Stoner, G. L. (1984) *J. Neurochem. 43*, 433–447.
Stryer, L. (1965) *J. Mol. Biol. 13*, 482–495.
Stryer, L. (1978) *Annu. Rev. Biochem. 47*, 819–846.
Vacher, M., Nicot, C., Pflumm, M., Luchins, J., Beychok, S., & Waks, M. (1984) *Arch. Biochem. Biophys. 231*, 86–94.
Weber, G., & Daniel, E. (1966) *Biochemistry 5*, 1900–1907.
Weber, L. D., Tulinsky, A., Johnson, J. D., & El-Bayoumi, M. A. (1979) *Biochemistry 18*, 1297–1303.
York, S. S., Lawson, R. C., & Worah, D. M. (1978) *Biochemistry 17*, 4480–4486.

An Engineered Disulfide Bond in Dihydrofolate Reductase†

Jesus E. Villafranca,*,‡ Elizabeth E. Howell,‡ Stuart J. Oatley,§ Nguyen-huu Xuong,§ and Joseph Kraut§

The Agouron Institute, La Jolla, California 92037, and Department of Chemistry, University of California, San Diego, La Jolla, California 92093

Received August 18, 1986; Revised Manuscript Received December 12, 1986

ABSTRACT: Substitution of cysteine for proline-39 in *Escherichia coli* dihydrofolate reductase by oligonucleotide-directed mutagenesis positions the new cysteine adjacent to already existing cysteine-85. When the mutant protein is expressed in the *E. coli* cytosol, the cysteine sulfur atoms are found, by X-ray crystallographic analysis, to be in van der Waals contact but not covalently bonded to one another. In vitro oxidation by dithionitrobenzoate results in formation of a disulfide bond between residues 39 and 85 with a geometry close to that of the commonly observed left-handed spiral. Comparison of 2.0-Å-refined crystal structures of the oxidized (cross-linked) and reduced (un-cross-linked) forms of the mutant enzyme shows that the conformation of the enzyme molecule was not appreciably affected by formation of the disulfide bond but that details of the molecule's thermal motion were altered. The disulfide-cross-linked enzyme is at least 1.8 kcal/mol more stable with respect to unfolding, as measured by guanidine hydrochloride denaturation, than either the wild-type or the reduced (un-cross-linked) mutant enzyme. Nevertheless, the cross-linked form is *not* more resistant to thermal denaturation. Moreover, the appearance of intermediates in the guanidine hydrochloride denaturation profile and urea-gradient polyacrylamide gels indicates that the folding/unfolding pathway of the disulfide-cross-linked enzyme has changed significantly.

Disulfide bonds are a common feature of many extracellular proteins, where they presumably serve to stabilize the native conformation by lowering the entropy of the unfolded form (Anfinsen & Scheraga, 1975). This stabilizing property has, prospectively, made disulfide bond cross-linking an attractive strategy for engineering additional conformational stability into proteins by site-directed mutagenesis (Villafranca et al., 1983; Perry & Wetzel, 1986). However, criteria for selecting the appropriate locus for disulfide cross-linking in a given protein have not been firmly established. Detailed analyses of disulfide bonds in proteins of known high-resolution X-ray structure have shown that, although many disulfides can be described simply as left-handed spirals or right-handed hooks, disulfide bridges in fact adopt a wide range of conformations and cross-link most kinds of secondary structure present in proteins (Richardson, 1981; Thornton, 1981). Thus, the set of geometric parameters on which one might base an effective disulfide cross-linking strategy is rather weakly constrained.

Another pertinent consideration for a disulfide-based stabilization strategy is the mechanism of formation of disulfide bonds. Disulfide-cross-linked cystosolic proteins are extremely rare, presumably because the intracellular environment is too reducing to allow cysteine oxidation to the disulfide state (Ziegler & Poulsen, 1977; Creighton et al., 1980). In those cases where cystosolic proteins do contain disulfide bonds, the proteins themselves are able to catalyze their formation (i.e., glutathione reductase and thioredoxin reductase). In vitro studies have shown that disulfide bond formation occurs under oxidizing conditions through disulfide interchange reactions that are driven to the final, correctly cross-linked species by protein folding (Haber & Anfinsen, 1962; Creighton, 1978). In vivo, this interchange reaction may be catalyzed by an enzyme, protein disulfide-isomerase, during the secretory posttranslational translocation process (Freedman & Hillson, 1980). Thus it appears that in vivo secretion may be required to produce a disulfide-cross-linked engineered protein without an additional chemical oxidation step.

To explore these concerns, we have engineered a disulfide bridge in the cystosolic enzyme dihydrofolate reductase (DHFR,[1] EC 1.5.1.3) from *Escherichia coli*. A mutation, proline-39 → cysteine, produced by site-directed mutagenesis, was designed to form an intramolecular disulfide bond between Cys-85 and Cys-39, thus cross-linking the N-terminus of α-helix E to the start of β-strand B (Villafranca et al., 1983). We report here on the effect of this disulfide cross-link on the mutant enzyme's structural stability and compare the high-resolution X-ray structures of the reduced (un-cross-linked) and oxidized (cross-linked) forms of this mutant enzyme, which we refer to as Cys-39 DHFR. Our results indicate that the disulfide bridge does not form spontaneously in the cytosol of

† This work was supported by Office of Naval Research Contract N0014-85-K-0663 and National Institutes of Health Contract GM-10928, both awarded to J.K.

* Author to whom correspondence should be addressed.
‡ The Agouron Institute.
§ University of California, San Diego.

[1] Abbreviations: DHFR, dihydrofolate reductase; DTT, dithiothreitol; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); EDTA, ethylenediaminetetraacetate; Gdn-HCl, guanidine hydrochloride.

0006-2960/87/0426-2182$01.50/0 © 1987 American Chemical Society

E. coli. Once formed in vitro, however, it enhances structural stability but reduces cooperativity of folding.

Materials and Methods

Oligonucleotide-directed mutagenesis of the cloned wild-type *E. coli fol* gene, as well as overexpression and purification of the wild-type and mutant enzymes, has been previously described (Villafranca et al., 1983). Enzyme concentrations were determined by the biuret method (Itzhaki & Gill, 1964) and by methotrexate titration (Huennekens et al., 1976). Quantitative determination of cysteines was carried out by the procedure of Ellman (1959).

Unfolding of DHFR by guanidine hydrochloride or by elevated temperature was monitored by UV absorbance changes at 292 nm in a Perkin-Elmer Lambda 3A spectrophotometer. Split cuvettes were used in the Gdn-HCl experiments. The reference cell contained the enzyme (0.4 mg/mL) in one compartment and the appropriate concentration of Gdn-HCl in the other, while the sample cell contained a mixture of enzyme and Gdn-HCl at the same concentrations as in the reference cell. After being mixed, the samples were allowed to equilibrate for 30 min prior to recording difference spectra. No changes in the spectra were observed after 30 min. The cuvettes were thermostated at 25 °C throughout the experiment by a recirculating water bath. Reversibility was established by recovery of both enzyme activity and UV absorbance.

For the thermal unfolding experiments, standard cells were used with the reference cell containing only buffer (0.1 M imidazole, 1 mM EDTA, pH 9.0). The sample cell contained enzyme at a concentration of 0.75 mg/mL. Absorbance readings were taken after 10 min at a variety of temperatures.

Quantitative yields (>95%) of disulfide-cross-linked (39 to 85) Cys-39 DHFR were obtained as follows: a solution of 2.5 mg/mL Cys-39 DHFR in 0.1 M Tris-HCl, 1 mM EDTA, and 10 mM DTT was made 20 mM in dithionitrobenzoate. After 2 h, this solution was passed over a Sephadex G-25 column equilibrated in the same buffer. Fractions containing protein were made 1 mM in DTT and, after 30 min , passed over the same Sephadex G-25 column. All procedures were performed at 4 °C. The number of free sulfhydryls in the protein fractions was determined by Ellman's method. Urea-gradient polyacrylamide gel electrophoresis was performed as described by Creighton (1979).

Crystals of Cys-39 DHFR in both oxidized and reduced forms were obtained under conditions almost identical with those for MB1428 DHFR (Bolin et al., 1982). The new crystals were isomorphous with the MB1428 crystals to within 0.5% in lattice parameters (space group $P6_1$, two molecules per asymmetric unit). X-ray data were collected to 2.0-Å resolution (1.9 Å for the wild type) on the Xuong–Hamlin Multiwire area detector/diffractometer (Xuong et al., 1978) from a single crystal in each case. R_{sym} values for each data set were <0.06. Initial atomic coordinates were readily obtained by model fitting into difference Fourier maps. In the case of the wild type, for which a glutamic acid was modeled into residue 154 (lysine in MB1428), phases from the refined MB1428 DHFR structure were used, whereas phases from the subsequently refined wild-type structure were applied for calculation of the mutant difference maps. The programs of Hendrickson and Konnert (1980) were used for least-squares refinement of atomic coordinates. After 10 cycles of restrained-parameter refinement, models for both the oxidized and the reduced forms of the Cys-39 enzyme were obtained that gave R factors of 0.18 and root mean square deviation of bond lengths from their dictionary values of 0.01 Å.

Results and Discussion

Sulfhydryls of Cys-39 DHFR: Reduced Form. We have already published a report on the mutagenesis and purification of the Cys-39 DHFR mutant (Villafranca, 1983). It now appears that in our initial purifications extensive air oxidation of the cysteines had occurred during the purification procedure. Nonreducing native polyacrylamide gels showed as many as four distinct bands corresponding to various oxidation states of the cysteines in Cys-39 DHFR. Quantitative determination of free cysteine according to Ellman's procedure showed that less than one of the three cysteines per protein molecule remained in the sulfhydryl form. It seemed probable that trace metals in our buffers were catalyzing the air oxidation of the cysteines, all of which are apparently quite labile. Ordinarily this problem can be overcome by including reducing agents in the purification buffers. In the present case, however, use of reducing agents could result in reduction of the desired Cys-39–Cys-85 disulfide bond. Reducing agents were therefore omitted. Instead, the problem was solved by simply decreasing the air exposure of our protein samples and including the chelating agent ethylenediaminetetraacetate (EDTA, 1 mM) in our buffer solutions. The revised purification conditions yielded fully reduced Cys-39 DHFR, with all three cysteines, Cys-39, Cys-85, and Cys-152, in the sulfhydryl form. We therefore conclude that a disulfide bond between Cys-39 and Cys-85 in this mutant DHFR does not form in the *E. coli* cystosol.

This finding complements a growing body of evidence indicating that disulfide bond formation in proteins is associated with transport across cellular membranes. In T_4 lysozyme, which is expressed in the *E. coli* cytosol, an engineered disulfide cross-link required oxidation in vitro to form (Perry & Wetzel, 1986). On the other hand, similarly engineered disulfide bonds in the secreted enzyme subtilisin are formed in vivo (Wells & Powers, 1986). In another study, mutations in the signal sequence of bacterial alkaline phosphatase have been described that interfere with the export of this enzyme (Michaelis et al., 1983). The absence of enzyme activity in this internalized alkaline phosphatase is thought to be due to improper folding caused by disruption of the intrachain disulfides. Improper folding may also explain the absence of enzymic activity in rat trypsinogen when it is expressed in the *E. coli* cytosol (Craik, personal communication).

Oxidation of Cys-39 DHFR. Model-building studies suggested that no significant conformational changes would be caused by the cysteine for proline substitution at position 39. The mobility of the mutant enzyme on native polyacrylamide gels and its behavior on a sieving (Sephadex G-75) column supported this notion. From these modeling studies, it appeared also that the sulfhydryls of Cys-39 and Cys-85 would be geometrically poised to form a disulfide bond. Indeed, as discussed below, the X-ray structure shows the two sulfhydryls are in van der Waals contact. Given the susceptibility of these sulfhydryls to oxidation by air to yield oxidation states higher than the disulfide (Villafranca, 1983), formation of the disulfide bond was instead attempted by a disulfide interchange reaction. It has been shown that when reduced ribonuclease or lysozyme is exposed to mixtures of oxidized and reduced forms of low molecular weight thiols such as glutathione or cysteine, the native protein disulfide bonds are formed almost quantitatively (Bradshaw et al., 1967; Saxena & Wetlaufer, 1970). In the case of Cys-39 DHFR, quantitative conversion to the disulfide-cross-linked form should leave only one free sulfhydryl, Cys-152, per molecule. Unfortunately, no fewer

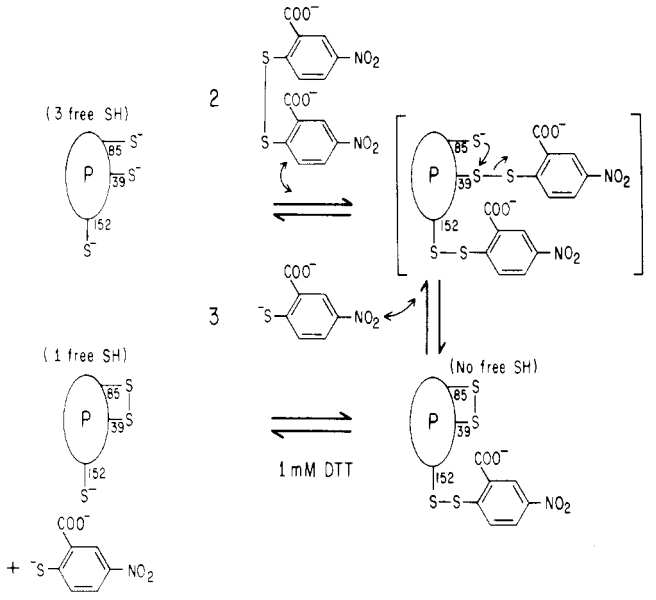

FIGURE 1: Probable reaction sequence for oxidation of Cys-39 DHFR by dithionitrobenzoate.

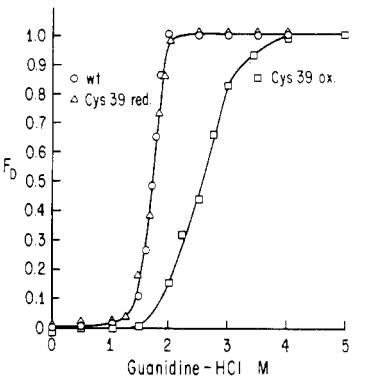

FIGURE 2: Gdn-HCl denaturation curves of (○) wild-type, (△) reduced Cys-39, and (□) oxidized Cys-39 DHFRs. The fraction of denatured protein, F_D, was measured as $(\epsilon_{obsd} - \epsilon_N)/(\epsilon_D - \epsilon_N)$, where ϵ_N and ϵ_D are the extinction coeficients at 292 nm for native and denatured forms of the protein.

than 1.5 free sulfhydryls (≤75% conversion) were found after oxidation by such mixtures of glutathione or cysteine (data not presented). It was observed, however, that treatment of the Cys-39 enzyme with dithionitrobenzoate (DTNB, Ellmans' reagent) yielded a species with only one thionitrobenzoate moiety bound per molecule, while the same treatment of the wild-type enzyme yielded a species with two thionitrobenzoate moieties bound per molecule. These results could be explained by three possibilities: (1) formation of a disulfide bond between Cys-39 and Cys-85, (2) formation of oligomers linked by intermolecular disulfide bonds, and (3) formation of an intramolecular disulfide bond between Cys-152 and either Cys-39 or Cys-85. Oligomers were not observed, so the second possibility was eliminated. Since Cys-152 is 28 Å from Cys-39 and Cys-85, an intramolecular disulfide bond between them would require a gross distortion of the native structure. Such a distortion in conformation was not detected by nondenaturing polyacrylamide gel electrophoresis of the oxidized Cys-39 DHFR and seems unlikely in light of the oxidized enzyme's full catalytic activity (see below). It seemed reasonable, therefore, to conclude that DTNB treatment did indeed specifically oxidize the Cys-39 and Cys-85 sulfhydryls to a disulfide bond.

Since the ultraviolet spectra of bound thionitrobenzoate and of the free anion are very different and readily allow quantification, one can easily follow the course of the DTNB reaction. The reaction was determined to be nearly complete with >95% conversion to the Cys-39–Cys-85 disulfide-cross-linked enzyme. The single remaining thionitrobenzoate group, presumably linked to Cys-152, is easily removed by mild reduction with 1 mM dithiothreitol (DTT), yielding a single free sulfhydryl per molecule. Reduction of the disulfide bond, on the other hand, requires more drastic conditions, such as 0.05 M DTT and 6 M urea at 32 °C, to achieve complete reduction. A probable reaction sequence for oxidation of Cys-39 DHFR by DTNB is depicted in Figure 1. We believe the two thionitrobenzoate moieties in the intermediate are linked to cysteine residues 39 and 152 since the X-ray structure shows that their SH groups are readily accessible to solvent, while that of Cys-85 is not.

Activity of Cys-39 DHFR. Both the reduced and oxidized forms of Cys-39 DHFR have kinetic parameters essentially identical with those of the wild-type enzyme (Howell et al., 1986). The k_{cat} values for wild-type and reduced and oxidized Cys-39 DHFR are 30, 28, and 27 s^{-1}, respectively. K_M values for dihydrofolate are 1.6, and 1.4 and 1.2 μM, and K_M values for NADPH are 1.4, 1.2, and 1.2 μM. The previously reported loss of activity following air oxidation of Cys-39 DHFR (Villafranca et al., 1983) was probably due to the formation of various extraneous byproducts rather than simple disulfide-bond formation.

Structural Stability. One of the most widely used approaches for estimation of protein structural stability is analysis of unfolding as a function of Gdn-HCl or urea concentration. These denaturants are capable of causing complete unfolding of proteins in a reversible manner that can be described by a two-state mechanism. Given the two-state assumption, one can make rather straight-forward estimates of structural stability in terms of the free energy change on denaturation (ΔG_D) (Greene & Pace, 1974).

We have used both Gdn-HCl and urea to evaluate the effect of the Cys-39 mutation, in both its reduced and oxidized forms, on the structural stability of DHFR. In the case of Gdn-HCl, unfolding of DHFR was monitored by measuring the spectral shift of tryptophan absorbance at 292 nm as a function of denaturant concentration. A blue shift at this wavelength is known to correspond to an increase in the polarity of the environment surrounding the tryptophan indole, as occurs when a buried tryptophan side chain is exposed to aqueous solvent, and thus indicates the extent to which protein unfolding has occurred (Donovan, 1972). This technique is particularly applicable to DHFR since the enzyme contains five tryptophan residues that are spatially distributed throughout the molecule. We obtain a $\Delta\epsilon^{292}$ value of −4350 M^{-1} cm^{-1} for all three maximally unfolded DHFRs (wild type and both forms of Cys-39 mutant). On the basis of model compounds (Donovan, 1969), this $\Delta\epsilon^{292}$ corresponds to exposure of 2.7 buried tryptophan residues, in good agreement with the expected number of 3.0 obtained by solvent-accessibility calculations (Richards, 1977) on the 2-Å crystal structures, and indicates complete unfolding has been attained in each case by Gdn-HCl denaturation. Figure 2 shows the denaturation curves for wild-type and reduced and oxidized Cys-39 DHFRs. Whereas the wild-type and the reduced forms of Cys-39 DHFR show a sharp transition between 1 and 2 M Gdn-HCl, with complete unfolding by 2.0 M, the oxidized form of Cys-39 DHFR shows a more gradual transition between 1.5 and 4 M, and complete unfolding does not occur until a concentration of 4.0 M Gdn-HCl is reached. The shift of the denaturation curve to higher Gdn-HCl concentration for the oxidized Cys-39 DHFR

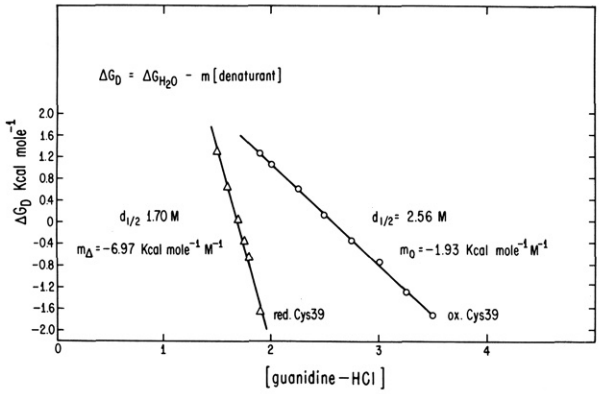

FIGURE 3: Plot of ΔG_D, the free energy of unfolding, vs. Gdn-HCl concentration; $d_{1/2}$ is the midpoint of the unfolding transition where $\Delta G_D = 0$.

indicates an increase in structural stability. However, the shallower slope of the transition represents a diminished folding/unfolding cooperativity, probably due to the appearance of stable intermediates (Pace, 1975).

Estimates of the free energy of unfolding, ΔG_D, can be obtained from analysis of the denaturation curves with

$$K_D = f_D/f_N = e^{-\Delta G_D/RT} \quad (1)$$

where f_D and f_N represent the fraction of protein in the denatured (D) and native (N) forms and K_D is the equilibrium constant. One can obtain a measure of the free energy of unfolding in the absence of denaturant ($\Delta G_D^{H_2O}$) by linear extrapolation (Aune & Tanford, 1969; Greene & Pace, 1974) according to

$$\Delta G_D = \Delta G_D^{H_2O} - m[\text{denaturant}] \quad (2)$$

where m is the slope of the plot and the quantity in brackets represents the concentration of denaturant. Such an analysis of the denaturation curves for reduced and oxidized Cys-39 DHFRs is presented in Figure 3. Extrapolation yields a $\Delta G_D^{H_2O}$ for the reduced form of 11.8 kcal/mol and 5.0 kcal/mol for the oxidized form, a difference in the energy of unfolding between the oxidized and reduced form ($\Delta\Delta G_D^{H_2O}$) of −6.8 kcal/mol, implying that the oxidized Cys-39 DHFR is *less* stable than either the reduced mutant or the wild-type enzyme, this despite the fact that a higher concentration of denaturant is required to cause the oxidized mutant to unfold. Clearly, because of the decreased slope of the curve, this estimate of the $\Delta G_D^{H_2O}$ of the oxidized mutant is much too low. Presumably the decreased slope in turn results from the existence of stable intermediates in the unfolding transition, leading to breakdown of the two-state assumption (Pace, 1975). The problem can be partially circumvented, and a reasonable estimate of the difference in stability between the reduced and oxidized forms can be made nevertheless. It has been shown that, even when intermediates are present, the region near the midpoint of the transition will yield a value of the unfolding equilibrium constant (K_D) expected for a two-state mechanism (Cupo & Pace, 1983). Consequently, by taking the difference in the Gdn-HCl concentrations at the midpoint of the reduced and oxidized unfolding transitions (0.86 M) and multiplying it by the slope of the oxidized form transition (m = −1.93 kcal mol^{-1} M^{-1}), one can obtain a value for the *lower limit* of the stability difference between the two forms. On the basis of this analysis, the oxidized form appears to be at least 1.8 kcal/mol more stable than the reduced form. From theoretical considerations (Poland & Scheraga, 1965), the increase in conformational stability due to the entropy effect of a disulfide bond cross-linking a 46-residue loop, as in the oxidized Cys-39 DHFR, can be estimated at 2.6 kcal/mol.

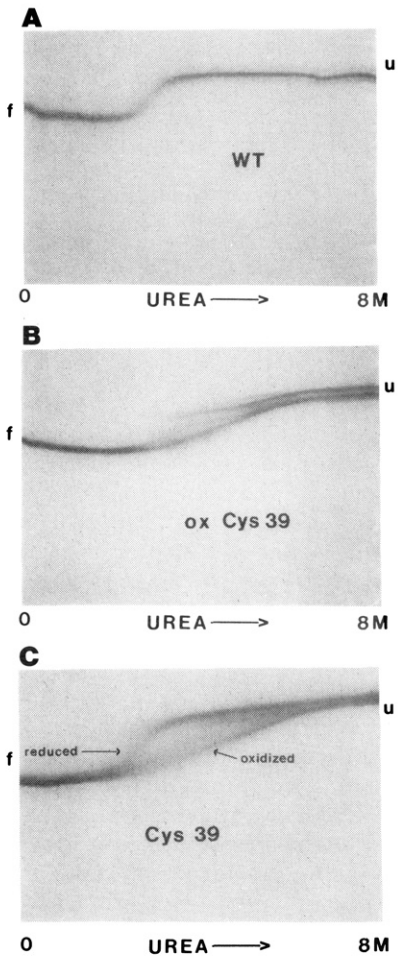

FIGURE 4: Urea-gradient polyacrylamide gels showing the unfolding profiles of (A) wild-type DHFR, (B) oxidized Cys-39 DHFR, and (C) a mixture of reduced and oxidized Cys-39 DHFR. The letters f and u indicate the position on the gel of the folded and unfolded forms of the protein.

As a check on our results with Gdn-HCl denaturation, experiments with urea-gradient polyacrylamide gels (Creighton, 1979) were also conducted. Samples of wild-type, reduced, and oxidized Cys-39 DHFRs were applied individually to polyacrylamide gels containing urea gradients (0–8 M) perpendicular to the direction of the electrophoretic potential. Electrophoresis was carried out with the gel temperature maintained at 15 °C. Under these conditions electrophoretic mobility is dependent on protein conformation and charge. Protein bands were visualized with Coomassie blue stain.

The urea-gradient gels confirmed and extended the conclusions from the Gdn-HCl experiments. The unfolding profiles of wild-type and reduced and oxidized Cys-39 DHFRs on urea gels are shown in Figure 4. Again, the reduced form shows a sharp reversible transition indicative of two-state behavior, indistinguishable from that of the wild type. The oxidized form, on the other hand, clearly shows a shift in the transition to a higher concentration of denaturant and a marked decrease in its slope.

One of the advantages of urea-gradient gels is their ability to display those stable conformational intermediates present in significant amounts. In the case of the oxidized form of Cys-39 DHFR, at least five distinct conformational species appear, two of which occur under maximally unfolding conditions. It is important to note that these intermediates are unlikely to be products of disulfide bond scrambling since identical Gdn-HCl denaturation profiles are obtained for all

three enzymes when the free sulfhydryls are blocked by treatment with iodoacetamide. Inspection of the relative positions of Cys-39, Cys-85, and Cys-152 in the structure (see Figure 6) reveals that a likely consequence of such disulfide interchange would be formation of oligomers. In the case of the urea-gradient gels, these oligomers would have a much lower mobility and thus would be seen above the unfolded form on the gel. In Figure 4C, the additional urea-unfolded species in the oxidized Cys-39 DHFR profile can be seen as a band having higher (not lower) mobility than the unfolded reduced Cys-39 DHFR. Evidently, the folding/unfolding pathway of the oxidized form of Cys-39 DHFR has been significantly affected by the disulfide cross-link.

Thermal Stability. DHFR is denatured by temperatures higher than 40 °C. This denaturation is only partially reversible and is accompanied by aggregation and precipitation. We studied thermal denaturation of DHFR by measuring recovery of enzymic activity after heat treatment and by spectroscopic monitoring of tryptophan absorbance at 292 nm. In either case, we observed no stabilization by the cysteine for proline substitution or by the disulfide cross-link. The activity half-life of wild-type DHFR and both forms of the Cys-39 DHFR is about 60 min at 70 °C, and the midpoint of the thermal unfolding transition is 44, 42, and 41 °C for wild-type DHFR, reduced Cys-39 DHFR, and oxidized Cys-39 DHFR, respectively. Similar results have been obtained by differential scanning calorimetry (Sturtevant, unpublished results). However, the maximum extent of the unfolding by heating was found to be significantly less than that for Gdn-HCl unfolding. While $\Delta\epsilon^{292}$ for thermal unfolding is −1510, −1190, and −1020 M^{-1} cm^{-1} for wild-type and reduced Cys-39 and oxidized Cys-39 DHFRs, the $\Delta\epsilon^{292}$ for maximal Gdn-HCl unfolding is −4300 M^{-1} cm^{-1} for wild-type DHFR and both forms of Cys-39 DHFR. Thus it appears that DHFR denaturation by heating yields a denatured form that is less completely unfolded than that produced by Gdn-HCl. Similar phenomena have also been observed for other proteins where it was determined that some secondary structure is retained at the temperature of maximum unfolding (Tanford, 1970). This absence of additional stabilization against thermal denaturation by the engineered disulfide bond in DHFR suggests that the thermally unfolded form of DHFR retains some structural integrity in the region near the disulfide bond. Thus, this cross-link would not contribute to entropic destabilization of the thermally unfolded form.

The finding that the disulfide-cross-linked DHFR was more stable toward Gdn-HCl or urea but was not more stable toward heating highlights the importance of first defining the molecular basis of the denaturation phenomenon against which one intends to stablize by protein engineering. For example, subtilisin has been specifically stabilized against H_2O_2 oxidation by making amino acid substitutions at the position (Met-126) known to be sensitive to oxidation (Estell et al., 1985).

Crystal Structures. The high-resolution X-ray structures of wild-type DHFR and both reduced and oxidized forms of Cys-39 DHFR were determined to elucidate the molecular basis for their observed properties. For example, knowing the precise structure of reduced Cys-39 DHFR helps to distinguish whether it is the reducing environment of the *E. coli* cytosol that prevents the disulfide bond from forming or whether some peculiar geometry at cysteine-39 makes for an energetically unfavorable process. Further, if the cysteine sulfurs are in van der Waals contact in the reduced Cys-39 enzyme, the distance between them must decrease from 4 to 2 Å when forming the disulfide bond. Does this shortening require an energetically unfavorable structural distortion? Would such a distortion in this region propagate to other parts of the molecule?

Although X-ray structure determination is often a long and arduous process, it was in this case greatly facilitated by the close similarity between the structures of the wild-type and mutant enzyme on the one hand and that of the previously determined *E. coli* MB1428 DHFR (Bolin et al., 1982) on the other. Because all three of the new crystals were isomorphous with the MB1428 DHFR crystals, their structures could be determined by difference Fourier methods.

Disulfide Bond Geometry. As expected, the refined structures of both forms of Cys-39 DHFR showed only slight changes in conformation when compared to the wild-type structure, the most significant changes occurring at the site of substitution. In the case of reduced Cys-39 DHFR, the C_α–C_β dihedral angle of residue 39 changes by −64°, placing the sulfhydryl at the exterior of the molecule where it is exposed to the solvent. The sulfhydryl of Cys-85 remains buried, 4.0 Å from the new sulfhydryl of residue 39. The two sulfhydryls are thus in van der Waals contact but are not covalently bonded. In the oxidized form, the Cys-39 C_α–C_β bond twists by an additional −82°, burying the sulfur atom beneath the side chain of Lys-58. The Cys-85 sulfhydryl remains buried in the oxidized form, having moved only slightly from its position in the reduced form, but now forms a covalent bond (2.0 Å) with the sulfur of Cys-39. The resulting inaccessibility of the disulfide bond to solvent probably explains the aforementioned difficulty in reducing this bond. The geometry of the disulfide is that of a left-handed spiral with $^{39}\chi_1 = -147°$, $^{39}\chi_2 = -149°$, $\chi_3 = -86°$, $^{85}\chi_2' = -63°$, and $^{85}\chi_1' = -78°$ in molecule 1 and $^{39}\chi_1 = -159°$, $^{39}\chi_2 = -158°$, $\chi_3 = -76°$, $^{85}\chi_2' = -65°$, and $^{85}\chi_1' = -81°$ in molecule 2. Although this geometry is not identical with that of any of the previously known protein disulfide bonds, the latter show considerable variation among themselves, and the dihedral angles found here are closest to those of known left-handed spirals (Richardson, 1981). Our engineered disulfide thus has adopted a more or less common conformation.

Modeling studies suggest that both left-handed spiral and right-handed hook geometries should be sterically allowed for this disulfide cross-link. The geometry of the modeled right-handed hook would, in fact, be close to those observed in other proteins. Assuming that the left- and right-handed disulfide geometries are equivalent in their conformational energy, why is the left-handed spiral preferred in this structure? One could argue that in vitro formation of the disulfide bond involves a mixed-disulfide intermediate (see Figure 1) in the fully folded molecule, imposing a particular stereochemistry on the course of the oxidation reaction that favors the left-handed disulfide conformation. Analysis of the structure shows that, in fact, the local geometry does constrain nucleophilic attack by the Cys-85 sulfide to just the one side of the putative Cys-39–thionitrobenzoate disulfide intermediate that yields a left-handed disulfide spiral. In addition, a substantial conformational change would appear to be necessary to rearrange this left-handed spiral into a right-handed conformation. Moreover, the most sterically acceptable right-handed disulfide we can model between Cys-39 and Cys-85 leaves their α-carbons 6.0 Å apart, whereas known right-handed disulfides have an average α-carbon distance of 5 Å (Richardson, 1981). Thus, it may be that the left-handed spiral is kinetically favored in the oxidation reaction and that its conversion into a right-handed geometry would be thermodynamically unfavorable.

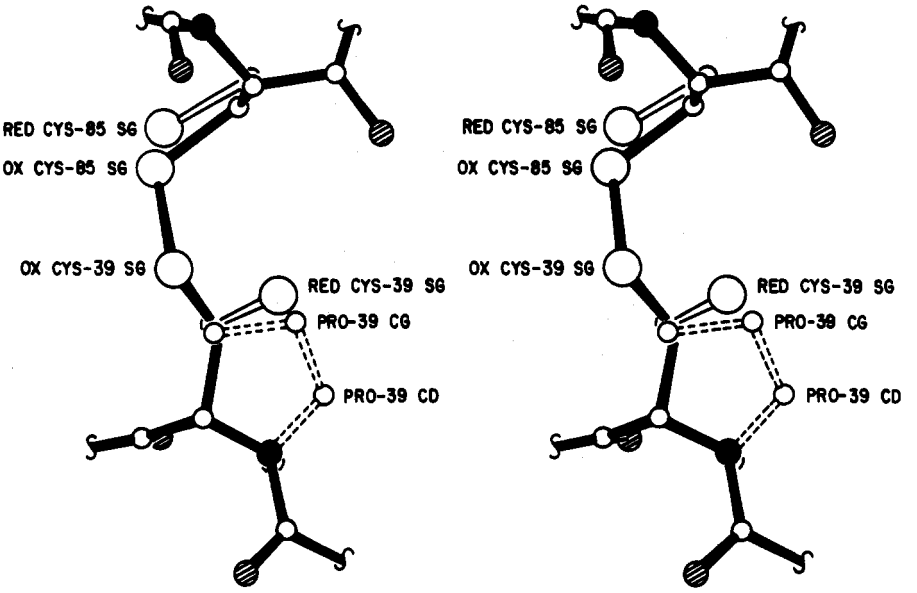

FIGURE 5: Stereoscopic depiction of atoms near the engineered disulfide bridge. Solid bonds represent oxidized Cys-39 DHFR, open bonds represent reduced Cys-39 DHFR, and dashed bonds represent wild-type enzyme.

In neither reduced nor oxidized Cys-39 DHFR are any significant backbone conformational changes propagated to other regions of the molecule. A plot of the differences in torsion angles ϕ and ψ between the wild-type and the mutant structures vs. residue number (excluding loop regions) shows backbone conformational changes only in the region of the substitution. The largest changes occur at residue 38, $\Delta\psi$ = 12°, and at residue 39, $\Delta\phi$ = 11°, in the wild type vs. reduced Cys-39 DHFR plot (not shown). Similarly, a comparison of the α-carbon positions showed no atomic displacement larger than 0.25 Å anywhere in the molecule. It should be noted that this analysis includes the α-carbons of Cys-39 and Cys-85, which might have been expected to move closer together in the oxidized Cys-39 enzyme. However, the distance between these two atoms were 6.1, 6.0, and 6.1 Å for the wild-type and reduced and oxidized Cys-39 DHFRs, respectively. The structure analysis thus shows that the major conformational adjustment caused by the Cys-39 substitution or the disulfide bond cross-linking is a simple twisting of the C_α–C_β bond of residue 39 (see Figure 5).

Molecular Dynamics. As is the case for an unfolded protein, a disulfide cross-link obviously can also affect the dynamics of the fully folded protein. There are two points to be considered in this regard when engineering a stabilizing disulfide bond. First, the protein's biological activity may be significantly dependent on its molecular motion. Although one could conceivably enhance the protein's activity by changing its dynamics, it is more likely that the molecular dynamics has evolved to optimize functionality and that tampering with it will only reduce its efficiency. Second, a cross-link that decreases the freedom of motion of the folded molecule will decrease its entropy, and thus, the relative stability of the folded form will be less than it would be if only the entropy of the unfolded form were decreased.

Since crystallographic temperature factors are a measure of molecular motion (Artymiuk et al., 1979; Frauenfelder et al., 1979), we can observe the effect of the Cys-39 mutation and of the introduction of a disulfide cross-link on the dynamics of the molecule by comparing the temperature factors of the wild-type and reduced and oxidized Cys39 DHFR structures. In the wild type, the temperature factors of α-carbons close to residue 39 are considerably lower (average B = 22 Å^2 for α-carbons 36–41) than those close to the Cys-85 position (average B = 41 Å^2 for α-carbons 83–87). It was anticiptated that the disulfide cross-link would reduce the motion of the more conformationally flexible Cys-85 region, which is at the molecular periphery, while not affecting the mobility of the region around residue 39, which is both buried and part of a parallel β-sheet (see Figure 6). It is important to note that the higher temperature factors of the Cys-85 region probably do indicate atomic motion rather than static disorder. Evidence for this conclusion is that the two molecules DHFR in the crystal asymmetric unit (which are thus in different environments) show nearly identical temperature factor distributions (Bolin et al., 1982). In addition, this region is also highly mobile in the structure of the *Lactobacillus casei* DHFR (Bolin et al., 1982).

Figure 7 shows isotropic temperature factors for the α-carbons of wild-type DHFR and reduced and oxidized Cys-39 DHFR in the region around Cys-85. The temperature factors of the Cys-39 DHFRs were normalized to the scale of the wild type since the X-ray data for the three enzymes extended to somewhat different resolutions. The root mean square differences between the wild-type temperature factors and those of the reduced Cys-39 DHFR and oxidized Cys-39 DHFR were 0.16 and 0.21 Å^2, respectively, showing that the temperature factors of the three molecules were, in general, nearly identical. However, as shown in Figure 7, the region around Cys-85 in the oxidized Cys-39 DHFR shows a considerable increase in temperature factors, and thus in mobility, over the corresponding region in either the wild-type DHFRs or reduced Cys-39 DHFRs. These results indicate that this cross-linking of the Cys-39 DHFR molecule introduces new interactions that loosen the structure near the disulfide bridge, the effects of which are evidently not compensated by constraints imposed by the cross-link.

A careful analysis of the refined structures was made to determine the molecular basis for these changes in the mobility of the polypeptide chain around Cys-85. This region, residues 80–90 (see Figure 6), comprises a short α-helix and a six-residue loop (85–90). These structures are associated with the rest of the molecule primarily through hydrophobic and van der Waals interactions. In the wild-type enzyme, the Pro-39 side chain interacts closely with the side chains of Cys-85 and Val-88. In reduced Cys-39 DHFR, the area of contact between Cys-39 and Val-88 appears to be somewhat

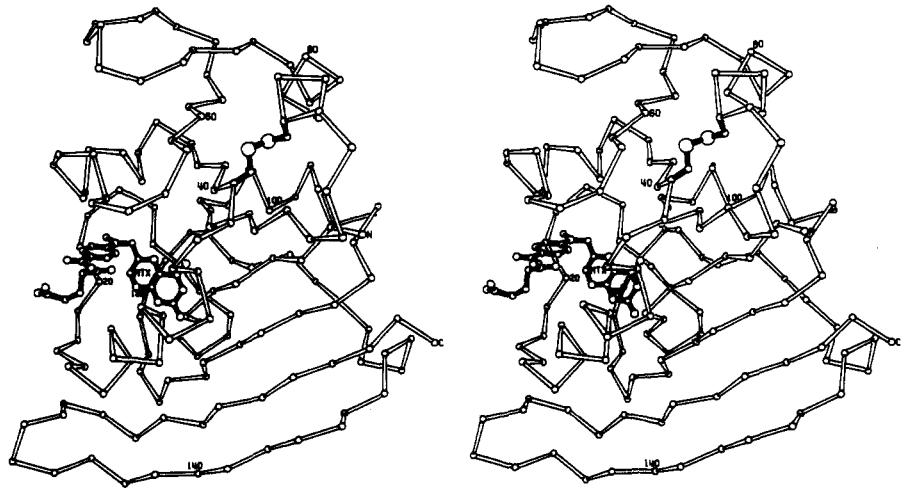

FIGURE 6: Stereoscopic α-carbon representation of oxidized Cys-39 DHFR showing the position of the engineered disulfide bond relative to secondary structural elements of the molecule.

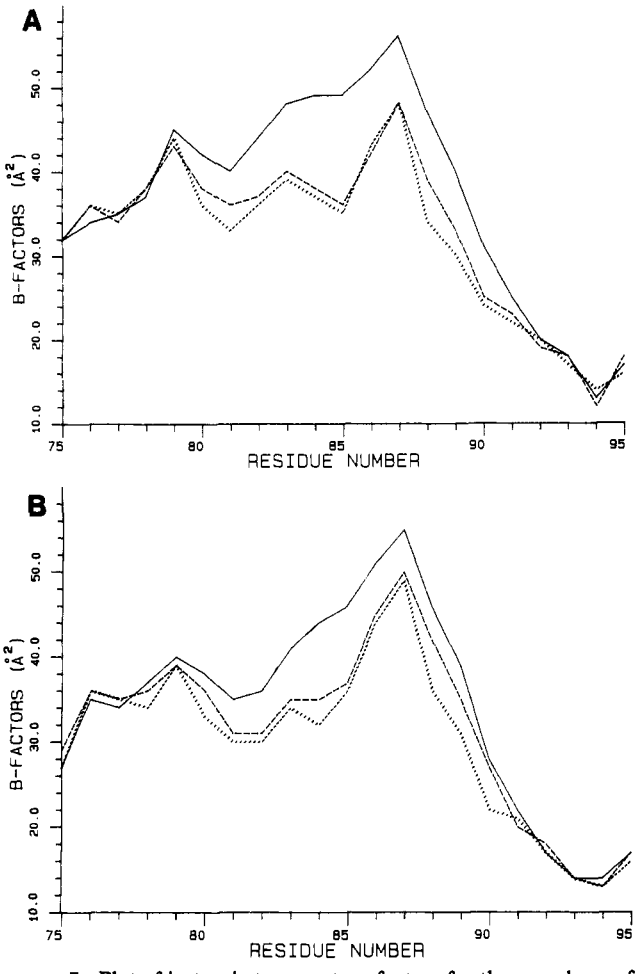

FIGURE 7: Plot of isotropic temperature factors for the α-carbons of residues flanking Cys-85 in wild-type DHFR (slashed line), reduced Cys-39 DHFR (dashed line), and oxidized Cys-39 DHFR (solid line). Panel A represents temperature factors (B factors) for molecule 1 in the asymmetric unit and panel B those for molecule 2.

smaller than that between Pro-39 and Val-88 in the wild-type structure, but contact between Cys-39 and Cys-85 is maintained. Additionally, in reduced Cys-39 DHFR, the Cys-39 sulfhydryl forms a hydrogen bond with the carbonyl oxygen of Cys-85. These small compensating differences could conceivably result in little net change in the interactions involving the Cys-85 region and thus produce little change in the molecular dynamics of this region. However, in contrast, the structure of the oxidized Cys-39 DHFR shows that formation of the disulfide bond causes the Cys-39 sulfhydryl to pull away from the Cys-85 loop, leaving a large van der Waals gap between Val-88 and Cys-39. Thus, the increased mobility of chain segment 80–90 in the disulfide-bridged mutant DHFR could be a result of the loss of this interaction.

But why does introduction of a disulfide cross-link fail to compensate for the loosening effect of the van der Waals gap? The notion that mobility will be restricted by a disulfide cross-link presupposes that the disulfide bond is rather inflexible. However, a survey of the torsion angles about disulfide bonds in proteins of known high-resolution structure shows considerable variation in χ_2 and χ_2' angles, and it has been calculated that even the more constrained χ_3 can twist by ±35° at a cost of less than 0.5 kcal/mol (Pullman & Pullman, 1974). Our results suggest that disulfide bonds are indeed geometrically flexible and consequently may not confer a significant reduction in local thermal mobility on the folded protein molecule.

ACKNOWLEDGMENTS

We are grateful to Robert Aust for his excellent technical assistance and to Robert Matthews and Nick Pace for helpful discussions.

Registry No. DHFR, 9002-03-3; L-Cys, 52-90-4; L-Pro, 147-85-3.

REFERENCES

Anfinsen, C. B., & Scheraga, H. A. (1975) *Adv. Protein Chem. 29*, 205–299.

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C., & Sternberg, M. J. E. (1979) *Nature (London) 280*, 563–568.

Aune, K., & Tanford, C. (1969) *Biochemistry 8*, 4586–4590.

Bolin, J. T., Filman, D. J., Matthews, D A, Hamlin, R. C, & Kraut, J. (1982) *J. Biol. Chem. 257*, 13650–13662.

Bradshaw, R A., Kanarek, L., & Hill, R. L. (1967) *J. Biol. Chem. 129*, 3789–3798.

Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.

Creighton, T. E. (1979) *J. Mol. Biol. 129*, 235–264.

Creighton, T. E., Hillson, D. A., & Freedman, R. B. (1980) *J. Mol. Biol. 142*, 43–62.

Cupo, J. F., & Pace, C. N. (1983) *Biochemistry 22*, 2654–2658.

Donovan, J. W. (1969) *J. Biol. Chem. 244*, 1961–1967.

Donovan, J. W. (1972) *Methods Enzymol. 27D*, 497–525.

Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Estell, D. A., Graycar, T. P., & Wells, J. A. (1985) *J. Biol. Chem. 260*, 6518–6521.
Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. C. (1979) *Nature (London) 280*, 558–563.
Freedman, R. B., & Hillson, D. A. (1980) in *The Enzymology of Post-Translational Modification of Proteins* (Freedman, R. B., & Hawkins, H. C., Eds.) Vol. 1, pp 167–212, Academic, New York.
Greene, R. F., & Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388–5393.
Haber, E., & Anfinsen, C. B. (1962) *J. Biol. Chem. 237*, 1839–1844.
Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 13.1–13.23, Indian Institute of Science, Bangalore, India.
Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., & Kraut, J. (1986) *Science (Washington, D.C.) 231*, 1123–1128.
Huennekens, F. M., Vitols, K. S., Whiteley, J. M., & Neef, V. G. (1976) in *Methods in Cancer Research* (Busch, H., Ed.) pp 199–223, Academic, New York.
Itzhaki, R. F., & Gill, D. M. (1964) *Anal. Biochem. 9*, 401–410.
Michaelis, S., Inouye, H., Oliver, D., & Beckwith, J.(1983) *J. Bacteriol. 154*, 366–374.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1–43.
Perry, L. J., & Wetzel, R. (1986) *Biochemistry 25*, 733–739.
Poland, D. C., & Scheraga, H. A. (1965) *Biopolymers 3*, 379–399.
Pullman, B., & Pullman, A.(1974) *Adv. Protein Chem. 28*, 347–526.
Richards, F. M. (1977) *Annu. Rev. Biophys. Biochem. 6*, 151–175.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 167–339.
Saxena, V. P., & Wetlaufer, D. (1970) *Biochemistry 19*, 5015–5022.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.
Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261–287.
Villafranca, J E., Howell, E. E., Voet, D. H., Strobel, M. S., Ogden, R. C., Abelson, J. N., & Kraut, J. (1983) *Science (Washington, D.C.) 222*, 782–788.
Wells, J. A., & Powers, D. B. (1986) *J. Biol. Chem. 261*, 6564–6570.
Xuong, N.-h., Freer, S. T., Hamlin, R., Nielson, C., & Vernon, W. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 289–295.
Ziegler, D. M., & Poulsen, L. L. (1977) *Trends Biochem. Soc. (Pers. Ed.) 2*, 79–81.

Amino Acid Sequence of Sialic Acid Binding Lectin from Frog (*Rana catesbeiana*) Eggs†

Koiti Titani,*,‡ Koji Takio,‡ Manabu Kuwada,‡,§ Kazuo Nitta,‖ Fusao Sakakibara,‖ Hiroaki Kawauchi,‖ Giichi Takayanagi,‖ and Sen-itiroh Hakomori*,⊥

Department of Biochemistry, University of Washington, Seattle, Washington 98195, Program of Biochemical Oncology/Membrane Research, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, and Department of Biochemistry, Cancer Research Institute, Tohoku College of Pharmaceutical Sciences, Sendai 983, Japan

Received June 2, 1986; Revised Manuscript Received December 17, 1986

ABSTRACT: The complete amino acid sequence of sialic acid binding lectin from frog (*Rana catesbeiana*) egg is presented. The 111-residue sequence was determined by the analysis of peptides generated by digestion of the S-carboxymethylated protein with *Achromobacter* protease I, chymotrypsin, or cyanogen bromide. The sequence is unique and not homologous to any known protein sequence. The protein may represent a new type of lectin.

Lectins are multivalent carbohydrate binding proteins with the ability to agglutinate erythrocytes, other normal and cancer cells, and microorganisms. They were first discovered in plants, but recently they have also been found in animals and other organisms [cf. review by Barondes (1981)].

Kawauchi et al. (1975) found that eggs of frog (*Rana catesbeiana*) contain two lectins that show different agglutinating activities, one with human blood group A erythrocytes and the other with mouse Ehrlich ascites carcinoma cells or rat ascites hepatoma cells (AH-109A). These two lectins were purified to homogeneity, and their binding specificity, agglutinatability, and physical and chemical properties have been studied. One of the lectins showed a preferential agglutination of cancer cells and displayed a specific binding to sialyl glycoprotein (Sakakibara et al., 1977), and consequently, it was designated sialic acid binding protein.

In the present paper, we present the complete amino acid sequence of 111 residues of the frog egg sialic acid binding lectin and discuss the structural features in relation to the biological function.

Materials and Methods

Sialic acid binding lectin was prepared as described by Sakakibara et al. (1977). After extraction from acetone powder of eggs of *Rana catesbeiana* with saline, it was purified by chromatography on Sephadex G-75, DEAE-cellulose, hy-

† This work was supported by Research Grants HL-29595 (to K. Titani), GM-15731 (to K. Walsh), and GM-23100 (to S.H.) from the National Institutes of Health. K. Titani is also a visiting professor of the Fujita-Gakuen Health University, Toyoake-shi, Japan. K. Takio is a senior associate of the Howard Hughes Medical Institute Laboratory at the University of Washington, Seattle.

* Address correspondence to K. Titani at the School of Medicine, Fujita-Gakuen Health University, Toyoake-shi, Japan, or to S. Hakomori at the Fred Hutchinson Cancer Research Center.
‡ University of Washington.
§ Present address: Eisai Co. Tsukuba Research Laboratories, 1-3, Tokodai 5-chome, Toyosato-machi, Tsukuba-gun, Ibaragi, Japan.
‖ Tohoku College of Pharmaceutical Sciences.
⊥ Fred Hutchinson Cancer Research Center.

0006-2960/87/0426-2189$01.50/0 © 1987 American Chemical Society