Identification of the Type I Trimethoprim-Resistant Dihydrofolate Reductase Specified by the *Escherichia coli* R-Plasmid R483: Comparison with Procaryotic and Eucaryotic Dihydrofolate Reductases

CHRISTIAN C. SIMONSEN,* ELLSON Y. CHEN, AND ARTHUR D. LEVINSON

Department of Molecular Biology, Genentech, Inc., South San Francisco, California 94080

We have isolated and determined the nucleotide sequence of a 1,626-base-pair fragment from R-plasmid R483 which encodes a trimethoprim-resistant dihydrofolate reductase. Analysis of the nucleotide sequence of this fragment revealed the presence of two open reading frames, each sufficient to encode polypeptides of approximately 17,000 daltons. Both open regions are preceded by sequences conforming closely to the canonical description of procaryotic promoters. A 490-base-pair *Hpa*I fragment spanning one of the potential coding regions was inserted into a plasmid vector under the transcriptional control of the *trp* promoter. Cells transformed with this plasmid were trimethoprim resistant and produced dihydrofolate reductase activity which in vitro was resistant to moderate levels of trimethoprim. Analysis of the predicted amino acid sequence of this protein indicated that the R483-encoded trimethoprim-resistant enzyme was distantly related to the trimethoprim-sensitive bacterial homologs. The conserved amino acids were localized primarily to the region of the enzyme previously shown to comprise the hydrophobic substrate binding pocket.

Dihydrofolate reductase (Dhfr; EC 1.5.1.3) catalyzes the reduction of dihydrofolate to tetrahydrofolate, a cofactor required in purine, pyrimidine, and glycine biosynthesis (10). Several folate analogs, such as methotrexate and trimethoprim, which powerfully inhibit Dhfr and result in cell death, have been described previously (10, 11). These agents bind tightly to the catalytic site of the enzyme in a region found, from X-ray diffraction analysis, to comprise a hydrophobic substrate binding pocket (25, 41). However, the use of these agents as antineoplastic or antibacterial compounds is compromised by the emergence of cells resistant to the effects of the drugs. Resistance can result from any of three mechanisms which may act separately or in concert. Overproduction of the enzyme (2, 4, 24), caused by increases in the level of transcription (30) or by amplification of the Dhfr gene (1, 30), results in a drug-resistant phenotype in bacterial and mammalian cells. In addition, the effects of the folate analogs can be overcome by mutations which affect the transport of the folate analogs across the cell membrane (22) or by the synthesis of Dhfrs having reduced affinities for the inhibitors (18, 22). Analyses of the mutant Dhfrs found in a trimethoprim-resistant strain of *Escherichia coli* (5) and in a methotrexate-resist-

ant line of mouse 3T6 cells (32) have shown that the altered Dhfr found in each organism differs from the wild-type enzyme at a single amino acid in the region of the enzyme thought to comprise the hydrophobic binding pocket essential for inhibitor binding (25, 41).

In 1972, Fleming and his colleagues (16) reported the presence of bacterial R-plasmids which conferred high levels of trimethoprim resistance upon their hosts. Later reports (3, 33) demonstrated that the R-plasmids encoded a Dhfr which was resistant to the antifolate drugs. Two forms of R-factor Dhfr based on differential susceptibility to trimethoprim and related antifolate compounds have been described (27). The type I Dhfr, typified by the R483-encoded enzyme, is inhibited by trimethoprim at levels 1,000-fold greater than those which inhibit the chromosomally encoded bacterial enzyme and has a native molecular weight of 35,000, which is twice that of the chromosomally encoded Dhfr (27). Expression of the type I enzyme in an *E. coli* minicell system has demonstrated that the subunit molecular weight is approximately 18,000, suggesting that the active form of the enzyme is a dimer (17). Type II R-factor Dhfr is completely insensitive to the effects of trimethoprim. The native molecular weight of type II R-

factor Dhfr, like the type I enzyme, is 35,000 (27). Since the subunit molecular weight of the type II enzyme is 9,000 (35), the active form is likely a tetramer. Amino acid (39) and nucleotide (40) sequence analyses of the type II Dhfr show that it is not related to the chromosomally encoded, trimethoprim-sensitive bacterial enzymes. In this study, the trimethoprim resistance marker encoded by R-plasmid R483 has been isolated and sequenced. Expression of this gene under control of the *E. coli trp* promoter results in the production of a 17,500-dalton polypeptide. The Dhfr activity obtained from this gene is inhibited by trimethoprim at levels identical to those described for the type I plasmid-encoded Dhfr (27). This enzyme is most homologous to the wild-type bacterial Dhfrs in the regions thought to be involved in substrate binding but is only slightly homologous in the areas lying outside the active site.

MATERIALS AND METHODS

Materials. Trimethoprim, NADPH, and dihydrofolate were purchased from Sigma Chemical Co., St. Louis, Mo. Calf alkaline phosphatase was obtained from Boehringer Mannheim Corp., New York, N.Y. All restriction endonucleases were obtained from commercial suppliers and used in accordance with the instructions of the manufacturers. Radiolabeled dCTP (PB10165) was obtained from Amersham Corp., Arlington Heights, Ill.

Bacterial strains and plasmids. *E. coli* J53 F-containing R-plasmid R483 was obtained from Esther Lederberg at the Stanford Plasmid Reference Center (PRC no. 392). *E. coli* K-12 MM294 (6) was used as the recipient strain for all subsequent experiments. Cells were grown in either L broth or in M9 medium without supplements as described previously (20). The derivation of the *trp* expression plasmid pSRCex16 has been previously reported (26).

DNA manipulations. Small-scale plasmid DNA was prepared by the method of Birnboim and Doly (8). The methods of digesting DNA, ligating DNA fragments, and transforming plasmid DNA into *E. coli* have been previously described (20, 21). DNA was digested and electrophoresed on acrylamide or agarose gels for preparation of fragments and analysis. Large DNA fragments were isolated from 0.6% low-melting-temperature agarose gels in Tris-acetate buffer (13). DNA fragments were stained with ethidium bromide and visualized by UV light, and the bands were cut out. DNA fragments from agarose gels were diluted in 0.1% agarose with water and then heated to 65°C. After the fragments were cooled to 37°C, restriction or ligation buffer was added, and further manipulations were performed as described above. DNA fragments from acrylamide gels were electroeluted as described previously (20, 21).

DNA sequencing. Restriction fragments encompassing the type I Dhfr gene were inserted into single-strand phage M13 vectors mp8 and mp9 and sequenced by the dideoxynucleotide chain termination method (29) as described previously (12).

Dhfr assays. Bacterial extracts were prepared by the method of Baccanari et al. (4) from cells grown in M-9 medium to an optical density at 550 nm of 1.0. Enzyme activity was measured by monitoring the reduction of dihydrofolic acid at 340 nm (23). The inhibition of activity by trimethoprim was measured by plotting fractional inhibition versus trimethoprim concentration and interpolating to obtain the 50% inhibitory concentration (27).

RESULTS

Isolation of a DNA fragment encompassing type I Dhfr. The trimethoprim-resistant *E. coli* strain containing plasmid R483 was originally isolated from a calf undergoing trimethoprim treatment by Fleming and his colleagues (16). Subsequent work showed that the gene encoding the type I Dhfr is contained within transposon Tn7 and is closely linked to a streptomycin resistance marker (7, 11). To isolate the Dhfr gene, R-factor DNA isolated from *E. coli* J53F(R483) was cleaved with *Bam*HI and *Eco*RI, and ligated to pBR322 which had been similarly cleaved. Competent *E. coli* 294 cells were transformed with the ligation mix and were plated onto L broth plates containing ampicillin and trimethoprim. No colonies arose from control ligations lacking R483 inserts; however trimethoprim-resistant colonies did arise from ligations containing R483 DNA digested with *Eco*RI plus *Bam*HI. Plasmid DNA was isolated from these colonies and analyzed, using several restriction endonucleases. All trimethoprim-resistant colonies harbored plasmids carrying identical inserts of 35 kilobases (kb), consisting of a 30-kb *Eco*RI fragment and a 5.4-kb *Eco*RI-*Bam*HI fragment. This plasmid, designated pRB483, is shown in Fig. 1a.

To further localize the trimethoprim resistance gene, extraneous sequences were removed from pRB483 by treatment of the plasmid with *Pvu*II under conditions which produced incomplete digestion of the DNA. The partially cleaved DNAs were mixed with T4 DNA ligase at low DNA concentrations and used to transform competent *E. coli* 294. Plasmid DNA was isolated from trimethoprim-resistant colonies and analyzed. The smallest plasmid isolated, pTM4 (Fig. 1b), retained the 5.4-kb *Eco*RI-*Bam*HI fragment and contained a 6-kb *Eco*RI fragment. To determine which of the two insert fragments encoded the trimethoprim resistance marker, the 6-kb *Eco*RI and the 5.4-kb *Eco*RI-*Bam*HI fragments of pTM4 were separately subcloned into pBR322. Trimethoprim-resistant colonies were obtained from cells containing the 6-kb *Eco*RI fragment inserted into the vector but not from a vector containing the 5.4-kb *Eco*RI-*Bam*HI fragment.

To obtain a smaller fragment sufficient for sequence analysis of the DNA fragment encom-
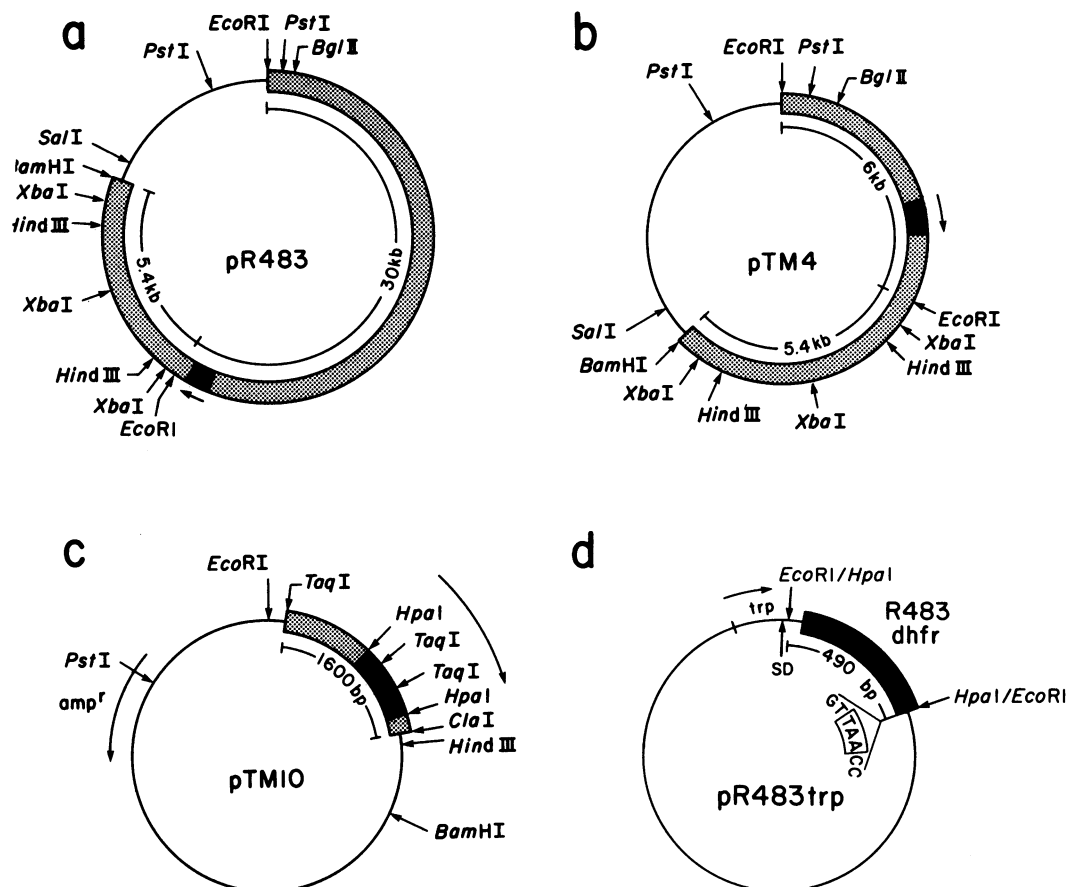
FIG. 1. Plasmid constructions. The derivations of the pBR322-R483 recombinant plasmids are described in the text. DNA derived from R483 is shaded; the type I Dhfr gene is shown in black. SD, Shine-Dalgarno sequence of *trp* promoter fragment. (a) pR483; (b) pTM4, (c) pTM10, (d) pR483trp.

passing the type I Dhfr, we digested the 6-kb DNA insert from pTM4 with *Taq*I under conditions which produced a spectrum of cleavage products ranging from completely digested to full-length 6-kb inserts. The DNA was then ligated into the *Cla*I site of pBR322. Trimethoprim-resistant colonies arising after transformation of competent *E. coli* with these mixtures were grown, and the plasmid DNA was isolated. Plasmids containing inserts ranging in size from approximately 1,600 to 3,500 base pairs (bp) were identified. The plasmid containing the smallest insert fragment, termed pTM10, was selected for sequence analysis (Fig. 1c).

DNA sequence analysis of a 1,626-bp fragment containing the type I Dhfr gene. The strategy used to sequence the insert is indicated in Fig. 2a. Briefly, the 1,626-bp fragment was digested with suitable restriction nucleases, and the products were inserted into phage M13 vectors mp8 and mp9 (12). The inserts were then sequenced

by the dideoxynucleotide chain termination method (12, 29). The resulting sequence of the 1,626-bp fragment is shown in Fig. 2b. Inspection of the sequence revealed that there were two nonoverlapping open reading frames capable of encoding polypeptides the size of the type I enzyme. One such open frame, in an orientation opposite to the sequence presented in Fig. 2b, begins at position 589 and ends at position 124; its translational initiation and termination codons are depicted in Fig. 2b. The second open reading frame initiates with a methionine codon at position 994 and terminates with an ochre codon at position 1449.

Identification of the gene encoding Dhfr. To identify which of the two open reading frames encoded the type I Dhfr, we inserted specific DNA sequences spanning these putative genes into plasmid vectors capable of expressing heterologous coding sequences and determined the ability of the expressed product to confer bacte-

a

```
                                                          TaqI
TaqI              RsaI    NcoI              HpaI |NcoI         TaqI  HpaI  TaqI
```

b

TaqI
TCGATAAGCAGAAGGGTATTTGTGATCTAAAGCAAAAGGCAGCGATGGCCCTACGCCTTGTAAGTTGTCGTCTTGCTGAATAAGCCGCGCTTGCTCAATGAGTTATTTTATTGCTGGGA 120

OC
TTAGGCGCGTGGGCAGTAGGCTGTTTCTGCTTTTCCCACCCTTACCGTCATGCACAGTGATGCAGCCATTATCAAAATCAAAATCTTTAACCCGCAAACGCAAGCATTCATTAATGCGC 240

AAACCTGCACCATACAGCAGCGTAAAAATAACTTGGTTGCGAGTATCCATAACCTGCAAAATGCGTTGCACTTCATTTGCAGAGATAACAGAGGGTAGCCGTCTAGGCTTGCTTGCAGG 360

RsaI
GATATAATCAATATCGCCCAACGGCTGTTGTAAAAACCTGTTGTACAAAAAAGCTAGGGCATTTAAAGCGATTTTCTGCGTGTTTATGGCTACATGTCTGCTGTTTGCTAAGCTGGATA 480

AAAACAGCCTGACCTCTTCACTGCCCATGGTCTGAGGATGACGTTTTTTGTGAAACAGAATAAAACGCTTAATCCAGTGCAGGTAAGTTTTTTCAGTTTTCAGCGCATAACCTTTTTGT 600
SD   -10   -35   ←————————————— ATG

CGCATATCCGTGCGTATAGAATTTAAAAATGGACTGTTAGACATAAAACGCTCCTTGTCTTGCAACTGTCTGCCTATACAGCCTATTCTAGCTGGGATTTAAAAAAGTGCCTGTTTTTT 720

ACGCCTAGAGATGCTTGTTTACCGGTAGAGTTTTAATTTAATGCTAAATAAATTAAAATGTTATGAGTTCTTTGGGTGAGATAATGTGCATCGTGCAAGCAGGATAGACGGCATGCACG 840

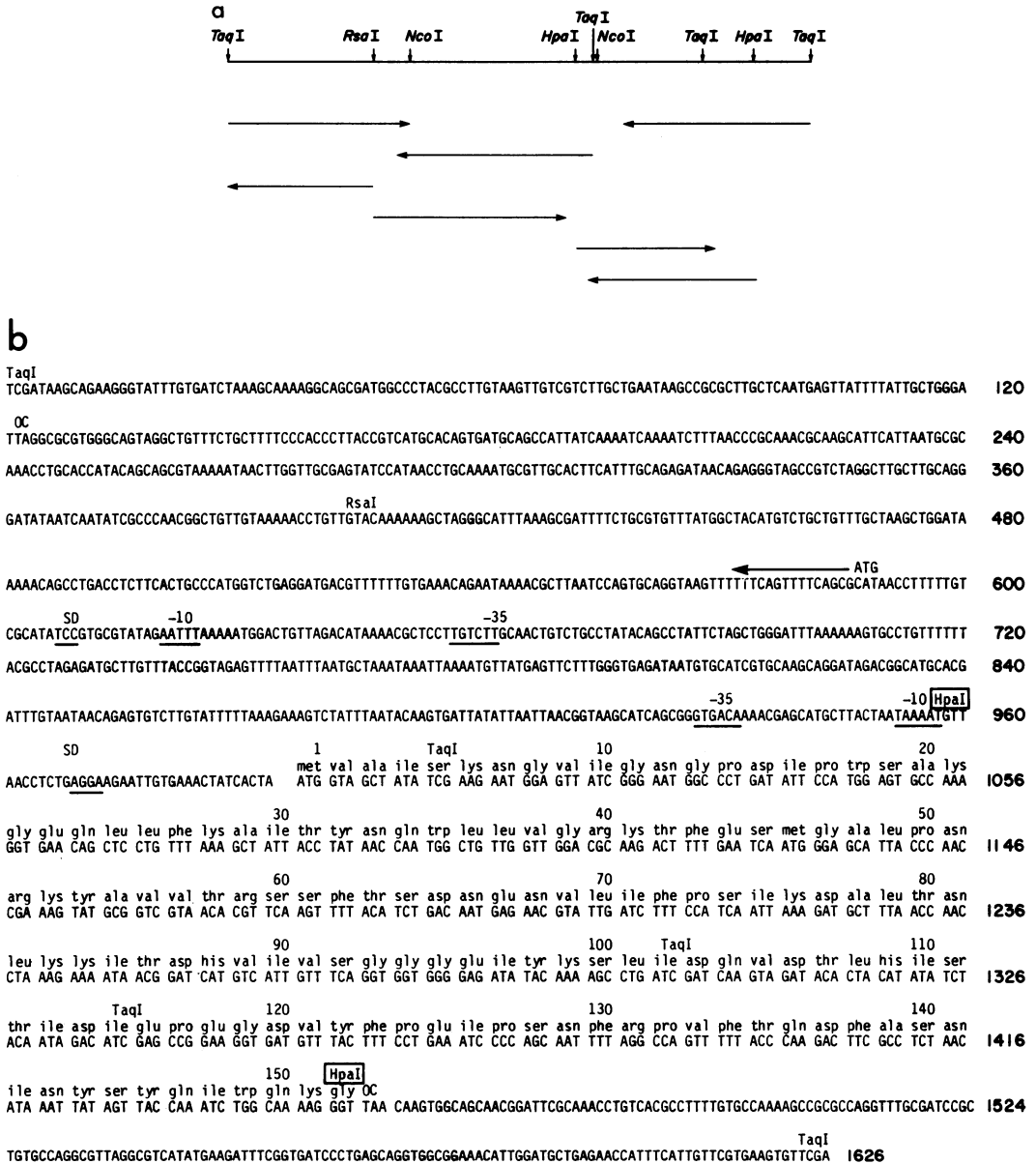   -35   -10 Hpal
ATTTGTAATAACAGAGTGTCTTGTATTTTTAAAGAAAGTCTATTTAATACAAGTGATTATATTAATTAACGGTAAGCATCAGCGGGTGACAAAACGAGCATGCTTACTAATAAAATGTT 960
SD   1   TaqI   10   20
                                 met val ala ile ser lys asn gly val ile gly asn gly pro asp ile pro trp ser ala lys
AACCTCTGAGGAAGAATTGTGAAACTATCACTA ATG GTA GCT ATA TCG AAG AAT GGA GTT ATC GGG AAT GGC CCT GAT ATT CCA TGG AGT GCC AAA 1056

              30   40   50
gly glu gln leu leu phe lys ala ile thr tyr asn gln trp leu leu val gly arg lys thr phe glu ser met gly ala leu pro asn
GGT GAA CAG CTC CTG TTT AAA GCT ATT ACC TAT AAC CAA TGG CTG TTG GTT GGA CGC AAG ACT TTT GAA TCA ATG GGA GCA TTA CCC AAC 1146

              60   70   80
arg lys tyr ala val val thr arg ser ser phe thr ser asp asn glu asn val leu ile phe pro ser ile lys asp ala leu thr asn
CGA AAG TAT GCG GTC GTA ACA CGT TCA AGT TTT ACA TCT GAC AAT GAG AAC GTA TTG ATC TTT CCA TCA ATT AAA GAT GCT TTA ACC AAC 1236

              90   100   TaqI   110
leu lys lys ile thr asp his val ile val ser gly gly gly glu ile tyr lys ser leu ile asp gln val asp thr leu his ile ser
CTA AAG AAA ATA ACG GAT CAT GTC ATT GTT TCA GGT GGT GGG GAG ATA TAC AAA AGC CTG ATC GAT CAA GTA GAT ACA CTA CAT ATA TCT 1326

        TaqI   120   130   140
thr ile asp ile glu pro glu gly asp val tyr phe pro glu ile pro ser asn phe arg pro val phe thr gln asp phe ala ser asn
ACA ATA GAC ATC GAG CCG GAA GGT GAT GTT TAC TTT CCT GAA ATC CCC AGC AAT TTT AGG CCA GTT TTT ACC CAA GAC TTC GCC TCT AAC 1416

                       150   Hpal
ile asn tyr ser tyr gln ile trp gln lys gly OC
ATA AAT TAT AGT TAC CAA ATC TGG CAA AAG GGT TAA CAAGTGGCAGCAACGGATTCGCAAACCTGTCACGCCTTTTGTGCCAAAAGCCGCGCCAGGTTTGCGATCCGC 1524

TaqI
TGTGCCAGGCGTTAGGCGTCATATGAAGATTTCGGTGATCCCTGAGCAGGTGGCGGAAACATTGGATGCTGAGAACCATTTCATTGTTCGTGAAGTGTTCGA 1626

FIG. 2. (a) Restriction map of 1,626-bp region and sequencing strategy. The 1,626-bp *Eco*RI and *Hind*III fragment contained in pTM10 and in various subfragments created by digestion with *Taq*I, *Rsa*I, and *Hpa*I was isolated using acrylamide gels, inserted into M13 vectors mp8 and mp9 (12), and sequenced by the dideoxynucleotide chain termination method (29). The direction sequenced and the extent of sequence information obtained from the various fragments are indicated by the arrows. (b) Sequence of the 1,626-bp region. The nucleotide sequence of the 1,626-bp fragment and the amino acid sequence of the open reading frame spanning the 490-bp *Hpa*I fragment are shown. The second open reading frame (positions 589 to 124) is indicated by the translational initiation (ATG) and termination (TAA) codons and is oriented in a direction opposite to that of the translated sequence. Consensus regulatory sequences located at 35 and 10 bp 5' of the presumed start sites of translation of both reading frames are underlined. SD, Shine-Dalgarno-like sequences; OC, ochre (TAA) termination codon.

rial resistance to trimethoprim. The open reading frame beginning at position 994 spans three *Taq*I sites and requires the coding information contained in all four *Taq*I fragments which constitute the 1,626-bp region. The other potential gene is contained entirely within a 1-kb *Taq*I

TABLE 1. Growth of *E. coli* 294 transformed with plasmids containing type I plasmid-encoded Dhfr[a]

Trimethoprim (mg/ml)	Growth[b] with:			
	pRB483	pR483trp	pSRCex16	No plasmid
0	+	+	+	+
0.0005	+	+	+	+
0.005	+	+	−	−
0.05	+	+	−	−
0.5	+	+	−	−
5.0	−	−	−	−

[a] Little or no effect on the number of colonies was observed until levels of trimethoprim were sufficient to totally inhibit growth.

[b] +, Colony formation observed; −, no colonies observed.

fragment and therefore should have been isolated as a 1-kb *Taq*I fragment capable of conferring trimethoprim resistance were it the type I gene; it thus seemed most likely that the open reading frame lying between positions 994 and 1449 encoded the type I Dhfr. This open reading frame is conveniently located between two *Hpa*I restriction endonuclease sites at positions 958 and 1448. A 490-bp fragment generated by *Hpa*I digestion of pTM10 was inserted into a plasmid following the *trp* promoter (contained in plasmid pSRCex16 [26]) at an *Eco*RI site made blunt, using Klenow DNA polymerase I (20).

The resulting plasmid, pR483trp (Fig. 1d), contains a Shine-Dalgarno sequence (31), a −35 region (14), and a −10 region (28) from the *trp* promoter (20) fused directly to the R483 sequences. By inserting the *Hpa*I fragment into a filled-in *Eco*RI site (5-AATTC...) following the *trp* promoter, the termination codon of the gene was restored. Plasmid pR483trp was introduced into competent cells, which were then plated onto medium containing various levels of trimethoprim. Only cells transformed with pR483trp produced trimethoprim-resistant colonies. Cells transformed with plasmid pR483trp grew in the presence of 0.5 mg of trimethoprim per ml (Table 1). In contrast, bacteria harboring either no plasmid or pSRCex16 failed to grow in the presence of even 0.005 mg of trimethoprim per ml. Plasmid pR483trp conferred a resistance to trimethoprim identical to the trimethoprim resistance conferred by pRB483, pTM4, or pTM10, thereby localizing the trimethoprim-resistant Dhfr to this 490-bp DNA sequence.

To determine the sensitivity of the plasmid-encoded Dhfr to trimethoprim, cells transformed with pR483trp were grown in minimal medium under conditions which induced high levels of transcription from the *trp* promoter (20). An extract of these induced cells was prepared;

when assayed for the reduction of dihydrofolate in the presence of NADPH, we observed that a level of 56 μM trimethoprim inhibited the Dhfr activity by 50% (Fig. 3), a value virtually identical to that previously reported for the R483-encoded enzyme (27). Control cultures induced under the same conditions did not produce detectable levels of Dhfr. Proteins from these same induced cell extracts were analyzed on sodium dodecyl sulfate-polyacrylamide gels; a protein with an apparent molecular weight of 17,000 was observed in induced cultures containing plasmid pR483 but not in induced control cultures (data not shown).

We have been unable to determine the identity of the second putative gene contained within the 1,626-bp insert of pTM10. We have observed that the streptomycin resistance marker, which is contained in plasmids pRB483 and pTM4, is not found in either pTM10 or pR483trp (data not shown), and no significant homologies were noted when a computer-assisted search of known protein sequences was performed.

Comparison of the predicted protein sequence of the type I enzyme with those of Dhfrs. The amino acid sequence of the type I enzyme, as derived from the DNA sequence, was compared (36) with the reported amino acid sequences of Dhfrs from *E. coli* (38), *Lactobacillus casei* (9), *Streptococcus faecium* (19), and methotrexate-resistant murine tissue culture cells (13) (Fig. 4). Since there is no available protein sequence for

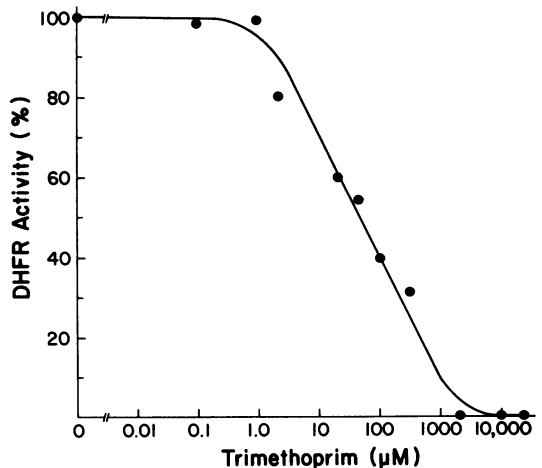FIG. 3. Inhibition of Dhfr activity by trimethoprim. Cells containing pR483trp were induced by growing in minimal M-9 medium to an optical density at 550 nm of 1.0 (20). Extracts (0.5 ml each) were prepared from 10 ml of induced cells as described previously (4). The reduction of dihydrofolate to tetrahydrofolate at 340 nm (23) was measured in increasing levels of trimethoprim (22, 26) and plotted as shown.

R483 Type 1

```
                                    1                              10                          20
R483 Type 1                         met val ala ile ser lys asn gly val ile gly asn gly pro asp ile pro trp ser ala lys gly glu gln
                                              *       *         *   *   *               *                 *   *
E. coli RT500          met ile ser leu ile ala asp leu ala val asp arg val ile gly met glu asn ala met pro trp asn leu pro ala asp leu
                                              *               *           *   *               *                 *   *
L. casei               thr ala phe leu trp ala gln asn arg asp gly leu ile gly lys asp gly his leu pro trp his leu pro asp asp leu
                         *         *         *   *   *   *         *   *               *             *   *
S. faecium             met phe ile ser met trp ala gln asn lys asn gly leu ile gly lys asp gly leu leu pro trp arg leu pro asn asp met
                         *   *         *   *   *       *             *   *               *             *   *
Murine S-180      val arg pro leu asn cys ile val ala val ser gln asn met gly ile gly lys asn gly asp leu pro trp         pro pro leu arg
                                     ■         ○ ○     ■ ■ ○       ○       ○ ■ ■     ○ ●       ○
                                     |———————————————————————A———————————————————————|
```

```
                  30                                40
R483 Type 1       leu leu phe lys ala ile thr tyr asn gln trp     leu leu val     gly arg lys thr phe glu ser met gly ala
                    *   *       *     *   *   *               *                       *   *   *   *   *       *   *
E. coli RT500     ala trp phe lys arg asn thr leu asn lys pro     val ile met     gly arg his thr trp glu ser ile gly
                    *   *       *   *               *                               *   *   *   *   *       *   *
L. casei          his tyr phe arg ala gln thr val gly lys ile     met val val     gly arg arg thr tyr glu ser phe pro lys
                        *       *                                                   *   *   *   *   *       *   *
S. faecium        arg phe phe arg glu his thr met asp lys ile     leu val met     gly arg lys thr tyr glu gly met gly lys
                    *   *       *                                                   *   *   *   *   *       *   *
Murine S-180      asn glu phe lys tyr phe gln arg met thr thr thr ser ser val glu gly lys gln asn leu val ile met gly arg lys thr trp phe
                    ■ ○         ●         ○                         ●   ■ ●     ●         ○   ○ ○
                                                                              |———————————————
```

```
                              50                                60                              70
R483 Type 1                   leu pro asn arg lys tyr ala val val thr arg ser ser phe thr ser asp asn glu asn val leu
                                *   *   *               *   *   *   *   *                 *
E. coli RT500                 arg pro leu pro gly arg lys asn ile ile leu ser ser gln pro gly thr asp asp arg val thr trp val
                                *   *   *   *                 *   *   *   *                     *                       *
L. casei                      arg pro leu pro glu arg thr asn val val leu thr his gln glu asp tyr gln ala gln gly ala val val
                                *   *   *   *             *   *   *   *                   *                               *
S. faecium                    leu ser leu pro tyr arg his ile ile leu thr thr gln lys asp phe lys val glu lys ala ala glu
                                    *   *       *             *   *   *       *
Murine S-180             ser ile pro glu lys asn arg pro leu lys asp arg ile asn ile val leu ser arg glu leu lys glu pro pro arg gly ala his phe
                              ○ ○   ■ ●       ■         ○ ○   ● ● ○       ○
                         |———————————B———————————|
```

```
                                    80                                        90
R483 Type 1       ile phe pro ser ile lys asp ala leu thr             asn leu lys lys ile thr asp his val ile val ser gly gly gly
                                        *                                                                     *       *   *   *
E. coli RT500     lys ser val asp glu ala ile ala                     ala cys gly asp val pro glu ile met val ile gly gly gly
                                                                                                                 *   *   *
L. casei          val his asp val ala ala val phe ala tyr             ala lys gln his leu asp gln glu leu val ile ala gly gly gly
                    *   *                                                    *                                     *   *   *   *
S. faecium        val leu his ser ile asp glu leu leu ala             tyr ala lys asp ile pro glu asp ile tyr val ser gly gly ser
                                          *   *   *                          *                 *
Murine S-180      leu ala lys ser leu asp asp ala leu arg leu ile glu gln pro glu leu ala ser lys val asp met val trp ile val gly gly ser
                              ○             ○ ○                                                                 ○     ■ ■
                                                                                                       |——————————————C——
```

```
                              100                             110                                 120
R483 Type 1       glu ile tyr lys ser leu ile asp gln val asp thr leu his ile ser     thr     ile asp ile glu pro glu gly asp val tyr phe
                            *                               *                             *     *   *   *   *   *   *   *       *   *
E. coli RT500     arg val tyr glu gln phe leu pro lys ala gln lys leu tyr leu           thr his ile asp ala glu val glu gly asp thr his phe
                        *   *       *           *   *   *   *                             *           *   *   *           *   *
L. casei          gln ile phe thr ala phe lys asp asp val asp thr leu leu val           thr arg leu ala gly ser phe glu gly asp thr lys met
                        *                     *                                           *   *   *   *   *   *       *   *   *
S. faecium        arg ile phe gln ala leu leu pro glu thr lys ile ile trp arg           thr leu ile asp glu glu phe glu gly asp thr phe ile
                        *   *                                                             *   *   *   *   *       *   *   *   *
Murine S-180      ser val tyr gln glu ala met asn gln pro gly his leu arg leu phe val thr arg ile met gln glu phe glu ser asp thr phe phe
                          ○ ○                               ●                             ■   ● ○       ● ○ ■ ● ■ ●             ○
                  |———————————————————————|                                              |——————————————————D——
```

```
                              130                                 140                             150
R483 Type 1       pro glu ile pro ser asn phe arg pro val phe thr gln asp phe ala ser asn ile asn tyr ser tyr gln ile trp gln lys gly
                    *                                               *
E. coli RT500     pro asp tyr glu pro asp asp trp glu ser val phe ser glu phe his asp ala asp ala gln asn ser his ser tyr cys phe glu ile
                                                                          *                           *
L. casei          ile pro leu asn trp asp asp phe thr lys val ser ser arg thr val glu asp thr asn pro ala leu thr his thr tyr glu val trp
                        *
S. faecium        gly glu ile asp phe thr ser phe glu leu val glu glu his gly gly ile val asn gln glu asn gln tyr pro his arg phe gln lys
                    *   *   *
Murine S-180      pro glu ile asp leu gly lys tyr lys leu leu pro glu tyr pro gly val leu ser glu val gln glu glu lys gly ile lys tyr lys cys
                    ○ ○ ○                                 ○
                  |———————————————|
```

```
R483 Type 1
E. coli RT500     leu glu arg arg
L. casei          gln                   lys lys ala
S. faecium        trp gln lys met ser lys val val
Murine S-180      phe glu val tyr glu lys lys asp
```
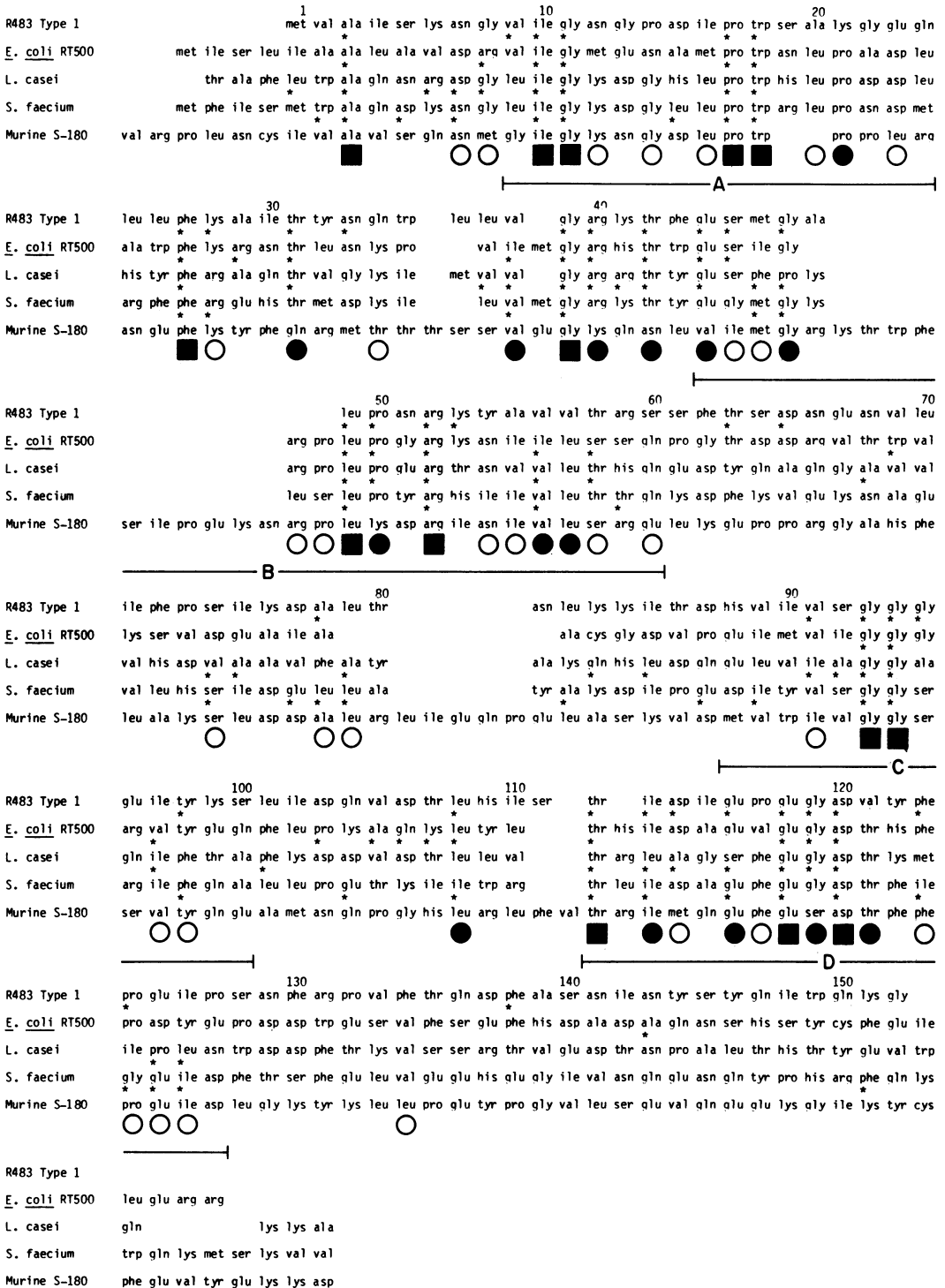
FIG. 4. Amino acid comparison of the type I Dhfr with different Dhfrs. The sequences were aligned, using the Smith-Waterman algorithm (36), and were manually adjusted to give the best fit relative to the type I enzyme. The sequence of the mouse S-180 cell line enzyme has been deduced from its nucleotide sequence (13), and is virtually identical to that of the mouse L1210 enzyme (37). The sequence of the *E. coli*, *S. faecium*, and *L. casei* enzymes were obtained by direct sequence analysis of the purified polypeptides (9, 19, 38). Symbols: *, residues homologous to the R483 enzyme; ■, residues which are common to all five enzymes; ●, homologous residues which are common to four of the Dhfrs; ○, homologous residues found in three-fifths of the Dhfrs. The regions of the enzyme forming the hydrophobic binding pocket (24) are underlined.

the type I Dhfr, our assignment of the N-terminal residues is based upon utilization of the ATG codon located at position 994. It should be noted that a slightly longer polypeptide may represent the primary translational product of the R483-encoded Dhfr in the event that the GTG codon, located 15 nucleotides 5' of the presumed initiator ATG at position 994, is utilized as the translational initiation codon. If so, such a polypeptide would be identical to that proposed in Fig. 2b except for the presence of five additional amino acids (Val-Lys-Leu-Ser-Leu-) at the amino terminus. The comparison in Fig. 4 shows that there is a significant degree (22 to 29%) of overall homology between the bacterial trimethoprim-sensitive enzymes and the type I enzyme. Surprisingly, the type I enzyme displays as much homology to the murine enzyme as to the procaryotic enzymes.

DISCUSSION

We have isolated and sequenced a 1,626-bp region from the drug resistance plasmid R483 which encodes a type I trimethoprim-resistant Dhfr. The gene was localized to a 490-bp restriction enzyme fragment, which encompassed an open reading frame sufficient to encode a polypeptide of 152 amino acids. This size is in accord with previous estimates for the monomeric subunit obtained by expression of the type I enzyme in an *E. coli* minicell system (17). These results suggest that the native form of the type I enzyme, which has a molecular weight of 35,000 (27), is composed of two identical subunits. Preceding the open reading frame are sequences conforming to the canonical description of bacterial promoters, such as the −35 region (5'TTGACA) (14) and the −10 region (5'-TATAAT) (28). Upstream from the proposed translational initiation codon is a sequence (5'-AGGA) which may act as a ribosome binding site (31), although the distance between this presumed Shine-Dalgarno site and the initiator methionine is longer than that usually observed (22 bp) (14). In this context it is possible that the in-phase GTG codon (valine), located 7 bp after this potential ribosome binding site, serves as the translational initiator codon; such a polypeptide would possess an additional five amino acids at the amino terminus. Resolution of this issue must await direct sequence analysis of the enzyme.

The type I enzyme is of particular interest in that it is inhibited by levels of folate analog intermediate between the exquisitely sensitive chromosomally located bacterial Dhfrs and the totally insensitive type II plasmid-encoded Dhfr found on R-plasmid R-388 (27). A comparison of the amino acid sequence of the type I Dhfr with other procaryotic and eucaryotic Dhfrs (36) reveals that the R483-encoded Dhfr has 22 to 29% of its amino acids in common with other Dhfrs, a homology sufficient to indicate a genuine relationship (15). It is notable that the degree to which the various bacterial Dhfrs are related is much less than that observed for Dhfrs from avian species and a wide variety of mammalian species (41). Approximately 70% of the residues which are conserved (to a level of 60% or greater) among the five different Dhfrs are found in the four regions (A, B, C, and D, Fig. 4) which have been shown to form the hydrophobic binding site of the enzyme (25). X-ray crystallographic studies have shown that methotrexate, and by analogy, trimethoprim and the substrate dihydrofolate, occupies a cleft formed by these regions (41).

Although it is possible that the trimethoprim resistance marker of Tn7 originated from an *E. coli* chromosomal gene, such a conclusion is not supported by a comparison of the amino acid sequences of the various Dhfrs, in that the type I Dhfr is no more similar to the *E. coli* enzyme than to other bacterial Dhfrs or even to the murine enzyme. Furthermore, an examination of the nucleotide sequences comprising the 1,626-bp fragment encoding the type I enzyme reveals little homology to the sequences which constitute and surround the chromosomally encoded *E. coli* Dhfr (34). In view of the observation that the drug sensitivity of a bacterial (5) and a mammalian (32) Dhfr can be profoundly affected by a single amino acid substitution, the extensive structural dissimilarities between the type I Dhfr and the bacterial enzymes make it difficult to assess which changes in the binding site reflect the functional changes in the properties of the enzyme. Clarification of this issue, as well as an appreciation of the phylogeny of the type I Dhfr, must await analyses of similar enzymes from additional sources.

LITERATURE CITED

1. **Alt, F. W., R. E. Kellems, J. R. Bertino, and R. T. Schimke.** 1978. Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. J. Biol. Chem. **253:**1357–1370.
2. **Alt, F. W., R. E. Kellems, and R. T. Schimke.** 1976. Synthesis and degradation of folate reductase in sensitive and methotrexate-resistant lines of S-180 cells. J. Biol. Chem. **251:**3063–3074.
3. **Aymes, S. G. B., and J. T. Smith.** 1974. R-factor trimethoprim resistance mechanism: an insusceptible target site. Biochem. Biophys. Res. Commun. **58:**412–418.
4. **Baccanari, D., A. Phillips, S. Smith, D. Sinsk, and J. Burchall.** 1975. Purification and properties of *Escherichia coli* dihydrofolate reductase. Biochemistry **24:**5267–5273.

5. **Baccanari, D. P., D. Stone, and L. Kuyper.** 1981. Effect of a single amino acid substitution on *E. coli* dihydrofolate reductase catalysis and ligand binding. J. Biol. Chem. **256:**1738–1747.

6. **Backman, K., M. Ptashne, and W. Gilbert.** 1976. Construction of plasmids carrying the cI gene of bacteriophage lambda. Proc. Natl. Acad. Sci. U.S.A. **73:**4174–4178.

7. **Barth, P. T., N. Datta, R. W. Hedges, and N. J. Grinter.** 1976. Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. J. Bacteriol. **125:**800–810.

8. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.

9. **Bitar, K. G., D. T. Blankenship, K. A. Walsh, R. B. Dunlop, A. V. Reddy, and J. H. Freisham.** 1977. Amino acid sequence of dihydrofolate reductase EC-1.5.1.3 from an amethopterin resistant strain of lactobacillus casei. FEBS Lett. **80:**119–122.

10. **Blakley, R. L.** 1969. The biochemistry of folic acid and related pteridines, p. 139. Elsevier North-Holland Publishing Co., Amsterdam.

11. **Burchall, J. J., and G. H. Hitchins.** 1965. Inhibitor binding analysis of dihydrofolate reductases from various species. Mol. Pharmacol. **1:**126–136.

12. **Chen, E. Y., P. M. Howley, A. D. Levinson, and P. H. Seeburg.** 1982. The primary structure and genetic organization of the bovine papillomavirus type 1 genome. Nature (London) **299:**529–534.

13. **Crouse, G. F., C. C. Simonsen, R. N. McEwan, and R. T. Schimke.** 1982. Structure of amplified normal and variant dihydrofolate reductase genes in mouse sarcoma S180 cells. J. Biol. Chem. **257:**7887–7897.

14. **De Boer, H. A., L. J. Comstock, D. G. Yansura, and H. L. Yansura.** 1981. Construction of a tandem *trp-lac* promoter and a hybrid *trp-lac* promoter for efficient and controlled expression of the human growth hormone in *Escherichia coli*, p. 462–481. *In* M. J. Chamberlain and R. Rodriguez (ed.), Promoters, structure, and function. Praeger Publishers, Division of Holt Rinehart Winston/CBS, New York.

15. **Doolittle, R. F.** 1981. Similar amino acid sequences: chance or common ancestry? Science **214:**149–159.

16. **Fleming, M. P., N. Datta, and R. Gruneberg.** 1972. Trimethoprim resistance determined by R factors. Brit. Med. J. **1:**726–728.

17. **Fling, M. E., and L. P. Elwell.** 1980. Protein expression in *Escherichia coli* minicells containing recombinant plasmids specifying trimethoprim-resistant dihydrofolate reductases. J. Bacteriol. **141:**779–785.

18. **Flintoff, W. F., S. V. Davidson, and L. Siminovitch.** 1976. Isolation and partial characterization of three methotrexate-resistant phenotypes from Chinese hamster ovary cells. Somatic Cell Genet. **2:**245–261.

19. **Gleisner, J. M., D. L. Peterson, and R. L. Blakley.** 1974. Amino acid sequence of dihydrofolate reductase from a methotrexate-resistant mutant of *Streptococcus faecium* and identification of methionine residues at the inhibitor binding site. Proc. Natl. Acad. Sci. U.S.A. **71:**3001–3005.

20. **Goeddel, D. V., H. L. Heyneker, T. Hozumi, R. Arentzen, K. Itakura, D. G. Yansura, M. J. Ross, G. Miozzari, R. Crea, and P. H. Seeburg.** 1979. Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. Nature (London) **281:**544–548.

21. **Gray, P. W., D. W. Leung, D. Pennica, E. Yelverton, R. Najarian, C. C. Simonsen, R. Derynck, P. J. Sherwood, D. M. Wallace, S. L. Berger, A. D. Levinson, and D. V. Goeddel.** 1982. Expression of human immune interferon cDNA in *E. coli* and monkey cells. Nature (London) **295:**503–508.

22. **Haber, D. A., and R. T. Schimke.** 1981. Unstable amplifi-

cation of an altered dihydrofolate reductase gene associated with double-minute chromosomes. Cell **26:**355–362.

23. **Hillcoat, B. L., and R. L. Blakley.** 1966. Dihydrofolate reductase of Streptococcus faecalis. I. Purification and some properties of reductase from the wild strain and from strain A. J. Biol. Chem. **241:**2995–3001.

24. **Kellems, R. E., F. W. Alt, and R. T. Schimke.** 1976. Regulation of folate reductase synthesis in sensitive and methotrexate-resistant sarcoma 180 cells. J. Biol. Chem. **251:**6987–6993.

25. **Matthews, D. A., R. A. Alden, J. T. Bolin, S. T. Freer, R. Hamlin, N. Xuong, J. Kraut, M. Poe, M. Williams, and K. Hoogsteen.** 1977. Dihydrofolate reductase: X-ray structure of the binary complex with methotrexate. Science **197:**452–455.

26. **McGrath, J. P., and A. D. Levinson.** 1982. Bacterial expression of an enzymatically active protein encoded by RSV *src* gene. Nature (London) **295:**423–425.

27. **Pattishall, K. H., J. Acar, J. J. Burchall, F. W. Goldstein, and R. J. Harvey.** 1977. Two distinct types of trimethoprim-resistant dihydrofolate reductase specified by R-plasmids of different compatibility groups. J. Biol. Chem. **252:**2319–2323.

28. **Pribnow, D.** 1975. Bacteriophage T7 promoters: nucleotide sequences of two RNA polymerase binding sites. J. Mol. Biol. **99:**419–443.

29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74:**5463–5467.

30. **Sheldon, R., and S. Brenner.** 1976. Regulatory mutants of dihydrofolate reductase in *Escherichia coli* K-12. Mol. Gen. Genet. **147:**91–97.

31. **Shine, J., and L. Dalgarno.** 1974. The 3-prime terminal sequence of *Escherichia coli* 16S ribosomal RNA complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. **71:**1342–1346.

32. **Simonsen, C. C., and A. D. Levinson.** 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. Proc. Natl. Acad. Sci. U.S.A. **80:**2495–2499.

33. **Skold, O., and A. Widh.** 1974. A new dihydrofolate reductase with low trimethoprim sensitivity induced by an R factor mediating high resistance to trimethoprim. J. Biol. Chem. **249:**4323–4325.

34. **Smith, D. R., and J. M. Calvo.** 1980. Nucleotide sequence of the *E. coli* gene coding for dihydrofolate reductase. Nucleic Acids Res. **8:**2255–2274.

35. **Smith, S. L., D. Stone, P. Novak, D. Baccanari, and J. J. Burchall.** 1979. R plasmid dihydrofolate reductase with subunit structure. J. Biol. Chem. **254:**6222–6225.

36. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. J. Mol. Biol. **147:**195–197.

37. **Stone, D., S. J. Paterson, J. H. Raper, and A. W. Phillips.** 1979. The amino acid sequence of dihydrofolate reductase from the mouse lymphoma L1210. J. Biol. Chem. **254:**480–488.

38. **Stone, D., A. W. Phillips, and J. J. Burchall.** 1977. The amino acid sequence of the dihydrofolate reductase EC-1.5.1.3 of a trimethoprim resistant strain of *Escherichia coli*. Eur. J. Biochem. **72:**613–624.

39. **Stone, D., and S. L. Smith.** 1979. The amino acid sequence of the trimethoprim-resistant dihydrofolate reductase specified in *Escherichia coli* by R-plasmid R67. J. Biol. Chem. **254:**10857–10861.

40. **Swift, G., B. J. McCarthy, and F. Heffron.** 1981. DNA sequence of a plasmid encoded dihydrofolate reductase. Mol. Gen. Genet. **181:**441–447.

41. **Volz, K. W., D. A. Matthews, R. A. Alden, S. T. Freer, C. Hansch, B. T. Kaufman, and J. Kraut.** 1982. Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH. J. Biol. Chem. **257:**2528–2536.