid residues involved in the interaction with promoter DNA. In this case one would characterize mutations within the ada-13 gene which prevent the Ada-13 protein from acting as a strong inducer of ada or alkA expression.

In conclusion, this study on the four E. coli mutants which constitutively express the adaptive response in the absence of alkylation damage has provided new insights into the mechanism by which the Ada protein acts as a regulator of gene expression. The use of a gene-activating mutant Ada protein constructed in this work in future structural studies may prove to be extremely helpful in the characterization of the potentially novel mode of DNA binding employed by this regulatory protein.
REFERENCES
REFERENCES

Bamborschke, S., O'Connor, P. J., Margison, G. P., Kleihues, P., and Maru, G. B.


**References added in proof**


PUBLISHED WORK
The Adaptive Response to Alkylation Damage

CONSTITUTIVE EXPRESSION THROUGH A MUTATION IN THE CODING REGION OF THE ada GENE

(Received for publication, June 26, 1989)

Stephen J. Hughes and Barbara Sedgwick
From the Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Potters Bar, Herts, EN6 3LD, United Kingdom

We have shown by genetic mapping, molecular cloning, and DNA sequencing that four Escherichia coli mutants, which express the adaptive response to alklylation damage constitutively, are mutated in the ada gene. All four mutant ada genes have two GC to AT transition mutations in the coding region and encode altered Ada proteins with two amino acid substitutions in the N-terminal domain. E. coli carrying the mutated ada genes on recombinant plasmids overexpressed both the mutated ada gene and the chromosomal alkA gene. This observation indicates that the mutant Ada proteins act as strong positive regulators of the ada and alkA genes in the absence of DNA alkylation. One mutant protein, Ada-11, was shown to be a strong activator of ada gene expression in a cell-free system. An altered pattern of tryptic digestion of the Ada-11 protein compared with the wild-type Ada protein suggested that it has a different conformation. One amino acid substitution, namely methionine residue 126 replaced by isoleucine, occurred in all four mutant Ada proteins, and this mutation alone was sufficient to convert the Ada protein into a strong activator of ada and alkA gene expression in vivo.

The adaptive response to alkylation damage is induced in E. coli on exposure to methylating agents such as N-methyl-N′-nitro-N-nitrosoguanidine (MNNG) and N-methyl-N-nitrosourea (MNU). Induced cells are more resistant to the mutagenic and cytotoxic effects of alkylating agents (1, 2) as a result of the increased expression of at least four genes, ada, alkA, alkB, and aidB (for recent reviews, see Refs. 3 and 4).

The ada gene product has two major roles. The first is to act as the positive regulator of expression of the inducible genes of the adaptive response (5, 6), and the second is to repair the major mutagenic lesions produced by methylating agents, O6-methylguanine and O4-methylthymine in DNA. The Ada protein, O6-methylguanine-DNA methyltransferase, demethylates O6-methylguanine and O4-methylthymine and transfers the methyl groups on to one of its own cysteine residues, Cys-321. It also demethylates innocuous methylphosphodiester in DNA and transfers the methyl groups onto a different cysteine residue, Cys-69 (7–13). Self-methylation at cysteine 69 by repair of a methylphosphodiester converts the Ada protein from a weak to a strong activator of transcription. This modified protein binds to a specific nucleotide sequence in the promoter regions of the inducible genes of the adaptive response and thereby stimulates their transcription (5, 6, 14). An altered sensitivity of the methylated Ada protein to proteolytic digestion supports the hypothesis that the modified protein undergoes a conformational change which enhances its affinity for the regulatory DNA sequences (13, 15). The primary structure of the 39-kDa Ada protein has been deduced from the nucleotide sequence of the cloned ada gene and from partial amino acid sequencing data (9, 16). The protein has two major functional and structural domains (13). The 20-kDa N-terminal domain participates in specific DNA binding and the repair of methylphosphodiesterases (13, 17), whereas the 19-kDa C-terminal domain demethylates O6-methylguanine and O4-methylthymine (8, 10).

The ada gene product controls its own expression and also that of the alkA, alkB, and aidB genes. The alkA gene product, 3-methyladenine-DNA glycosylase II, excises the cytotoxic lesions 3-methyladenine and 3-methylguanine and also O4-methylated pyrimidines from alkylated DNA (10, 18–20). The functions of the alkB (21, 22) and aidB (23) gene products are unknown.

Four independently derived constitutive mutants of the adaptive response have been isolated. An E. coli B strain was mutagenized with MNNG, and the mutants were selected by six cycles of treatment with a toxic dose of MNU. They are even more resistant to alkylating agents than the wild-type strain in which the adaptive response has been induced (24). Three of the mutants, BS11, BS31, and BS41, constitutively overproduce both the Ada and AlkA proteins. The fourth mutant, BS21, constitutively synthesizes Ada protein but has only slightly elevated cellular levels of the 3-methyladenine-DNA glycosylase II activity (19, 25). These mutants have been employed as the source of overexpressed enzymes in previous biochemical studies (3). The mutation responsible for the phenotype of BS21 has been mapped at, or very close to, the ada gene at 47 min on the E. coli genetic map (26). The constitutive phenotypes could be a result of mutations in the ada promoter, the ada structural gene, or an unidentified repressor gene. An altered ada promoter would allow constitutive expression of the ada gene, an altered Ada protein would act as a strong positive regulator of the ada and alkA genes in the absence of methylation damage, and an inactive repressor would fail to down-regulate the response.

In this paper we show that the constitutive phenotype of the four mutants is a consequence of a single transition mutation within the coding region of the ada gene. An amino acid substitution in the Ada protein which converts the unmethylated protein into a strong positive autoregulator has been identified.
Gene Activation by Mutated Ada Proteins

MATERIALS AND METHODS

Bacterial Strains and Plasmids—The E. coli K12 strains used were HB101 [F', hisd205 lacI22(2) metE4 (B. Sedgwick, unpublished data)], F' proA::lac2 ara-14 galE2 lacY1 [37], and E. coli B strains used were F' (his thy supE lacZ mbl 166), DHL101 (HB101/F' pro lac: Tn3), GW7101 (F' pro lac: Tn3 lacZ metE4 (B. Sedgwick, unpublished data)), and E. coli BL21(DE3) (Stratagene). Bacterial media, cultivation conditions, and plasmids were described previously (38). Media and culture conditions—Bacteria were cultured at 37°C on LB medium supplemented with 20 μg of tetracycline/ml for F'-derived strains (33). Ampicillin was added to broth and agar at 50 and 100 μg/ml, respectively, to select for strains carrying pEMBL-derived plasmids. Strains transformed with the temperature-sensitive plasmid pSHG415-derived plasmids were grown in the presence of 20 μg/ml ampicillin at 30°C.

Enzymes and Reagents—Restriction endonucleases, T4 DNA ligase, and Klenow fragment of DNA polymerase I were purchased from New England Biolabs and Bethesda Research Laboratories. calf intestinal phosphatase was from Boehringer Mannheim. The Sequenase DNA sequencing kit was obtained from United States Biochemical Corporation. The prokaryotic DNA-directed translation kit radiolabeled deoxyribonucleotides and L-[35S]methionine were supplied by Amersham International PLC.

Recombinant DNA Techniques—Recombinant DNA techniques were as described by Maniatis et al. (34). Nucleotide sequencing was by the Maxam and Gilbert method (35) or the chain termination method (36). For Maxam-Gilbert sequencing, the plasmid DNA was digested with restriction endonuclease HindIII or EcoRI and the products end-labeled with [α-32P]dCTP or [α-32P]dATP, respectively, using the Klenow fragment of E. coli DNA polymerase I. The end-labeled DNA was digested with a second restriction endonuclease, Sall or HindIII, and the fragment to be sequenced was purified by polyacrylamide gel electrophoresis. Both DNA strands were sequenced. For chain termination sequencing, an EcoRI-Smal fragment of the ada structural gene was subcloned into the pEMBL8-1 plasmid. Single-stranded DNA was prepared as described by Dente et al. (30), and the coding strand of the ada gene sequences using Sequenase enzyme (27) and a set of five oligonucleotide primers synthesized on an Applied Biosystems 380B DNA synthesizer. Wherever the sequence required clarification, the opposite DNA strand was examined by double-stranded chain termination sequencing.

Immunological Screening for High Cellular Levels of Ada Protein—Replica-plated transformants were transferred to nitrocellulose filters and suspended in a chloroform vapor tank for 60 min (38). Each filter was placed in 15 ml of 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM MgCl2, containing 3% (w/v) bovine serum albumin, 1 μg/ml DNase I and 40 μg/ml lysozyme, and agitated gently for 5 h. All incubations were at room temperature. After washing in 50 ml Tris-HCl, pH 7.4, 200 mM NaCl, 0.05% (v/v) Nonidet P-40 (TSN buffer), the filters were slowly agitated for 16 h in TSN, 10% horse serum, and 0.05% (v/v) hydrogen peroxide for 10 min.

Enzyme Assays—Bacteria were harvested from 20-ml mid-exponential cultures and sonicated in 150 μl of 50 mM Tris-HCl, pH 8.0, 10 mM dithiothreitol, and 1 mM EDTA. The protein concentration of the cell extracts was determined by the method of Bradford (39). The O6-adenosine DNA methyltransferase activity of the extracts was assayed by monitoring the demethylation of O6-[3H]methylguanine residues in [3H]MNU-treated DNA (40). Methylosphorothriester-DNA methyltransferase activity was measured by following the transfer of radioactivity from [3H]MNU-treated poly(dT)-poly(dA) to a protease-sensitive form (11). The 3-methyladenine-DNA glycosylase II activity in extracts of BK2012 (tag ada) transformants was assayed by following the release of etheno-oligodeoxyribonucleotide material from [3H]DMS-treated M. luteus DNA (41).

DNA-directed Ada Protein Synthesis—The wild-type Ada protein was purified as described previously (11). The Ada-11 protein from the mutant strain BS11 was purified by the same procedure except that a higher salt concentration (0.5 M NaCl) was found to elute Ada-11 from the DNA cellulose column. Methylated Ada protein was prepared by incubating 3 μg of Ada protein with 7.5 μg of NNU-treated DNA in 6 μl of 70 mM Hepes-KOH, pH 7.5, 10 mM dithiothreitol, 1 mM EDTA, and 5% glycerol at 22°C for 20 min. The pCS70 plasmid carries the ada operon, comprised of the ada and adaB genes, cloned into the vector pAT153 (32). This plasmid was used to direct protein synthesis in a cell-free coupled transcription-translation system (Amersham International PLC). 30 μl of commercial reaction mixtures were supplemented with 2 μg of plasmid DNA. 20 μCi of L-[35S]methionine, and 0.6 μg of unmodified or methylated Ada protein and incubated at 37°C for 30 min. After a 5-min wash with unlabeled methionine, the newly synthesized proteins were resolved by 12% polyacrylamide-SDS gel electrophoresis and visualized by autoradiography.

RESULTS

Cloning the Mutant ada Genes—The mutants BS11, BS21, BS31, and BS41 express the adaptive response to alkylating damage constitutively (24). The mutation in BS21, which confers this phenotype, has been mapped to or close to the ada gene on the E. coli genetic map (26). We have now also mapped the mutations in BS11, BS31, and BS41 to the region of the ada gene. To determine the nature of the mutations which confer the constitutive phenotype, the ada genes were cloned from the mutant strains. The entire ada gene is located on a 3.1-kb HindIII fragment or a 1.3-kb HindIII-Smal fragment of chromosomal DNA (31, 32). Initial attempts to isolate the ada genes from the four constitutive mutants by purifying 3.1-kb HindIII chromosomal DNA fragments, digesting with Smal, and ligating into a pEMBL vector were unsuccessful. In each case 500-1000 ampicillin-resistant D1H101 transformants were screened immunologically for high cellular levels of Ada protein, but no such transformants were detected. The wild-type ada gene has been successfully cloned into the pEMBL vector. It is, therefore, possible that the ada genes from the mutant strains are deleterious to the cell when expressed on the high copy number pEMBL vector. In subsequent attempts to clone the mutant ada genes, a low copy number vector, pHS415 (29), was used. 3.1-kb HindIII fragments of chromosomal DNA isolated from the constitutive mutants were ligated into the HindIII site of pHS415 and the ligated products transformed into strain HB101. Immunological screening of ampicillin-resistant transformants identified between 1 in 100 to 1 in 30 transformants producing high cellular levels of Ada protein. The recombinant plasmids isolated from these transformants were shown by restriction endonuclease analysis to contain single 3.1-kb inserts carrying ada structural genes from different chromosomal locations, but no such transformants were detected. The wild-type ada gene has been successfully cloned into the pEMBL vector.
Gene Activation by Mutated Ada Proteins

Ada Proteins—The nucleotide sequences of the promoter regions (5) and the first 76 nucleotides of the structural genes of ada-11, ada-21, ada-31, and ada-41 were determined by the Maxam-Gilbert method. The ada-11, ada-31, and ada-41 genes were found to have a G to A transition mutation at nucleotide +17 and, therefore, encode a mutated Ada protein containing tyrosine (TAC) in place of cysteine (TGC) at amino acid residue 6 (Fig. 1). None of the genes contained a mutation in the promoter region. To determine the remaining nucleotide sequence of the ada genes, a fragment of the gene, which lacks the promoter region, was subcloned into a pEMBL vector. These constructs do not permit synthesis of Ada protein, and therefore any possible problems of inviability due to overexpression of the mutant ada genes from this high copy number plasmid were avoided. The 1.2-kb EcoRI-Smal ada fragments of pSH11, pSH21, and pSH41 were ligated into pEMBL 8(—) cleaved with EcoRI and HindIII restriction endonuclease. Nucleotide sequencing by the chain termination method showed that the ada-11, ada-21, and ada-41 genes have a G to A transition mutation at nucleotide +378. Double-stranded sequencing of pSH31 has shown that ada-31 has the same transition mutation. All four mutant Ada proteins, therefore, have methionine (ATG) residue 126 replaced by isoleucine (ATA). The three independent isolates BS11, BS31, and BS41, which have a similar phenotype, therefore have the same two point mutations in the ada gene. The ada-21 gene, which was reported to confer a different phenotype with respect to alkA gene expression, has a second G to A transition mutation at nucleotide +298, which results in the replacement of glutamic acid (GAAT) residue 97 by lysine (AAA). Fig. 2 summarizes the amino acid substitutions in the altered Ada proteins.

The Mutant Ada Proteins Are Strong Positive Autoregulators in Vivo and in Vitro—The ability of the mutant Ada proteins to induce ada and alkA gene expression constitutively was examined by assaying the products of these genes in extracts of cells carrying the recombinant plasmids. The ada-21 and ada-41 genes were taken to represent the two types of mutant ada genes. The ada genes in pSH21 and pSH41 were inserted into the kan gene of the vector, and both genes were transcribed in the same direction (Fig. 1). To avoid synthesis of Ada protein from transcripts initiated at the promoter kan gene, the ada gene inserts were inverted. The resulting plasmids were designated pSHR21 and pSHR41. The wild-type ada gene on a 3.1-kb HindIII fragment was subcloned from pCS2 (31) into pHS415 to form pSHR1 with the ada+ gene in the opposite orientation to the kan gene. The O4-methylguanine-DNA methyltransferase activity of the ada gene product was assayed in transformants of strain GW7101 (Δada-25). The 3-methyladenine-DNA glycosylase activity of the alkA gene product was assayed in BK2012 (tag ada) transformants. The absence of the constitutive 3-methyladenine-DNA glycosylase I activity from extracts of the tag mutant ensures that the activity measured represents 3-methyladenine-DNA glycosylase II activity. The level of both enzyme activities in extracts of pSHR1 (ada+) transformants was comparable to that of the untransformed recipient strain (Table II). The Ada+ protein synthesized from the low copy number plasmid therefore did not detectably increase cellular expression of the ada and alkA genes. Strains transformed with pSHR21 or pSHR41, however, were found to have constitutively high activities of both enzymes. In fact, the methyltransferase activity was observed to be even greater than that in the original mutants BS21 and BS41 (Table I). Also, the high level of glycosylase activity in BK2012/pSH21 contrasted with the moderately elevated levels previously observed in the original BS21 mutant (19, 25). The observations suggest that the altered Ada-21 and Ada-41 proteins were acting as strong activators of the ada and alkA genes in vivo.

Self-methylation of the Ada protein at cysteine residue 69 was previously shown by DNA-directed protein synthesis and run-off transcription experiments to convert Ada from a weak to a strong transcriptional activator (5). The ability of the purified Ada-11 protein, in its unmethylated form, to act as an autoregulator in vitro was examined by DNA-directed protein synthesis. The template was plasmid pCS70, an ada

Fig. 1. Schematic map of the recombinant plasmid pSH11. The plasmid was constructed by insertion of the 3.1-kb HindIII fragment of BS11 chromosomal DNA containing the ada gene into the HindIII site of the low copy number vector pSH415. The insert DNA is indicated by a thick line and the vector DNA by a thin line. The direction of transcription from the promoter (Pr, hatched shading) of the ada-11 gene (dark shading) is indicated. The DNA downstream of the ada-11 gene contains several other genes. The ampicillin (bla), chloramphenicol (cat), and kanamycin (kan) resistance genes of the vector are shown, and the direction of their transcription is indicated. The inserted DNA lies within the kan gene which is denoted by a dotted line.

Fig. 2. Diagram of the Ada protein and its mutated forms. The location of the active cysteine residues, the major proteolytic cleavage site, and the amino acid substitutions in the mutant proteins are shown.
Gene Activation by Mutated Ada Proteins

Table 1

<table>
<thead>
<tr>
<th>Strain</th>
<th>Amino acid substitutions in plasmid-encoded Ada protein</th>
<th>O(^*)-methylguanine-DNA methyltransferase activity</th>
<th>3-Methyladenine-DNA glycosylase II activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>units/mg protein(^*)</td>
<td>units/mg protein(^*)</td>
</tr>
<tr>
<td>BS21</td>
<td></td>
<td>17, 21</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>BS41</td>
<td></td>
<td>13, 9</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>Host/pSHR1</td>
<td>None</td>
<td>&lt;1.0, &lt;1.0</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>Host/pCS70</td>
<td>Glu-97 to Lys, Met-126 to Ile</td>
<td>29, 28</td>
<td>14, 22</td>
</tr>
<tr>
<td>Host/pCS70</td>
<td>Cys-6 to Tyr, Met-126 to Ile</td>
<td>32, 46</td>
<td>38, 24</td>
</tr>
<tr>
<td>Host/pCS70</td>
<td>Cys-6 to Tyr, Met-126 to Ile</td>
<td>&lt;1.0, &lt;1.0</td>
<td>&lt;1.0, &lt;1.0</td>
</tr>
<tr>
<td>Host/pCS70</td>
<td>Cys-6 to Tyr, Met-126 to Ile</td>
<td>24, 28</td>
<td>17, 24</td>
</tr>
</tbody>
</table>

Two estimates using two different extracts are presented. One unit of activity repairs 1 pmol of methylated base under standard conditions.

\(^*\)The host strain for the pSHR plasmids was GW7101 (Ada+) for methyltransferase determinations and BK2012 (tag ada) for DNA-glycosylase determinations.

![Fig. 3. Stimulation of DNA-directed Ada protein synthesis by the Ada, methylated Ada, and Ada-11 proteins. Different forms of the Ada protein were added to a cell-free DNA-dependent protein synthesis system containing the plasmid pCS70 as template. This plasmid is an ada\(^*\) derivative of the vector pAT153 (Ada tet). Newly synthesized \(^\text{\textsuperscript{35}}\)S-labeled proteins resolved on a 12\% SDS-polyacrylamide gel and visualized by autoradiography. The additions to the reaction mixtures were: lane 1, pAT153, as template; lane 2, pCS70; lane 3, pCS70 and Ada protein; lane 4, pCS70 and Ada-11 protein; lane 5, pCS70 and methylated Ada protein. \(^{13}\)C-labeled molecular weight markers in lane M were trypsin inhibitor (21.5 kDa), ovalbumin (30 kDa), bovine serum albumin (68 kDa), phosphorylase b (92.5 kDa), and myosin (200 kDa).

![Fig. 4. Trypsin digestion of the 39-kDa Ada and Ada-11 proteins. Two different preparations of the wild-type Ada protein and one preparation of Ada-11 were digested with varying amounts of trypsin at 37°C for 20 min. Lane 1, 0 ng. Lane 2, 0.6 ng. Lane 3, 1.2 ng. Lane 4, 2.0 ng. Lane 5, 2.8 ng. Lane 6, 3.6 ng. The tryptic fragments were resolved on a 15\% SDS-polyacrylamide gel and visualized by immunoblotting. The molecular mass markers were American rainbow markers; ovalbumin (46 kDa), carbonic anhydrase (30 kDa), a small fragment of trypsin inhibitor (21.5 kDa), and lysozyme (14.3 kDa).

Two of these are shown in Fig. 4. The major proteolytic fragments of 20 and 19 kDa result from cleavage within the central hinge region of the protein (see Fig. 2) (13). Larger fragments of approximately 27, 32, 33, and 34 kDa were also produced. A difference in the pattern of appearance and further degradation of these larger fragments was repeatedly observed between the Ada\(^+\) and Ada-11 proteins. This was most pronounced for the 27-kDa fragment. These different digestion patterns suggest that the Ada-11 protein has a different conformation to the Ada\(^+\) protein. Attempts to label the tryptic fragments retaining active cysteine 69 or cysteine 321 residues with DNA substrates containing either [\(^3\)H]methylphosphotriesters or O\(^-\)\[^3\]H)methylguanine lesions, respectively, have indicated that the 27-kDa fragment, and possibly also the larger fragments, arise as a result of cleavages in the N-terminal domain (data not shown). These observations suggest that the mutated N-terminal domain of the Ada-11 protein has a different conformation to that of the Ada\(^+\) protein.

The Ada-11 Protein May Have an Altered Conformation—A difference in the sensitivity of the Ada protein in its methylated and unmethylated forms to proteolytic digestion suggested that methylation results in a conformational change of Ada (13, 15). To examine the possibility that the Ada-11 protein also has a different conformation to the wild-type protein, their sensitivities to trypsin digestion were compared (Fig. 4). Similar patterns of appearance of proteolytic fragments were observed for three different Ada\(^+\) preparations.
Gene Activation by Mutated Ada Proteins

Adaptation is the evolutionary induction of new specific DNA binding sites. The Ada protein, a major regulator of the adaptive response, is composed of two subunits, Ada-12 and Ada-13, which bind specifically to the operator DNA sequence. Mutant Ada proteins, which have amino acid substitutions in their N-terminal domains, can still bind to the DNA, but with lower affinity. This lower affinity allows for the induction of the adaptive response, even with a single mutation in the Ada protein. The Ada protein is also involved in the induction of alkA and ada genes, which regulate the repair of methylphosphotriesters.

**DISCUSSION**

We have shown that the mutations responsible for constitutive expression of the adaptive response in four independently isolated mutants occur in the coding sequence of the regulatory ada gene. The mutant ada genes each have two GC to AT transition mutations and encode altered Ada proteins with two amino acid substitutions in their N-terminal domains. The induction of GC to AT transition mutations was consistent with the use of MNN or MNNU as the mutagenic agent (43, 44). One of the amino acid substitutions, methionine 126 to isoleucine, was common to all four mutated proteins, and this single mutation was sufficient to convert the Ada protein into a strong positive regulator of the ada and alkA genes. The reason for selection of a second amino acid substitution in the mutant Ada proteins remains unclear. MNN is known to induce clustered mutations. However, the cysteine 6 to tyrosine change occurred in three of the four mutants and is, therefore, unlikely to have occurred by chance in the heavily mutagenized cultures. In the fourth mutant, glutamic acid residue 97 was substituted by lysine. The presence of a second mutation caused neither an increase nor a decrease in the activity of the Met^2-Ile-substituted protein to induce ada and alkA gene expression. Also, the Cys^1-Tyr substitution alone did not stimulate the regulatory function of Ada. Two observations, however, on the growth and survival of the pSHR transformants suggest that the second amino acid substitutions may have an effect. The pSHR21 and pSHR41 transformants both grew less well than the pSHR13 transformants, and the pSHR41 transformants were also more sensitive to methylmethanesulfonate. The Cys^1-Tyr and Glu^1-Lys substitutions may have occurred as secondary alterations to increase the stability of the Met^2-Ile-substituted protein. Alternatively, they may enhance the ability of the Ada protein to induce the expression of genes other than ada and alkA.

The mutant BS21 has only moderately elevated cellular levels of 3-methyladenine-DNA glycosylase II (19, 25). However, the product of the ada-21 gene, which was cloned from this mutant, was a strong activator of alkA gene expression. This anomaly could possibly be explained by the higher cellular level of the Ada-21 protein in the plasmid carrying strain compared with BS21 or by a mutation in the alkA gene of BS21.

We have now demonstrated two ways in which the Ada protein can be converted into a strong transcriptional activator, either by methylation of residue cysteine 69 (5) or by substitution of methionine 126 by isoleucine. Methylation of cysteine 69 increases the DNA binding affinity of the Ada protein to the promoter regions of the ada and alkA genes (5, 17). Methylation also alters the sensitivity of Ada to proteolytic digestion which supports the hypothesis that a conformational change results in the enhanced promoter DNA binding affinity (13, 15). It is presently unclear whether the mutated Ada protein has an increased promoter DNA binding affinity. We have observed an increase in nonspecific DNA binding by the mutated Ada-11 protein, during its purification a higher salt concentration was required for elution from DNA-cellulose than is normally used for the wild-type protein. The mutant λ repressor with an enhanced operator DNA binding affinity also has an increased non-operator DNA
binding affinity (45), and this could be the case with the mutant Ada proteins. We have demonstrated that the mutated Ada-11 protein has a changed pattern of tryptic cleavage when compared with the wild-type protein. However, at present it is uncertain whether the methylated Ada protein and the Ada-11 proteins show a similar tryptic digestion pattern. These two proteins cannot be compared directly because the methylated Ada protein has not been separated from the unmethylated form. The proposed transition in conformation of the transcriptional activating Ada proteins could possibly expose the region of the N-terminal domain which is involved in DNA binding. The most pronounced alteration in the tryptic cleavage pattern of Ada-11 compared with Ada was the production of a 27-kDa proteolytic fragment. The site cleaved to yield this fragment must occur close to the isoleucine 126 residue. The distortion produced by this amino acid substitution may therefore expose a site in the N-terminal domain which is involved in DNA binding. The question of whether the methylated cysteine 69 and/or isoleucine 126 residues are directly involved in DNA binding remains to be answered.

Several other mutated Ada proteins with an enhanced gene activating function have been obtained by molecular engineering of the ada gene on multicopy plasmids. All these mutations have involved alterations to the C-terminal domain of the Ada protein. Substitution of the active cysteine 321 residue by alanine was found to increase its positive regulatory function (46). Also, a number of fusion proteins, which derive their N termini from Ada and their C termini from transcribed and translated vector DNA sequences were found to be strong transcriptional activators. The nature of the fused C termini strongly influenced the degree of gene activation (47). These observations suggest that the C-terminal domain of the Ada protein may also have a role in gene activation as was first suggested by Demple (48). The constitutive mutants of the adaptive response which were selected in vivo have alterations to the N-terminal and not the C-terminal domain of the Ada protein. It is possible that amino acid substitutions induced by MNNG or MNU in the C-terminal domain do not produce powerful gene activating Ada proteins.

Our inability to clone the mutant ada genes into a high copy number vector suggests that excessive overproduction of the gene activating mutant Ada proteins is cytotoxic. Even strains carrying these genes on the low copy number vector, pPHSG415, which has four to six copies/chromosome (29), resulted in poor growth, and the original mutants with one or two copies/cell were suspected of having a survival disadvantage (24). The strongly activating Ada fusion proteins were also considered to be cytotoxic at a high cellular disadvantage (24). The strongly activating Ada fusion proteins were also considered to be cytotoxic at a high cellular disadvantage (24). The strongly activating Ada fusion proteins were also considered to be cytotoxic at a high cellular disadvantage (24).

An unexpected observation was that the pSHR13 and pSHR41 transformants, which have constitutively high levels of 3-methyladenine-DNA glycosylase II, had a similar resistance to methylmethanesulfonate toxicity as the pSHR1 transformants. The pSHR13 and pSHR41 strains also have a high level of the DNA binding mutant Ada proteins which could block access of the glycosylase to the alkylated chromosomal DNA and, thereby, inhibit the repair of toxic 3-methyladenine-DNA lesions. However, a related phenomenon has been observed by Kaasen et al. (49) who found that overproduction of the glycosylase from the alkA gene on a multicopy plasmid sensitized cells to methylmethanesulfonate.

A more detailed understanding of the mechanism of transcriptional activation by the Ada protein may come from x-ray crystallographic studies. The mode of DNA recognition of the protein remains unknown. It does not contain an apparent helix-turn-helix motif but it does have pairs of Cys/Ile residues which could possibly form metal binding fingers. The methylated Ada protein is not available in sufficient amounts for physical studies. The use of the Ada-13 transcriptional activator, with a single Met/Ile substitution, could therefore facilitate a crystallographic analysis of the Ada-promoter complex. Also, such analyses may verify whether the Met/Ile substitution results in an altered three-dimensional structure.

Acknowledgments—We thank T. Lindahl for advice and discussion throughout this work, M. Olsson for purifying the Ada-11 protein, P. Karran, P. Calouz, and S. G. Sedgwick for comments on the manuscript, and J. Green for typing the manuscript.

REFERENCES

Gene Activation by Mutated Ada Proteins