Bacillus subtilis ada operon encodes two DNA alkyltransferases

Fumiko Morohoshi, Kenshi Hayashi[1] and Nobuo Munakata

Radiobiology Division and [1]Oncogene Division, National Cancer Center Research Institute, Tsukiji, 5-1-1, Chuo-ku, Tokyo 104, Japan

ABSTRACT

By prophage transformation and subcloning, we have obtained *Bacillus subtilis* DNA fragments that could complement the hypersensitivity of *ada* (adaptive response deficient) mutants to N-methyl-N'-nitro-N-nitrosoguanidine (MNNG). The nucleotide sequence contained two open reading frames that were assigned to the genes *adaA* and *adaB*, encoding methylphosphotriester-DNA methyltransferase and O^6-methylguanine-DNA methyltransferase, respectively. These two genes overlap by 11 bp and comprise a small operon. The 1.6 Kb transcripts derived from the operon were detected in *ada*+ cells cultured in the presence of MNNG but not in control *ada*+ cells. From analysis of the syntheses of DNA alkyltransferases in the *ada* mutant cells harboring the plasmid carrying the complete or partial fragment, we conclude that the *adaA* gene product functions as a transcriptional activator of the *ada* operon, while the *adaB* gene product specializes in repair of mutagenic O^6-methylguanine residues. Comparison with *Escherichia coli ada* operon showed that the two genes correspond to portions of the *E. coli ada* gene, implicating gene fusion or splitting as the origin of the difference in the organizations of the genes.

INTRODUCTION

In 1977, Samson and Cairns reported that *Escherichia coli* cells acquired resistance to alkylating chemicals such as N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) when they were cultured in the presence of these compounds (1). This 'adaptive response' has been intensively studied in *E. coli*, leading to the discovery of DNA alkyltransferase, which collects alkyl groups from alkylated DNA in a suicidal manner (2–4). The *ada* gene coding for DNA alkyltransferase (Ada protein) has been cloned and sequenced (5–9). The adaptive response depends on the unique activities of the Ada protein. The protein is composed of two domains, an N-terminal one that accepts a methyl group from methylphosphotriesters and a C-terminal one that accepts a methyl group from mutagenic O^6-methylguanine residues (4,6,10). When the Ada protein has collected a methyl group from methylphosphotriesters, it can activate the transcription of the *ada* operon itself and of other genes including *alkA* coding for

3-methyladenine-DNA glycosylase II (11,12). The inducible syntheses of the repair enzymes make the cells resistant to the lethal and mutagenic effects of MNNG.

This response is not universally distributed, but seems to be confined to several prokaryotic species (4,13,14). We showed that *Bacillus subtilis* cells exhibit a similar response to alkylation; however, several differences between the *E. coli* and *B. subtilis* system have been noted (15–19). Most of the *ada* mutants isolated in *E. coli* seem to be pleiotrophic, and defective in multiple phenotypes of the response to various degrees (20,21), while the *ada* mutants of *B. subtilis* isolated belong to two distinct types; those totally deficient in the response (*ada-1* type) and those defective only in the O^6-methylguanine-DNA methyltransferase activity (*ada-4* type) (17,18). This may be explained from the finding that the inducible DNA methyltransferase activities in *B. subtilis* resided in two separate molecular species; 27 Kd methylphosphotriester-DNA methyltransferase and 22 Kd O^6-methylguanine-DNA methyltransferase (18). However, we could not completely rule out the possibility that these two activities were produced by post-translational proteolysis, as seen in the *E. coli* Ada protein (22). Another phenotypic difference noted was that *B. subtilis ada* mutants were not hypersensitive to methylmethanesulfonate (MMS) unlike *E. coli* mutants (16,20).

We thought that cloning of the genes complementing the defects in the *B. subtilis ada* mutants should provide answers to these questions and shed light on the generality and variations of the molecular mechanisms of this response. We report here that the genes complementing the defects of the *ada* mutations encode two DNA alkyltransferases with partial sequence similarities to the *E. coli* Ada protein.

MATERIALS AND METHODS

Bacteria, phage and plasmids

Bacillus subtilis strains 168T (*thyA thyB*) (23), TKJ1922 (*thyA thyB hisH101 metB101 leuA8 lys-21*)(16), TKJ0922 (the same as TKJ1922 but *ada-1*)(16) and TKJ2924 (the same as TKJ1922 but *ada-4*)(17) were used. *B. subtilis* strains UOTO994 (*hisA1 metB5 φCM*) (24) and *B. subtilis* template phage φ105 were obtained from Fujio Kawamura (Institute of Applied Microbiology, University of Tokyo). pHY300PLK, a shuttle vector between *E. coli* and *B. subtilis* (25) was obtained from Toyobo (Tokyo). *E. coli* plasmid, Bluescript SK+ and its host strain, XL1-blue, were from Stratagene. *E. coli* strains, K802

and HB101 were used as hosts of plasmids (26). Lambda EMBL3a and its host ED8654 and NM535 were obtained from Hans Lehrach (European Molecular Biology Laboratory)(27). Lambda EMBL3 and a packaging kit were obtained from Promega (Madison).

Transformation and assay of sensitivity

Competent cells of *B. subtilis* prepared in transformation medium (28) were mixed with chromosomal DNA or plasmids, and incubated for 90 min at 37°C with shaking. Chloramphenicol- and tetracycline- resistant transformants were selected by plating the cells on Nutrient agar medium (Difco) containing 5 μg/ml of chloramphenicol and 12.5 μg/ml of tetracycline, respectively. For the selection of MNNG^r colonies, drug resistant colonies were aligned on nutrient agar plates containing the antibiotic and after overnight-incubation at 37°C, they were replica plated on Spizizen medium (28) containing 30 μg/ml of MNNG and required nutrients. For examination of the *ada*$^+$ transforming activities of DNA fragments in phage or plasmids, they were digested with appropriate restriction enzymes to separate chromosomal DNA from vector sequences. The digested DNA (0.2 μg) together with a small amount (0.01 μg) of the DNA from *met*$^+$ *ada* cells carrying the same *ada* mutation as the recipient cells was used to transform *metB101 ada* mutant cells. The MNNG sensitivity of *met*$^+$ transformants was examined by replicating the *met*$^+$ colonies on plates containing MNNG as described above.

Construction of phage library

A *B. subtilis* DNA library from strain 168T was constructed with lambda EMBL3a and a DNA library from a CM^r MNNG^r transformant was constructed with lambda EMBL3 as described (26).

Restriction enzyme digestion and subcloning

Standard procedures were followed (26).

Assay of DNA alkyltransferases

Crude sonic extracts (15) were incubated for 20 min with calf thymus DNA that had been treated with [^{3}H] MNU (1.6 Ci/mM, NEN) and heated at 80°C for 16 hours to remove N-methylated purines. After the addition of gel-loading buffer (18), extracts were boiled for 3 min and subjected to SDS gel electrophoresis (12.5%). The fluorogram of the gel was prepared using EN^{3}HANCE (NEN).

Sequence determination

Isolated DNA fragments (2.2 Kb *EcoRI-SalI* and 5.2 Kb *EcoRI-SalI*) were inserted into the *EcoRI-SalI* site of the polylinker of Bluescript SK+. DNA fragments with a deletion from the *EcoRI* or *SalI* site were obtained by digesting the plasmids with either *SacI* and *EcoRI* or *KpnI* and *SalI*. Unidirectional deletions were introduced with Exonuclease III (BRL) and Exonuclease VII (BRL) according to the protocol of Yanish-Perron et al.(29). Their sequences were determined by the dideoxy chain termination method (30) with Sequenase (USB).

Northern blotting and primer extension analysis

Total RNA was extracted from logarithmically growing cells according to Gilman and Chamberlin (31), and Northern blotting was carried out as described (32). Probe I was a 0.59 Kb *EcoT221-PvuII* fragment located in the coding region of the *adaA*
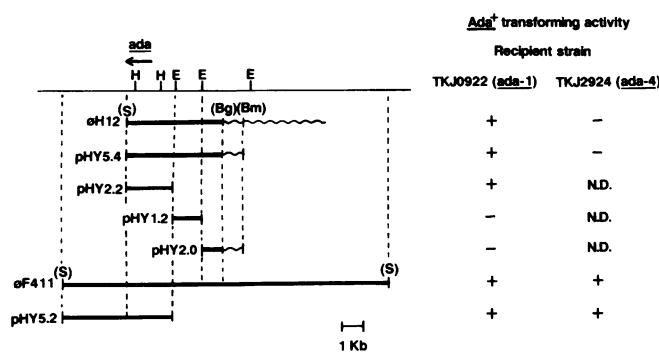

Fig. 1. *Ada*$^+$ transforming activities of DNA fragments derived from the *ada* region. Thick horizontal and wavy lines show *B. subtilis* DNA and ϕCM DNA, respectively. All *EcoRI* (E) sites in the region and two *HindIII* (H) sites close to the *ada* operon are shown. Restriction sites in brackets are in the linker sequences of the phage or plasmid. S: *SalI*, Bg: *BglII*, Bm: *BamHI*. The arrow indicates the locus of the *ada* operon and the direction of *ada* transcription. When more than 1% of the *met*$^+$ colonies (about 1,000) could grow on the medium containing MNNG, the DNA employed was judged to be positive (+). N.D. means 'not done'.

and *adaB* genes (nucleotide 621 to 1212 in Fig.2). The following four probes were synthesized by the polymerase chain reaction (33) using wild type DNA as a template and 20mer synthetic nucleotides from the 5'-end of each strand of the fragments as primers: probe II (nucleotide 44 to 367 in Fig.2), probe III (nucleotide 329 to 680), probe IV (nucleotide 641 to 990) and probe V (nucleotide 951 to 1329). The transcription start site of the *ada* operon was determined by primer extension analysis. A 20mer oligodeoxynucleotide that corresponds to the 8th to 27th nucleotides from the translational initiation site of the *adaA* gene was synthesized and used for the analysis. A mixture of total RNA (10 or 50 μg) and 2 pmole of the primer in 40 μl of buffer (60 mM Tris-HCl pH 8.3, 90 mM KCl, 3.6 mM MgCl$_2$ and 120 μg/ml gelatin) was heated at 95°C for 2 min and then incubated for 30 min at 37°C. To the mixture, 1 μl of 0.5 M dithiothreitol, 2 μl of RNasin (Promega, Madison), 5 μl of 5 mM 4 dNTPs and 0.4 units of MMLV-reverse transcriptase (BRL) were added and the mixture was further incubated for 30 min at 37°C. After the incubation, it was treated with RNase A (1 mg/ml) for 30 min at 37°C in the presence of 20mM EDTA. Then the mixture was shaken with phenol-chloroform and the DNA was precipitated with ethanol. An equal volume of formamide-dye solution (26) was added to the DNA dissolved in 10 mM Tris-HCl (pH 8.0)-1 mM EDTA, and the mixture was heated for 2 min at 80°C. The size of the DNA was analyzed with sequencing gel.

RESULTS

Cloning of DNA fragment conferring MNNG-resistance on *ada-1* cells

B. subtilis DNA fragments capable of correcting the high sensitivity to MNNG of *ada-1* cells were cloned by the prophage transformation method using lysogenic phage vector ϕCM. The vector was derived from ϕ105 and carried the chloramphenicol acetyltransferase (*cat*) gene as a selection marker (24). The *BglII*-digested 168T (*ada*$^+$) DNA was ligated into the ϕCM DNA digested with the same enzyme. This ligation mixture was used to transform cells of strain TKJ0922 (*ada-1*) lysogenized with ϕ105. Chloramphenicol-resistant (CM^r) colonies were picked

```
1
TGGTTGAAATTAGTTTATCACATAATCATCTTTACTCATTTTGTTATCTTGCTATTACATTCTAACCACCTTATTTTTTTCTATTTATGA
     100        ↓
GGTTATAGTGTAGTTATCAAGAATGCTAAACGGGAGTAGATAAAAATGGCCCAAGATATAAAAAACGCGTGTCGAAAATTTAAATTGGGA
          200
TTCAATTCAAAGAATTAGATGAGCATTGTAAGGAGAGGTGAGATAAATGCCGGATAGTATCAATAACGGACATAAAGAGAGCCATGATCA
                                        MetProAspSerIleAsnAsnGlyHisLysGluSerHisAspHi
          300
TAGAATTTCGAATGATGCAGAAATGATAACAGATGAAAAGTGGCAAGCAATTATAAATAATGATGCAGCGTACAATAATCAATTTTTCTA
sArgIleSerAsnAspAlaGluMetIleThrAspGluLysTrpGlnAlaIleIleAsnAsnAspAlaAlaTyrAsnAsnGlnPhePheTy
          400
CGCTGTTAAATCGACAGGGATATTCTGTAAACCATCCTGTAAATCTCGCGTTCCGAAAAAAGAAAATGTATGTATCTTTCCAAACACAGA
rAlaValLysSerThrGlyIlePheCysLysProSerCysLysSerArgValProLysLysGluAsnValCysIlePheProAsnThrGl
          500
ACAAGCTCTCCGCGCAAATTTTCGCCCTTGTAAACGTTGCAAGCCCACTAATGAAAAAATGCCTGATAGCGAGTGGGTTGATTTAATTAC
uGlnAlaLeuArgAlaAsnPheArgProCysLysArgCysLysProThrAsnGluLysMetProAspSerGluTrpValAspLeuIleTh
          600
TGAATACATTGATAAAAATTTCACAGAAAAATTAACGCTAGAATCGTTAGCAGATATTTGTCATGGGAGTCCGTATCATATGCATCGAAC
rGluTyrIleAspLysAsnPheThrGluLysLeuThrLeuGluSerLeuAlaAspIleCysHisGlySerProTyrHisMetHisArgTh
          700
ATTTAAAAAAATTAAAGGCATTACGCTGGTTGAGTATATACAACAAGTTAGAGTACACGCGGCTAAAAAGTATTTGATTCAGACAAATAA
rPheLysLysIleLysGlyIleThrLeuValGluTyrIleGlnGlnValArgValHisAlaAlaLysLysTyrLeuIleGlnThrAsnLy
          800
AGCGATTGGAGATATCGCTATATGTGTGGGTATAGCTAACGCGCCTTATTTTATTACTTTATTTAAAAAGAAAACTGGACAGACACCAGC
sAlaIleGlyAspIleAlaIleCysValGlyIleAlaAsnAlaProTyrPheIleThrLeuPheLysLysLysThrGlyGlnThrProAl
          900
ACGATTTCGTCAAATGAGTAAAATGGAGGAAACGTACAATGGAAACAAATAAACCAACCCTTTATTGGTCTTTACTAATGTTCAAGGATT
aArgPheArgGlnMetSerLysMetGluGluThrTyrAsnGlyAsnLys
                                        MetGluThrAsnLysProThrLeuTyrTrpSerLeuLeuMetPheLysAspT
GGAATTTTTATATTGCTTCAACTTTAAAGGGGCTTGTGTTTGTAGGTTCACAGAACAAACCAATCGAGGAATTGTTCGAATGGGCTAGGA
rpAsnPheTyrIleAlaSerThrLeuLysGlyLeuValPheValGlySerGlnAsnLysProIleGluGluLeuPheGluTrpAlaArgL
          1000
AACGCTTTCCGGGAAGTCTTCTTGTTGAAGATGATGATAAGCTTGAACCCTATGCCGTTGAAATCACTCAATATCTAGAAGGAAAGCGGA
ysArgPheProGlySerLeuLeuValGluAspAspAspLysLeuGluProTyrAlaValGluIleThrGlnTyrLeuGluGlyLysArgL
          1100
AAAACTTTACTGTTCCAGTTGAGTACGCCGGTACGCAATTTCAGCTAGCCGTCTGGAATGCACTTTGTGAAATTCCTTATGGACAGACGA
ysAsnPheThrValProValGluTyrAlaGlyThrGlnPheGlnLeuAlaValTrpAsnAlaLeuCysGluIleProTyrGlyGlnThrL
          1200
AATCCTATTCCGACATTGCAAATGATATAAATAAACCAGCAGCTGTTCGTGCTGTAGGTGCGGCTATTGGGGCTAATCCGGTATTAATTA
ysSerTyrSerAspIleAlaAsnAspIleAsnLysProAlaAlaValArgAlaValGlyAlaAlaIleGlyAlaAsnProValLeuIleT
          1300
CTGTACCGTGCCATCGTGTAATAGGAAAGAATGGCTCATTAACTGGCTATCGGGGCGGATTTGAAATGAAGACACTGCTCCTCGATCTGG
hrValProCysHisArgValIleGlyLysAsnGlySerLeuThrGlyTyrArgGlyGlyPheGluMetLysThrLeuLeuLeuAspLeuG
          1400
AAAAGCGAGCTTCATCTGAAATGGATGTTCCTCATTGAAAGTGAATCATCTATAATGGGTTGGATTCTTTTAGAGGAAAGTAAAGACCAG
luLysArgAlaSerSerGluMetAspValProHis
          1500
TTTACTGATAAACTTGACAACTTTTAAAATTTAGCATATCTTATACGTAACTTAATAATCGATGCAGATGTTTATTTTTCCTTTTAAAAT
```

Fig. 2. Nucleotide sequence of *B. subtilis adaA* and *adaB* loci and the deduced amino-acid sequences of two ORFs. The nucleotides from 227 to 859 and from 849 to 1385 code for AdaA and AdaB proteins, respectively. The dashed underlines indicate possible SD sequences, and the arrow shows the transcription start site. Possible '−10' and '−35' sequences are boxed. Open arrows indicate inverted-repeat sequences.

up. Then, MNNG-resistant clones were selected among the CM^r clones. We selected five doubly-resistant (CM^r MNNG^r) clones, but on treatment with mitomycin C, none of them yielded phage capable of complementing the *ada-1* mutation, suggesting that they were defective in induction. Therefore, we examined the CM^r MNNG^r clones to determine if the two resistant markers were closely linked in DNA-mediated transformation. Two of the clones exhibiting more than 90% linkage were chosen for subsequent experiments. Both of them regained the ability of adaptive response to the lethal and mutagenic effects of MNNG, and the extract from the cells pretreated with MNNG contained as high DNA alkyltransferase activities as the extract from the pretreated *ada*$^+$ cells (data not shown).

A DNA library was constructed in lambda EMBL3 from one of the CM^r MNNG^r strains. This library was screened with a 1.0 Kb *Bgl*II-*Bam*HI fragment from ϕCM located next to the cloning site as a probe (Fig. 1). DNAs from all three independent clones hybridizing with the probe could confer MNNG-resistance upon transformation of the *ada-1* cells. The restriction sites of one of the phage clones, H12, are shown in Fig. 1. The insert in the phage DNA consisted of 4.4 Kb *B. subtilis* DNA and 10.4 Kb ϕCM DNA. The activity conferring MNNG resistance to *ada-1* cells was found to be located in the terminal 2.2 Kb *Sal*I-*Eco*RI fragment of the 4.4 Kb fragment. However, this DNA could not confer MNNG resistance on another mutant *ada-4*. Since these two mutations are closely linked in the chromosome

(17), we searched for extended clones that could also complement the *ada-4* mutation by screening a *B. subtilis* DNA library made with lambda EMBL3a using the 2.2 Kb *Sal*I-*Eco*RI fragment as a probe. Most of the phages hybridizing to the fragment did not multiply well for some unknown reason. However, one clone, F411, that contained a 15.2 Kb insert, grew relatively well and its DNA could complement both the *ada-1* and *ada-4* mutation. These activities resided in a terminal 5.2 Kb *Sal*I-*Eco*RI fragment including the 2.2 Kb *Sal*I-*Eco*RI fragment (Fig. 1).

DNA sequence of *ada* genes

The DNA sequence of the 2.2 Kb *Eco*RI-*Sal*I fragment in the phage H12 was determined. Two open reading frames (ORF1 and ORF2) were recognized. ORF1, located in the middle of the fragment could encode a protein consisting of 211 amino acid residues. ORF2 is initiated at the 3' terminus of ORF1, and the two ORFs overlap by 11 bp. ORF2 continues to the *Sal*I site in the linker sequence of the plasmid without a termination codon. To extend the 3' sequence, we inserted the 5.2 Kb *Eco*RI-*Sal*I fragment in phage F411 into Bluescript SK+ and determined its sequence. ORF2 can code for a protein consisting of 179 amino acid residues. The nucleotide sequence of the region containing ORF1 and ORF2 and their deduced amino acid sequences are shown in Fig. 2. Two sequences, GAGGTGA and AATGGAGG conforming to the Shine-Dalgarno consensus are seen immediately upstream of the initiation codons of ORF1 and

```
AdaA   1    M P D S I N N G H K E S H D H R I S N D A E M I T

AdaA  26    D E K W Q A I I N N D A A Y N N Q F F Y A V K S T
Ada   10    D Q R W Q S V L A R D P N A D G E F V F A V R T T

AdaA  51    G I F C K P S C K S R V P K K E N V C I F P N T E
Ada   35    G I F C R P S C R A R H A L R E N V S F Y A N A S

AdaA  76    Q A L R A N F R P C K R C K P T N E K M P D S E W
Ada   60    E A L A A G F R P C K R C Q P E K A N A Q Q H R -

AdaA 101    V D L I T E Y D - D K N F T - E K - L T L E S L A
Ada   84    L D K I T H - A - C R L E Q E T P V T L E A L A
AraC 179    V R E A C Q Y I S D H L A D - S N - F D I A S V A

AdaA 123    D I C H G S P Y H M H R T F K I K G I T - L V E
Ada  107    D Q V A M S P F H L H R L F K A T T G M T P K A W
AraC 202    Q H V C L S P S R L S H L F R Q Q L G I S V L S W

AdaA 147    Y I Q Q V R V H A A K Y L I Q T N K A G D I A
AraC 227    R E D Q - R I S Q A K L L L S T T R M P I A T V G

AdaA 172    I C V G I A N A P Y F I T L F K K K T G Q T P A R
AraC 251    R N V G F D D Q L Y F S R V F K K C T G A S P S E

AdaA 197    F R Q M S K M E E T Y N G N K L S
AraC 276    F R A G C E E K V N D V A V K L S
```

Fig. 3. Sequence comparison of *B. subtilis* AdaA, *E. coli* Ada (N-terminal portion) and *E. coli* AraC (C-terminal portion). Amino acid residues that are identical in at least two proteins are boxed. A possible helix-turn-helix motif is underlined.

```
AdaB  61   K L E P Y A V E I T Q Y L E G - K R K N F T V P V
Ada  240   L M F Q Q H V R E V I A S L N Q R D T P L T L P L
Dat   50   N L A E A K Q Q L Q E Y F A G - E R K T F S L P L
Ogt   58   N P G G L S D K L R D Y F A G N L S I I D T L P T
MGMT  64   A W L N A Y F H Q P E A I E E F P V P A L H H P V

AdaB  85   E Y A G T Q F Q L A V W N A L C E I P Y G L Q T K S
Ada  265   D I R G T A F Q Q Q V W Q A L R T I I P C G E T V S
Dat   74   S Q K G T P F Q Q K V W Q A L E R I I P Y G E S R S
Ogt   83   A T G G T P F Q R E V W K T L R T I I P C G G Q V M H
MGMT  89   F Q Q E S F T R Q V L W K L L K V V K F G E V I S

AdaB 110   Y S D I A N D I N K P A A V R A V G A A I G A N P
Ada  290   Y Q Q L A N A I G K P K A V R A V A S A C A A N K
Dat   99   Y A D I A A A V G S P K A V R A V G Q A N K R N D
Ogt  108   Y G Q L A E Q L G R P G A A R A V G A A N G S N P
MGMT 114   Y Q Q L A A L A G N P K A A R A V G G A M R G N P

AdaB 135   V L I T V P C H R V I G K N G S L T G Y R G - G F
Ada  315   L A I I T P C H R V V R G D G T L S G Y R W - G V
Dat  124   L P I F V P C H R V T G K N S A L T G Y A G S K T
Ogt  133   I S I V V P C H R V I G R N G T M T G Y A G - G V
MGMT 139   V P I L T P C H R V V C S S G A V G N Y S G - G L

AdaB 159   E M K T L L L D L E K R A S S E M D V P H
Ada  339   S R K A Q L L R R E A E N E E R
Dat  149   E I K A F L L N I E R I S Y K E K
Ogt  157   Q R K E W L L R H E G Y L L L
MGMT 163   A V K E W L L A H E G H R L G K P G L G G + + + +
```

Fig. 4. Sequence comparisons of O⁶-methylguanine-DNA methyltransferases. Amino acid residues that are identical in at least three of the five proteins are boxed.

ORF2, respectively. There is an inverted repeat of about 40 bp from 60 bp downstream of the termination codon of ORF2. This sequence may be the signal terminating the transcriptions of these genes.

Calculated molecular masses of the protein encoded by ORF1 and ORF2 were 24,3 Kd and 20.1 Kd, respectively. It was considered likely that they corresponded to the two inducible DNA alkyltransferases identified by SDS-polyacrylamide gel electrophoresis; 27 Kd methylphosphotriester-DNA methyltransferase and 22 Kd O⁶-methylguanine-DNA methyltransferase (18). Since our previous studies have demonstrated that the functions of the *ada* gene(s) in *B. subtilis* are similar to those of the *ada* gene in *E. coli*, we first compared the amino acid sequences specified by the two ORFs with that of the *E. coli* Ada protein. The predicted amino acid sequences of ORF1 and the N-terminal region of the *E. coli* Ada protein exhibit similarity as shown in Fig. 3. Contiguous residues from 43 to 137 of ORF1 are similar to the residues from 27 to 121 of the Ada protein; 44 (46%) amino acid residues are identical in the two sequences. A stretch of 7 amino acid residues, -Phe-Arg-Pro-Cys-Lys-Arg-Cys-, located in the middle of the aligned region, is found in both proteins (6,7). The central cysteine residue in this stretch of the *E. coli* Ada protein has been shown to accept a methyl group from methylphosphotriesters (10). From the sequence similarity between these two proteins and also from the similarity of the size of the protein encoded by ORF1 and that of methylphosphotriester-DNA methyltransferase estimated previously, we concluded that ORF1 was the structural gene coding for methylphosphotriester-DNA methyltransferase and named this gene *adaA*.

On the other hand, the amino acid sequence of ORF2 shows extensive similarity to that of the C-terminal portion of the *E. coli* Ada protein carrying O⁶-methylguanine-DNA methyltransferase activity (6,7). We previously recognized the similarity of the sequences of Ada and Ogt of *E. coli* (35) and Dat of *B. subtilis* (34,36). Recently, c-DNA coding for human O⁶-methylguanine-DNA methyltransferase (MGMT) was cloned. The predicted amino acid sequence of the MGMT also showed similarity to the bacterial proteins with the same activity (37). From these comparisons and from the similarity of the size of the protein, we conclude that ORF2 is a structural gene coding for O⁶-methylguanine-DNA methyltransferase, and named this

gene *adaB*. The AdaB protein was aligned with the four other O⁶-methylguanine-DNA methyltransferases (Fig. 4). Sequence similarity is obvious within a contiguous sequence of 88 amino acid residues (from amino acid 81 to 168 in AdaB) in which 24 residues are identical in all five alkyltransferases and 32 residues are identical in the four prokaryotic alkyltransferases. Comparison of pairs of sequences showed identities of about 52% between prokaryotic transferases and of about 40% between bacterial and human proteins. There are no significant differences in the degrees of sequence conservation in the prokaryotic enzymes; neither a species differences (*E. coli* vs. *B. subtilis*) nor a difference in the mode of synthesis (constitutive vs. inducible) seems to affect the extent of the similarity. The four prokaryotic enzymes seem to have diverged or converged independently. A stretch of five amino acid residues, -Pro-Cys-His-Arg-Val- is conserved in all five proteins and this sequence is flanked by either valine or isoleucine. In *E. coli* Ada, the cysteine residue in this stretch has been shown to accept a methyl group from O⁶-methylguanine (6). Residues that are identical in all five proteins are distributed on both sides of this active center.

Transcriptions of *adaA* and *adaB*

We previously showed that two alkyltransferases were synthesized when *ada⁺* cells were cultured in the presence of MNNG (18). To determine whether this inducible synthesis is regulated at transcription, we examined total RNAs from cells of strain TKJ1922 (*ada⁺*) and strain TKJ0922 (*ada-1*) by Northern blot analysis using probe I, which is a 0.59 Kb *EcoT*221-*Pvu*II fragment located in the coding region of the *adaA* and *adaB* genes. Three bands migrating to positions corresponding to 1.6, 1.2 and 0.9 Kb were observed in the RNA from the MNNG-treated cells, as shown in Fig. 5. The signal of the largest RNA of 1.6 Kb appeared to be the strongest. In contrast, no band was observed in the RNA from the *ada⁺* cells without pretreatment or in RNAs from *ada-1* cells with or without pretreatment. This clearly demonstrates that inducible synthesis of the two transferases is regulated primarily at the transcription level. The finding that *ada* transcripts were not detectable in pretreated *ada-1* cells is consistent with our previous observation that neither of the two enzymes was induced in this mutant strain (18).

To characterize the nature of the three transcripts, we repeated
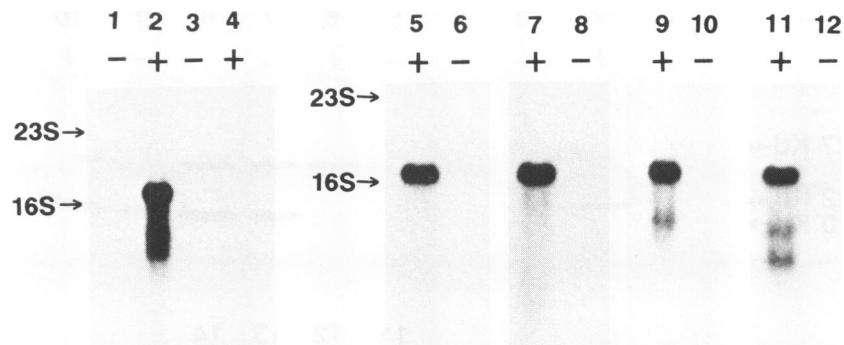
Fig. 5. Northern blot of *ada* m-RNA. Total RNA from control (−) and MNNG-treated (+) cells were subjected to electrophoresis in 1.2% agarose-formamide gel, transferred to a nylon filter and hybridized to the probes described in MATERIALS AND METHODS. Probe I (lanes 1−4), probe II (lanes 5−6), probe III (lanes 7−8), probe IV (lanes 9−10) and probe V (lanes 11−12). Lanes 1, 2 and 5−12, RNAs extracted from TKJ1922 (*ada*+) cells; lanes 3 and 4, RNAs from TKJ0922 (*ada-1*) cells.

Northern blottings with four probes of different regions of the *adaA* and *adaB* genes (Fig. 5). Probe II covered the 5′-terminal and upstream non-coding region of *adaA*, probe III, the middle part of *adaA*, probe IV, the 3′-terminal part of *adaA* and 5′-terminal part of *adaB* and probe V covered the middle part of *adaB* (see MATERIALS AND METHODS). Probes II and III detected only the largest 1.6 Kb transcript, probe IV detected 1.6 Kb and 1.2 Kb transcripts, and probe V detected all three types of the transcripts. These findings demonstrate that the syntheses of the two smaller transcripts with the same 3'sequence as the largest transcript start within *adaA* or *adaB* gene (18).

The start site for the largest transcript was determined by primer extension analysis. As shown in Fig.6, it was identified as an adenine residue located 121 bp upstream of the initiation codon of the *adaA* gene. TATAGT and TTATTT located in the promoter region of the *adaA* gene (boxed in Fig. 2) may be the '−10' and '−35' sequences, respectively. From the size of the main transcript (1.6 Kb), no additional ORF seems to exist in the region covered by the transcript. Therefore, we conclude that the *adaA* and *adaB* genes comprise a small operon, the expression of which is induced by treatment with MNNG.

Functions of the two genes

Genetic analyses suggested that methylphosphotriester-DNA methyltransferase was responsible for transcriptional activation of the *ada* operon upon accepting a methyl group from methylphosphotriesters. To substantiate this proposition and also to confirm that *adaA* gene directed the synthesis of 27 Kd methylphosphotriester-DNA methyltransferase and *adaB* gene directed 22 Kd O⁶-methylguanine-DNA methyltransferase, we introduced two kinds of DNA fragments, a 1.2 Kb *Hin*dIII fragment containing the *adaA* gene and a 5.2 Kb *Sal*I-*Eco*RI fragment containing both the *adaA* and *adaB* gene (Fig. 1) into a shuttle vector pHY300PLK and termed the plasmids pHY1.2 and pHY5.2, respectively. We used two mutant strains, TKJ0922 (*ada-1*) and TKJ2924 (*ada-4*) as hosts of the plasmids. The *ada-1* mutant cells are deficient in all the inducible phenotype of the adaptive response and defective in transcription of the *ada* operon, as shown in the previous section, while the *ada-4* mutant cells are deficient only in the inducible synthesis of the 22 Kd
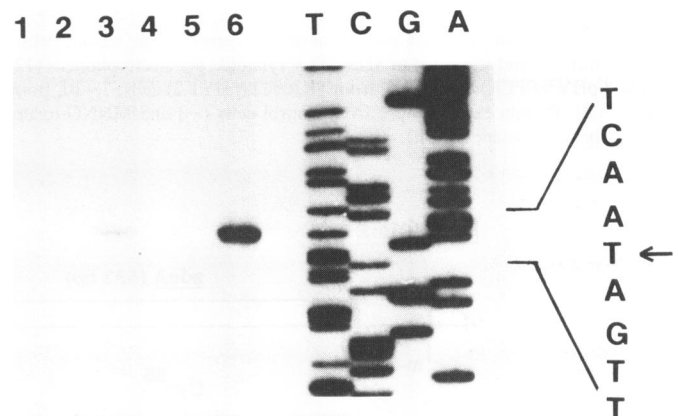

Fig. 6. Primer extension analysis of the start site of *ada* transcription. Total RNA (10 or 50 μg) and 20mer synthetic primer (nucleotide 234−253 of the sequence in Fig. 2) were used for reverse transcription. The DNA sequence of the complementary strand of the corresponding region is shown in lanes T, C, G and A. Template RNAs: lanes 1 and 4, t-RNA; lanes 2 and 5, RNAs from control cells of strain TKJ1922 (*ada*+); lanes 3 and 6, RNAs from MNNG-treated cells of TKJ1922 (*ada*+). Sample size: lanes 1−3, 10 μg RNA; lanes 4−6, 50 μg RNA.

O⁶-methylguanine-DNA methyltransferase (18). Both mutant cells constitutively synthesize 20 Kd O⁶-methylguanine-DNA methyltransferase (Dat), its synthesis being independent of the adaptive response (34).

The DNA alkyltransferase activities of extracts of cells harboring plasmid pHY1.2 or plasmid pHY5.2 were assayed either by SDS-polyacrylamide gel fluorography (Fig. 7) or by proteinase K digestion (3) (data not shown). Extracts from the untreated TKJ0922 (pHY1.2) and TKJ2924 (pHY1.2) cells contained neither 22 Kd nor 27 Kd transferase. But the extract from MNNG-pretreated TKJ0922 (pHY1.2) cells contained both 27 Kd and 22 Kd transferases, the former activity being predominant. The total alkyltransferase activity in this extract was several times higher than that in control cells of the same strain. In contrast, the extract from MNNG-treated TKJ2924
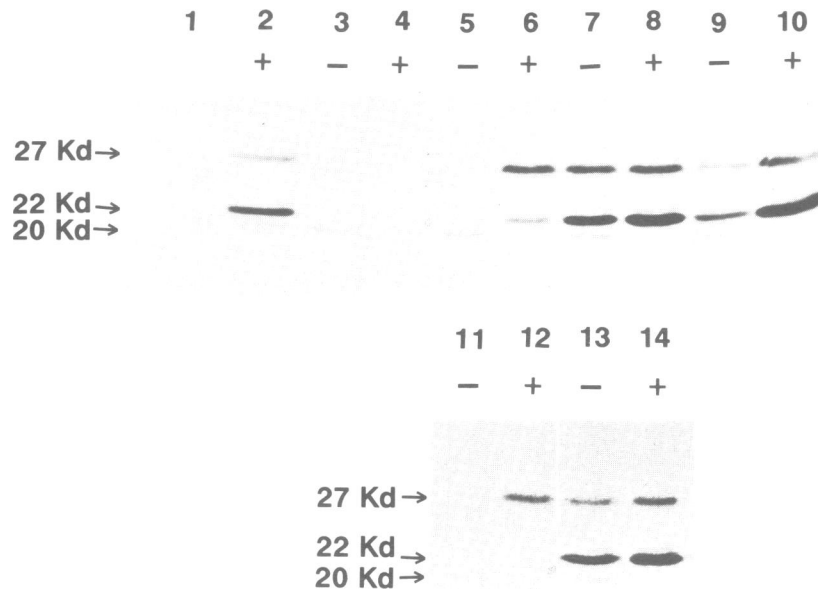
Fig. 7. Expression of cloned methyltransferase genes in *ada* mutant cells. Crude extracts were incubated with [^{3}H]-MNU treated and heated DNA (3,500 cpm) for 20 minutes and subjected to SDS-polyacrylamide gel electrophoresis (12.5%). Lane 1, without extract; lane 2, from TKJ1922 (*ada$^+$*) cells; 3−4, from TKJ0922 (*ada-1* pHY300PLK) cells; 5−6, from TKJ0922 (pHY1.2) cells; 7−10, from TKJ0922 (pHY5.2) cells; 11−12, from TKJ2924 (pHY1.2) cells; 13−14, from TKJ2924 (pHY5.2). Protein extracts were from control cells (−) and MNNG-treated cells (+). Samples of 150 μg protein were applied to lanes 2−8 and 11−14, and of 50 μg protein to lanes 9−10.
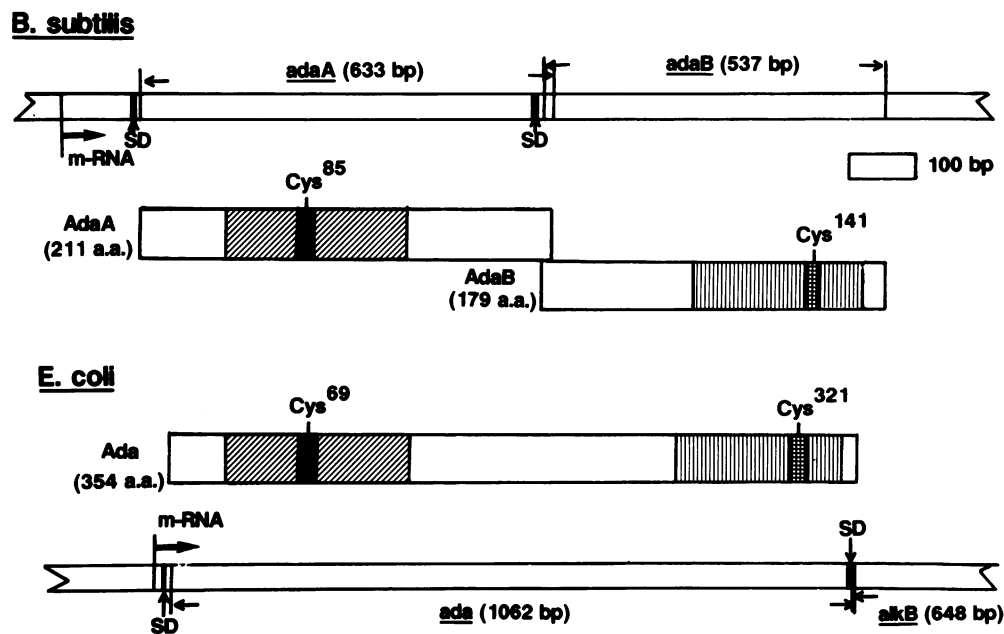
Fig. 8. Comparison of the structures of *B. subtilis* and *E. coli ada* operons. The regions of amino-acid sequence similarity (shaded and striped) and the regions of sequence identity (dark shaded and grated) are shown. Methyl-accepting cysteine residues in *E. coli* Ada and corresponding residues in *B. subtilis* AdaA and AdaB are shown.

(pHY1.2) cells contained only 27 Kd transferase. These results were consistent with the MNNG sensitivities of these transformed mutant strains. Namely, TKJ0922 (pHY1.2) cells were as resistant to MNNG as wild type cells. They were about 2.8 times more resistant than TKJ0922 cells when the concentrations of the drug that reduced the survival to 37 percent were compared.

TKJ2924 (pHY1.2) cells were as sensitive as the parental TKJ2924 cells. The extracts from control TKJ0922 (pHY5.2) cells and TKJ2924 (pHY5.2) cells contained both 22 Kd and 27 Kd transferase activities. Their total alkyltransferase activities were about 1.5 to 2 times higher than that of the adapted cells of the wild-type *ada$^+$* strain (data not shown). Similar

constitutive expression was observed when the *E. coli ada* gene was introduced into a multicopy plasmid (5). The activity was further enhanced about six times by adaptive pretreatment. In the extracts of both control and MNNG-pretreated cells of *ada* mutant strains carrying the pHY5.2, AdaB activity was higher than AdaA activity, as observed in the extract of pretreated cells of the wild type strain (18). These two transformed strains were about 3.2 times more resistant to MNNG than the wild type strain.

These results can be interpreted as follows; pHY1.2 can direct the synthesis of AdaA transferase but not of AdaB since the C-terminal coding sequence is missing, while pHY5.2 can direct the syntheses of both AdaA and AdaB transferases. When pHY1.2 is introduced into the *ada-1* mutant strain with a defect presumably in the *adaA* gene, methylated AdaA protein derived from the plasmid gene can activate the transcription of the chromosomal *ada* operon, leading to the synthesis of 22 Kd O⁶-methylguanine-DNA methyltransferase (AdaB protein), which is intact in the *ada-1* mutant. However, when the same plasmid is introduced into the *ada-4* mutant strain with a defect presumably in the *adaB* gene, it cannot complement the defect and the active AdaB protein is not synthesized.

DISCUSSION

Using the prophage transformation method, we cloned DNA fragments that can correct the MNNG-hypersensitivity of *ada* mutants. The cloned DNA carried two genes, *adaA* and *adaB* in tandem with an overlap of 11 bp, comprising a small operon. The *adaA* gene is assigned as the structural gene for methylphosphotriester-DNA methyltransferase and the *adaB* as that of O⁶-methylguanine-DNA methyltransferase.

The main object of this work was to compare the adaptive response of *E. coli* and *B. subtilis* at a molecular level. Our results showed intriguing differences in the structures of the *ada* operons as schematized in Fig.8. In both bacteria, all the inducible DNA alkyltransferase activities are derived from small operons, and the products of the operons act as a transcriptional activator upon collection of a methyl group from methylphosphotriesters and as a suicidal repair enzyme for mutagenic O-alkylated bases (4,6,9,10,37). These alkyltransferase activities reside in a product (Ada) of one gene in *E. coli* (3,6,10,22), while they are split in products (AdaA and AdaB) of two genes in *B subtilis*. The striking similarities in amino acid sequences of the proteins and overall functions of the operons indicate that the sequences of functional importance are conserved. The conserved sequences comprise the two domains for DNA alkyltransferase activities on methylphosphotriesters and O⁶-methylguanine residues. We do not know whether they were originally in one gene or two genes. However, the events causing the reorganization seem to have left a region of sequence divergence between the two domains, since the sequences in the middle part (corresponding to residues 128 to 241 of the *E. coli* Ada protein) are totally dissimilar.

Another difference in operon organization is that, in *E. coli*, there is an additional gene *alkB* that overlaps the termination codon of the *ada* gene by one bp (38). There is no large ORF within about 500 bp downstream region of the *B. subtilis adaB* gene. Since the largest transcript should terminate within about 300 bp from the end of the *adaB* coding sequence, we conclude that a gene corresponding to *alkB* does not exist in the *ada* operon of *B. subtilis*. The mutations in the *alkB* gene cause hypersensitivity to MMS, though the biochemical activity of the

gene product has not yet been identified (39,40). The absence of a similar gene in *B. subtilis* seems to explain why none of the six *B. subtilis ada* mutants exhibits hypersensitivity to MMS, unlike *E. coli ada* mutants (16,20).

We searched the NBRF protein data bank for sequence similarity, and found that a portion of AdaA is similar to the AraC proteins of *E. coli* (42,43) and *Salmonella typhimurium* (44). As shown in Fig. 3, in a stretch of 79 amino acids from residue 120 of AdaA, 27 amino acid residues (34%) are identical to those of the *E. coli* AraC protein. AraC protein is well characterized as a prokaryotic transcriptional regulator (45), so this similarity may be meaningful. In particular, the 20 amino acid residues starting at residue 197 of *E. coli* AraC form a helix-turn-helix structure that is implicated in binding to DNA (46). The corresponding sequence in the *B. subtilis* AdaA protein also yields a helix-turn-helix structure when analyzed by the algorithm of Garnier et al. (47). The sequence of the *E. coli* Ada protein aligned to this motif contains several residues identical to those in AdaA (12 of 20). However, it does not yield a typical helix-turn-helix structure by the same algorithm.

The AdaA protein also shows similarity to another classical DNA-binding protein, lambda CI repressor; the sequence of five amino acid residues, -Glu-Ser-Leu-Ala-Asp-, in the N-terminal helix (helix2) of the AdaA is almost identical to the corresponding sequence, -Glu-Ser-Val-Ala-Asp-, of the repressor. This sequence similarity is intriguing, because in this sequence of the CI repressor, two acidic amino acids (glutamic acid and aspartic acid) are considered to be crucial for transcriptional activation (48).

In *E. coli*, methylated Ada protein binds to the immediate 5'-sequence of the start sites of the *ada* operon and *alkA* gene (11). In this region, the two genes have a common sequence, AAAGCGCA, and this sequence ('ada box') has been shown to be important for regulation (11,12,49). The *B. subtilis ada* promoter does not, however, have a similar sequence. Thus the regulatory sequences in this operon remain to be identified.

In *B. subtilis*, the AdaA protein seems to be sufficient for transcriptional activation and the intact AdaB protein is not required. However, in *E. coli*, C-terminal truncation of the Ada protein modified its function as a transcriptional activator, indicating the C-terminal portion is also required for its proper function (50). Since proteolytic cleavage at the middle part of the protein leaves the activity of O⁶-methylguanine-DNA methyltransferase intact, this may be a way to economize the response by facilitating repair and terminating the response (4). In *B. subtilis*, however, such regulation is not possible. The present finding that in adapted cells transcription also started inside the *adaA* coding region suggests a way to produce the repair enzyme in larger amount than the activator protein. This may also explain our previous finding that in the adapted cells, the alkyltransferase activity on O⁶-methylguanine is consistently higher than that on methylphosphotriesters (18).

In the current model of the adaptive response (11), constitutive production of Ada protein (or AdaA protein in *B. subtilis*) at a low level is required. In our assays, neither the AdaA activity nor the *ada* transcript has been detected so far. A more sensitive assay of the transcript may be required to determine the exact level of constitutive expression. It is also imperative to define DNA-binding and activation functions of the AdaA protein by *in vitro* assays and mutational analyses. Finally, for obtaining information on the phylogenic relationship of *E. coli* and *B. subtilis* systems, extension of comparisons to other organisms seems indicated.

ACKNOWLEDGMENTS

REFERENCES

1. Samson, L. and Cairns, J. (1977) Nature, 267, 281–283.
2. Karran, P., Lindahl, T. and Griffin, B. (1979) Nature (London), 280, 76–77.
3. McCarthy, T. V. and Lindahl, T (1985) Nucleic Acids Res., 13, 2683–2698.
4. Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y. (1988) Ann. Rev. Biochem., 57, 133–157.
5. Sedgwick, B. (1983) Mol. Gen. Genet., 191, 466–472.
6. Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D. and Lindahl, T. (1985) Proc. Natl. Acad. Sci. USA, 82, 2688–2692.
7. Nakabeppu, Y., Kondo, H., Kawabata, S., Iwanaga, S. and Sekiguchi, M. (1985) J. Biol. Chem., 260, 7281–7288.
8. LeMotte, P. K. and Walker, G. C. (1985) J. Bacteriol., 161, 888–895.
9. Margison, G. P., Cooper, D. P. and Brennand, J. (1985) Nucleic Acids Res., 13, 1939–1952.
10. Sedgwick, B., Robins, P., Totty, N. and Lindahl, T. (1988) J. Biol. Chem. 263, 4430–4433.
11. Teo, I., Sedgwick, B., Kilpatrick, M. W., McCarthy, T. V. and Lindahl, T. (1986) Cell, 45, 315–324.
12. Nakabeppu, Y. and Sekiguchi, M. (1986) Proc. Natl. Acad. Sci. USA, 83, 6297–6301.
13. Kimball, R. F. (1980) Mutation Res., 72, 361–372.
14. Ceccoli, J., Rosales, N., Goldstein, M. and Yarosh, D. B. (1988) Mutation Res., 194, 219–226.
15. Morohoshi, F. and Munakata, N. (1983) Mutation Res., 110, 23–37.
16. Morohoshi, F. and Munakata, N. (1985) J. Bacteriol., 161, 825–830.
17. Morohoshi, F. and Munakata, N. (1986) Mol. Gen. Genet., 202, 200–206.
18. Morohoshi, F. and Munakata, N. (1987) J. Bacteriol., 169, 587–592.
19. Jeggo, P., Defais, M., Samson, L. and Schendel, P. (1977) Mol. Gen. Genet., 157, 1–9.
20. Jeggo, P. (1979) J. Bacteriol. 139, 783–791.
21. Demple, B. (1986) Nucleic Acids Res., 14, 5575–5589.
22. Teo, I., Sedgwick, B., Demple, B., Li, B. and Lindahl, T. (1984) EMBO J., 3, 2151–2157.
23. Makino, F. and Munakata, N. (1977) J. Bacteriol., 131, 438–445.
24. Seki, T., Miyachi, H., Yoshikawa, H., Kawamura, F. and Saito, H. (1986) J. Gen. Appl. Microbiol., 32, 73–79.
25. Ishiwa, H. and Shibahara-Sone, H. (1986) Jpn. J. Genet., 61, 515–528.
26. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory.
27. Frischauf, A. M., Lehrach, H., Poustka, A. and Murray, N. (1983) J. Mol. Biol., 170, 827–842.
28. Spizizen, J. (1958) Proc. Natl. Acad. Sci. USA, 44, 1072–1078.
29. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene, 33, 103–119.
30. Sanger, F., Coulson,A. R., Barrell, B. G., Smith, A. J. H. and Poe, B. A. (1980) J. Mol. Biol., 143, 161–178.
31. Gilman, M. Z. and Chamberlin, M. J. (1983) Cell, 35, 285–293.
32. Davis, L. G., Dibner, M. D. and Battey, J. F. (1986) Basic Methods in Molecular Biology. Elsevier, 276–284.
33. Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1989) Genomics, 5, 874–879.
34. Morohoshi, F., Hayashi, K. and Munakata, N. (1989) Nucleic Acids Res., 17, 6531–6543.
35. Potter, P. M., Wilkinson, M. C., Fitton J., Carr, F. J., Brennand, J., Cooper, D. P. and Margison, G. P. (1987) Nucleic Acids Res., 15, 9177–9193.
36. Kodama, K., Nakabeppu, Y. and Sekiguchi, M. (1989) Mutation Res. 218, 153–163.
37. Tano, K., Shiota, S., Collier, J., Foote, R. S., Mitra, S. (1990) Proc. Natl. Acad. Sci. USA, 87, 686–690.
38. Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S. and Sekiguchi, M. (1986) J. Biol. Chem. 261, 15,772–15777.
39. Kataoka, H., Yamamoto, Y. and Sekiguchi, M. (1983) J. Bacteriol. 153, 1301–1307.
40. Kataoka, H. and Sekiguchi, M. (1985) Mol. Gen. Genet., 198, 263–269.
41. Demple, B., Jacobsson, A., Olsson, M., Robins, P and Lindahl, T. (1982) J. Biol. Chem., 257, 13776–13780.
42. Wallace, R. G., Lee, N. and Fowler, A. V. (1980) Gene, 12, 179–190.
43. Miyada, C. G., Horwitz, A. H., Gass, L. G., Timko, J. and Wilcox, G. (1980) Nucleic Acids Res. 8, 5267–5274.
44. Clarke, P., Lin, H.-C. and Wilcox, G. (1982) Gene, 18, 157–163.
45. Greenblatt, J. and Schleif, R. (1971) Nature (London) New Biol. 233, 166–170.
46. Pabo, C. O. (1984) Ann. Rev. Biochem. 53, 293–321.
47. Garnier, J., Osguthorpe, D. J. and Robson, B. (1978) J. Mol. Biol., 120, 97–120.
48. Bushman, F. D., Shang, C. and Ptashne, M. (1989) Cell, 58, 1163–1171.
49. Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, J., Nakabeppu, Y. and Sekiguchi, M. (1988) J. Mol. Biol., 202, 483–494.
50. Shevell, D. E., Lemotte, P. K. and Walker, C. G. (1988) J. Bacteriol., 170, 5263–5271.