JOURNAL OF BACTERIOLOGY, Jan. 1990, p. 164–171
0021-9193/90/010164-08$02.00/0

Vol. 172, No. 1

# Sequence Analysis and Expression of the Bacterial Dichloromethane Dehalogenase Structural Gene, a Member of the Glutathione *S*-Transferase Supergene Family

SALOME D. LA ROCHE AND THOMAS LEISINGER*

*Department of Microbiology, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich, Switzerland*

Received 23 June 1989/Accepted 24 September 1989

**The nucleotide sequence of a cloned 2.8-kilobase-pair *Bam*HI-*Pst*I fragment containing *dcmA*, the dichloromethane dehalogenase structural gene from *Methylobacterium* sp. strain DM4, was determined. An open reading frame with a coding capacity of 287 amino acids (molecular weight, 37,430) was identified as *dcmA* by its agreement with the N-terminal amino acid sequence, the total amino acid composition, and the subunit size of the purified enzyme. Alignment of the deduced dichloromethane dehalogenase amino acid sequence with amino acid sequences of the functionally related eucaryotic glutathione *S*-transferases revealed three regions containing highly conserved amino acid residues and indicated that *dcmA* is a member of the glutathione *S*-transferase supergene family. The 5′ terminus of in vivo *dcmA* transcripts was determined by nuclease S1 mapping to be 82 base pairs upstream of the GTG initiation codon of *dcmA*. Despite a putative promoter sequence with high resemblance to the *Escherichia coli* −10 and −35 consensus sequences, located at an appropriate distance from the transcription start point, *dcmA* was only marginally expressed in *E. coli*. The strong induction of dichloromethane dehalogenase in *Methylobacterium* sp. by dichloromethane was abolished by deleting the 1.3-kilobase-pair upstream region of *dcmA*. Plasmid constructs devoid of this region directed expression of dichloromethane dehalogenase at a constitutively induced level.**

*Methylobacterium* sp. strain DM4 is a pink-pigmented facultative methylotrophic bacterium utilizing dichloromethane (DCM) as the sole source of carbon and energy (16, 27). The first step in the degradation of this xenobiotic compound is catalyzed by DCM dehalogenase. This glutathione (GSH)-dependent enzyme dechlorinates DCM to formaldehyde and inorganic chloride (31). DCM dehalogenase has a hexameric structure with a subunit molecular weight of approximately 33,000. Its catalytic activity is low, but the enzyme is strongly inducible by DCM. Upon growth on DCM, 15 to 20% of the total soluble protein consist of DCM dehalogenase. The purified enzymes of four different facultative methylotrophic bacteria, including *Methylobacterium* sp. strain DM4, have been shown to be catalytically and immunologically similar and to have identical N-terminal amino acid sequences and subunit molecular weights (16). These observations suggested that the genes for DCM utilization were horizontally distributed among facultative methylotrophic bacteria. The DCM utilization genes of *Methylobacterium* sp. strain DM4 were isolated by complementation of the DCM-nonutilizing ($DCM^-$) mutant strain DM4-2cr with a clone of a cosmid library of DM4 wild-type DNA in the broad-host-range vector pVK100 (10). It was also demonstrated that the genes for DCM utilization are located on the chromosome or on a megaplasmid and that none of the three cryptic plasmids carried by strain DM4 is involved in DCM metabolism (10).

The evolutionary origin of DCM dehalogenase and its potential relatedness to known bacterial enzymes are a matter of speculation. To explore these questions and to provide a basis for studies on the expression of the enzyme, we determined the nucleotide sequence of the *dcmA* gene, the structural gene of DCM dehalogenase, and its flanking regions. We found that this enzyme had regions of amino acid sequences that were conserved in eucaryotic GSH *S*-transferases and that it exhibited the same percentage of amino acid sequence similarity to eucaryotic GSH *S*-transferases as the eucaryotic GSH *S*-transferases among themselves.

* Corresponding author.

## MATERIALS AND METHODS

**Bacterial strains, plasmids, and growth conditions.** The bacteria and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown at 37°C in nutrient yeast broth (NYB) or on nutrient agar (30) containing the appropriate antibiotics at the following concentrations: ampicillin, 100 to 200 μg/ml; kanamycin, 25 μg/ml; tetracycline, 25 μg/ml. *Methylobacterium* strains were grown at 30°C on liquid minimal medium containing DCM or methanol as a carbon source as previously described (16). Solid media were methylamine agar (10) and DCM indicator agar (10) to test for DCM utilization.

**Construction of plasmids.** Plasmids constructed in this work are shown in Fig. 1 and 5 or listed in Table 1. Plasmids pME1526, pME1527, pME1528, pME1530, pME1547, pME 1548, and pME1549 are pUC18 derivatives containing various restriction fragments of the 4.2-kilobase (kb) *Bam*HI insert of pME1518. Plasmids pME1529 and pME1531 were constructed in pUC19 by insertion of restriction fragments from pME1518 (Fig. 1B).

Plasmids pME1532 (Table 1), pME1541, pME1542, pME 1543, pME1544, and pME1545 (see Fig. 5) were constructed by cloning *Bgl*II-*Hin*dIII or *Bam*HI-*Hin*dIII fragments into pVK100, which contains no multiple cloning site (15). To create the fragments for cloning, the original restriction fragments of pME1518 were first cloned into pUC18, such that its multiple cloning site provided *Bgl*II-*Hin*dIII or *Bam*HI-*Hin*dIII fragments suitable for cloning into pVK100.

Plasmid pVK100-E (Table 1) represents a pVK100 derivative whose single *Eco*RI site was removed by *Eco*RI

TABLE 1. Bacterial strains, phages, and plasmids used in this work

| Strain, phage, or plasmid | Relevant genotype or phenotype[a] | Source or reference[b] |
|---|---|---|
| *E. coli* | | |
| JM105 | *thi rpsL endA sbcB15 hsdR4* Δ(*lac-proAB*)(F′*traD36 proA*$^+$*B*$^+$ *lacI*$^q$ *lacZ*Δ*M15*) | 37 |
| RR28 | *pheS12* F$^-$ *thi leu pro lac gal ara mtl xyl supE44 endA* r m *thyA*$^+$, *recA* | 13 |
| S17-1 | *thi pro hsdR hsdM*$^+$ *recA*, chromosomally integrated RP4-2 (Tc::Mu, Km::Tn*7*) | 28 |
| *Methylobacterium* strains | | |
| DM4 | DCM$^+$ | 9 |
| DM4-2cr | DCM$^-$ Sm$^r$, derived from DM4 | 10 |
| DM4-2cr(pME1510) | DCM$^+$ Sm$^r$ Tc$^r$ | 10 |
| DM4-2cr(pME1523) | DCM$^+$ Sm$^r$ Tc$^r$ Km$^r$ | |
| Phage | | |
| M13mp18, M13mp19 | | Amersham |
| Plasmids | | |
| pUC18, pUC19 | Amp$^r$, pBR322-derived expression vector | 36 |
| pBLS, bluescript KS (+) | Amp$^r$, pBR322-derived expression vector containing an intergenic region of phage M13 | Stratagene, San Diego, Calif. |
| pVK100 | Km$^r$ Tc$^r$, mobilizable cloning vector derived from RK2 | 15 |
| pVK100-E | pVK100, deleted in its single *Eco*RI site | |
| pME1510 | Tc$^r$ DCM$^+$ recombinant plasmid of pVK100 containing a 21-kb *Hin*dIII fragment from DM4 | 10 |
| pME1518 | DCM$^+$, 4.2-kb *Bam*HI fragment of pME1510 in pUC18 | |
| pME1523 | DCM$^+$, 4.2-kb *Bam*HI fragment of pME1510 in pVK100 | |
| pME1532 | DCM$^+$, 2.2-kb *Bgl*II-*Pst*I fragment of pME1518 in pVK100 | |
| pME1533 | DCM$^-$, 1.7-kb *Bgl*II-*Eco*RI fragment of pME1518 in pVK100-E | |

[a] DCM$^+$, DCM utilizing; DCM$^-$, DCM nonutilizing.
[b] From this work if not otherwise indicated.

digestion, filling up of sticky ends with Klenow polymerase, and religation. It was used as a vector in the construction of pME1533. This plasmid was created by cloning the 2.2-kb *Bgl*II-*Pst*I fragment of pME1518 (as a polylinker-adapted *Bgl*II-*Hin*dIII fragment; see above) into pVK100-E, digestion of the construct with *Eco*RI and *Hin*dIII, filling up of the ends with Klenow polymerase, and religation.

**Mobilization of plasmids.** For expression studies, recombinant plasmids were mobilized from *E. coli* into *Methylobacterium* sp. strain DM4-2cr. Colonies of *E. coli* S17-1 containing the plasmid to be mobilized were replicated onto a lawn of strain DM4-2cr on nutrient agar (Difco Laboratories, Detroit, Mich.). After 2 days of coculture at 30°C, selection for transconjugants was carried out by replication onto methylamine agar containing the appropriate antibiotic. Methylamine as the sole carbon source provided contraselection against the *E. coli* donor strain.

**Assay for polypeptides with DCM dehalogenase antigenic determinants.** The *E. coli* strains to be analyzed for the production of material cross-reacting with DCM dehalogenase antiserum were grown on nutrient agar, and single colonies were subjected to the previously described assay procedure (10). The same procedure was also used to detect cross-reacting material in crude extracts blotted on nitrocellulose membranes (BA 85/20; Schleicher and Schüell).

**Preparation of crude extract.** *Methylobacterium* sp. was grown in 1-liter fed-batch cultures contained in 5-liter Erlenmeyer flasks closed with rubber stoppers. A total of 32 mmol of DCM was added in four portions of 8 mmol each, after the pH had been adjusted with 1 M sodium hydroxide to 7.0. Alternatively, 30 mmol of methanol, added in three portions of 10 mmol, served as the carbon source. *E. coli* cells were grown overnight in 1 liter of NYB containing ampicillin and 5 mM isopropyl-β-D-thiogalactopyranoside. After centrifugation (4,000 × *g* at 4°C for 50 min), cells were suspended at 30% (wt/vol) in 50 mM Tris sulfate (pH 7.5)–2 mM dithiothreitol–0.5 mM EDTA and passed twice through a French pressure cell at a pressure of 35 MPa. Cell debris was removed by centrifugation (16,000 × *g* at 4°C for 30 min). Enzyme assays with *Methylobacterium* extracts were carried out the same day, since freezing the crude extract resulted in a loss of activity.

**Enzyme assay.** DCM dehalogenase activity was determined by assaying the rate of formaldehyde production. Formaldehyde was detected colorimetrically after reaction with acetylacetone to 2,6-dimethyl-3,5-diacetyl-1,4-dihydropyridine (31). The assay mixture (2 ml) contained 120 to 200 μl of crude extract (10 to 15 μg of total protein, determined by the method of Bradford [4]) in 0.1 mM Tris sulfate (pH 8.2)–8 mM reduced GSH–8 mM DCM. Enzyme activity is expressed in millikatals (mkat); 1 mkat corresponds to the amount of activity catalyzing the conversion of 1 mmol of substrate per s.

**DNA sequence analysis.** The DCM dehalogenase structural gene and its upstream region were sequenced by the dideoxy-chain termination method (26) with the M13mp18 and M13mp19 system of Amersham (Amersham, United Kingdom) with Sequenase (U.S. Biochemical Corp., Cleveland, Ohio) by following the protocol of the manufacturer. Two sets of progressive deletions of the 2.2-kb *Bgl*II-*Pst*I fragment (Fig. 1B and 2) were constructed with exonuclease III and mung bean nuclease in the vector pBLS as recommended by Stratagene (San Diego, Calif.) and subcloned into M13mp18 or M13mp19. Appropriate restriction fragments were also cloned into these M13 derivatives. The nucleotide sequence was analyzed on the VAX by using the Genetics Computer Group program package (University of Wisconsin, Madison) (6) and some programs of the PC/Gene pack-

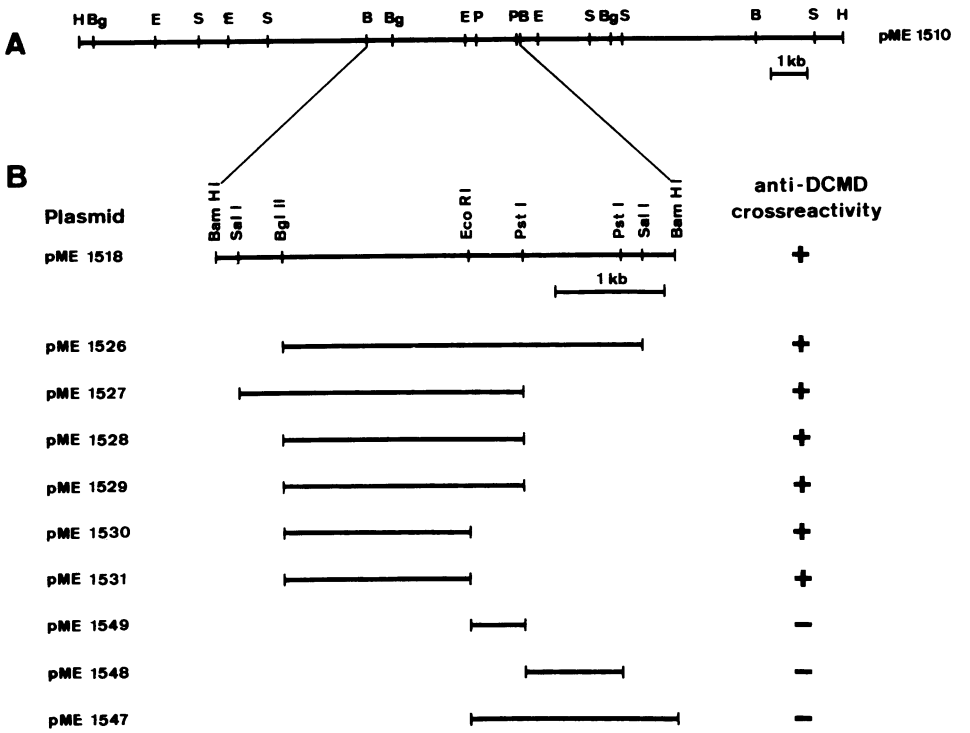


FIG. 1. Subcloning of the DCM utilization genes of *Methylobacterium* sp. strain DM4. (A) Restriction map of pME1510, the recombinant plasmid that complemented the DCM-nonutilizing mutant strain DM4-2cr (10). B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. (B) Subcloned restriction fragments of the 4.2-kb *Bam*HI fragment of plasmid pME1518 into the expression vector pUC18 or pUC19. The cross-reactivities of the polypeptides produced by the subclones with DCM dehalogenase antiserum are indicated.

age (Genofit S.A., Geneva, Switzerland). The amino acid sequence alignment was carried out with the programs GAP and LINE UP (Genetics Computer Group, University of Wisconsin, Madison), and the similarity was calculated by the GAP program of Gribskov and Burgess (11).

**Isolation of cellular RNA.** *Methylobacterium* cells were grown on 50 mM methanol or on 24 mM DCM (3 times 8 mM, with pH correction) in 30-ml cultures with gas-tight screw caps (Precision Sampling, Baton Rouge, La.) to an $A_{600}$ of 0.8. Cell lysis and isolation of RNA were performed as described by Machlin and Hanson (20), and the RNA was stored at −20°C as an isopropanol precipitate (0.1 volume 3 M sodium acetate, 1 volume isopropanol).

**Nuclease S1 mapping.** Samples of 20 μg of cellular RNA were mixed with $2 \times 10^5$ cpm of 5′-$^{32}$P-end-labeled DNA (approximately 0.1 μg) and incubated at 85°C for 15 min in 30 μl of hybridization solution (8). The mixture was transferred to 57°C, and hybridization was continued by incubation overnight. Hybridization was stopped by dilution with 300 μl of ice-cold S1 digestion buffer (21), and the mixture was chilled on ice. Nuclease S1 (100 U) was added, and digestions were performed at 37°C for 30 min. The reaction was terminated by adding 75 μl of termination solution (8), and the nucleic acids were precipitated with isopropanol after the addition of 50 μg of carrier yeast RNA. After an additional isopropanol precipitation, the length of the nuclease S1-resistant hybrids was analyzed on a denaturing sequencing gel with a Maxam-Gilbert (A+G reaction) sequencing ladder (22) of the 5′-end-labeled DNA fragment used in the experiment.

## RESULTS

**Confining the DCM dehalogenase structural gene *dcmA*.** The expression in *E. coli* of DCM dehalogenase from low-copy-number recombinant plasmids carrying the *Methylobacterium* sp. strain DM4 DCM utilization genes has previously been found to lie below the detection limit (10). However, the multicopy nature of the pUC18 and pUC19 vectors used in the present study ensured the formation of sufficient DCM dehalogenase for detection of the enzyme by the antibody assay. It was thus possible to use the assay for polypeptides with DCM dehalogenase antigenic determinants in assigning the *dcmA* gene to a particular fragment of the 4.2-kb insert contained on plasmid pME1518. The results of such subcloning experiments in *E. coli* are shown in Fig. 1. They revealed that the 1.7-kb *Bgl*II-*Eco*RI insert of pME1530 and of pME1531 is the smallest fragment leading to the production of material cross-reacting with DCM dehalogenase antiserum. This analysis thus placed the entire *dcmA* gene or the major part of it on this fragment. Since the *Bgl*II-*Eco*RI fragment cloned in either direction (pME1530 and pME1531) conferred production of cross-reacting material to the *E. coli* host, transcription seems to be initiated from a promoter on the insert. To determine the direction of transcription of *dcmA*, crude extracts of isopropyl-β-D-thiogalactopyranoside-induced *E. coli* strains containing the pUC derivatives pME1528, pME1529, pME1530, and pME1531 were prepared, and their cross-reactivity with DCM dehalogenase antiserum was tested. Only the *E. coli* strains containing pME1528 or pME1531, in which transcription by the *lac* promoter proceeded from *Bgl*II toward *Eco*RI and *Pst*I (Fig. 1B), produced large amounts of cross-reacting material. This indicated that *dcmA* is transcribed in the direction from left to right in Fig. 1.

To verify the assignment of the *dcmA* gene, a set of progressively smaller fragments derived from the original cosmid pME1510 were subcloned in pVK100, mobilized into the DCM$^-$ *Methylobacterium* sp. strain DM4-2cr, and as-

TABLE 2. Specific enzyme activity of DCM dehalogenase, measured in crude extracts of different *Methylobacterium* strains

| *Methylobacterium* strain | Size of cloned fragment (kb) | DCM dehalogenase sp act (mkat/kg of protein) | |
|---|---|---|---|
| | | Induced[a] | Noninduced[b] |
| DM4 | | 4.0 | 0.05 |
| DM4-2cr(pME1510) | 21 | 9.4 | 0.20 |
| DM4-2cr(pME1523) | 4.2 | 9.0 | 0.15 |
| DM4-2cr(pME1532) | 2.2 | 8.2 | 1.90 |
| DM4-2cr(pME1533) | 1.7 | ND[c] | <0.005 |
| DM4-2cr | | ND | <0.005 |

[a] Cells grown on DCM.
[b] Cells grown on methanol.
[c] ND, Not determined.

sayed for the DCM dehalogenase activity conferred on the host. The 2.2-kb *Bgl*II-*Pst*I fragment of plasmid pME1532, but not the 1.7-kb *Bgl*II-*Eco*RI fragment of pME1533, restored dehalogenase activity (Table 2). The fact that the *Bgl*II-*Eco*RI fragment encoded an antigenically active but catalytically inactive *dcmA* polypeptide (Fig. 1B) and the direction of transcription of *dcmA* suggested that the C-terminal region of DCM dehalogenase covers the *Eco*RI site at the insert-vector junction of pME1533. Upon introduction of the cloned DCM utilization genes on the vector pVK100 into strain DM4-2cr, a twofold increase in DCM dehalogenase specific activity over the wild-type level was observed (Table 2). Due to this gene dosage effect, 30 to 40% of the total soluble protein of the transconjugants consisted of DCM dehalogenase.

**Nucleotide sequence of *dcmA* and flanking regions.** The nucleotide sequence of the 2.8-kb *Bam*HI-*Pst*I fragment (Fig. 1B) with the deduced amino acid sequence of DCM dehalogenase is shown in Fig. 2, and the sequencing strategy is summarized in Fig. 3. The open reading frame extending from nucleotides 1508 to 2374 encodes a protein of 287 amino acids. This coding sequence was assigned to *dcmA*, since the deduced amino acid sequence of its 5′ end corresponded to the N-terminal amino acid sequence of the purified DCM dehalogenase (16). The molecular weight of the protein encoded by *dcmA* amounted to 37,430, which compared well with the molecular weight of 33,000 determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (27), and the previously determined amino acid composition of DCM dehalogenase (27) was identical to the composition deduced from the nucleic acid sequence. Translation of *dcmA* started at the GTG codon at position 1508. Screening of the amino acid sequence upstream of position 1508 revealed no evidence for a signal peptide, and a hydropathy plot (not

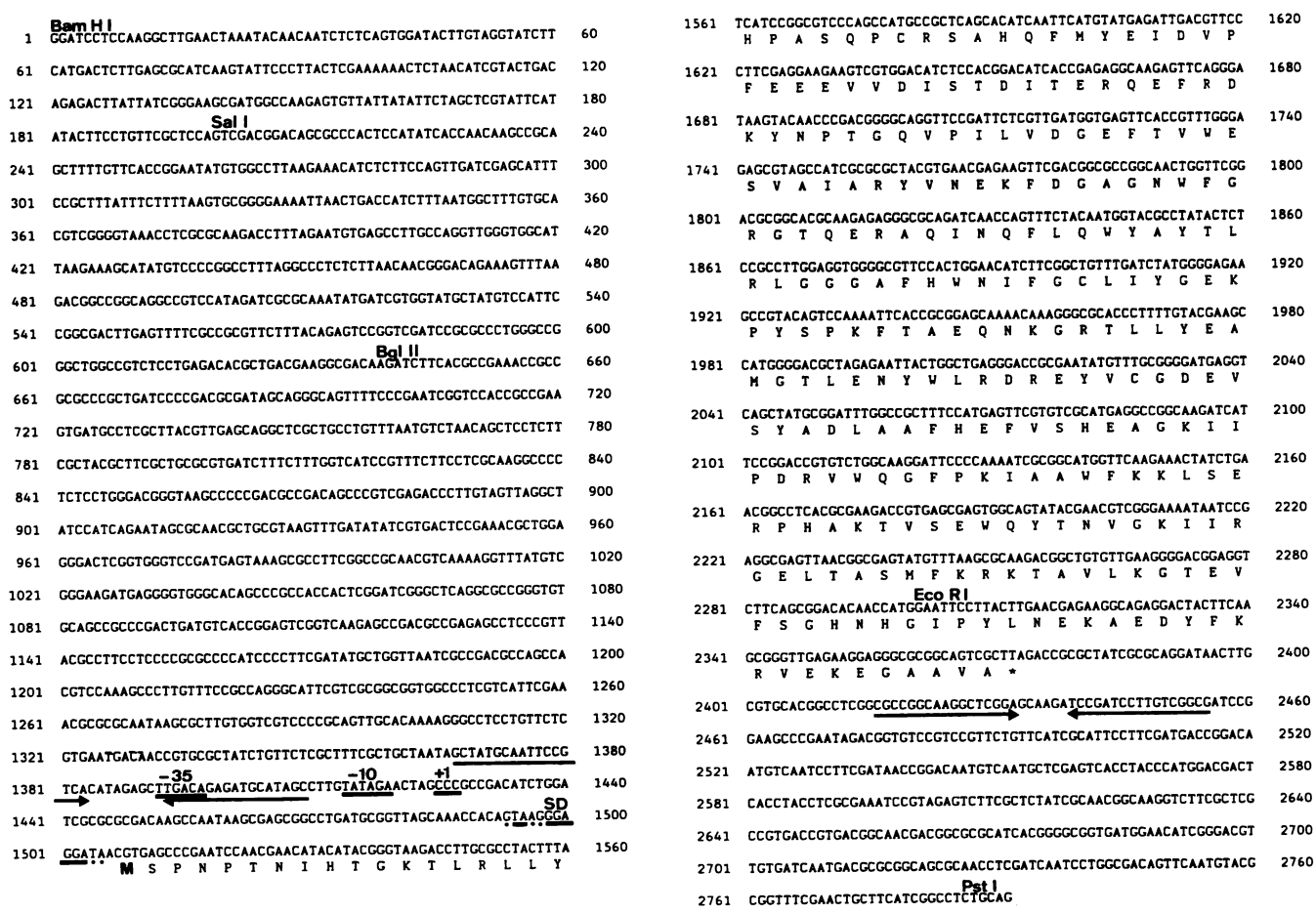


FIG. 2. Nucleotide sequence of the 2,790-bp *Bam*HI-*Pst*I fragment containing the DCM dehalogenase structural gene *dcmA* and its upstream region. The deduced amino acid sequence of DCM dehalogenase is written below the nucleotide sequence. Homologies with *E. coli* in a possible Shine-Dalgarno sequence (SD) are marked by continuous underlining. Two putative stem-loop structures are denoted with arrows, and the transcription start point (+1) and the putative promoter region (−10 and −35) are underlined.

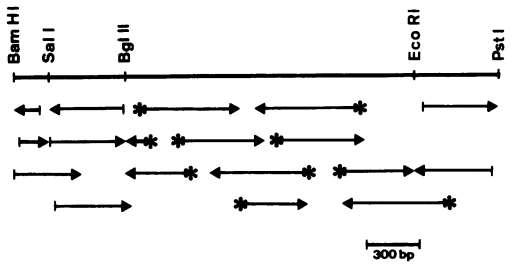


FIG. 3. Restriction map of the 2.8-kb *Bam*HI-*Pst*I fragment containing the DCM dehalogenase structural gene *dcmA* and summary of the sequencing strategy. The arrows indicate the direction of sequencing and the extent of sequence obtained from individual clones. Arrows starting with an asterisk indicate clones obtained after subcloning of deletion fragments (exonuclease III and mung-bean nuclease) from pBLS into M13mp18 or M13mp19. Arrows without aterisks represent restriction fragments cloned from pUC derivatives into M13mp18 or M13mp19.

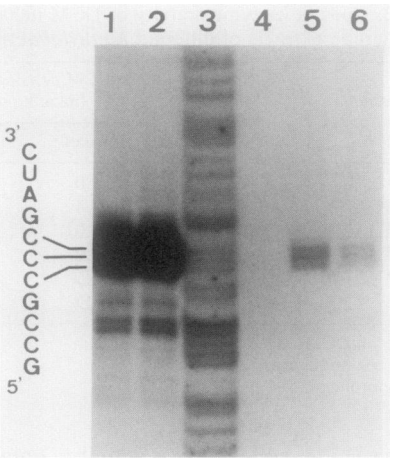


FIG. 4. Nuclease S1 mapping of the 5′ terminus of the *dcmA* transcript. RNA (20 μg per experiment and lane) was taken from *Methylobacterium* sp. strain DM4 (wild type) grown on DCM (lane 1) or methanol (lane 5) *Methylobacterium* sp. strain DM4-2cr(pME1523) grown on DCM (lane 2) or on methanol (lane 6), or *Methylobacterium* sp. strain DM4-2cr grown on methanol (lane 4). Lane 3 shows the Maxam and Gilbert sequencing ladder (A+G reaction) of the 5′-end-labeled probe.

shown) obtained by using the Genetics Computer Group program package indicated an overall hydrophilic character of DCM dehalogenase. In accordance with our earlier observations (17), the enzyme thus appears to be located in the cytoplasm.

The open reading frame encoding DCM dehalogenase was preceded by a potential ribosome-binding site (Fig. 2). A 34-base-pair (bp) G+C-rich sequence of imperfect dyad symmetry with the potential to form a stem-loop structure ($\Delta G^0 = -43$ kcal/mol) was located 40 bp downstream of the *dcmA* termination codon. Approximately 100 bp upstream of the translation start point another imperfect inverted repeat ($\Delta G^0 = -36$ kcal/mol) was situated, covering the −35 region of the potential *dcmA* promoter (Fig. 2).

**Transcriptional start point of *dcmA*.** The 5′ end of the in vivo *dcmA* transcripts was determined by nuclease S1 mapping as described in Materials and Methods. Cellular RNA was isolated from *Methylobacterium* sp. strains DM4 and DM4-2cr(pME1523), each grown on methanol or on DCM. The 5′-$^{32}$P-end-labeled DNA probe used for hybridization was a 414-bp fragment reaching from positions 1694 to 1280 (Fig. 2). Three strong signals at positions 1435, 1436, and 1437 were observed with RNA originating from DCM-grown cells (Fig. 2; Fig. 4, lanes 1 and 2). The same bands at much lower intensity were observed with RNA from methanol-grown cells (Fig. 4, lanes 5 and 6), a result reflecting transcriptional control of *dcmA* expression. Since the DCM utilization genes are deleted in the DCM$^-$ strain DM4-2cr (10), no signals were observed when the nuclease S1 mapping was carried out with RNA from this organism (Fig. 4, lane 4). Identical results confirming a single transcription start site around position 1436, 82 bp upstream of the translation start, were obtained with a hybridization probe extending 868 bp upstream from the GTG codon at position 1508 (not shown).

At 6 bp upstream of the transcription start site is a hexamer with considerable homology (four of six bases) to the −10 consensus sequence of *E. coli*, and 16 bp further upstream is a perfect −35 *E. coli* consensus sequence. This putative *dcmA* promoter is indicated by underscoring in Fig. 2. Its 6-bp distance from the transcription start site and the 16-bp spacing between the −10 and −35 sequences are within the *E. coli* norm (12).

**Regulation of *dcmA* expression.** The DCM dehalogenase activity encoded by some of the recombinant plasmids shown in Table 2 was induced by DCM. Intact regulation of *dcmA* was conferred on the DCM$^-$ host only by plasmids pME1510 and pME1523 but not by pME1532 (Table 2). This made it likely that the 4.2-kb fragment carried by plasmid pME1523 encodes, in addition to *dcmA*, a regulatory gene(s) governing induction of DCM dehalogenase. Alternatively, the *trans*-acting protein(s) involved in the specific induction of the enzyme by DCM could be encoded at some other location(s) on the genome of strain DM4-2cr. To distinguish between these two possibilities, plasmid pME1523 was subjected to deletion analysis. Deletion of the 1.4-kb *Pst*I-*Bam*HI fragment downstream of *dcmA* did not affect the regulation of dehalogenase formation (Fig. 5). Figure 5 also shows a series of pME1523 derivatives with progressive deletions reaching into the 1.5-kb region upstream of *dcmA* and the inducibility of DCM dehalogenase conferred by these constructs. Deletion of the first 200 bp of the *dcmA* upstream region (pME1542) was without effect on DCM dehalogenase inducibility. Deletions covering the 0- to 920-bp segment of this region (pME1543 and pME1544) resulted in partial constitutivity of the enzyme, and a deletion extending from 0 to 1,280 bp (pME1545) completely abolished the regulated expression of *dcmA*, leading to a strain with a constitutively induced enzyme level. When pME1545 was mobilized into the DM4 wild-type strain, partially constitutive DCM dehalogenase expression upon growth on methanol was observed (1.3 mkat/kg of protein). This suggests the presence of a *trans*-acting regulatory protein in the wild type but not in strain DM4-2cr. Interaction of this protein with pME1545 would explain the partial constitutivity of DCM dehalogenase in DM4(pME1545).

## DISCUSSION

The nucleotide sequence of the DCM dehalogenase structural gene from *Methylobacterium* sp. strain DM4 presented in this report is, to our knowledge, the first nucleotide sequence of a bacterial gene encoding an enzyme specifically and exclusively involved in a dehalogenation reaction. It is therefore not possible to compare this sequence with se-

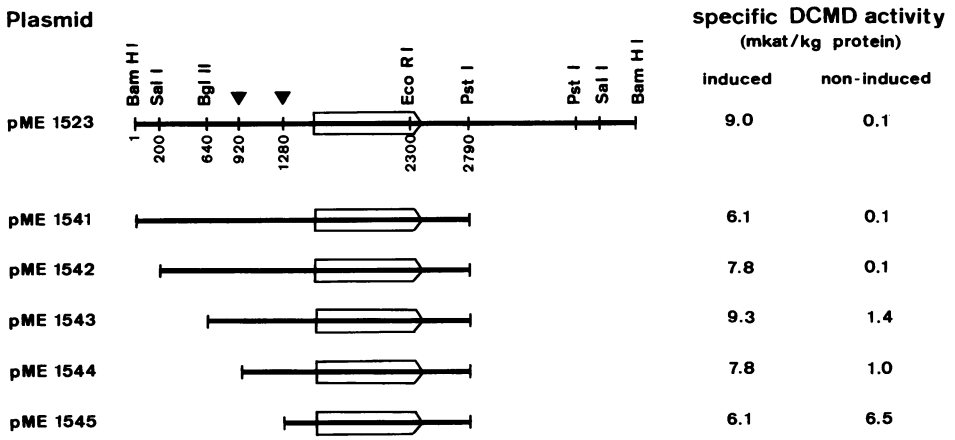


FIG. 5. Effect of deletions in the *dcmA* upstream region on the regulation of DCM dehalogenase expression. The depicted plasmids were constructed in pVK100 and mobilized from *E. coli* S17-1 into *Methylobacterium* sp. strain DM4-2cr, a DCM-nonutilizing mutant. Crude extracts of methanol- and DCM-grown cells were prepared, and the specific activity of DCM dehalogenase (DCMD) was measured. Nucleotide positions (Fig. 2) are indicated on the restriction map of pME1523. Positions marked with a triangle are clones originally constructed by deletion with exonuclease III and mungbean nuclease. The arrow-shaped boxes indicate the position of *dcmA*.

quences of functionally related bacterial genes. A computer search for sequences homologous to *dcmA* in the DNA sequence libraries available by the Genetics Computer Group yielded no results. However, since the reaction mechanism of DCM dehalogenase is postulated (17) to be similar to the reaction mechanism of GSH *S*-transferases (EC 2.5.1.18), we have compared its deduced amino acid sequence with the amino acid sequences of members of the GSH *S*-transferase supergene family (24). Like *Methylobacterium* DCM dehalogenase, rat liver cytosolic GSH *S*-transferase dehalogenates DCM with GSH to formaldehyde (1).

A search in the Swissprot databank yielded amino acid sequences of GSH *S*-transferases from humans, rats, the helminth parasite *Schistosoma japonicum*, and maize. Bacterial GSH *S*-transferases have been detected in crude extracts (2, 17), and three forms of GSH *S*-transferases from *Proteus mirabilis* have been purified (7), but there are no amino acid sequences of the bacterial enzymes available. The alignment of the deduced DCM dehalogenase amino acid sequence with the amino acid sequences of four representative eucaryotic GSH *S*-transferases (Fig. 6) shows three regions containing highly conserved amino acid residues. These regions extend from amino acid positions 64 to 106, 171 to 182, and 205 to 220. Within the first region lies a heptadecapeptide (positions 77 to 93) that has been hypothesized to represent part of a GSH-binding site common to all GSH *S*-transferases (14, 34, 35). The arginine residue proposed as a point of GSH anion binding is present at position 83 of the *Methylobacterium* DCM dehalogenase.

Table 3 lists the percent similarities between the amino acid sequences of human, rat, helminth, maize, and bacterial GSH *S*-transferases. The amino acid sequences used for this comparison were chosen to represent GSH *S*-transferases from different species and from the three multigene families that have been defined in mammals (24). The three representatives of the α-class multigene family had similarities above 80% among themselves (Table 3). The similarities of the remaining sequences listed in Table 3 ranged between 32 and 63%, and the similarities of the bacterial DCM dehalogenase (38 to 51%) were well within this range. This analysis, like the alignment presented in Fig. 6, demonstrates that the bacterial DCM dehalogenase is as related to eucaryotic GSH *S*-transferases as all of the eucaryotic enzymes from different multigene families are related among themselves. It thus confirms the functional relationship of DCM dehalogenase to eucaryotic GSH *S*-transferases (16) at the structural level. This is remarkable, since the *Methylobacterium* DCM dehalogenase differs in at least two important properties from eucaryotic GSH *S*-transferases: it does not react with 1-chloro-2,4-dinitrobenzene, a compound serving as the "universal" substrate for GSH *S*-transferases, and its hexameric structure with a subunit molecular weight of 37,430 differs from the dimeric structure with molecular weights between 23,000 and 26,000 reported for all eucaryotic GSH *S*-transferases (14).

Our understanding of the elements controlling the expression of *dcmA* in its proper host and in *E. coli* is fragmentary. A putative promoter sequence with high resemblance to the *E. coli* −10 and −35 consensus sequences (12) is located appropriately spaced from the transcription start point (Fig. 2). This element (TTGACA-16 bp-TATAGA-5 to 7 bp), despite its deviation in only two bases of the −10 region from

TABLE 3. Percent similarity[a] between the amino acid sequence of DCM dehalogenase and the amino acid sequences of human, rat, helminth, and maize GSH *S*-transferases

| GSH *S*-transferase | % similarity with GSH *S*-transferase[b]: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | H2 | Ya | Yc | Yb | Yp | S | M | D |
| H2 | **100** | **87** | **86** | 44 | 49 | 47 | 36 | 47 |
| Ya | **87** | **100** | **81** | 44 | 47 | 48 | 35 | 47 |
| Yc | **86** | **81** | **100** | 42 | 50 | 46 | 34 | 45 |
| Yb | 44 | 44 | 42 | 100 | 54 | 63 | 32 | 38 |
| Yp | 49 | 47 | 50 | 54 | 100 | 54 | 34 | 51 |
| S | 47 | 48 | 46 | 63 | 54 | 100 | 33 | 43 |
| M | 36 | 35 | 34 | 32 | 34 | 33 | 100 | 42 |
| D | 47 | 47 | 45 | 38 | 51 | 43 | 42 | 100 |

[a] Percent similarity was calculated by the method of Gribskov and Burgess (11) with the GAP program. Similar amino acids were Ile, Leu, Val, Met/Lys, His, Arg/Asp, Glur, Asn, Gln/Phe, Trp, Tyr/Ser, Thr.

[b] Abbreviations and references of the GSH *S*-transferases analyzed: H2: GSH *S*-transferase 2 of *Homo sapiens* (3); Ya, Yc, Yb, and Yp, GSH *S*-transferase subunits of *Rattus norvegicus* (18, 19, 32, 33); S, GSH *S*-transferase of *Schistosoma japonicum* (29); M: GSH *S*-transferase III of *Zea mays* (23); D, DCM dehalogenase of *Methylobacterium* sp. strain DM4 (this work). The values in boldface type are similarities among members of the GST α-class multigene family.

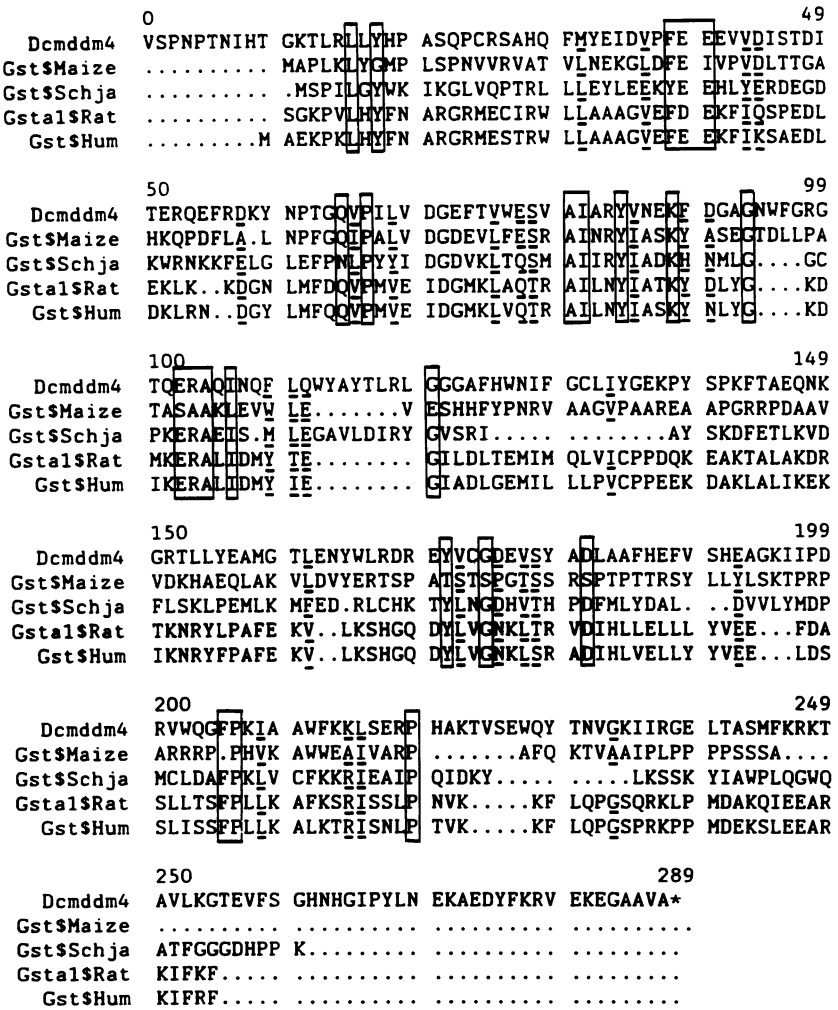


FIG. 6. Alignment of the amino acid sequences of DCM dehalogenase and GSH *S*-transferases from humans, rats, helminths, and maize. The sequence comparison was carried out using the programs GAP and LINE UP, and similarity was calculated by the method of Gribskov and Burgess (11) with the GAP program (Table 3). Similar amino acids are Ile, Leu, Val, Met/Lys, His, Arg/Asp, Glu, Asn, Gln/Phe, Trp, Tyr/Ser, Thr. Identical amino acids with at most one mismatch are boxed, and similar amino acids with at most one mismatch are underlined. Abbreviations and references of the sequences used: Dcmddm4, DCM dehalogenase of *Methylobacterium* sp. strain DM4 (this work); Gst$Maize, GSH *S*-transferase III of *Z. mays* (23); Gst$Schja, GSH *S*-transferase (SJ26 antigen) of *S. japonicum* (29); Gsta1$Rat, GSH-*S*-transferase subunit $Y_a$ of *R. norvegicus* (19); Gst$Hum: GSH *S*-transferase 2 of *H. sapiens* (3).

the *E. coli* consensus, permitted only marginal expression of *dcmA* in *E. coli*. Strains of *E. coli* carrying plasmids with the *dcmA* gene and its upstream region produced DCM dehalogenase at a low level, detectable only by the qualitative antibody assay. Weak expression of *dcmA* in *E. coli* may be due to one or more of the following factors: (i) to the deviation of the putative *dcmA* promoter in the 3′ terminus of the −10 hexamer from the *E. coli* consensus promoter (T in this position has been observed to be the most strongly conserved base among positively controlled and constitutive *E. coli* promoters [12, 25]); (ii) to the requirement for a positive regulatory element; (iii) to restrictions in tRNA availability imposed by differences in codon usage between *E. coli* (5) and *dcmA*; and (iv) to failure of *dcmA* expression due to the binding of a putative repressor encoded in the *dcmA* upstream region. The toxicity of DCM for *E. coli* has precluded induction experiments with *E. coli* strains harboring *dcmA* recombinant plasmids to refute or support this possibility. However, functional regulation of *dcmA* in *E. coli* would be testable by measuring DCM dehalogenase expression in *E. coli* from plasmid pME1545 (Fig. 5), which directs constitutive enzyme formation in *Methylobacterium* sp. strain DM4-2cr.

In *Methylobacterium* sp. strain DM4, DCM dehalogenase is induced 50- to 80-fold by DCM and about 5-fold by the gratuitous inducer 1,2-dichloroethane (16). Since the concentration of RNA hybridizing to *dcmA* is strongly increased in cells grown on DCM as compared with that in methanol-grown cells (Fig. 4), part or all of this regulation seems to occur on the transcriptional level. Truncation of the *dcmA* upstream region led to partially constitutive expression and its complete removal led to fully constitutive expression of DCM dehalogenase (Fig. 5). This observation makes it seem likely that *dcmA* expression is subject to negative control. *trans*-acting and/or *cis*-acting regulatory elements encoded in the upstream region of *dcmA* (nucleotides 200 to 1280 in Fig. 2) remain to be explored. This region contains four possible open reading frames that range between 44 and 123 amino acids (data not shown).

## ACKNOWLEDGMENTS

We thank R. Gälli for providing plasmids pME1518 and pME1523, and we are grateful to S. Reithaar for performing some of the cloning experiments.

This research was supported by a grant from the Swiss Federal Institute of Technology, Zürich, Switzerland.

## LITERATURE CITED

1. **Ahmed, A. E., and M. W. Anders.** 1978. Metabolism of dihalomethanes to formaldehyde and inorganic halide. II. Studies on the mechanism of the reaction. Biochem. Pharmacol. **27**:2021–2025.
2. **Arca, P., M. Rico, A. F. Braña, C. J. Villar, C. Hardisson, and J. E. Suárez.** 1988. Formation of an adduct between fosfomycin and glutathione: a new mechanism of antibiotic resistance in bacteria. Antimicrob. Agents Chemother. **32**:1552–1556.
3. **Board, P. G., and G. C. Webb.** 1987. Isolation of a cDNA clone and localization of human glutathione S-transferase 2 genes to chromosome band 6p12. Proc. Natl. Acad. Sci. USA **84**:2377–2381.
4. **Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
5. **de Boer, H. A., and R. A. Kastelein.** 1986. Biased codon usage: an exploration of its role in optimization of translation, p. 225–285. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworths, Boston.
6. **Devereux, I., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
7. **Di Ilio, C., A. Aceto, R. Piccolomini, N. Allocati, A. Faraone, L. Cellini, G. Ravagnan, and G. Federici.** 1988. Purification and characterization of three forms of glutathione transferase from *Proteus mirabilis*. Biochem. J. **255**:971–975.
8. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis of two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
9. **Gälli, R., and T. Leisinger.** 1985. Specialized bacterial strains for the removal of dichloromethane from industrial waste. Conserv. Recycling **8**:91–100.
10. **Gälli, R., and T. Leisinger.** 1988. Plasmid analysis and cloning of the dichloromethane utilization genes of *Methylobacterium* sp. DM4. J. Gen. Microbiol. **134**:943–952.
11. **Gribskov, M., and R. R. Burgess.** 1986. Sigma factors from *E. coli*, *B. subtilis*, phage SPO1 and phage T4 are homologous proteins. Nucleic Acids Res. **14**:6745–6763.
12. **Harley, C. B., and R. R. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15**:2343–2361.
13. **Hennecke, H., J. Günther, and F. Binder.** 1982. A novel cloning vector for the direct selection of recombinant DNA in *E. coli*. Gene **19**:231–234.
14. **Ketterer, B., D. J. Meyer, and A. G. Clark.** 1988. Soluble glutathione transferase isozymes, p. 73–135. *In* H. Sies and B. Ketterer (ed.), Glutathione conjugation mechanisms and biological significance. Academic Press, Inc., London.
15. **Knauf, V. C., and E. W. Nester.** 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid **8**:45–54.
16. **Kohler-Staub, D., S. Hartmans, R. Gälli, F. Suter, and T. Leisinger.** 1986. Evidence for identical dichloromethane dehalogenases in different methylotrophic bacteria. J. Gen. Microbiol. **132**:2837–2843.
17. **Kohler-Staub, D., and T. Leisinger.** 1985. Dichloromethane dehalogenase of *Hyphomicrobium* sp. strain DM2. J. Bacteriol. **162**:676–681.
18. **Lai, H.-C. J., G. Grove, and C.-P. Tu.** 1986. Cloning and sequence analysis of a cDNA for a rat liver glutathione S-transferase $Y_b$ subunit. Nucleic Acids Res. **14**:6101–6114.
19. **Lai, H.-C. J., N.-Q. Li, M. J. Weiss, C. C. Reddy, and C.-P. D. Tu.** 1984. The nucleotide sequence of a rat liver glutathione S-transferase subunit cDNA clone. J. Biol. Chem. **259**:5536–5542.
20. **Machlin, S. M., and R. S. Hanson.** 1988. Nucleotide sequence and transcriptional start site of the *Methylobacterium organophilum* XX methanol dehydrogenase structural gene. J. Bacteriol. **170**:4739–4747.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
23. **Moore, R. E., M. S. Davies, K. M. O'Connell, E. I. Harding, R. C. Wiegand, and D. C. Tiemeier.** 1986. Cloning and expression of a cDNA encoding a maize glutathione-S-transferase in *E. coli*. Nucleic Acids Res. **14**:7227–7235.
24. **Peterson, L. A., and F. P. Guengerich.** 1988. Comparison of and relationships between glutathione S-transferase and cytochrome P-450 systems, p. 193–233. *In* H. Sies and B. Ketterer (ed.), Glutathione conjugation mechanisms and biological significance. Academic Press, Inc., London.
25. **Raibaud, O., and M. Schwartz.** 1984. Positive control of transcription initiation in bacteria. Annu. Rev. Genet. **18**:173–206.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
27. **Scholtz, R., L. P. Wackett, C. Egli, A. M. Cook, and T. Leisinger.** 1988. Dichloromethane dehalogenase with improved catalytic activity isolated from a fast-growing dichloromethane-utilizing bacterium. J. Bacteriol. **170**:5698–5704.
28. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Bio/Technology **1**:784–790.
29. **Smith, D. B., K. M. Davern, P. G. Board, W. U. Tiu, E. G. Garcia, and G. F. Mitchell.** 1986. $M_r$ 26,000 antigen of *Schistosoma japonicum* recognized by resistant WEHI 129/J mice is a parasite glutathione S-transferase. Proc. Natl. Acad. Sci. USA **83**:8703–8707.
30. **Stanisich, V. A., and B. W. Holloway.** 1972. A mutant sex factor of *Pseudomonas aeruginosa*. Genet. Res. **19**:91–108.
31. **Stucki, G., R. Gälli, H.-R. Ebersold, and T. Leisinger.** 1981. Dehalogenation of dichloromethane by cell extracts of *Hyphomicrobium* DM2. Arch. Microbiol. **130**:366–371.
32. **Suguoka, Y., T. Kano, A. Okuda, M. Sakai, T. Kitagawa, and M. Muramatsu.** 1985. Cloning and the nucleotide sequence of rat glutathione S-transferase P cDNA. Nucleic Acids Res. **13**:6049–6057.
33. **Telakowski-Hopkins, C. A., K. A. Rodkey, C. D. Bennett, A. Y. H. Lu, and C. B. Pickett.** 1985. Rat liver glutathione S-transferase: construction of a cDNA clone complementary to a $Y_c$ mRNA and prediction of the complete amino acid sequence of a $Y_c$ subunit. J. Biol. Chem. **260**:5820–5825.
34. **Telakowski-Hopkins, C. A., G. S. Rothkopf, and C. B. Pickett.** 1986. Structural analysis of a rat liver glutathione S-transferase Yα gene. Proc. Natl. Acad. Sci. USA **83**:9393–9397.
35. **Tu, C.-P. D., and B. Qian.** 1986. Human liver glutathione S-transferases: complete primary sequence of an $H_a$ subunit cDNA. Biochem. Biophys. Res. Commun. **141**:229–237.
36. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
37. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**:103–119.