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

Received 10 March 1983/Accepted 20 June 1983

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 HpaI 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 Dhf rs 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-resistant 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 153 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 to 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 BamHI and EcoRI, 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 EcoRI plus BamHI. 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 EcoRI fragment and a 5.4-kb EcoRI-BamHI 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 PvuII 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 EcoRI-BamHI fragment and contained a 6-kb EcoRI fragment. To determine which of the two insert fragments encoded the trimethoprim resistance marker, the 6-kb EcoRI and the 5.4-kb EcoRI-BamHI fragments of pTM4 were separately subcloned into pBR322. Trimethoprim-resistant colonies were obtained from cells containing the 6-kb EcoRI fragment inserted into the vector but not from a vector containing the 5.4-kb EcoRI-BamHI fragment.

To obtain a smaller fragment sufficient for sequence analysis of the DNA fragment encom-
IDENTIFICATION OF TRIMETHOPRIM-RESISTANT Dhfr

passing the type I Dhfr, we digested the 6-kb DNA insert from pTM4 with TaqI 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 Clal 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-
FIG. 2. (a) Restriction map of 1,626-bp region and sequencing strategy. The 1,626-bp EcoRI and HindIII fragment contained in pTM10 and in various subfragments created by digestion with TaqI, Rsal, and HpaI 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 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 HpaI 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.

bial resistance to trimethoprim. The open reading frame beginning at position 994 spans three TaqI sites and requires the coding information contained in all four TaqI fragments which constitute the 1,626-bp region. The other potential gene is contained entirely within a 1-kb TaqI
fragment and therefore should have been isolated as a 1-kb TaqI 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 Dfr. This open reading frame is conveniently located between two HpaI restriction endonuclease sites at positions 958 and 1448. A 490-bp fragment generated by HpaI digestion of pTM10 was inserted into a plasmid following the trp promoter (contained in plasmid pSRCex16 (26)) at an EcoRI 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 HpaI fragment into a filled-in EcoRI 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 Dfr to this 490-bp DNA sequence.

To determine the sensitivity of the plasmid-encoded Dfr 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 Dfr 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 Dfr. 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 Dhrs.** 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 Dhrs from *E. coli* (38), *Lactobacillus casei* (9), *Streptococcus faecium* (19), and methotrexate-resistant murine tissue culture cells (13) (Fig. 4). Since there is not available protein sequence for

---

**TABLE 1. Growth of *E. coli* 294 transformed with plasmids containing type I plasmid-encoded Dfr**

<table>
<thead>
<tr>
<th>Trimethoprim (mg/ml)</th>
<th>pRB483</th>
<th>pR483trp</th>
<th>pSRCex16</th>
<th>No plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>0.0005</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>0.005</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>0.05</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>0.5</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>5.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

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

* +, Colony formation observed; -, no colonies observed.

---

**FIG. 3. Inhibition of Dfr 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.
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; O, 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 Dhrf, 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 Dhrf in the event that the GTG-encoded Dhrf, 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 Dhrf. The gene was localized to a 490-bp restriction 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' TAAT) (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 Dhrfs and the totally insensitive type II plasmid-encoded Dhrf found on R-plasmid R-388 (27). A comparison of the amino acid sequence of the type I Dhrf with other procaryotic and eucaryotic Dhrfs (36) reveals that the R483-encoded Dhrf has 22 to 29% of its amino acids in common with other Dhrfs, a homology sufficient to indicate a genuine relationship (15). It is notable that the degree to which the various bacterial Dhrfs are related is much less than that observed for Dhrfs 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 Dhrfs 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 Dhrfs, in that the type I Dhrf is no more similar to the E. coli enzyme than to other bacterial Dhrfs 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 Dhrf (34). In view of the observation that the drug sensitivity of a bacterial (5) and a mammalian (32) Dhrf can be profoundly affected by a single amino acid substitution, the extensive structural dissimilarities between the type I Dhrf 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 Dhrf, must await analyses of similar enzymes from additional sources.

ACKNOWLEDGMENTS

We thank Fran Fields for expert technical assistance, Alan Gray for preparation of the illustrations, and Doug Smith for aiding in DNA sequencing.

LITERATURE CITED


