Bacillus subtilis gene coding for constitutive O^6-methylguanine-DNA alkyltransferase

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

Received June 5, 1989; Revised and Accepted July 17, 1989 EMBL accession no. X15659

ABSTRACT

We have cloned a *Bacillus subtilis* DNA fragment that could correct the defect in a constitutive O^6-methylguanine-DNA alkyltransferase (Dat1). This fragment also corrected the hypersensitivity of the strain TKJ6951(*ada-1 dat-1*) to N-methyl-N'-nitro-N-nitrosoguanidine (MNNG). In the fragment, the gene activity resides in a region of about 850 bp which contains an open reading frame capable of coding for a protein of 165 amino acid residues. The amino acid sequence of this protein exhibits striking similarity to those of *E. coli* O^6-methylguanine-DNA alkyltransferases (Ogt and Ada proteins). We conclude that this is a structural gene for the Dat1 protein, which is distinct from inducible DNA alkyltransferases involved in the adaptive response. The *dat-1* mutation was shown to be caused by a structural rearrangement affecting the coding region, and the 0.8 kb transcripts of this gene were detected in dat^+ cells but not in *dat* mutant cells.

INTRODUCTION

O^6-Methylguanine residues are highly mutagenic and possibly carcinogenic lesions produced by methylating agents such as N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) and N-methyl-N-nitrosourea (MNU) (1, 2). We have shown that *Bacillus subtilis* cells contain two species of O^6-methylguanine-DNA alkyltransferase which collect methyl groups from O^6-methylguanine residues in DNA (3). The 20 Kd protein (Dat1) is synthesized constitutively, whereas the 22 Kd protein (Dat2) is synthesized inducibly by cells cultured in the presence of low concentrations of methylating agents (termed 'adaptive treatment'). Adaptive treatment also induces the synthesis of a 27 Kd protein (Dat3) with the methylphosphotriester-DNA alkyltransferase activity (3). We have isolated six mutant strains (*ada*) sensitive to the lethal and mutagenic effects of methyl and ethyl nitroso compounds (4). None of them show an adaptive response to MNNG or MNU or inducible synthesis of Dat2. However, all these mutant cells have almost the same level of the Dat1 activity as wild-type cells (3). Furthermore, we have recently isolated a double-mutant strain, TKJ6951 (*ada-1 dat-1*), defective in both Dat1 and Dat2, and supersensitive to MNNG. From this strain, a *dat-1* single-mutant strain has been constructed (submitted for publication). These results are consistent with the notion that Dat1 and Dat2 are coded by separate genes that are regulated in different ways; one constitutively and one inducibly on adaptive treatment.

Recent studies suggested that *Escherichia coli* cells also have a similar dual defense system against DNA O-alkylation. The 37 Kd Ada protein, which has alkyltransferase activities on O^6-methylguanine and methylphosphotriester and is synthesized upon adaptive treatment (5), was considered to be the sole DNA alkyltransferase, but recently another

constitutive species of O^6-methylguanine-DNA alkyltransferase has been demonstrated (6−8).

We thought that further insight into this repair system could be obtained by comparisons of these two organisms at a molecular level. Here, we report the cloning, sequence determination and the expression of the *dat* gene coding for the constitutive DNA alkyltransferase in *B. subtilis* as a basis for analysis of the mechanism regulating expression of the gene.

MATERIALS AND METHODS

Bacterial strains, phage and plasmids

Bacillus subtilis strains 168T (*thyA thyB*) (9), TKJ1922 (*thyA thyB hisH101 metB101 leuA8 lys-21*) (4), TKJ0922 (the same as TKJ1922 but *ada-1*) (4), TKJ6951 (the same as TKJ0922 but *dat-1*) (submitted) were used. *B. subtilis* strain UOTO994 (*hisA1 metB5* ϕCM) (10) and *B. subtilis* temperate phage ϕ105 were obtained from Fujio Kawamura (Institute of Applied Microbiology, University of Tokyo). pHY300PLK (11), a shuttle plasmid between *E. coli* and *B. subtilis* was purchased from Toyobo (Tokyo) and pGEM2 from Promega (Madison, Wisconsin), respectively. *E. coli* strains HB101 and K802 used as hosts of plasmids, were as described (12).

Induction and purification of phage

B. subtilis cells lysogenized with ϕ105 or ϕCM were grown in LB broth containing 0.3% glucose and 5 mM $MgCl_2$ (13) until the absorbtion at 565 nm (A_{565nm}) of the culture reached 0.5. Cells lysogenized with recombinant ϕCM were grown until the A_{565nm} of the culture reached 0.2 to 0.3. At this point, mitomycin C (Kyowa Hakko, Tokyo) was added at a concentration of 0.5 μg/ml and after culture for 15 minute, the cells were collected by centrifugation and resuspended in the original volume of fresh LB broth. The cells were cultured until lysis (4 to 6 hours). Phages were purified as described (13).

Prophage transformation, transduction and selection of chloramphenicol resistant (Cm^r) and MNNG resistant ($MNNG^r$) colonies

Transformation and transduction were carried out as described (9, 14). Cm^r colonies were selected by plating the cells on nutrient agar plates containing 5 μg/ml of chloramphenicol. For selection of $MNNG^r$ colonies, clones were picked up and aligned on Spizizen minimum plate (15) containing required nutrients, and grown overnight at 37° C. They were then replica-plated on the same medium containing concentrations of MNNG of 12 to 20 μg/ml. After overnight culture, resistant colonies that grew on plates containing the drug were distinguishable.

Assay of survival and mutagenicity

These assays were carried out as described previously (4, 16).

Assay of DNA-alkyltransferase activity

Crude sonic extracts (16) were incubated with calf thymus DNA that had been treated with [^{3}H] MNU. The [^{3}H]-DNA was heated at 80°C for 16 hours before used in the assay [2]. The reaction mixtures were subjected to the following three procedures to assay DNA alkyltransferase activity. [1]. The mixtures were subjected to SDS-polyacrylamide gel (12.5%) electrophoresis, and a fluorogram of the gel was prepared and analysed as described before (3). [2]. The mixture was precipitated with an equal volume of 0.8 M TCA and the precipitate was suspended in 0.1 M Tris HCl (pH 8.0)-1 mM EDTA and treated with 300 μg/ml of proteinase K (Boehringer Mannheim) for 2 hours at 37°C. It was then mixed with an equal volume of 0.8 M TCA, and the radioactivity in the TCA-soluble fraction

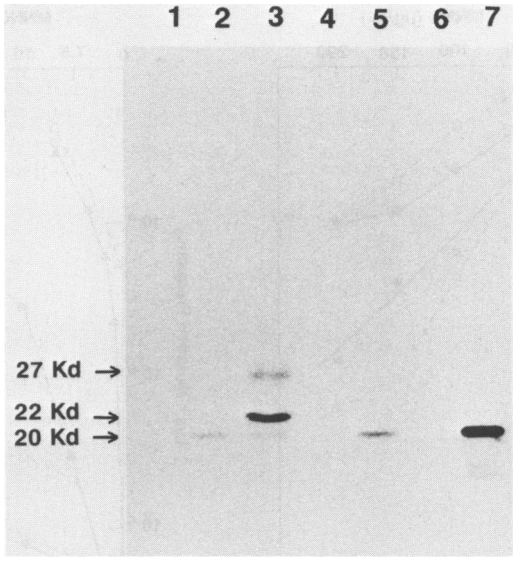

Fig. 1. Separation of O^6-methylguanine-DNA alkyltransferases of *B. subtilis* cells carrying cloned *dat* gene. A crude extract (120 μg protein) was incubated with [^{3}H]-MNU treated and heated DNA (2,700 cpm) for 20 min and subjected to SDS-polyacrylamide gel electrophoresis (12.5%). The fluorogram of the gel was prepared as described (3). Lane 1, without extract; 2, extract of 168T (*ada*$^+$ *dat*$^+$); 3, extract of adapted 168T cells; 4, extract of TKJ6951 (*ada-1 dat-1*)(φCM); 5, extract of TKJ6951(φCM 9 Kb); 6, extract of TKJ6951(pHY300PLK); 7, extract of TKJ6951(pHY*dat*$^+$).

after proteinase K treatment was determined (17). [3]. O^6-Methylguanine and 7-methylguanine remaining in [^{3}H] MNU treated DNA after incubation with extracts and acid hydrolysis were assayed by paper chromatography (16).

Restriction enzyme digestion and subcloning

Standard procedures were used (12). Restriction enzymes were purchased from Toyobo.

Introduction of deletions in subcloned fragments and sequence determination

The 2 Kb *Eco*RI fragment carrying the *dat* gene was inserted into the *Eco*RI site of pGEM2 in both orientations (pGem*dat*$^+$). For production of progressive unidirectional deletions, pGem*dat*$^+$ was cut at the *Hin*dIII site in the polylinker and was incubated with α-phosphorothioates of 4 dNTPs (Promega, Madison) and Klenow fragments of DNA polymerase I (Boehringer Mannheim) to fill the 5′-protruding ends by the procedure recommended by Promega. Linearized plasmids were digested with *Bam*HI and, from the *Bam*HI site, unidirectional deletions were produced using Exonuclease III (Bethesda Research Laboratories) and Exonuclease VII (Bethesda Research Laboratories) by the method of Yanisch-Perron et al.(18). The DNA sequences of the partially deleted plasmids were determined by the dideoxy chain termination method (19).

Extraction of RNA, Northern blotting and S1 mapping

Total RNA was extracted from logarithmically growing cells as described by Gilman and Chamberlin (20). Northern blotting and S1 nuclease mapping were carried out as described (12, 21). The probe used for Northern blot hybridization was a 0.36 Kb *Eco*47I-*Hin*cII fragment located in the coding region of the *dat* gene. The probe used for S1 nuclease

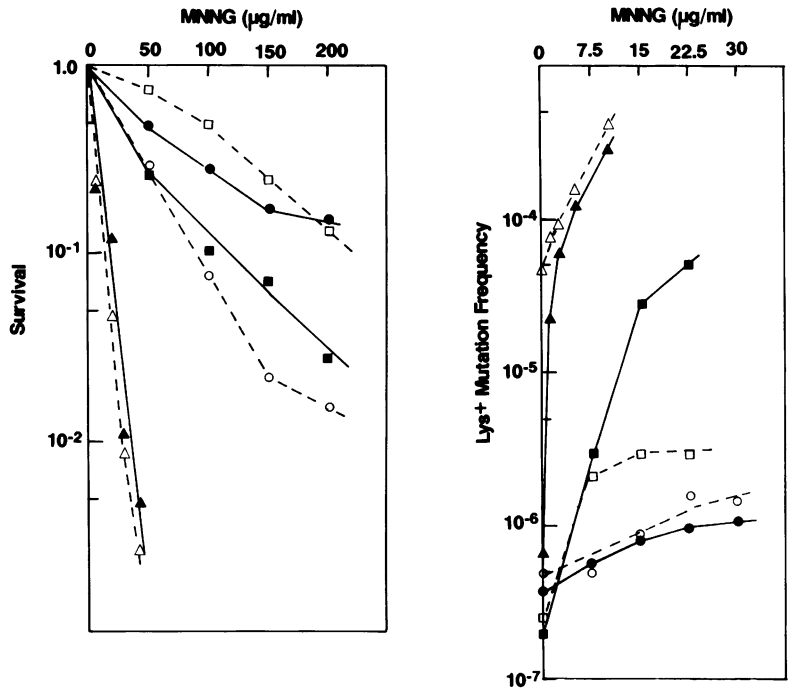

Fig. 2. Survival (left) and mutation frequency (right) of MNNG-treated cells harboring plasmids. Logarithmically growing cells (adapted and control) were treated with MNNG for 15 min in Spizizen minimum medium containing required nutrients and tetracycline. Adaptive pretreatment was for 70 min with 1 μg/ml of MNNG. Symbols: TKJ6951(pHY300PLK)(▲,△); TKJ6951(pHY*dat*$^+$)(●,○); TKJ1922(pHY300PLK)(■,□); closed symbols, control cells; open symbols, adapted cells.

mapping was prepared as follows. pGem218 (described in Fig. 4) was digested with *Eco*47I and labeled with γ-^{32}P ATP (10 Ci/ml, 3,000 Ci/mM, Amersham) using a 5′-end labeling kit according to the manual supplied by the manufacture (Boehringer Mannheim). The labeled fragments were digested with *Pvu*II and the digests were separated by neutral polyacrylamide gel (5%) electrophoresis. A radioactive band corresponding to the 0.4 Kb *Pvu*II-*Eco*47I fragment was recovered from the gel. Hybridization was carried out in 75% formamide at 48°C for 16 hours.

RESULTS

Cloning of a DNA fragment conferring MNNG resistance by the prophage transformation method

A *B. subtilis* DNA fragment that rendered the strain TKJ6951 (*ada-1 dat-1*) MNNG resistant (MNNG^r) was obtained by the prophage transformation method using a lysogenic phage vector ϕCM. This vector was a derivative of ϕ105 and carried the chloramphenicol resistance (Cm^r) gene (10). DNAs from strain 168T and from ϕCM were digested with *Bgl*II and ligated, and used to transform competent TKJ6951 cells lysogenized with ϕ105. The prophage genome provided the integration site for recombinant DNA molecules. Chloramphenicol resistant (Cm^r) transformants were picked up, grown on agar and replica

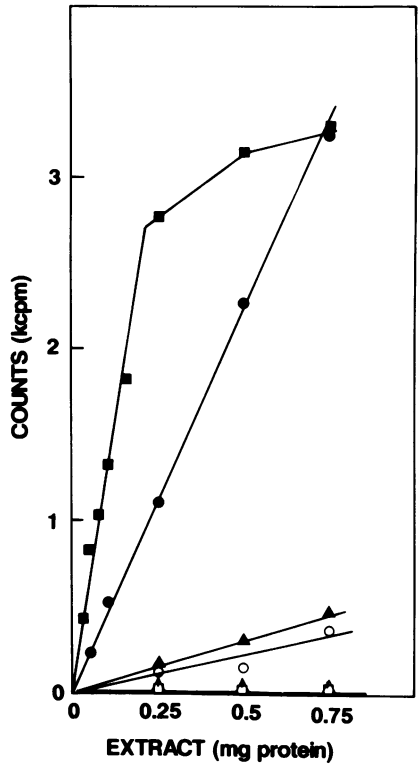

Fig. 3. Alkyltransferase activity in extracts of *B. subtilis* cells carrying the cloned *dat* gene. Various amounts of crude extracts were incubated with [^{3}H]-MNU treated and heated calf thymus DNA (4,500 cpm) for 20 min. Radioactivity transferred to protein was determined. Extract: 168T (*ada*$^+$ *dat*$^+$) (○); adapted 168T (●); TKJ6951 (*ada-1 dat-1*)(φCM) (△); TKJ6951(φCM 9 Kb) (▲); TKJ6951(pHY300PLK) (□); TKJ6951(pHY*dat*$^+$) (■).

plated on agar medium containing MNNG. Cells from nineteen Cm^r MNNG^r colonies were grown independently and treated with mitomycin C to induce the prophage. The phage were used to infect the TKJ6951 cells together with wild-type helper phage, and Cm^r MNNG^r transductant colonies were selected as above. All the recombinant φCMs recovered from three clones of Cm^r MNNG^r transductants contained the same 9 Kb *Bgl*II fragment of chromosomal origin. Cells of strain TKJ6951 lysogenized with φCM(*Bgl*II 9 Kb) did not show an adaptive response and were as MNNG-sensitivie as *ada-1* cells (data not shown). The extract of the cells contained DNA alkyltransferase activity of 20 Kd, but not of 22 Kd or 27 Kd (Fig. 1). Therefore, the 9 Kb *Bgl*II fragment was concluded to contain a gene complementing the *dat-1* mutation, but not the *ada-1* mutation.

For more exact determination of the region with activity, the 9 Kb insert was digested with *Eco*RI and ligated to pHY300PLK digested with the same enzyme. The ligation mixture was then used to transform *E. coli* cells of strain K802. Plasmids containing foreign DNA were isolated, from the tetracycline-resistant (Tet^r) transformants, and used to transform *B. subtilis* cells of strain TKJ6951 (*ada-1 dat-1*). MNNG^r Tet^r transformants were selected. All plasmids obtained from Tet^r MNNG^r transformant strains contained a 2 Kb *Eco*RI fragment. The cells harboring this plasmid (pHY*dat*$^+$) showed almost the same

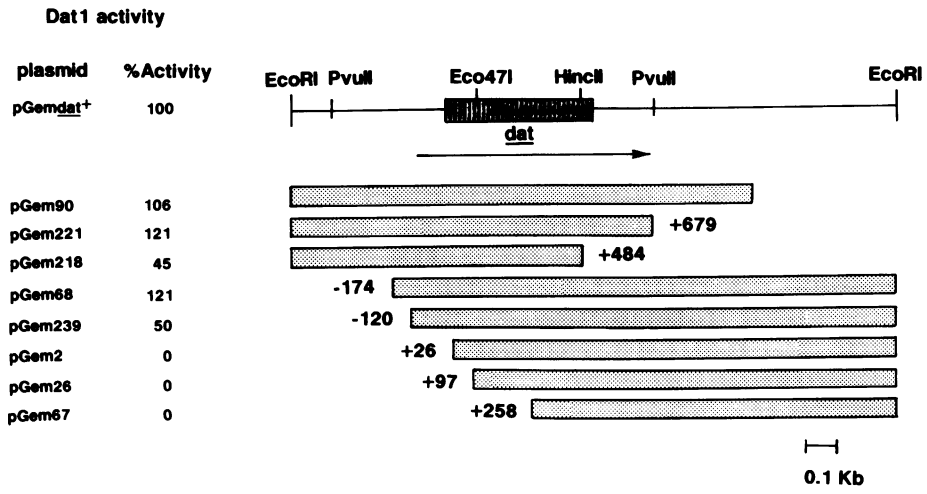

Fig. 4. Dat1 activity in *E. coli* harboring various deletion plasmids of pGEM2. Various unidirectional deletions were produced in a 2 Kb *Eco*RI fragment introduced into pGEM2 as described in Materials and Methods. The arrow under the restriction map indicates the *dat* transcript. Activities are expressed as percentages of the activity of pGem(*dat*$^+$). Numbers on the left or right of bars show terminal nucleotide numbers of the deletion fragments taking translation initiation A as +1.

level of MNNG resistance as wild type cells, and the frequency of *lys-21* reversion induced by MNNG became even lower than that in adapted cells (adaptive treatment for 70 min with 1 μg/ml of MNNG) of the wild type strain TKJ1922 (Fig. 2). The extract from strain TKJ6951(pHY*dat*$^+$) contained a high level of 20 Kd O^6-methylguanine-DNA methyltransferase (Dat1) activity (Fig. 1). These results showed that the 2 Kb fragment was sufficient to complement the *dat-1* mutation.

Alkyltransferase activity in cells harboring the dat gene

The activity of DNA alkyltransferase was quantitated by a method employing proteinase K as described in the Materials and Methods. No activity was detectable in cells of strain TKJ6951(φCM) or TKJ6951(pHY300PLK). The extract from cells of strain TKJ6951(φCM*dat*$^+$) exhibited a low level of activity which was equivalent to that of the wild type strain 168T. On the other hand, the extract from cells of strain TKJ6951(pHY*dat*$^+$) and from adapted cells of wild type strain 168T exhibited high levels of activity. On increase in protein concentration in the reaction mixture, the activity of the extract of strain TKJ6951(pHY*dat*$^+$) increased almost linearly to a plateau with over 0.25 mg protein, whereas that of the extract of the adapted 168T cells increased linearly to at least the highest concentration tested (0.75 mg protein). DNA alkyltransferases are used only once and are inactivated by the reaction (22). In the linear portions of the activity curves, the amounts of protease-sensitive radioactivity reflect the amounts of DNA alkyltransferase molecules, so the relative contents of enzyme in each extract could be estimated from the slopes of these linear portions in Fig. 3. In this way, the Dat1 content of the TKJ6951(pHY*dat*$^+$) extract was estimated to be about 40 times that in the wild-type (strain 168T) extract without adaptive treatment. It was also three times that of the extract from adapted wild-type cells. In the latter extract, the activity is attributable to

```
 -200
GACAGAAGAT AAATTAAAGC GACTTGGTGA ACCGTTTTAT ACGACAAAGG AACGGGGAAC
                                                -100
TGGACTTGGG CTTATGGTAA GCTATAAAAT TATTGAAGAA CATCAAGGCG AGATTATGGT

AGAAAGTGAA GAAGGGAAGG GCACCGTTTT CCATATTACG CTTCCTGTCA GACAGAATGC
                          1
TGAAGAAAGA AGGAATGATG AATGAACTAC TATACGACAG CCGAAACGCC GCTCGGTGAA

CTTATCATTG CCGAAGAGGA GGACCGGATC ACTCGTCTAT TTCTCAGTCA GGAAGATTGG
100
GTGGATTGGA AAGAAACGGT TCAGAATACT GAGCATAAGG AAACACCTAA TCTTGCAGAA
                                                200
GCGAAACAAC AGCTTCAAGA ATATTTCGCA GGCGAAAGGA AGACATTCTC CCTGCCGCTC

AGCCAAAAGG GCACTCCTTT TCAGCAAAAA GTGTGGCAAG CGCTGGAGAG GATTCCATAT
                          300
GGCGAATCCC GAAGCTATGC GGATATTGCC GCTGCTGTCG GCAGTCCGAA AGCGGTGCGC

GCTGTCGGGC AGGCTAATAA ACGTAACGAC CTGCCGATTT TTGTCCCGTG CCACAGAGTG
400
ATCGGCAAAA ACAGTGCGTT AACAGGATAC GCCGGAAGCA AAACGGAGAT CAAAGCATTT
                                                500
TTGCTAAACA TCGAGCGAAT CTCCTATAAA GAAAAATAAA ACATATGGCA CGTTCCCTTT

TTTCACGGCC GACATATCAT TTTATTAAGT TAAAAAAGTT GGTCTGATGA AAAAGGAGGA
                          600
CAAAGCAATT GAAGCCGAGC CGCTCTGAGA AACTTGCCGT
```

Fig. 5. Nucleotide sequence of *B. subtilis dat* locus. Underlines and the broken underline show predicted protein coding regions and the ribosome binding site, respectively. Predicted promoter regions are boxed. The arrow indicates the transcription initiation site determined by S1 nuclease mapping.

Dat2 (O^6-methylguanine-DNA alkyltransferase) and Dat3 (methylphosphotriester-DNA alkyltransferase) in a ratio of about 2 : 1 (3). Therefore, the activity of TKJ6951 (pHY*dat*$^+$) cells for repair of O^6-methylguanine residues should be about 4.5 times that of adapted, wild type cells. This is probably the reason for the better protection afforded by this plasmid against MNNG challenge.

Sequence of the dat gene

To determine the base sequence of the 2 Kb *Eco*RI fragment, we inserted it into pGEM2 in both orientations. Unidirectional deletions of various sizes were produced from the *Bam*HI site in the polylinker sequence. As *E. coli* HB101 cells harboring pGem*dat*$^+$ expressed Dat1 activity, the cells harboring the deletion plasmids were tested to determine whether they retained this activity (Fig. 4). About 0.8 Kb from the 3′-*Eco*RI site to the *Pvu*II site shown in Fig. 4 could be deleted without affecting the activity (pGem221). The significance of the small increase (20%) was not known. However, further deletion of about 0.2 Kb from this *Pvu*II site decreased the activity about 55% (pGem218). Deletion of 0.3 Kb in the opposite direction from the 5′-*Eco*RI site did not affect the activity (pGem68). But when an additional sequence of 50 bp were deleted, the activity was reduced 50% (pGem239), and a further deletion of 145 bp abolished the activity (pGem2). These results indicated that a region of about 850 bp inside the two *Pvu*II sites was sufficient for expression of Dat1 activity in *E. coli* cells. The nucleotide sequence of this region was determined and is shown in Fig. 5. This region contains an open reading frame (ORF) coding for a protein consisting of 165 amino acid residues. This ORF is preceded by a possible Shine-Dalgarno consensus sequence (AGAAAGAAGG) for ribosome binding. As the size of the protein was consistent with that predicted for Dat1, and as extensive similarity to the

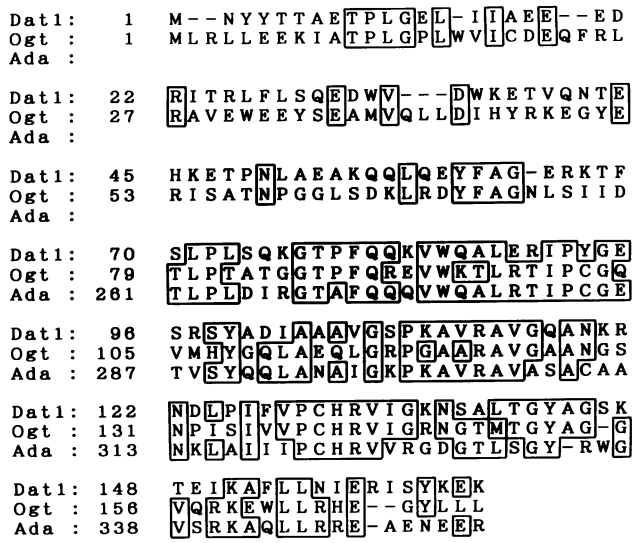

Fig. 6. Comparison of amino acid sequence of *B. subtilis* Dat1 and *E. coli* Ogt and Ada. Only carboxy terminal 94 amino acids of the sequence of Ada (24) are shown.

amino acid sequences of *E. coli* DNA alkyltransferases was found, as described below, we conclude that this is the structural gene for the Dat1 protein.

The predicted amino acid sequence of this ORF was compared with those of two species of *E. coli* DNA alkyltransferases, Ogt and Ada, by the MICROGENIE program (Beckmann). The Ogt protein is composed of 171 amino acid residues and has O^6-methylguanine-DNA alkyltransferase activity (6). Ada protein is composed of 354 amino acid residues and is multifunctional, the carboxy-terminal domain having O^6-methylguanine-DNA alkyltransferase activity (23, 24). The three sequences were aligned for maximal matching as shown in Fig. 6. Clear sequence similarity is observed between the entire sequence of this Dat1 protein and that of the Ogt protein (37% identity), characterized by many identical blocks of one to eight amino acid residues. A stretch of 96 amino acid residues in the carboxy-terminal portion of Dat1 is also remarkably similar to the corresponding portions of the Ogt and Ada proteins. In these stretches of the three proteins, 34 amino acid residues are identical, and the identities are 47% between Dat1 and Ogt, 49% between Dat1 and Ada, and 48% between Ogt and Ada. The similarity extends to the amino-terminal portions of Dat1 and Ogt, but, a corresponding sequence is not found in the Ada sequence.

Nature of the dat-1 mutation

To determine the molecular basis of the *dat-1* mutation, we compared the restriction fragments of DNAs from cells of TKJ1922 (*ada*$^+$ *dat*$^+$), TKJ0922 (*ada-1*) and TKJ6951 (*ada-1 dat-1*) by Southern blot analysis (Fig. 7). The 1.2 Kb *Pvu*II fragment and 0.36 Kb *Eco*47I-*Hin*cII fragment containing the coding region of the *dat* gene were used as probes. When the DNAs of the *dat*$^+$ strains were digested with the same restriction enzymes used for the probes, radioactive bands were observed in positions corresponding to the sizes of the probes. On the other hand, the *Pvu*II-digest of DNA of the *dat* strain gave two bands in positions corresponding to 1.4 Kb and 0.7 Kb instead of the band in

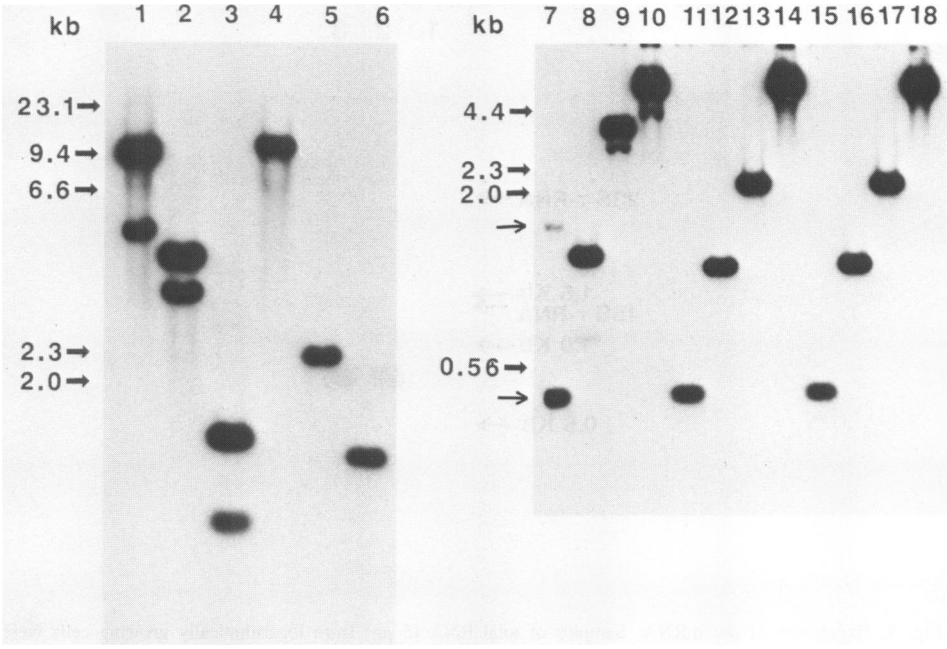

Fig. 7. Southern blot analysis of the *dat* locus of *dat*$^+$ and *dat* strains. DNA (1 μg) was digested with *Bgl*II (lanes, 1, 4, 10, 14, 18), *Eco*RI (lanes, 2, 5, 9, 13, 17) *Pvu*II (lanes, 3, 6, 8, 12, 16) or *Eco*47I and *Hin*cII (lanes, 7, 11, 15) and was separated by electrophoresis in 0.5% (for lanes 1 to 6) or 0.8% (for lanes 7 to 18) agarose gel. DNA was transferred to nylon membranes, fixed and hybridized with the probes. Source of DNA: lanes 1−3 and 7−10, TKJ6951(*ada-1 dat-1*), lanes, 4−6 and 15−18; TKJ1922(*ada*$^+$ *dat*$^+$); lanes 11−14 TKJ0922 (*ada-1 dat*$^+$); the probe used for lanes 1−6 was 1.2 Kb *Pvu*II fragment and that for lanes 7−18 was 0.36 Kb *Eco*47I-*Hin*cII fragment.

a position corresponding to 1.2 Kb. When the 0.36 Kb *Eco*47I-*Hin*cII fragment was used to hybridize the DNA digested with the same two enzymes, one smaller and one larger band than the probes were detected as indicated by arrows. These results suggest that in strain TKJ6951 the coding region of the *dat* gene is split by DNA rearrangement. This might be expected as this mutant was obtained after severe exposure of the spores to gamma radiation and we have not succeeded in transferring the mutation to other strains. This result, together with analysis of the transcript described in the next section, is consistent with our conclusion that the *dat-1* allele is null and completely lacks the Dat1 protein.

Transcription of the dat gene

Transcripts of the *dat* gene were examined by Northern blot analysis of total RNAs extracted from log-phase cells of strains TKJ1922 (*ada*$^+$ *dat*$^+$), TKJ0922 (*ada-1 dat*$^+$) and TKJ6951 (*ada-1 dat-1*). The 0.36 Kb *Eco*47I-*Hin*cII fragment was used as a probe. RNAs from the *dat*$^+$ cells but not from the *dat-1* cells gave a radioactive band in a position corresponding to about 800 bases (Fig. 8). S1 nuclease mapping with RNA from *dat*$^+$ cells demonstrated that the transcription started at the sequence ATCAA, which is located about 100 bp upstream of the initiation codon of the protein (Fig. 9). RNA from cells of the *dat-1* strain gave faint bands in the same region, indicating that initiation of transcription could take place at a lower frequency in this strain, but that transcription

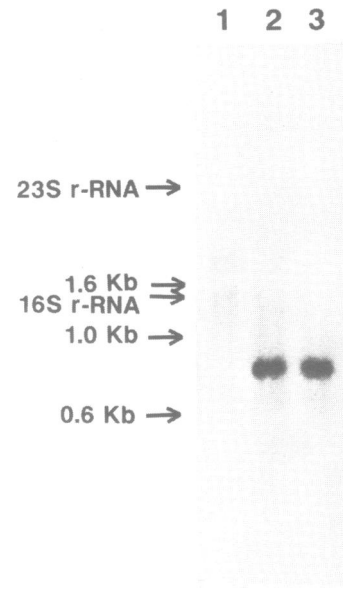

Fig. 8. Expression of *dat* mRNA. Samples of total RNA (5 μg) from logarithmically growing cells were electrophoresed in 1.4% agarose-formaldehyde gel, transferred to a nylon membrane, fixed and hybridized with the ^{32}P-labeled 0.36 Kb *Eco*47I-*Hin*cII fragment. Source of RNA, lane 1, TKJ6951 (*ada-1 dat-1*), lane 2, TKJ0922 (*ada-1 dat*$^+$), lane 3, TKJ1922 (*ada*$^+$ *dat*$^+$).

was aberrant and did not yield the 800 base form. Consensus promoter sequences, TGGACT (−35) and TAAAAT (−10), are seen from 41 bp and 17 bp, respectively, upstream of the initiation site of transcription.

DISCUSSION

We have obtained a DNA fragment of *B. subtilis* that confers MNNG resistance on MNNG-supersensitive cells of strain TKJ6951 (*ada-1 dat-1*) by prophage transformation. A 9 Kb *Bgl*II fragment restored the defect in the synthesis of Dat1 protein, but not of Dat2 or Dat3. By subcloning, the gene activity was found to reside in a 2 Kb *Eco*RI fragment. A region of about 850 bp in this fragment was essential for the synthesis of Dat1 in *E. coli* host cells. Sequence analysis of this portion revealed the presence of an ORF encoding 165 amino acid residues with a calculated molecular mass of 18.8 Kd. As this size is consistent with that of Dat1 protein (20 Kd) estimated by SDS-polyacrylamide gel electrophoresis (3), and as the sequence of the protein is remarkably similar to those of the Ogt and Ada proteins of *E. coli* (6, 23, 24), we conclude that this DNA sequence is the structural gene for the Dat1 protein, a constitutive O^6-methylguanine-DNA alkyltransferase in *B. subtilis* and term this gene *dat*. The mutation *dat-1* was found to be caused by a rearrangement affecting this gene.

This is the first DNA alkyltransferase gene cloned from an organism other than *E. coli*. The *dat* gene of *B. subtilis* is probably a functional homologue of the *ogt* gene of *E. coli* (6). The two genes encode O^6-methylguanine-DNA alkyltransferases of similar molecular masses and similar amino acid sequences and are expressed constitutively (6), although

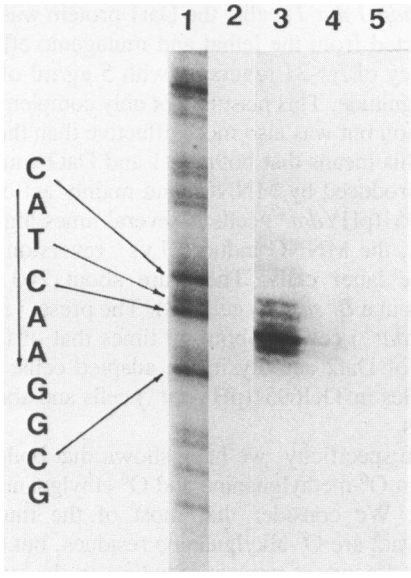

Fig. 9. S1 nuclease analysis of *dat* mRNA. The same RNA as used for Northern blotting was hybridized to the 5'-end labeled 0.4 Kb *Pvu*II-*Eco*47I fragment and treated with S1 nuclease as described in Materials and Methods. Lane 1, G + A of the probe fragment determined by the method of Maxam and Gilbert, lane 2, without RNA, 3. RNA from strain TKJ1922 (*ada*$^+$ *dat*$^+$), lane 4, RNA from TKJ6951 (*ada-1 dat-1*), 5, *E. coli* t-RNA.

the cellular function of the *ogt* gene is not known exactly as no mutation of this gene has yet been reported.

Transcription of the *dat* gene starts about 100 bp upstream of the initiation codon. The *dat* promoter sequences TGGACT (−35) and TAAAAT (−10) with a space of 18 bp between them, are likely to be recognized by a vegetative form of a sigma factor (σ^{43}) (25). There is an ORF which could encode a polypeptide of 29 amino acids residues in this region and the termination codon of this polypeptide overlaps the initiation codon of Dat1. Although we have no evidence for the synthesis of this polypeptide, expression of the *dat* gene might be controlled by translation of this proximal gene by translational coupling (26, 27).

It has been shown that Cys^{321} of the Ada protein accepts a methyl group from an O^6-methylguanine residue (23). A pentapeptide surrounding this cysteine (Pro-Cys-His-Arg-Val) is perfectly conserved in the Ada, Ogt and Dat1 proteins (6, 23, 24). In Dat1, Cys^{129} in this pentapeptide is the only cysteine residue and presumably functions as a methyl acceptor, too. Demple et al. noticed that the reaction center for O^6-methylguanine-DNA alkyltransferase is composed of the sequence Pro-Cys-His-Arg flanked by a few hydrophobic amino acid residues on the amino terminal side (23). It now appears that this tetrapeptide is flanked by two hydrophobic residues on the carboxy-terminal side and three on the amino-terminal side (Fig. 6). We do not know the functional significance of the conserved sequences other than this reaction center, but the conservation of the residues in phylogenically distant organisms suggests that they are important for the activity of this unique enzyme.

When a multicopy plasmid (pHY*dat*$^+$) harboring this gene was introduced into MNNG-

supersensitive TKJ6951 (*ada-1 dat-1*) cells, the Dat1 protein was synthesized actively and the cells were fully protected from the lethal and mutagenic effects of MNNG (Fig. 2). For example, the frequency of *lys-21* reversion with 5 μg/ml of MNNG was reduced to less than two orders of magnitude. This plasmid not only complemented the Dat1 deficiency caused by the *dat-1* mutation but was also more effective than the capacity of the adaptive response in *ada*$^+$ cells. This means that both Dat1 and Dat2 can repair most of the lethal and mutagenic damages produced by MNNG, and mainly act on the same damages. As the Dat1 content in TKJ6951(pHY*dat*$^+$) cells is several times higher than the Dat2 content in adapted wild type cells, the MNNG-induced Lys$^+$ reversion frequency in the former is lower than that in the latter cells. There are about 240 O^6-methylguanine-DNA alkyltransferase molecules in a *B. subtilis* cell (28). The present results show that the Dat1 activity of TKJ6951(pHY*dat*$^+$) cells is about 40 times that of un-treated wild type cells and about 4.5 times that of Dat2 activity in the adapted cells, indicating that there are about 1×10^4 Dat1 molecules in TKJ6951(pHY*dat*$^+$) cells and about 2,000 Dat2 molecules in adapted wild type cells.

With regard to substrate specificity, we have shown that both Dat1 and Dat2 proteins collected alkyl groups from O^6-methylguanine and O^6-ethylguanine residues, but not from alkylphosphotriesters (3). We consider that most of the mutagenic lesions that are counteracted by Dat1 or Dat2 are O^6-alkylguanine residues, but the contributions of other alkylated products are not known at present. Studies on the substrate specificity of the Dat1 protein should be facilitated by the availability of the overproducing plasmid.

Constitutive activity of DNA alkyltransferase may be ubiquitous in various prokaryotic species (3, 6–8, 28, 29) and in mammalian tissues (30–32). These enzymes are all similar in size, although they could vary greatly in activity in different species or tissues (3, 6, 30, 31). On the other hand, the capacity for an adaptive response seems to be restricted to a relatively few bacterial species (29). This means that different organisms may each have evolved their own way to cope with the highly mutagenic stress posed by environmental alkylating agents. We are currently trying to determine the structure of the *ada* gene in *B. subtilis*, and the results, together with those on the *dat* gene, should clarify the strategy employed by this organisms. We hope that the results will also provide a basis for further elucidation of this repair system in various organisms.

ACKNOWLEDGMENTS

We are grateful to Dr. Fujio Kawamura (Institute of Applied Microbiology, University of Tokyo) for phage vector and advice and to Miss Reiko Makino of this institute for technical instruction. We also thank Dr. Hiroshi Tanooka of this laboratory for financial support. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Schendel, P. F. and Robins, P. E. (1978) Proc. Natl. Acad. Sci. 75, 6017–6020.
2. Zarbl, H., Sukumar, S., Arthur, A. V., Martin-Zanca, D. and Barbacid, M. (1985) Nature 315, 382–385.
3. Morohoshi, F. and Munakata, N. (1987) J. Bacteriol. 169, 587–592.
4. Morohoshi, F. and Munakata, N. (1985) J. Bacteriol. 161, 825–830.
5. Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y. (1988) Ann. Rev. Biochem. 57, 133–157.
6. Potter, P. M., Wilkinson, M. C., Fitton, J., Carr, F. J., Brennand, J., Cooper, D. P. and Margison, G. P. (1987) Nucleic Acid Res. 15, 9177–9193.
7. Rebeck, G. W., Coons, S., Carroll, P. and Samson, L. (1988) Proc. Natl. Acad. Sci. USA 85, 3039–3043.
8. Shevell, D., Abon-Zamzam, A. M., Demple, B. and Walker, G. C. (1988) J. Bacteriol. 170, 3294–3296.

9. Morohoshi, F. and Munakata, N. (1977) J. Bacteriol. 131, 438–445.
10. Seki, T., Miyachi, H., Yoshikawa, H., Kawamura, F. and Saito, H. (1986) J. Gen. Appl. Microbiol. 32, 73–79.
11. Ishiwa, H. and Shibahara-Sone, H. (1986) Jpn. J. Genet. 61, 515–528.
12. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory.
13. Kawamura, F. and Ito, J. (1974) Virology 62, 414–425.
14. Iijima, T., Kawamura, F., Saito, and Ikeda, Y. (1980) Gene 9, 115–126.
15. Spizizen, J. (1958) Proc. Natl. Acad. Sci. USA 44, 1072–1078.
16. Morohoshi, F. and Munakata, N. (1983) Mutation Res. 110, 23–37.
17. McCarthy, T. V. and Lindahl, T. (1985) Nucleic Acid Res. 13, 2683–2698.
18. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103–119.
19. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl Acad. Sci. USA 74, 5463–5467.
20. Gilman, M. Z. and Chamberlin, M. J. (1983) Cell 35, 285–293. 21. Davis, L. G., Dibner, M. D. and Battey, J. F. (1986) Basic Methods in Molecular Biology. Elsevier 276–284.
22. Lindahl, T., Demple, B. and Robins, P. (1982) EMBO J. 1, 1359–1363
23. Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D. and Lindahl, T. (1985) Proc. Natl. Acad. Sci. USA 82, 2688–2692.
24. Nakabeppu, Y., Kondo, H., Kawabata, S., Iwanaga, S. and Sekiguchi, M. (1985) J. Biol. Chem. 260, 7281–7288.
25. Moran , C. P., Lang, N., LeGrice, S., Lee, G., Stephens, M., Sonenshein, A., Pero, J. and Losick, R. (1982) Mol. Gen. Genet. 186, 339–334.
26. Das, A. and Yanofsky, C. (1984) Nucleic Acids Res. 12, 4757–4768.
27. Aksoy, S., Squires, C. L. and Squires, C. (1984) J. Bacteriol. 157, 363–367.
28. Hadden, C. T., Foote, R. S. and Mitra, S. (1983) J. Bacteriol. 153, 756–762.
29. Ceccoli, J., Rosales, N., Goldstein, M. and Yarosh, D. B. (1988) Mutation Res. 194, 219–226.
30. Harris, A. L., Karran, P. and Lindahl, T. (1983) Cancer Res. 43, 3247–3252.
31. Yagi, T., Yarosh, D. B. and Day, III, R. F. (1984) Carcinogenesis 5, 593–600.
32. Grafstrom, R. C., Pegg, A. E., Trump, B. F. and Harris, C. C. (1984) Cancer Res., 44, 2855–2857.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.