of differences in the expression of particular genes. In contrast, the representation of coding sequences in cDNA libraries reflects differences in cellular mRNA levels. Thus, expression libraries containing genomic DNA rather than complementary DNA may be more suitable for a thorough examination of coding capacity of eukaryotes with small genome sizes.

References and Notes

1. P. Chambon, *Annu. Rev. of Biochem.* **44**, 613 (1975).
2. R. G. Roeder, in *RNA Polymerase*, R. Losick and M. Chamberlin, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1976), p. 285.
3. M. R. Paule, *Trends Biochem. Sci.* **6**, 128 (1981).
4. A. Sentenac and B. Hall, in *The Molecular Biology of the Yeast Saccharomyces*, J. N. Strathern, E. W. Jones, J. R. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), p. 561.
5. J.-M. Buhler, J. Huet, K. E. Davies, A. Sentenac, P. Fromageot, *J. Biol. Chem.* **255**, 9949 (1980).
6. J. Huet, A. Sentenac, P. Fromageot, *ibid.* **257**, 2613 (1982).
7. R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194 (1983).
8. R. A. Young, unpublished data.
9. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
10. R. A. Young, O. Hagenbuchle, U. Schibler, *Cell* **23**, 451 (1981).
11. P. W. Hartman and J. R. Roth, *Adv. Genet.* **17**, 1 (1973).
12. This article describes the isolation of genes for two of the approximately ten subunits of RNA polymerase II. Genes for the other subunits of this enzyme are being cloned in our laboratory and those for yeast RNA polymerases I and III are being isolated by J.-M. Buhler and A. Sentenac. The study of a complete set of genes for the classically defined RNA polymerase proteins should play an important role in defining the eukaryotic transcription apparatus.
13. D. J. Kemp, R. L. Coppel, A. F. Cowman, R. B. Saint, G. V. Brown, R. F. Anders, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3787 (1983).
14. T. St. John, personal communication.
15. Y1088 was derived from KM392 (itself a derivative of LE392) by transformation with plasmid pMC9, a pBR322 plasmid carrying *lac*I (*16*). Y1089 was derived from BNN103 (*7*) by transformation with pMC9. Y1090 was constructed by phage P1 tranduction of *sup*F from BNN99 (*7*) into BNN96 (*7*), followed by transformation with pMC9. The λgt11 lysogen BNN97 (*7*) and the strains Y1088, Y1089, and Y1090 are available through the American Type Culture Collection, Rockville, Md. 20852.
16. M. P. Calos, T. S. Lebkowski, M. R. Botchan, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3015 (1983).
17. M. Snyder, unpublished data.
18. W. H. Hunter and F. C. Greenwood, *Nature (London)* **194**, 495 (1962).
19. R. K. Mortimer and D. Schild, in *The Molecular Biology of the Yeast Saccharomyces*, J. N. Strathern, E. W. Jones, J. R. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), p. 11.
20. R. W. Davis, D. Botstein, J. R. Roth, in *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980), p. 174.

21. We thank D. Ruden and C. Parker for the essential contributions of purified yeast RNA polymerase II; J. Huet, J.-M. Buhler, and A. Sentenac for subunit antibodies; M. Snyder for the λgt11 yeast genomic DNA library; D. Manoli for technical assistance; M. Calos for pMC9; K. Moore for KM392; and A. Buchman, J.-M. Buhler, C. Mann, B. J. Meyer, C. Parker, A. Sentenac, S. Scherer, and T. St. John for enlightening discussions.

Directed Mutagenesis of Dihydrofolate Reductase

Jesus E. Villafranca, Elizabeth E. Howell, Donald H. Voet
Marjorie S. Strobel, Richard C. Ogden, John N. Abelson
Joseph Kraut

The enzyme dihydrofolate reductase (DHFR; E.C. 1.5.1.3) (*1*), is found in every kind of organism, from bacteria to mammals, and in large amounts in rapidly dividing cell lines. It catalyzes the NADPH-dependent (NADPH, reduced form of nicotinamide adenine dinucleotide phosphate) reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate, which in turn plays a central metabolic role as a carrier of one-carbon units in the biosynthesis of thymidylate, purines, and some amino acids (*2*). An unusual feature of thymidylate biosynthesis in particular is that tetrahydrofolate is oxidized to dihydrofolate in the course of a one-carbon transfer to uridylate. Thus blockade of the DHFR-catalyzed reduction of dihydrofolate back to tetrahydrofolate in a rapidly proliferating cell results in the depletion of its tetrahydrofolate pool, with consequent cessation of DNA synthesis and, ultimately, stasis and cell death.

Furthermore, DHFR is especially susceptible to inhibition by synthetic folate analogs (the antifolates), some of which show a high degree of species selectivity. These peculiarities of DHFR enzymology have assumed practical importance in the clinical treatment of a number of diseases, notably the leukemias and other cancers, and certain bacterial and protozoal infections, against which chemotherapy with antifolates such as methotrexate, trimethoprim, and pyrimethamine proves to be effective (*3*). Not surprisingly, therefore, DHFR has become an enzyme of considerable interest to pharmaceutical chemists and drug designers.

Being the smallest of the well-characterized nicotinamide-dependent oxidoreductases, DHFR is also of interest to the structural biochemist. For example, the *Escherichia coli* enzyme consists of just a single 159-residue polypeptide chain. Thus DHFR presents an opportunity to study the stereochemistry of nicotinamide activation by high-resolution x-ray diffraction methods.

Motivated by these considerations, Matthews and Kraut and their co-workers (*4, 5*) have during the past few years determined the x-ray crystal structures of DHFR's derived from three species—*E. coli*, *Lactobacillus casei*, and chicken (the last being representative of all vertebrate DHFR's, which are highly homologous—at the maximum practical resolution and degree of refinement. Examination of these structures reveals a characteristic, highly conserved backbone chain fold, even though their amino acid sequences are only about 25 percent homologous, and provides a catalog of enzyme-substrate interactions, about half of which are also conserved. One can guess how some of these interactions might contribute to the catalytic mechanism, but reasons for the evolutionary conservation of others remain obscure. In addition, a few residues that do not interact directly with the substrate or cofactor are also conserved in the known DHFR sequences. Thus, many new questions are raised by the x-ray structures; but what is of overriding importance is that these questions are specific and highly focused.

It has been somewhat easier to ask incisive questions about the function of some part of an enzyme, given its three-dimensional structure, than it has been to devise equally incisive experiments to

This research project was conducted at the Agouron Institute, 505 Coast Boulevard South, La Jolla, California 92037. All authors except D. H. Voet are members of the Agouron Institute. J. Kraut is also professor of chemistry and J. E. Villafranca is also a postgraduate research chemist at the University of California, San Diego, La Jolla 92093. J. N. Abelson is also professor of biology and M. Strobel is currently a graduate student at the California Institute of Technology, Pasadena 91125. D. H. Voet is professor of chemistry at the University of Pennsylvania, Philadelphia 19104.

Summary. Three mutations of the enzyme dihydrofolate reductase were constructed by oligonucleotide-directed mutagenesis of the cloned *Escherichia coli* gene. The mutations—at residue 27, aspartic acid replaced with asparagine; at residue 39, proline replaced with cysteine; and at residue 95, glycine replaced with alanine—were designed to answer questions about the relations between molecular structure and function that were raised by the x-ray crystal structures. Properties of the mutant proteins show that Asp-27 is important for catalysis and that perturbation of the local structure at a conserved *cis* peptide bond following Gly-95 abolishes activity. Substitution of cysteine for proline at residue 39 results in the appearance of new forms of the enzyme that correspond to various oxidation states of the cysteine. One of these forms probably represents a species cross-linked by an intrachain disulfide bridge between the cysteine at position 85 and the new cysteine at position 39.

answer them. In recent years, however, a powerful new experimental tool has been developed by molecular biologists. The contemporary techniques for manipulating DNA make it possible to alter a cloned gene in vitro so as to change any specific base in the sequence to any other base, to delete a given segment of the gene, or to insert a new segment wherever one wishes (*6, 7*). This technique, sometimes known as site-directed mutagenesis, has been increasingly used to investigate the function of genes themselves and of their RNA transcripts (*8*). But only recently has the realization begun to spread among those preoccupied with relations between structure and function in protein molecules—the gene products—that site-directed mutagenesis is prospectively a powerful experimental tool with which to answer their questions as well (*9*).

In this article, we describe the beginning of a series of in vitro site-directed mutagenesis experiments, the purpose of which is to increase our understanding of the DHFR molecule in particular, and of enzyme-protein structure in general. For several reasons the DHFR of *E. coli* is especially well suited as a subject for such experiments. It is a relatively simple, unprocessed water-soluble protein, so that expression from an *E. coli* compatible vector in an *E. coli* host should be straightforward. The gene has already been cloned and sequenced (*10*). A convenient in vivo selection for mutants with DHFR activity can be devised with the use of trimethoprim as an antibiotic. And finally, the three-dimensional structure of *E. coli* DHFR has been thoroughly characterized by high-resolution x-ray crystallography, and a number of mechanistic hypotheses have been proposed (*4, 5*). It is important to recognize that our ability to formulate precise questions about the DHFR molecule and to design relevant mutagenesis experiments depends entirely on the availability of precise geometrical information provided by the x-ray structure of the molecule.

Certainly the most practical and efficient procedure for in vitro generation of specific mutations at a given locus is that which uses synthetic oligodeoxyribonucleotides as mutagens, as outlined by Itakura and Riggs (*11*). We report here on the application of this method to the construction of three mutant *E. coli* DHFR genes and on our preliminary investigation into the properties of the resulting mutant enzymes. The three mutations targeted the residues of aspartic acid at position 27, proline at position 39, and glycine at position 95, which are approximately located as shown by the lower, upper, and middle circles, respectively, in Fig. 1. They were designed to illuminate both the enzymic mechanism and certain aspects of chain folding and dynamics. In the following sections we describe the rationale underlying the design of these particular mutations.

Design of Mutations

The puzzle of Asp-27 is typical of many encountered in structural enzymology. A crucial question in this field is how an enzyme stabilizes the activated transition state complex for the reaction being catalyzed. In the molecular structures of DHFR we find Asp-27 buried below the enzyme surface, sequestered from external solvent, and poised, we believe, to form a hydrogen-bonded salt linkage with the pteridine ring of the substrate. In any event, the x-ray results clearly establish that Asp-27 does interact in this way with 2,4-diaminoheterocyclic inhibitors and that the pteridine ring of a substrate would occupy the same site, but turned upside down. Because the DHFR-catalyzed reaction involves hydride ion transfer between the nicotinamide's C-4 and the substrate's pteridine ring at C-6, one can assume that the transition state must have some degree of carbonium ion character at C-6, presumably induced by protonation at N-5. Thus the role of Asp-27 may well be to serve as the ultimate proton donor and to stabilize the transition state by providing a negatively charged carboxylate counterion to hydrogen bond with the resulting positively charged pteridine ring.

The difficulty with this simple picture, however, is that although protonation ought to occur at N-5, the x-ray structure clearly shows that the Asp-27 side chain would instead be hydrogen-bonded to N-3 and the 2-amino group. But whether or not its function has been correctly surmised, an aspartic acid residue is conserved at this position in all of the bacterial DHFR's of known sequence, and it is replaced by a glutamic acid residue at a structurally equivalent position in the vertebrate DHFR's.

The first mutagenesis experiment described here was designed to explore this paradox. By mutating Asp-27 to an asparagine, we expect to leave the catalytic-site geometry essentially undisturbed while simultaneously removing the side chain as a potential proton source and negatively charged counterion. Several kinds of questions come to mind. Will the mutant enzyme still function? Will it still bind 2,4-diaminoheterocyclic inhibitors upside down? Will it even fold up in the native configuration?

The second mutagenesis experiment is designed primarily to shed light on chain folding and molecular dynamics and only indirectly on mechanism. In this experiment, we attempt to introduce a new disulfide bridge into the DHFR molecule. We know of no report of a species of naturally occurring DHFR that contains even a single intrachain disulfide bridge.

The basis for deciding where to insert the new disulfide bridge was quite straightforward. Focusing on Pro-39, we find that this amino acid residue lies at the beginning of the extended strand beta B, opposite Cys-85 in another chain segment. If Pro-39 could be mutated to a cysteine, the latter might form a disulfide bridge with Cys-85 upon oxidation. Both residues are sufficiently far removed from the substrate or cofactor binding areas that minor chemical modifications would not be expected to interfere with the activity of the enzyme, and indeed reaction of Cys-85 with 5,5′-dithiobis(2-nitrobenzoic acid) does not (*12*). Neither residue is conserved in other species.

Simple modeling experiments show that Pro-39 and Cys-85 are relatively positioned so that the distance between their alpha carbons is 6.0 angstroms, about right for a disulfide bridge (*12*).

Either a left-handed or right-handed bridge should be possible geometrically, and no other groups are close enough to interfere with its formation.

In an attempt to see if there is any indication that disulfide bridges tend to occur between backbone chain segments with certain favorable geometries or relative orientations, we briefly examined a sample of 11 disulfide bridges in four highly refined protein structures deposited in the Brookhaven Protein Data Bank. All were different, and none of them resembled our proposed bridge at Cys-85.

What questions can we ask with this mutagenesis experiment? First, can a disulfide bridge actually be formed at a site chosen by us? Do we know enough about the geometrical requirements for disulfide formation to select a suitable site? Will such a bridge stabilize the molecule with respect to thermal denaturation? Will it alter the activity of the enzyme? One might conceive that the activity would increase because certain ineffective equilibrium conformations are now eliminated. On the other hand, the activity might decrease because increased molecular rigidity interferes with some motion required by the mechanism. Conceivably, both changes could occur but with different temperature dependencies.

To explain the rationale behind our third mutagenesis experiment we must first present some background information. It is known from the x-ray structures that all three DHFR molecules contain the same Gly-Gly pair linked by an unusual *cis* peptide bond (Gly-95–Gly-96 in the *E. coli* enzyme) and that an equivalent pair of glycine residues also occurs in all the other known DHFR sequences. Thus it is probable that this bit of topography plays some role in the working of the molecule, possibly as a conformational switch of some kind. That such a switch might be involved in initiating hydride transfer is suggested by the following observations. The Gly-Gly pair lies at the NH_2-terminus of the strand beta E, constituting a very short loop connecting beta E with alpha F. However, Ile-94, the residue preceding Gly-95, is included as part of beta E; its carbonyl oxygen is twisted out of the ordinary parallel beta sheet hydrogen-bonding pattern (a structural feature sometimes called a beta bulge) and is pointing instead toward the C-4 atom of the cofactor's nicotinamide ring. Arguments have been presented that this carbonyl oxygen is part of an array of oxygen atoms surrounding the nicotinamide ring which helps to activate it with respect to hydride transfer [see Filman *et al.* (*5*)]. Our idea was that if there is any validity to this notion, disrupting this peculiar bit of local geometry might prevent the molecule from functioning.

A minimal change that should nevertheless alter the geometry of this chain segment would be the mutation of the first glycine, Gly-95, to an alanine. Inspection of the structure shows that the beta carbon of an alanine at position 95 would be about 2.4 angstroms from the carbonyl oxygen of Ile-94, 2.8 angstroms from the alpha carbon of Gly-96, and 2.6 angstroms from C-5 of the cofactor's nicotinamide ring, contacts that are all at least 1 angstroms too close to be accommodated by the unperturbed structure. Exactly what would happen to the structure as a result of this substitution is unpredictable, but some change is almost certain to result.

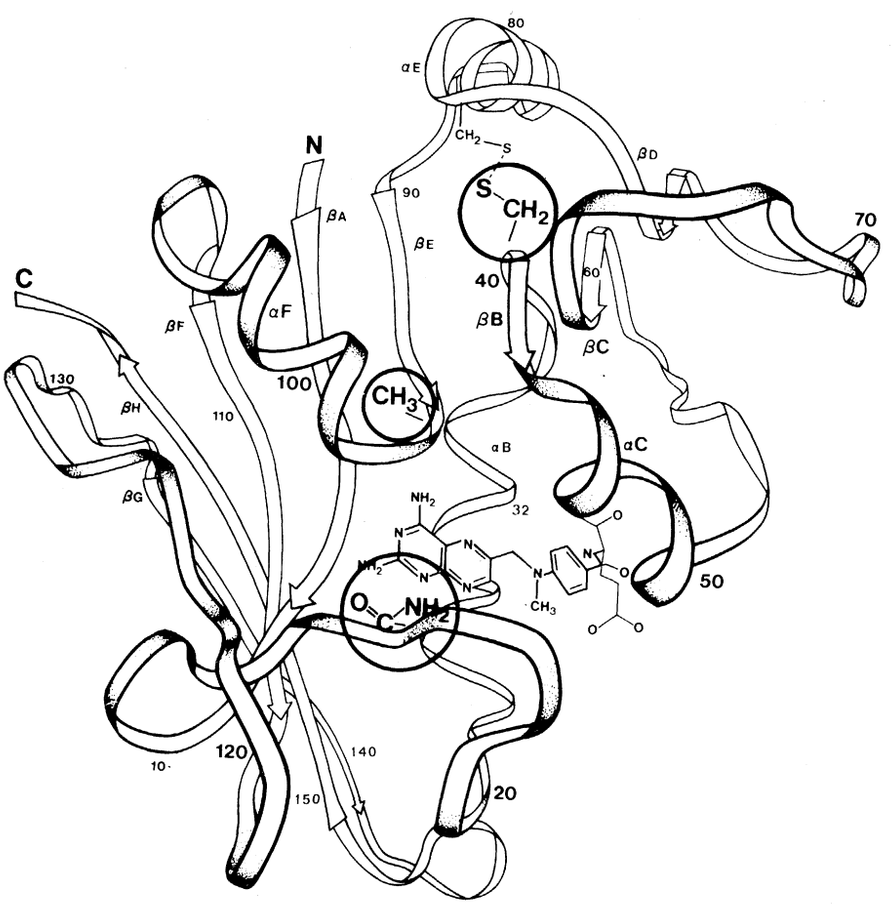

Fig. 1. Ribbon representation of the *E. coli* DHFR molecule showing the locations of three mutations made by oligonucleotide-directed mutagenesis. The upper circle shows the residue 39 mutation of proline to cysteine, the middle circle the residue 95 mutation of glycine to alanine, and the lower circle the residue 27 mutation of aspartic acid to asparagine. Beta strands (represented by arrows) and alpha helices are labeled. The approximate position of every tenth residue is indicated. A bound methotrexate molecule is depicted in the enzyme active site, but the cofactor NADPH is not.

Mutagenesis

Oligonucleotide-directed mutagenesis of the gene encoding *E. coli* DHFR was achieved with techniques similar to those described by Zoller and Smith (*7*). Restriction fragments (1.0 and 1.6 kilobases) from the plasmid pCV29 (*13*) containing the wild-type *E. coli* DHFR gene (*fol*) were cloned into bacteriophage M13mp8. Oligonucleotides designed to produce the desired mutations were then used as primers on the single-stranded phage DNA. Following extension and ligation by the large fragment of *E. coli* polymerase I and T4 DNA ligase, the resulting covalently closed heteroduplex DNA was isolated from agarose gels run in the presence of ethidium bromide. After transfection of *E. coli* strain JM103, plaques were picked and mutants identified by dot-blot hybridization (oligonucleotide hybridization with nitrocellulose-bound phage DNA) (*6, 7, 14*) with the corresponding ^{32}P-labeled mutagenic oligonucleotides, dissociation temperature differences being used to distinguish wild-type from mutant phage (see Fig.

2). Since at this stage a single plaque may contain a mixture of wild-type and mutant phage, dot-blot positives were plaque-purified to obtain homogeneous mutants. These were initially verified by detection of restriction site changes resulting from mutagenesis. The entire gene for each mutant DHFR was subsequently sequenced by the Sanger method (*15*) to ascertain that no changes other than those intended had been introduced.

Figure 3 shows the oligonucleotides used for DHFR mutagenesis with the corresponding region of the gene to which they anneal. In designing the oligonucleotides, we considered the following. (i) The primer should be sufficiently long that spurious priming is minimal under mutagenesis reaction conditions. (ii) The amino acid codon change should give a reasonable temperature difference (5 to 10 degrees) for dissociation of the mutagenic oligonucleotide from the wild-type and the mutant template (*14*). (iii) Since we anticipated cloning into plasmid vectors for enzyme overproduction, changes in restriction sites that could facilitate identification of mutants during subcloning should also be included. The restriction site changes did indeed prove to be quite helpful with respect to the foregoing, but predictions of the extent of spurious priming were not always reliable. In the mutagenesis of Gly-95 to alanine, for example, the mutagenic oligonucleotide was calculated to produce insignificant secondary priming. However, of three mutants that were positive by dot-blot hybridization, only one contained the desired sequence change. The other positives were mutations located,

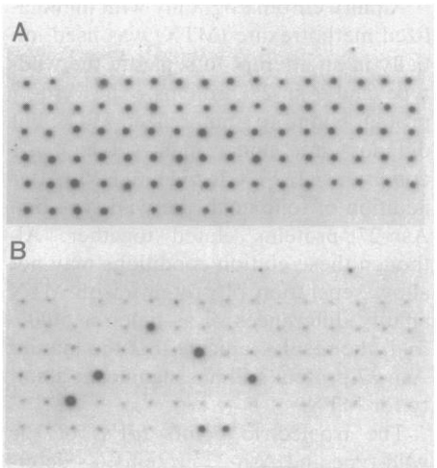

Fig. 2. Detection of residue 39 proline-to-cysteine mutants. (A) An autoradiogram resulting from hybridization of ^{32}P-labeled mutagenic oligonucleotide with single-stranded phage DNA blotted on a nitrocellulose filter at 20°C. (B) Autoradiogram after the filter was washed at 58°C for 10 minutes (*6, 7, 14*).

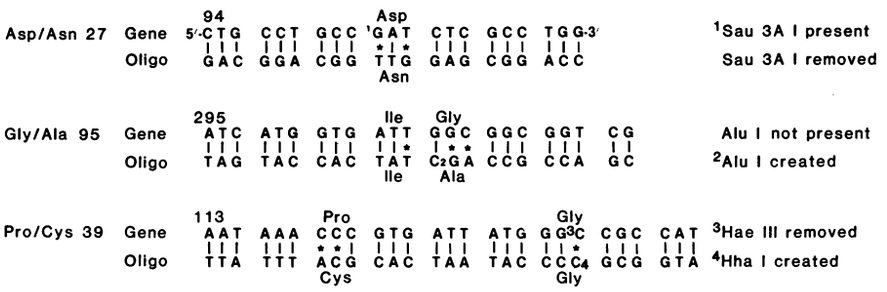

Fig. 3. Mutagenic oligonucleotides and their complementary *E. coli* DHFR gene sequences. Asterisks denote mismatches between the gene and the oligonucleotide. Superscripts and subscript numbers mark sites of cleavage by restriction enzymes indicated at the right of the sequence. Gene sequence numbering is as indicated in (*13*). The oligonucleotides were synthesized by the phosphotriester method (*26*).

in one case, outside the DHFR gene, and in the other case at the correct position but resulting in a 186-base deletion of the 3′ end of the structural gene. Similar problems were observed with a 17-base oligonucleotide that was used initially for mutagenesis of Asp-27 to asparagine but was later discarded in favor of the 21-base oligonucleotide shown in Fig. 3. We found that testing a candidate oligonucleotide as a primer for dideoxy sequencing provided a good prediction of its usefulness for mutagenesis. Oligonucleotides that gave a mixed pattern on a sequencing gel, indicating multiple priming sites, were less likely to produce the desired mutation.

With the exception of the mutation of Asp-27 to asparagine, mutagenesis of DHFR achieved efficiencies between 5 and 30 percent, comparable to those reported by others (*6, 7*). The low frequency of 0.3 percent observed for the Asp-27 mutation is consistent with the notion that DNA repair of M13 in *E. coli* occurs at or near asymmetrically A-methylated GATC sites (G, guanine; A, adenine; T, thymine; C, cytosine) (*16*). It has been shown that this repair results in marker recoveries having a strong bias in favor of the methylated strand, which is the wild-type strand in our case. In Fig. 3, the oligonucleotide mismatch for the Asp-27 mutation occurs precisely at a GATC sequence. The low frequency of this mutation suggests that repair occurs right at GATC sites and that even mismatched GATC sites are recognized.

Since we planned to study structure-function correlations in the mutant DHFR's in detail, we believed it was necessary to sequence the entire DHFR gene for each mutant to verify that only the intended mutation was present (Fig. 4). This precaution seemed to be especially prudent when we discovered that our first mutation of Pro-39 to cysteine did, in fact, contain an additional unintended point mutation changing Trp-30 to a serine. The Trp-30 mutation was already present in the single-stranded template originally used for in vitro extension.

DHFR Gene Expression

In addition to the DHFR gene, the plasmid pCV29 was shown by sequence analysis to contain a point mutation in the DHFR promoter that results in overproduction of the enzyme (*13*). The presence of this mutant promoter can be monitored by the loss of a Sal I restriction site. However, even with the mutant promoter, we observed a low level (less than 1 percent of the total protein) of DHFR production in cells persistently infected with M13mp8 containing *fol*. This contrasts with the high level of expression of the *Bacillus stearothermophilus* tyrosyl transfer RNA synthetase gene cloned into M13, where the synthetase protein level approaches 50 percent of total soluble protein (*9*). The reasons for this difference are not yet known.

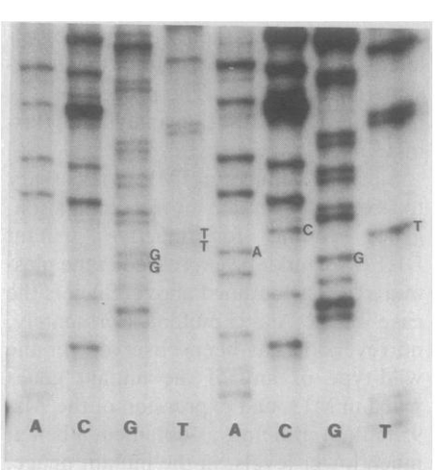

Fig. 4. Autoradiogram of dideoxy sequencing gel comparing the wild-type *E. coli* DHFR gene sequence with the Asn-27 mutant gene sequence. The sequence shown is the complementary strand. The codon change (reading from top to bottom) is GAT (wild type; Asp) to AAC (Asn).

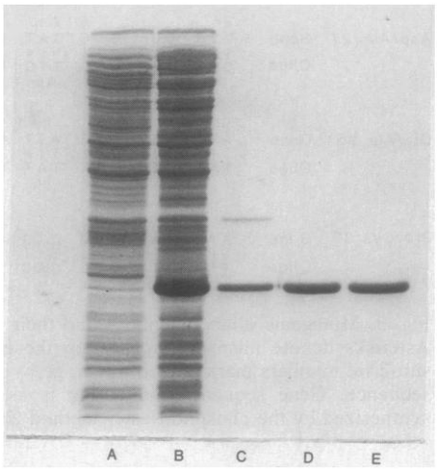

Fig. 5. Asn-27 mutant protein overproduction and purification visualized by Coomassie blue staining of SDS-PAGE. Overproduction of the mutant protein can be seen by comparing the lysates of *E. coli* containing pUC8 with (lane B) and without (lane A) the DHFR gene insert. In the purification of DHFR, 1 to 3 liters of cells in early stationary phase are lysed, either by sonication or Brij treatment (*27*). The lysate (lane B) was then applied to a 2.5 by 200 cm Sephadex G-75 column. The major peak following the void volume corresponds to DHFR (lane C). This step typically resulted in an initial purification of 50-fold (≥ 90 percent pure). Application of the pooled protein peak to an aminohexylagarose column (1 by 18.5 cm) and elution with a 0 to 0.6*M* KCl gradient futher purified the DHFR about threefold (lane D) (*25*). The last step in the purification was elution from a DEAE-Sephadex column (1.5 by 38 cm) with a 0 to 0.7*M* KCl gradient (lane E). The various mutant DHFR proteins were homogeneous as determined by PAGE. A typical yield when the pUC8 vector was used for expression was 8 mg of protein per liter, whereas purification of Ala-95 mutant expressed in M13mp8 resulted in < 0.3 mg of protein per liter. The bottom of the gel is the anode end.

Increased production of the wild-type DHFR, as well as the Asn-27 and Cys-39 mutant proteins, was achieved by cloning a Bam HI fragment containing the DHFR gene from M13mp8 into the plasmid pUC8 (*17*) maintained in *E. coli* strain SK383. However, we observed that plasmids that contained the mutant promoter and that were purified from cultures not kept under the selective pressure of trimethoprim (TMP) gave a restriction pattern indicating that partial reversion to the wild-type *fol* gene had occurred, presumably as a result of recombination between the plasmid and the chromosomal genes. This reversion seems to be related to the presence of the mutant promoter in *fol*, since it is not observed when the wild-type promoter is present. The problem was circumvented by growing starter cultures in the presence of TMP and then washing the pelleted cells before inoculating the large-scale growth media. The procedure was monitored by restriction analysis of plasmid DNA.

Since resistance to TMP is necessary to prevent reversion by recombination, genes coding for inactive mutant DHFR's cannot be expressed in the plasmid pUC8 as outlined above. Such is the case for the Ala-95 mutant. Fortunately, no reversion has been observed for the wild-type or any of the mutant genes when in M13, and expression of the Ala-95 mutant protein was obtained, albeit at much lower levels, in the mutant phage.

From the foregoing, it is evident that the degree of TMP resistance of cells carrying a given mutant DHFR gene provides a useful, preliminary in vivo assay for mutant enzyme activity. For example, at TMP concentrations of 10 μg/ml, we noticed that the growth rate of cells containing the Asn-27 plasmid was markedly decreased with respect to the growth rate of cells containing the wild-type plasmid or Cys-39 plasmid. This observation suggested that the Asn-27 mutant enzyme had reduced activity. Further, our inability to maintain the Ala-95 mutation in pUC8 suggested that the Ala-95 protein is inactive.

Protein Purification and Initial Characterization

Mutant protein production was detected by comparison of cellular lysates by sodium dodecyl sulfate (SDS)–reducing polyacrylamide gel electrophoresis (PAGE). The control lysate was obtained from cells containing the plasmid or phage without the DHFR gene insert. A prominent new protein band in the lysate from cells containing the mutant DHFR gene was presumed to be the mutant protein (see lanes A and B in Fig. 5).

Since our mutagenesis experiments are expected to generate some totally inactive DHFR's, we developed a purification procedure that is not dependent on enzyme binding of inhibitors or substrates but rather on general physical properties of the protein. The purification process was monitored by PAGE rather than by activity assays; this again allows the purification of inactive proteins. Details of the purification procedure are described in Fig. 5.

As a preliminary characterization, purified proteins were examined by nondenaturing PAGE, in which mobility depends on the shape and charge of the proteins. In addition, the nondenaturing gel system allows in situ enzyme assays by staining for activity (*18*). Thus we were able to observe, at least qualitatively, any gross differences in physical properties and catalytic activities resulting from the various mutations.

For mutants with very low enzyme activity, contamination with wild-type DHFR of chromosomal or recombined plasmid origin (or both) was a serious concern, and additional separation methods were used to further purify these mutant enzymes. In addition, in the case of the Asn-27 protein, the possibility of posttranslational deamidation (to aspartic acid) had to be considered. The degree of contamination with wild-type DHFR was monitored by activity stains on nondenaturing gels.

Asn-27 mutant protein. Nondenaturing gel electrophoresis of the Asn-27 protein shows that the mutant protein did not migrate as fast as the wild-type DHFR (lane D in Fig. 6). The lower electrophoretic mobility of the Asn-27 protein is consistent with the anticipated alteration in charge (Asp-27 → asparagine). Activity stains of the Asn-27 protein on a nondenaturing gel showed bands in both the wild-type and the mutant positions. This observation of distinct and separate bands on an activity-stained electrophoretic gel clearly shows that the mutant protein has its own intrinsic enzymic activity, although at a much lower level than the wild-type DHFR. In the isolated Asn-27 mutant protein preparation we found that, although wild-type contamination was less than 1 percent of the total protein, it accounted for more than 90 percent of the total activity.

Affinity chromatography with immobilized methotrexate (MTX) was used initially in an attempt to separate the wild-type and Asn-27 proteins. The Asn-27 protein bound to immobilized MTX, but did not elute under high salt–high *p*H conditions (1*M* KCl, *p*H 9.0) (*19*). After addition of folic acid, the wild-type and Asn-27 proteins eluted together. Although these elution conditions may not allow separation of proteins with MTX affinity differences of as much as 1000-fold, the results indicate that the mutant Asn-27 protein still has significant affinity for MTX.

The isoelectric points (*p*I's) of the wild-type and Asn-27 proteins, as determined by marker analysis (*20*) in analytical scale polyacrylamide tube isoelectric focusing gels, were found to be 4.5 and 4.8, respectively. This large a difference in *p*I values makes the wild-type and Asn-27 proteins easily separable on a

preparative scale by granulated bed isoelectric focusing when a *p*H gradient of short range (4 to 6) is used. A peak in the enzyme activity profile corresponds to the *p*I value of the mutant Asn-27 protein; no activity was detected at the *p*I value corresponding to the wild-type DHFR. This experiment again shows that the Asn-27 protein has its own intrinsic activity. The observation that the *p*I of the mutant Asn-27 protein is slightly higher than that of the wild type is in accord with the mutation of Asp-27 to asparagine. These *p*I values are comparable to the value of 4.6 previously reported for the RT500 form 1 of the *E. coli* DHFR (*21*).

Cys-39 mutant protein. As expected, the purified Cys-39 mutant protein was found to contain a total of three sulfhydryls (versus two in the wild type) by both Ellman's titration of the fully reduced protein (*22*) and by cysteic acid determination (*23*).

Nondenaturing gel electrophoresis of the Cys-39 protein in the absence of reducing agents showed that the mutant DHFR was present in several forms. After reduction with either dithiothreitol (10 m*M*) or 2-mercaptoethanol (100 m*M*), only one form remained; this indicates that the multiplicity of forms is probably due to various inter- and intramolecular disulfide bridgings and other oxidation states of the cysteine residues. Since the Cys-39 protein was purified under nonreducing conditions to prevent reduction of the anticipated Cys-39–Cys-85 disulfide, it is likely that the protein underwent cysteine oxidation to the sulfenic and disulfide states, both of which are readily reduced by thiols.

Ellman's titrations of the nonreduced protein showed that fewer than one free sulfhydryl was present. This result indicates that most of the cysteines have been oxidized and, although it seemed likely that one form would have the desired Cys-39–Cys-85 intrachain disulfide, we have not yet shown definitively that it has actually been formed.

Activity staining of the Cys-39 protein in nondenaturing gels revealed that all the DHFR species were active, with the fully reduced form showing the highest activity. Indeed, a comparison of the activities of nonreduced versus reduced Cys-39 mutants showed that activity almost doubled on complete reduction of the protein.

Ala-95 mutant protein. Nondenaturing gel electrophoresis of the Ala-95 DHFR shows that the mutant protein does not migrate as fast as the wild-type protein (lane D in Fig. 6). Since an alteration in charge is not anticipated for the Ala-95 DHFR (glycine → alanine), the different band position may be due to a conformational change. Activity staining of the Ala-95 protein in a nondenaturing gel shows a band in the wild-type position only. This suggests that the residual activity associated with the Ala-95 preparation is due to copurifying wild-type DHFR.

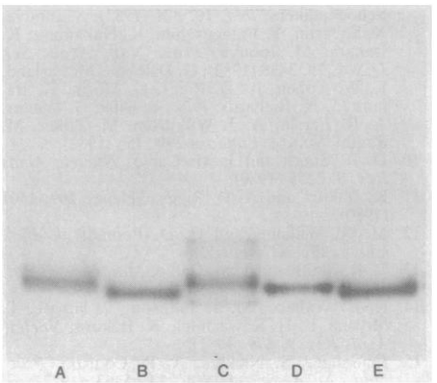

Fig. 6. A nondenaturing electrophoretic gel of wild-type and mutant DHFR proteins visualized by Coomassie blue staining. Lanes B and E are wild-type protein, lane A is the Cys-39 protein, lane C is the Ala-95 protein, and lane D is the Asn-27 protein. The bottom of the gel is the anode end.

Specific activities. Specific activities of the various mutant DHFR proteins and of the wild-type protein were determined at *p*H 7.0 and room temperature by the procedure of Baccanari and Joyner (*24*). The specific activities are 50 units per milligram of protein for both the wild-type and fully reduced Cys-39 and 0.05 unit per milligram of protein for the Asn-27. For these specific activity measurements, copurifying wild-type DHFR was removed from the Asn-27 protein only and not from the Cys-39 preparation, since the latter has full activity and the wild type constitutes less than 1 percent of the total protein. As a check on our activity measurement procedure, we found that the specific activity for the wild-type DHFR is comparable to the specific activity for the RT500 form 1 *E. coli* enzyme reported by Baccanari *et al.* (*21*).

Discussion

Using oligonucleotide-directed mutagenesis, we generated three different mutant DHFR proteins. Initial characterization of these mutant proteins indicates that substantial changes in protein character are associated with each modification.

For the Asn-27 protein, a low specific activity (0.1 percent of that of the wild type) was seen. This observation strongly supports the postulated role of the Asp-27 side chain in catalysis. Nevertheless, that there remains some residual activity suggests that the Asp-27 residue (wild type) is not the sole factor involved in stabilizing the protonated transition state.

The Cys-39 mutant shows that a cysteine at position 39 does not seem to affect enzymic activity of DHFR, since the mutant is fully active in the reduced form. Our preliminary data suggest, however, that activity is significantly diminished on oxidation of the protein, and although we have not unequivocally established the formation of the Cys-39–Cys-85 intrachain disulfide, its existence seems likely since oxidation of wild-type Cys-85 and Cys-152 with Ellman's reagent has no effect on enzymic activity (*12*). The diminished activity of the oxidized Cys-39 protein could result, then, from a loss of some dynamic flexibility in the molecule when the alpha E helix is tethered by a disulfide to the beta B strand, but a structural explanation for this hypothesis is not readily apparent from inspection of the DHFR structure.

Our preliminary tests indicate that substitition of an alanine for Gly-95 causes complete inactivation of DHFR. Also the lower mobility of the Ala-95 protein on nondenaturing gels suggests that a change in conformation has occurred. These two results are consistent with our expectation that an alanine at position 95 would disturb the conformation of the Gly-95–Gly-96 *cis* peptide and thereby affect the activation of the nicotinamide ring of NADPH. Extensive changes in structure have probably not occurred since the behavior of the Ala-95 protein on an aminohexylagarose column (*25*) during purification was identical to that of the wild-type DHFR, an indication that the general DHFR protein fold has been retained in the Ala-95 protein.

The above experiments establish that the technique of directed mutagenesis, in conjunction with accurate three-dimensional x-ray crystal structure determinations, constitutes a powerful tool for investigating relations between structure and function in enzymes. This approach might be of value in inducing useful functional changes in enzyme molecules.

References and Notes

1. The abbreviations utilized are dihydrofolate reductase, DHFR; trimethoprim, TMP; sodium dodecyl sulfate–reducing polyacrylamide gel electrophoresis, SDS-PAGE; methotrexate, MTX; aspartic acid, Asp; asparagine, Asn; proline, Pro; cysteine, Cys; glycine, Gly; alanine, Ala; Ile, isoleucine; and Trp, tryptophan.
2. R. L. Blakley, *The Biochemistry of Folic Acid and Related Pteridines* (North-Holland, Am-

sterdam, 1969); M. Friedkin, *Adv. Enzymol.* **38**, 37 (1973); J. E. Gready, *Adv. Pharmacol. Chemother.* **17**, 37 (1980).
3. F. M. Huennekens, K. S. Vitols, J. M. Whitely, V. G. Neef, *Methods Cancer Res.* **13**, 199 (1976); J. J. McCormack, *Med. Res. Rev.* **1**, 303 (1980); B. Roth and C. C. Cheng, *Prog. Med. Chem.* **19**, 269 (1982).
4. D. A. Matthews *et al.*, *Science* **197**, 452 (1977); D. A. Matthews *et al.*, *J. Biol. Chem.* **253**, 6946 (1978); D. A. Matthews, R. A. Alden, S. T. Freer, N. Xuong, J. Kraut, *ibid.* **254**, 4144 (1979); K. W. Volz *et al.*, *ibid.* **257**, 2528 (1982); J. T. Bolin, D. J. Filman, D. A. Matthews, R. C. Hamlin, J. Kraut, *ibid.*, p. 13650.
5. D. J. Filman, J. T. Bolin, D. A. Matthews, J. Kraut, *J. Biol. Chem.* **257**, 13663 (1982).
6. D. Shortel, D. DiMaio, D. Nathans, *Annu. Rev. Genet.* **15**, 265 (1981).
7. M. J. Zoller and M. Smith, *Nucleic Acids Res.* **10**, 6487 (1982); M. Smith, *Trends Biochem. Sci.* (December 1982), p. 440.
8. R. B. Wallace, P. F. Johnson, S. Tanaka, M. Schold, K. Itakura, J. Abelson, *Science* **209**, 1396 (1980).
9. R. B. Wallace, M. Schold, M. J. Johnson, P. Dembek, K. Itakura, *Nucleic Acids Res.* **9**, 3647 (1981); G. F. M. Simons, G. H. Veeneman, R. N. H. Konings, J. H. Van Boom, J. G. G. Schoenmakers, *ibid.* **10**, 821 (1982); S. Inouye, X. Soberon, T. Franceschini, K. Nakamura, K. Itakura, M. Inouye, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3438 (1982); G. Dalbadie-McFarland, L. W. Cohen, A. D. Riggs, C. Morin, K. Itakura, J. H. Richards, *ibid.*, p. 6409; G. Winter, A. R. Fersht, A. J. Wilkinson, M. Zoller, M. Smith, *Nature (London)* **299**, 756 (1982).
10. D. R. Smith and J. M. Calvo, *Nucleic Acids Res.* **8**, 2255 (1980).
11. K. Itakura and A. D. Riggs, *Science* **209**, 1401 (1980).
12. M. N. Williams and C. D. Bennett, *J. Biol. Chem.* **252**, 6871 (1977).
13. D. R. Smith and J. M. Calvo, *Mol. Gen. Genet.* **187**, 72 (1982).
14. R. B. Wallace, M. J. Johnson, T. Hirose, T. Miyake, E. H. Kawashima, K. Itakura, *Nucleic Acids Res.* **9**, 879 (1981).
15. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977); F. Sanger, A. R. Coulson, B. G. Barrell, A. J. H. Smith, B. A. Roe, *J. Mol. Biol.* **143**, 161 (1980).
16. R. Wagner and M. Meselson, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 4135 (1976); W. Kramer, K. Schughart, H. J. Fritz, *Nucleic Acids Res.* **10**, 6475 (1982); A. L. Lu, S. Clark, P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4639 (1983).
17. J. Messing and J. Vierra, *Gene* **19**, 259 (1982).
18. M. Hiebert, J. Gauldie, B. L. Hillcoat, *Anal. Biochem.* **46**, 433 (1972).
19. D. P. Baccanari, D. Stone, L. Kuyper, *J. Biol. Chem.* **256**, 1738 (1981).
20. T. Laas, I. Olsson, L. Soderberg, *Anal. Biochem.* **101**, 449 (1980).
21. D. P. Baccanari, D. Averett, C. Briggs, J. Burchall, *Biochemistry* **16**, 3566 (1979).
22. G. L. Ellman, *Arch. Biochem. Biophys.* **82**, 70 (1959).
23. S. Moore, *J. Biol. Chem.* **238**, 235 (1963).
24. D. P. Baccanari and S. S. Joyner, *Biochemistry* **20**, 1710 (1981).
25. S. S. Joyner and D. P. Baccanari, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **41**, 1283 (1982).
26. K. Miyoshi, T. Huang, K. Itakura, *Nucleic Acids Res.* **8**, 5491 (1980).
27. D. Baccanari, A. Phillips, S. Smith, D. Sinski, J. Burchall, *Biochemistry* **14**, 5267 (1975).
28. We thank R. Matthews for providing the *E. coli* strain CV634, which contained the plasmid pCV29; P. Price for the cysteic acid determinations; and R. Aust, E. Dyke, F. Lopez, and M. E. Ford for technical assistance. Supported in part by contract N00014-88-C-1482 (to J.K.) from the Office of Naval Research, Public Health Service fellowship F32 GM09375 (to E.E.H.), and NIH grants GM10928 and CA17374 (to J.K.).

Surface Molecules Identify Groups of Growing Axons

R. D. G. McKay, S. Hockfield, J. Johansen
I. Thompson, K. Frederiksen

Summary. Studies on vertebrate and invertebrate species have established that, during development, axons have the ability to choose particular paths over others. The chemical basis of this pathfinding is not clear but biochemical differences between neurons have long been postulated to account for the specificity of neuronal connections. Such subtle molecular differences between different cells in a single tissue are difficult to study with standard biochemical techniques but hybridoma technology has offered a potential solution to this type of problem. This technique has made possible the production of monoclonal antibodies for identifying and characterizing a family of glycoproteins which are expressed on the surface of specific axon bundles during the development of the leech nervous system. The results show that groups of growing axons do indeed carry chemically distinct surface molecules.

In contrast to those of vertebrates the nervous systems of many invertebrates contain relatively few neurons. In the leech, for example, the central nervous system is made up of a total of 10^4 neurons, but these are arranged in repeating similar ganglia, each composed of only 400 neurons. Within a ganglion many of the neurons are bilaterally symmetrical so the basic unit of the leech nervous system is a 200 cell half ganglion. This anatomical simplicity of the organization of neuronal cell bodies in the leech offers a great advantage as an experimental system (*1*). We made use of this anatomical simplicity to screen a large number of monoclonal antibodies, raised against the dissected nerve cord, for the ability to recognize antigens present in subsets of neurons (*2*). The initial fusions generated several hundred hybridoma cell lines.

Using a whole mount preparation of the leech ganglion as our immunohistochemical assay, we found that approximately 10 percent of these lines secreted antibody that bound to subsets of neurons. These monoclonal antibodies, which bound to subsets of neurons, often bound to overlapping subsets; but in no cases in our initial studies did we find two antibodies which recognized the same subset of neurons. As we saw more than 30 different staining patterns without a repeat, there is a high statistical probability that there are more than 200 differentially distributed antigens in the leech nerve cord. These data suggest that the nervous system is made up of a complex series of chemically differentiated cell types; they are consistent with measurements of a high degree of neuronal complexity obtained by different methods in vertebrate nervous systems (*3, 4*). These data are also consistent with models where synaptic networks and other differentiated physiological functions and anatomical features of the nervous system result from molecular differences between neurons.

Another important consequence of the production of antibodies to antigens in subsets of neurons is that these antibodies have allowed us to see surprising new features in the cellular organization of the leech nervous system and that of vertebrate (*5–7*). In the leech, we have made use of the simple anatomy of its central nerve cord to show that axons are organized into stereotyped groups (*5*). The leech central nervous system consists of a chain of segmental ganglia. Large bundles of axons, the connectives,

The authors work at the Cold Spring Harbor Laboratory, Post Office Box 100, Cold Spring Harbor, New York 11724. The present address of I. Thompson is the Department of Physiology, South Parks Road, Oxford, United Kingdom.