Volume 15 Number 22 1987 Nucleic Acids Research

Characterisation and nucleotide sequence of *ogt*, the O^6-alkylguanine-DNA-alkyltransferase gene of *E.coli*

P.M.Potter, M.C.Wilkinson, J.Fitton[1], F.J.Carr[1], J.Brennand*, D.P.Cooper and G.P.Margison

Department of Carcinogenesis, Paterson Institute for Cancer Research, Christie Hospital and Holt Radium Institute, Manchester M20 9BX and [1]Pharmaceutical division, ICI plc, Alderley Park, Cheshire, UK

Received September 22, 1987; Revised and Accepted October 28, 1987 Accession no. Y00495

ABSTRACT

The plasmid p061 that was isolated from an E.coli genomic DNA library and codes for O^6-alkylguanine (O^{6}AG) DNA alkyltransferase (ATase) activity (1) has been further characterised. Subclones of the 9 Kb insert of p061 showed that the ATase activity was encoded in a 2Kb Pst1 fragment but a partial restriction endonuclease map of this was different to that of the E.coli ada gene that codes for O^6-AG and alkylphosphotriester dual ATase protein. Fluorographic analyses confirmed that the molecular weight of the p061-encoded ATase was 19KDa i.e. similar to that of the O^{6}AG ATase function that is cleaved from the 39KDa ada protein but rabbit polyclonal antibodies to the latter reacted only very weakly with the p061-encoded protein. A different set of hybridisation signals was produced when E.coli DNA, which had been digested with a variety of restriction endonucleases was probed with 2Kb Pst 1 fragment or the ada gene. These results provided evidence for the existence of a second ATase gene in E.coli. The 2Kb Pst-1 fragment of p061 was therefore sequenced and an open reading frame (ORF) that would give rise to a 19KDa protein was identified. The derived amino acid sequence of this showed a 93 residue region with 49% homology with the O^{6}AG ATase region of the ada protein and had a pentamer and a heptamer of identical sequence separated by 34 amino acids in both proteins. The pentamer included the alkyl accepting cysteine residue of the ada O^{6}AG ATase. The hydrophobic domains were similarly distributed in both proteins. Shine-Dalgarno, -10 and -35 sequences were identified and the origin of transcription was located by primer extension and S1 nuclease mapping. The amino-terminal amino acid sequence of the protein was as predicted from the ORF.

INTRODUCTION

Alkylating agents can react at 12 sites in the purine and pyrimidine bases in DNA and with the phosphodiester linkage to produce alkylphosphotriesters (AP) (2,3). It has been shown that mammalian cells and bacteria possess repair enzymes that can eliminate certain of these modifications from DNA (4) and this can provide protection against some of the biological effects of these agents. Active removal of one such lesion O^6-alkylguanine (O^{6}AG) from DNA in E.coli (5) and rat liver (6) has been known for some time and the mechanism in both cases has been shown to involve the transfer of the alkyl group at the O^6 position to a cysteine residue in the repair protein itself (7,8), a process that is autoinactivating.

Having devised a sensitive and convenient assay for ATase activity we used it to screen an E.coli genomic DNA library cloned in the expression vector pUC8 with the intention of isolating the O^6AG ATase gene (1). Two bacterial clones expressing high levels of ATase activity were identified and the corresponding plasmids called p061 and p062 were isolated. Although extracts of bacteria harbouring these plasmids catalysed the transfer of radiolabelled methyl groups from N-[^{3}H]-methyl-N-nitrosourea (MNU)-methylated substrate DNA to protein, further characterisation showed that p061 coded for an O^6AG ATase function whilst p062 encoded both this and an additional ATase function acting on AP. The 1.1Kb gene responsible for the dual ATase activity was found to be a previously identified gene ada (9). The ada gene was known to control the "adaptive" response in E.coli in which exposure of bacteria to low doses of N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) increased their resistance to the mutagenic and toxic effects of a higher "challenging" dose of this agent (10).

Extracts of bacteria harbouring p061 contained an ATase protein of closely similar molecular weight to one of the breakdown products of the ada protein and transferred the methyl group from O^6-methylguanine (O^6-meG) in DNA to a cysteine residue in protein (1). For these and other reasons we originally assumed that p061 was a truncated or rearranged section of E.coli DNA in which the 3' region of the ada gene that codes for O^6AG ATase activity had come under the influence of the β-galactosidase promoter in the vector. This assumes that the ATase coding region is in frame with the translation initiation codon in the expression vector used (pUC8) and would give rise to a fusion protein. In the present paper we initially explored this possibility further by subcloning the original 9Kb insert in p061 and restriction endonuclease (RE) mapping; Southern analysis of E.coli DNA using p061 or ada DNA as probes and western blotting using an antibody to the O^6AG ATase region of the ada protein. The results were consistent with the existence of a second ATase gene in E.coli. p061 was subcloned to a 2Kb Pst-1 fragment and this was sequenced. Computer-derived amino acid sequence data, primer extension and S_1 nuclease mapping and amino terminal amino acid sequencing were used to confirm the presence of a novel O^6AG ATase gene of E.coli origin.

A preliminary report of some of this data has been published earlier (11)

MATERIALS AND METHODS

Alkyltransferase assay

Bacteria (usually 1ml of a stationary culture) were collected by centrifugation and sonicated (two pulses of 10 seconds at 10μ peak to peak

distance with cooling in ice in between) in 1ml of buffer I (50mM Tris HCl, 1mM EDTA, 3mM dithiothreitol pH 8.3). Phenylmethylsulphonyl fluoride was added to 0.5mM immediately after the last sonication. Sonicates were clarified by centrifugation and supernatants assayed for ATase activity essentially as described previously (1) except that the total incubation volume was 1.1ml and the substrate DNA (5µg) was prepared by reaction of purified calf thymus DNA with N-[^{3}H]-methyl-N-nitrosourea of 23Ci/mmole (Amersham International).

Extracts which were to be subjected to electrophoresis and fluorography after radiolabelling were placed on ice immediately after incubation and made 0.25M with respect to HCl. The precipitate was collected by centrifugation and washed twice in 90% ethanol, containing 25mM Tris pH 6.8 before redissolving in PAGE loading buffer (12).

Polyacrylamide gel electrophoresis

Protein extracts were analysed by discontinuous buffer SDS electrophoresis using an adaptation (13) of the method of Laemmli (12). Detection of the proteins was carried out using the silver-staining method of Heukeshoven and Dernick (14). If western analysis was to be carried out, proteins were transferred to nitrocellulose membranes (Hybond-C, Amersham International) using a Bio-Rad Trans-Blot Cell. Transfer was conducted at 30v overnight in a buffer containing 20% methanol, 192mM glycine and 25mM Tris pH 8.8.

Gels containing protein extracts which had been labelled by incubation with substrate DNA were first fixed in 40% methanol containing 10% acetic acid and then soaked in Amplify (Amersham International) for 30 mins. The gel was dried between cellophane and exposed to film (Kodak XAR-5).

Plasmids

pUC8, pUC9 and the sequencing plasmids M13mp18 and M13mp19 were obtained from Pharmacia.

Restriction endonuclease mapping and subcloning

RE were from BRL-Gibco and were used with the reaction buffers provided. Approximately 1µg of plasmid DNA (isolated by the alkaline lysis procedure (15)) and purified by Ultrogel A2 chromatography (16) was digested with the appropriate enzymes and subjected to electrophoresis in 1% agarose gels.

Southern analysis

Following denaturation and neutralisation, DNA was transferred to Hybond-N (Amersham International) using 6 x SSC (1 x SSC is 0.15M NaCl, 15mM Na citrate). Filters were prehybridised in 6 x SSC containing BLOTTO (17) for 12 hours prior to hybridisation with 2 x 10^{6} dpm per lane of [^{32}P]-labelled

nick translated probe which had been isolated from the appropriate plasmid by recovery from low melting point (LMP) agarose gels. Filters were washed in 0.1 x SSC at 65°C before exposure to Kodak XAR5 or Hyperfilm MP (Amersham International).

Western Analysis

Nitrocellulose membranes were incubated twice in 200ml of 20mM Tris-HCl, pH7.5, 0.5M NaCl (TBS) containing 1% BLOTTO for a total of 1 hour. After washing briefly in TBS containing 0.1% Nonidet P-40, the membrane was incubated for 1 hour in 20ml PBS containing the primary antibody which was rabbit anti-E.coli O^6-MeG DNA methyltransferase IgG (Applied Genetics Inc.) at a concentration of 5µg/ml. The membrane was washed twice in TBS/Nonidet P-40 before incubation with anti-rabbit IgG linked to alkaline phosphate (ICN Biochemicals). After washing briefly in TBS/Nonidet P-40 the membrane was developed with a solution of 4mg Fast Blue BBN (Sigma) in 19ml of 20mM Tris pH 8.8 mixed with 20mg naphthol AS-BI phosphate (Sigma) in 1ml 2-ethoxyethanol immediately before use. A positive reaction gave a blue colour.

Nucleotide Sequencing

The dideoxy chain termination method (18) was employed and 7-deazadeoxyguanosinetriphosphate was used in GC rich areas (19). When the maximum amount of sequence had been read with the M13 primers, oligonucleotides 17 bases long were employed as synthetic primers. Altogether 14 such oligonucleotides were synthesised. Both strands of the 2Kb Pst 1 fragment were sequenced completely and the results were analysed using HOMOL (20), LDNA (21) and a hydropathicity programme (22) on a Systime Mainframe computer.

Primer extension

RNA was extracted from bacteria harbouring p061SP1 using a phenol procedure and used in the primer extension reaction to determine the origin of transcription (23). The primer was a 132 bp Hind III-Hae III fragment isolated from the 2Kb Pst 1 insert of p061SP1 and 50 units of AMV reverse transcriptase (Amersham International) were used in the reaction.

S1 nuclease mapping

A 1304 bp Mnl 1 fragment of the 2Kb Pst 1 region of p061SP1 was cloned into Sma 1 cut pUC9. A 461 bp Hind III fragment was isolated from this and annealed to RNA (see above) after end-labelling. This was incubated with 2 or 8 units of S1 nuclease (BRL-Gibco) and the control was without nuclease (23).

RESULTS

p061 subcloning

Extracts of bacteria harbouring p061 contained high levels of an ATase that acted only on O^6-MeG in [^{3}H]-MNU methylated substrate DNA (1). A variety of RE were used to produce subclones of the original 9 Kb insert in p061 and these were ligated into pUC9 and used to transform E.coli. Extracts of bacteria harbouring plasmids containing the appropriate size inserts were assayed for ATase activity using the standard assay. The presence of only O^6AG ATase in the extracts was indicated by the amount of radioactivity transferred to protein under substrate limiting conditions (usually 700 cpm equivalent to 50 fmoles of methyl groups) and confirmed by parallel assays using extracts of bacteria harbouring plasmids containing the ada gene (p062 HX see ref 24) or a truncated version (p062SX (25)) expressing only the O^6AG ATase function of the ada gene. On no occasion did we observe activity corresponding to the dual ATase activity of the ada protein in extracts of bacteria harbouring p061 or subclones thereof.

The O^6AG ATase coding region of p061 was found to be contained within a 2Kb Pst 1 fragment and the plasmid was designated p061SP1. Extracts of bacteria harbouring p061SP1 showed only O^6AG ATase activity (Figure 1). Subclones of p061SP1 involving Hind III lost their ability to code for the ATase function. Since there is no Hind III site in the protein coding or

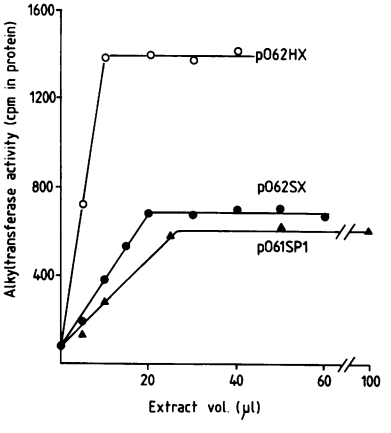

Figure 1 Alkyltransferase activity in extracts of E.coli harbouring various plasmids (O—O), p062HX; (●—●) p062SX (▲—▲), p061SP1.

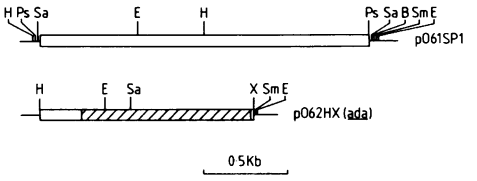

Figure 2 Restriction endonuclease sites in p061SP1 and p062HX. All sites for the enzymes are shown. R.E. used: B, Bam H1; E, Eco R1; H, Hind III; Pst 1; Sa, Sal 1; Sm, Sma 1; X, Xho II.

promoter region of ada this was already an indication that p061SP1 might contain a different ATase gene. Furthermore the Pst 1 fragment was at the 3' end (with respect to the β-galactosidase promoter in pUC8) of the p061 insert and therefore unlikely to be expressed under the influence of this promoter. Finally the distribution of sites for some common RE bore no resemblance to those of a 1.3Kb fragment of E.coli DNA spanning the ada gene (Figure 2).

These results suggested that p061SP1 contained an ATase gene different from ada. However, in the original E.coli genomic DNA library preparation, a rec A^+ host had been used (1) and it was considered possible that this procedure may have resulted in an internal rearrangement of the ada gene accounting for the above observations. If this was the case, then probing E.coli DNA with ada or the p061SP1 insert should result in an identical pattern of hybridisation signals. On the other hand, a different pattern would be expected if the genes were not the same.

Southern analysis

Under conditions of high stringency (0.1 x SSC, 65°C) the 2Kb Pst1 fragment of p061 hybridised strongly to E.coli DNA which had been digested with a variety of RE, electrophoresed in 1% agarose and transferred to a nylon membrane (Figure 3). Where this could be determined, the hybridising fragments were of the size expected from the results of subcloning the p061 insert. Thus the p061SP1 2Kb Pst I DNA hybridised to 0.6Kb and 3.0Kb Eco RI E.coli DNA fragments agreeing with the restriction enzyme map for the parent p061 clone. In contrast, the ada gene hybridised to different genomic fragments exemplified by the 3Kb Hind III and 2.3Kb Sma I signals.

In addition when p061 was digested with several RE and combinations and probed with the 1.3Kb insert of p062HX the only signals that were detected were of 2.8Kb or larger and due to the presence of vector DNA contaminating the 1.3Kb insert that was isolated from LMP agarose gels (Figure 4). These results thus confirmed that E.coli was the origin of the 9Kb insert in

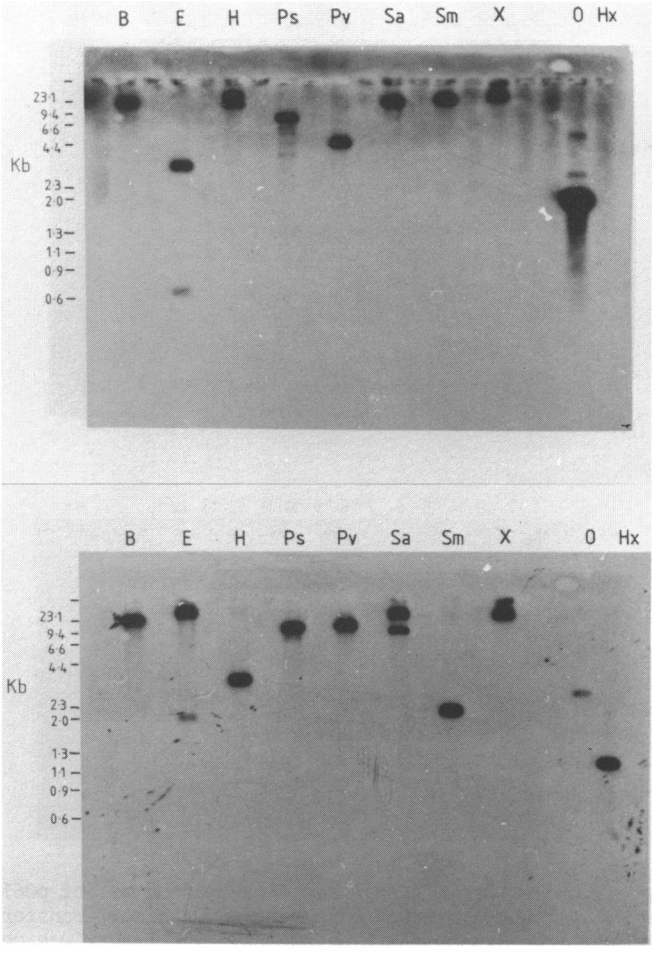

Figure 3 Autoradiographs of E.coli DNA following hybridisation with the 2Kb insert of p061SP1 1.3Kb insert of p062HX (upper) or the (lower). Pv, Pvu II., other R.E. as in Figure 2. O. and Hx control lanes containing p061 or p062HX DNA digested with Pst 1 or Hind III and Bam H1 respectively.

p061 and, together with the functional assays, demonstrate that E.coli contains at least 2 distinct ATase genes ada and ogt i.e. O^6-alkylguanine-DNA-alkyltransferase.

Western analysis

If the ogt protein is a novel ATase then polyclonal antibodies raised to the O^6AG ATase region of the ada protein might be expected not to react in Western analysis. Duplicate polyacrylamide gels were loaded with

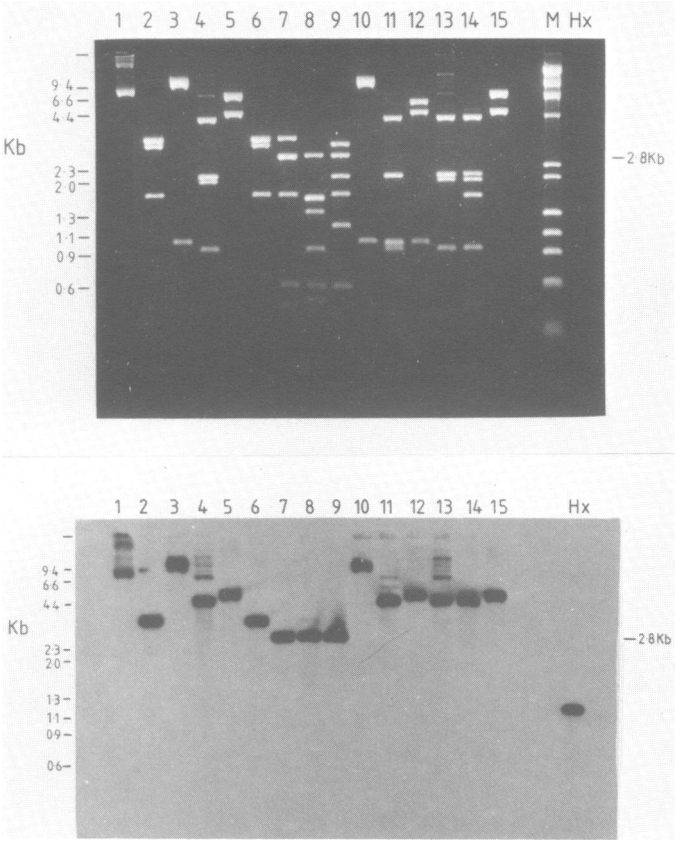

Figure 4 Upper; photograph of EBr-stained 1% agarose gel of p061 digested with various R.E. Lower; Autoradiograph of above after transfer to nylon and hybridisation with the 1.3Kb insert of p062HX. The insert was contaminated with a trace of vector DNA. R.E used. 1, Bam H1; 2 Bam H1/Eco R1; 3, Bam H1/Hind III; 4, Bam H1/Pst 1; 5, Bam H1/Sal 1; 6, Eco RI, 7, Eco R1/Hind III; 8, Eco R1/Pst 1;9, Eco R1/Sal 1; 10, Hind III; 11 Hind III/Pst 1; 12, Hind III Sal 1. 13, Pst 1; 14, Pst 1/Sal 1; 15, Sal 1. Hx, p062HX digested with Bam H1/Hind III.

approximately equivalent amounts of purified ada or ogt protein and silver stained or electroblotted onto a nitrocellulose membrane and stained with the antibody. Although the amounts of protein loaded were closely similar only a faint reaction occurred with the ogt protein in contrast to the heavy staining of the ada protein fragments that were generated by proteolysis during extraction (Figure 5). In view of the similarity of action of the two ATases, the slight cross reaction is probably due to common epitopes in the protein

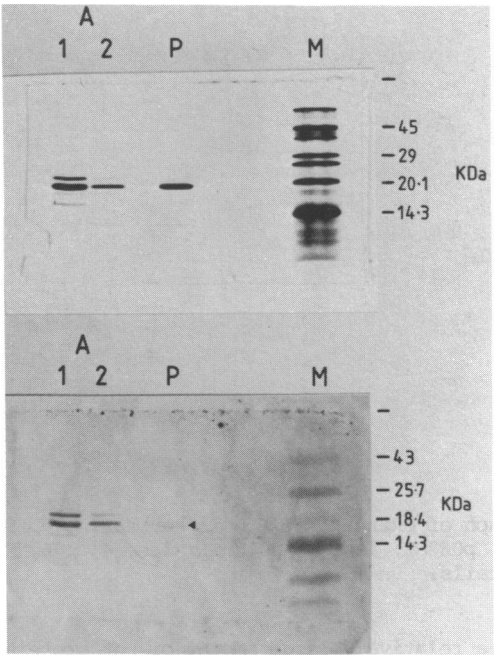

Figure 5 Upper, photograph of a silver-stained polyacrylamide gel. Lane 1A 5µg of purified dual function ada protein proteolysis cleavage fragments. Lane 2A, 4µg of the O^6AG ATase cleavage fragment of the ada protein. Lane P, 3µg of purified 061 ATase protein (Wilkinson et al, in preparation). Lower: Photograph of a western blot of the duplicate of the above gel except that different size markers were used. (lane M).

molecules and suggests that there may be a slight degree of homology at the amino acid level.

Fluorographic analysis

Extracts of E.coli harbouring p061SP1, p062SX, p062HX or p062HSR (a truncated version of p062HX that codes for AP ATase, K. Kleibl and G.P Margison, unpublished observations) were subjected to fluorographic analyses as described (see Materials and Methods and ref 1). In contrast to the full-length or truncated ada gene products the ogt ATase produced a single band on fluorography, even though in order to increase the ability to detect fragmentation products, more labelled protein had been loaded onto the gel (Figure 6). The multiple bands can be attributed to the proteolytic cleavage sites in the ada protein (1,26,27) together with the p062SX and p062HSR coded protein being fusion proteins having their translation initiation codon in the

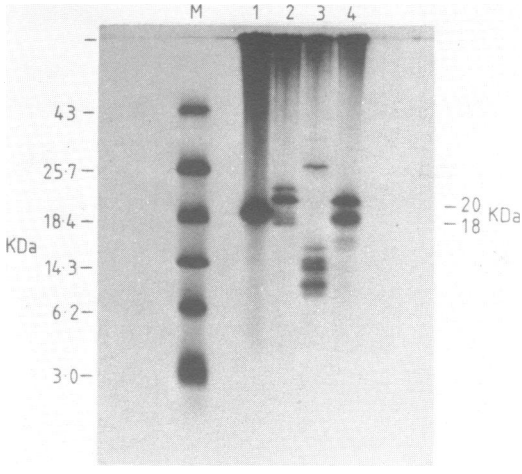

Figure 6 Fluorograph of extracts of E.coli harbouring various plasmids. Lane 1, p061SP1; lane 2, p062SX; lane 3, p062HSR; lane 4, p062HX. See Materials and Methods for details.

vector. Even when a relatively large amount of ogt protein was loaded there was little or no evidence for any fragmentation. This agrees with our previous results (1) and provides further evidence that ogt and ada ATases are different.

Nucleotide and derived amino acid sequence of p061SP1

Having accummulated overwhelming evidence that E.coli contains another ATase gene it was considered appropriate to undertake nucleotide sequence analysis. Since the 2Kb Pst 1 fragment lost its ability to code for ATase on further subcloning (data not shown) it was necessary to sequence the entire fragment.

The derived amino acid sequence of one of the ORFs found in positions 788 to 1303 of the 2Kb Pst 1 fragment had a high degree of homology with the O^6AG ATase coding region of ada (26). Figure 7 shows the sequence of this ORF and 200 bases of the 5' untranslated region. A derived amino acid heptamer and pentamer were found to be identical to sequences in the O^6AG ATase region of the ada protein. Furthermore the separation of these sequences was 34 residues in both proteins and many of the amino acid changes in this region were of a conservative nature. An homology plot of the amino acid sequences of the ada and ogt proteins indicate this area of high homology (Figure 8) which gives rise to a closely overlapping hydropathicity plot (Figure 9).

A decanucleotide 4 bases upstream of the putative initiation codon

```
      610         620         630         640         650         660         670         680         690         700
       "           "           "           "           "           "           "           "           "           "
TTCCACTGTTTCTTGGATTCCTGCAACGCTACAAACCAGACGCGAAACTGGGTACTTACTATTCGTTAGTCTTGCCCTATCCGCTTATCTTTTTGGTGG

      710         720         730         740         750         760         770         780         790
       "           "           "           "           "           "           "           "           "
TATGGCTGCTGATGTTGCTGGCGTGGTATCTTGTTGGTCTGCCGATAGGTCCGGGGATTTCCCCACGTTTGTCTTAAGAGAGAACGG ATG CTG AGA
 -35                     -10          RNA START                                S.D.          Met Leu Arg
                                                                                             *   *   *

  800         810         820         830         840         850         860         870
   "           "           "           "           "           "           "           "
TTA CTT GAA GAA AAA ATT GCC ACG CCA CTG GGT CCA CTG TGG GTG ATT TGC GAT GAA CAA TTT CGT CTG CGG GCG
Leu Leu Glu Glu Lys Ile Ala Thr Pro Leu Gly Pro Leu Trp Val Ile Cys Asp Glu Gln Phe Arg Leu Arg Ala
 *   *   *   *   *

       880         890         900         910         920         930         940
        "           "           "           "           "           "           "
GTT GAA TGG GAA GAG TAC AGC GAA GCC ATG GTG CAG CTG CTG GAC ATC CAT TAT CGC AAA GAA GGC TAT GAG CGC
Val Glu Trp Glu Glu Tyr Ser Glu Ala Met Val Gln Leu Leu Asp Ile His Tyr Arg Lys Glu Gly Tyr Glu Arg

  950         960         970         980         990         1000        1010        1020
   "           "           "           "           "           "           "           "
ATT TCT GCC ACC AAC CCA GGC GGT TTA AGC GAC AAG CTT CGT GAT TAT TTT GCC GGT AAT CTT AGC ATT ATT GAT
Ile Ser Ala Thr Asn Pro Gly Gly Leu Ser Asp Lys Leu Arg Asp Tyr Phe Ala Gly Asn Leu Ser Ile Ile Asp

       1030        1040        1050        1060        1070        1080        1090
        "           "           "           "           "           "           "
ACG CTT CCC ACT GCC ACA GGG GGG ACG CCA TTT CAG CGC GAA GTC TGG AAA ACA CTA CGC ACT ATC CCC TGC GGG
Thr Leu Pro Thr Ala Thr Gly Gly Thr Pro Phe Gln Arg Glu Val Trp Lys Thr Leu Arg Thr Ile Pro Cys Gly

 1100        1110        1120        1130        1140        1150        1160        1170
  "           "           "           "           "           "           "           "
CAG GTA ATG CAT TAC GGC CAA CTG GCT GAA CAA TTG GGC CGC CCT GGC GCG GCG CGT GCC GTT GGT GCG GCA AAC
Gln Val Met His Tyr Gly Gln Leu Ala Glu Gln Leu Gly Arg Pro Gly Ala Ala Arg Ala Val Gly Ala Ala Asn

       1180        1190        1200        1210        1220        1230        1240
        "           "           "           "           "           "           "
GGA TCG AAT CCC ATC AGC ATC GTC GTA CCT TGC CAT CGG GTT ATT GGC CGA AAC GGC ACC ATG ACC GGA TAT GCA
Gly Ser Asn Pro Ile Ser Ile Val Val Pro Cys His Arg Val Ile Gly Arg Asn Gly Thr Met Thr Gly Tyr Ala

 1250        1260        1270        1280        1290        1300        1310        1320
  "           "           "           "           "           "           "           "
GGC GGA GTT CAG CGA AAA GAG TGG TTA TTG CGC CAT GAA GGT TAT CTT TTG CTG TAA ACATTAAACAATTTGTG
Gly Gly Val Gln Arg Lys Glu Trp Leu Leu Arg His Glu Gly Tyr Leu Leu Leu TER
```

Figure 7 Nucleotide and derived amino acid sequence of a section of the 2Kb Pst 1 fragment of p061SP1 including 200 bases of 5'untranslated region. Amino acid residues underlined were homologous to the O^6AG ATase coding C-terminal region of the ada gene product. In addition, those with asterisk were determined by amino acid sequencing from the amino terminus. The origin of transcription is indicated by RNA START. The size of the predicted 171 residue protein is 19143 Da.

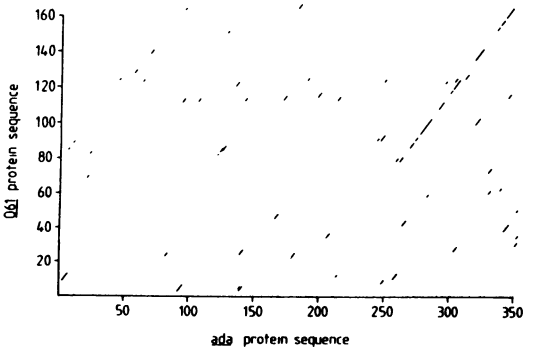

Figure 8 Graphic representation of amino acid and homology between ogt and ada ATases. A point is plotted when 3 homologous amino acids are found within any pentapeptide group.

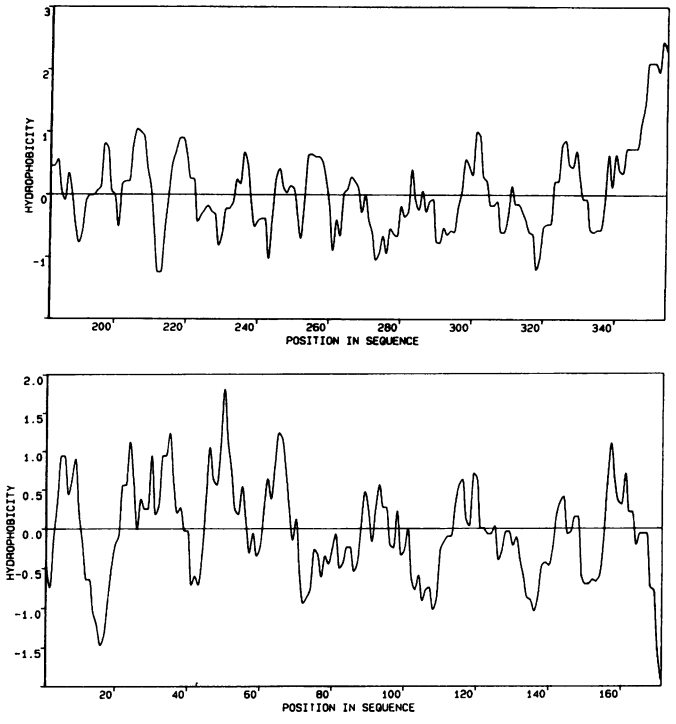

Figure 9 Hydrophobicity plots of (upper) the O^6AG ATase region of the ada protein and (lower) the ogt protein.

had a 50% homology with the consensus Shine-Dalgarno ribosome binding site (28). The origin of transcription was located by S_1 nuclease (Figure 10) and primer extension (Figure 11) reactions to be 48 bases upstream of the translation initiation codon. The putative -10 and -35 promoter binding domains had, respectively 80% and 50% homology to the consensus sequences (29).

Amino terminal amino acid sequence

To confirm that the initiation codon in the putative ORF was correct and to demonstrate that the ATase present in extracts of bacteria harbouring p061SP1 was encoded by the plasmid, the amino terminal acid sequence of the purified protein was determined. The purification procedure will be presented elsewhere (Wilkinson et al, in preparation). Residues 3 to 8 were identical to those predicted from the nucleotide sequence and although residues 1 and 2 were ambiguous, methionine and leucine were amongst the possible candidates.

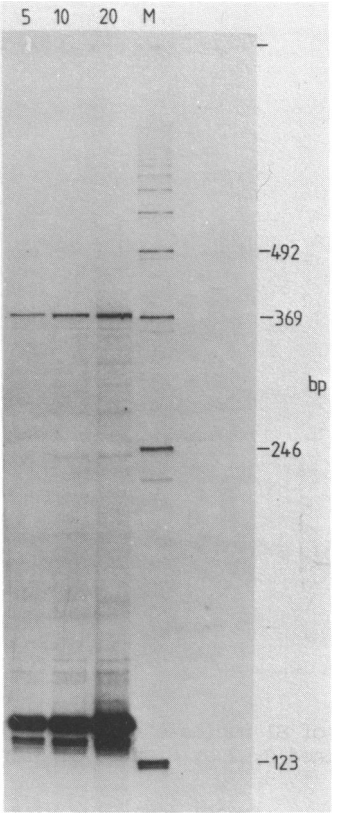

Figure 10 Autoradiograph of primer extension reaction polyacrylamide gel. The primer extension reaction mix (5, 10 or 20µl as indicated) was loaded onto the gel. M, end-labelled BRL Gibco 123 bp ladder.

DISCUSSION

The ability to generate the p062SX subclones of the *ada* gene that coded only for functional O^6AG ATase activity suggested that p061 might contain a truncated *ada* gene. Because the O^6AG coding region was at the 3' end of *ada* and hence without its own promoter it was assumed that the ATase coding region in p061 would be at the 5' end of the 9Kb insert and be expressed from the *lac* Z promoter giving rise to a fusion protein analogous to that coded by p062SX (25). However, subcloning of the insert in p061 showed that the ATase coding region was at the 3' end (with respect to the *lac* Z promoter) of the insert so this explanation seemed improbable. Furthermore the *Sau* 3a I sites in the *ada* gene that could account for truncation would give only a *ca*. 8KDa

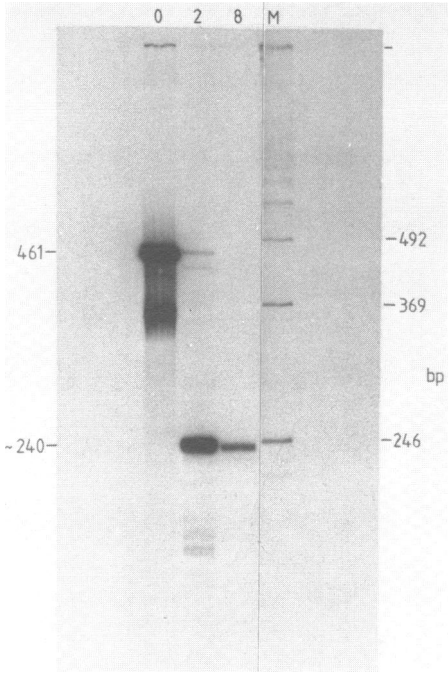

Figure 11 Autoradiograph of S1 nuclease reaction polyacrylamide gel. The incubation mixture contained 0, 2 or 8 units of S1 nuclease. M, see Fig 10.

protein. Also, no subclones of p061 employing Hind III gave rise to ATase activity in the host. These observations also imply the non-ada origin of the ATase gene in p061. The RE map supported these suggestions and the Southern and Western analysis provided final confirmation that ogt is a novel E.coli ATase gene. The relative resistance of the ogt protein to the proteolytic cleavage undergone by the ATases encoded by the ada gene or its subclones indicates the absence of similar cleavage sites in the former protein.

This systematic approach to exploring the origin of the ATase encoded by p061 provided conclusive evidence for the existence of another ATase gene in E.coli, the product of which acts on O^6AG and not AP in substrate DNA. Although an exhaustive search was not made, the inability to isolate subclones of p061SP1 that conferred ATase activity on the host necessitated sequencing the entire 2Kb fragment.

Computer analysis of the data revealed an ORF with a high degree of homology to the O^6AG ATase region of the ada protein (Figures 7-9). In view of the functional similarity of the two ATases this might have been expected.

One of the peptides common to both proteins is a pentamer which in the case of the ada protein includes the cysteine residue that becomes alkylated during the repair reaction. It seems likely that this will prove to have a similar function in the ogt protein and this is being examined. The other common peptide which also includes a cysteine residue may be involved in protein tertiary structure or in DNA binding and since the penta- and heptapeptides are separated by an identical number of amino acid residues, it is not unreasonable to suggest that this separation may be critical for functional activity.

The amino terminal amino acid sequence of the ogt ATase was as predicted from the nucleotide sequence and this confirms that ATase activity in extracts of bacteria harbouring p061SP1 (or p061) is not due to a plasmid coded protein that upregulates an endogenous O^6AG ATase but to the ATase encoded in the 2Kb insert. Further subcloning using the Eco R1 and Mnl 1 sites also confirms the ORF (data not shown).

Exposure of E.coli to low doses of MNNG triggers the adaptive response in which the expression of the ada gene is considerably increased resulting in the increased capacity to repair the promutagenic lesions O^6AG, O^4AT and the S stereoisomer of AP (26,30-32). Expression of alk A which codes for 3-methyladenine-DNA-glycosylase II is also upregulated and this increases the capacity to repair 3-meA, O^2-meC, 3-meG, 7-meG and 7-meA (3,4). The mechanism of upregulation of ada and alk A involves the binding of the alkylated form of the AP ATase of the ada protein to a specific control sequence, the ada box, in the promoter regions of the two genes (27). The 5' untranslated region of ogt contains sequences of reasonable homology to the Shine-Dalgarno, -10 and -35 consensus sequences but there does not appear to be a region that could be considered as an ada box.

E.coli is also known to contain another 3-meA-DNA-glycosylase gene tag I (4) the expression of which is constitutive and not affected by adaptation. In the present paper we have demonstrated the existence of another ATase gene in E.coli and this may be equivalent to the non-inducible tag I gene in evolutionary terms. By analogy it may be appropriate to call the single function protein ATase 1. Fluorographic analysis of ATase proteins in B. subtilis support the suggestions of inducible and non-inducible forms of ATase (33). Furthermore, O^6AG ATase activity is present in certain E.coli ada mutants (34) Margison, G.P. unpublished results) and the possibility that this is due the expression of ogt is being examined. It may be that the expression of the non inducible ATase is sufficient to deal with very low levels of

mutagenic alkylation damage generated in the DNA of bacteria by endogenous or exogenous factors.

Circumstantial evidence for the lack of induction of ogt expression during the adaptive response is that there are equal amounts of O^6AG and AP ATase activity in extracts of adapted E.coli (Margison, G.P. unpublished results). If ogt was upregulated, the specific activity of the total O^6AG ATase would exceed that of the AP ATase. However, the question of whether or not ogt is inducible by adaptation or under other circumstances is difficult to resolve because of the absence of a specific assay method. For this reason we are attempting to raise antibodies to the ogt gene product and to examine the substrate specificity and other biochemical characteristics of the ogt ATase.

ACKNOWLEDGEMENTS

The technical assistance of Brian C. Keenan, Sean Baker and Gillian McMaster is gratefully acknowledged. Our thanks to Linda Evans for preparing the MS. The work was supported by grants from the Cancer Research Campaign.

*Present address: Corporate Biosciences Laboratory, ICI plc, The Heath, Runcorn, Cheshire, UK

REFERENCES

1. Margison, G.P., Cooper, D.P. and Brennand, J. (1985). Nucleic Acids Res. 13 1939-1952.
2. Margison, G.P. and O'Connor, P.J. (1979). In Chemical Carcinogens and DNA, vol. 1 (ed. P.L. Grover). CRC Press, Boca, Raton, Florida pp. 111-159.
3. Saffhill, R., Margison, G.P. and O'Connor, P.J. (1985). Biochim. Biophys. Acta 823, 111-145.
4. Lindahl, T. (1982), Ann. Rev. Biochem. 51, 61-87.
5. Lawley, P.D., and D.J. Orr (1970). Chem.-Biol. Interact., 2, 154-157.
6. O'Connor, P.J., Capps, M.J. and A.W. Craig (1973). Br. J. Cancer, 27, 153-166.
7. Olsson, M. and Lindahl, T. (1980). J.Biol. Chem., 255, 1056 - 1071.
8. Mehta, J.R., Ludlum, D.B., Renard, A. and Verly, W.G. (1981). Proc. Natl. Acad. Sci. USA, 78, 6766-6770.
9. Sedgwick, B. (1982). J. Bacteriol. 105, 984-988.
10. Samson, L., and J. Cairns (1977). Nature (London), 267, 281-282.
11. Potter, P.M., Brennand, J. and Margison, G.P. (1986) Br. J. Cancer. 54, 366-367.
12. Laemmli, U.K. (1970). Nature (London), 227, 680-685.
13. Bury, A.F. (1981). J. Chromatog. 213, 491-500.
14. Heukeshoven, J. and Dernick, R. (1985). Electrophoresis, 6, 103-112.
15. Birnboim, H.C. and Doly, J. (1979). Nucleic Acids Res. 7, 1513-1525.
16. Miscard, D., Sobrier, M.L., Conderc, J.L. and Dastugne, B. (1985). Anal. Biochem. 148, 121-126.
17. Johnson, D.A., Gautsch, J.W., Sportsman, J.R. and Elder, J. (1984). Gene Anal. Techn. 1, 3-8.

18. Sanger, F., Nicklen, S. and Coulsen, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.
19. Mitsuzawa, S., Nishimura, S. and Seela, F. (1986). Nucleic Acids Res. 14, 1319-1324.
20. Taylor, P. (1985). Nucleic Acids Res. 13, 8764-8785.
21. Stockwell, P.A. (1982). Nucleic Acids Res. 10, 115-125.
22. Hopp, T.P. and Woods, K.R. (1981). Proc. Natl. Acad. Sci. USA, 78, 3824-3828.
23. Serizawa, H., and Fukuda, R. (1987). Nucleic Acids Res. 1153-1163.
24. Brennand, J. and Margison, G.P. (1986). Carcinogenesis 7, 185-188.
25. Brennand, J. and Margison, G.P. (1986). Carcinogenesis 7, 2081-2084.
26. Demple, B., Sedgewick, B., Robins, P., Totty. N., Waterfield, M.D. and Lindahl. T. (1985). Proc. Natl. Acad. Sci. USA, 82, 2688-2692.
27. Teo, I., Sedgwick, B. Kilpatrick, M.W., McCarthy, T.V. and Lindahl, T. (1986). Cell, 45, 315-324.
28. Shine, J. and Dalgarno, L. (1975). Nature (London) 254, 34-38.
29. Harley, L.B. and Reynolds, R.P. (1987). Nucleic Acids Res. 15, 2343-2361.
30. McCarthy, T.V., Karran, P. and Lindahl, T. (1984). EMBO J. 3, 545-550.
31. McCarthy, T.V. and Lindahl, T. (1985). Nucleic Acids Res. 13, 2683-2698.
32. Weinfeld, M., Drake, A.R., Saunders, J.K. and Paterson, M.C. (1985). Nucleic Acids Res. 13, 7067-7077.
33. Morohoshi, F. and Munakata, N. (1987). J. Bacteriol. 169, 587-592.
34. Demple, B. (1986). Nucleic Acids Res. 14, 5575-5589.