THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1986 by The American Society of Biological Chemists, Inc.

Vol. 261, No. 33, Issue of November 25, pp. 15480–15485, 1986
Printed in U.S.A.

The Crystallographically Determined Structures of Atypical Strained Disulfides Engineered into Subtilisin*

(Received for publication, April 30, 1986)

Bradley A. Katz‡ and Anthony Kossiakoff

From the Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143 and the Department of Biocatalysis, Genentech, Inc., South San Francisco, California 94080

The geometries of two disulfide bridges genetically engineered into subtilisin have been characterized by x-ray crystallography to determine the structural and energetic constraints involved in introducing disulfide bonds into proteins. Both disulfide bridges (Cys-24—Cys-87 and Cys-22—Cys-87) exhibit atypical sets of dihedral angles compared to those for other reported disulfide structures in proteins. The geometric trends for naturally occurring disulfides in protein crystal structures are examined. Comparison of the disulfide-containing mutant protein structures with the wild-type structure shows that, in both cases, disulfide incorporation is accommodated by relatively minor changes in local main-chain conformation. The Cys-22—Cys-87 disulfide has two high energy dihedral angles ($X2$ = 121°, $X2'$ = 143°). Both disulfides produce short non-bonded contacts with the main-chain.

A disulfide bond is one of several types of structural elements that impart conformational stability to proteins (1). A question being addressed by protein engineering studies is the degree to which engineered disulfides can produce the stabilizing effects of natural ones. Recently, a measurable increase in the stability of two proteins has been accomplished by introducing disulfides into their structures (2).[1]

Although it is thought that the location of insertion of a disulfide bridge into a protein is a principal determinant in its stabilizing effect, it was unknown how well a modeled disulfide group must conform to an optimal geometry or how closely it must adhere to the stereochemical trends observed for natural disulfides in proteins in order that the disulfide bond form. We felt that insight into these issues could be gained by comparing crystal structures of protein variants containing a disulfide whose modeled geometry was somewhat structurally incompatible with the native protein. Here we describe the modeled and crystallographically determined geometries of two atypical disulfides incorporated into subtilisin. A primary goal of the study was to determine how the resulting perturbations to the structure would be partitioned between the engineered disulfide and adjacent main- and side-chain groups. The structural characterization of such disulfide-containing mutants provides data regarding the relative importance of the several physicochemical factors involved in introducing disulfide bonds into proteins.

Recently Wells and Powers (3) produced disulfide-containing mutants of subtilisin, a protease of about 27,500 M_r containing no natural cyst(e)ine groups. We chose two of these mutants (expressed in *Bacillus subtilis*), each with a common residue (Cys-87), to study crystallographically. Wells and Powers (3) demonstrated by polyacrylamide gel electrophoresis that both mutants, Ser-24→Cys/Ser-87→Cys and Thr-22→Cys/Ser-87→Cys,[2] contain a disulfide bond that forms quantitatively during fermentation of the bacteria (3). In both mutants the disulfide is on the surface of the molecule, about 24 Å away from the active site, and does not affect enzymatic activity (3).

EXPERIMENTAL PROCEDURES

Potential sites for introducing disulfide bridges were identified by determining the locations in the structure where a modeled disulfide group of good stereochemistry could be inserted with a minimum of main-chain rearrangement.[3] The $X3$ dihedral angle was constrained to be within 10° of ±90°, while $X1$, $X2$, $X2'$, and $X1'$ were varied to best approximate staggered conformations (see diagram in legend of Table II for definitions of the dihedrals). Theoretical dihedral energies of disulfide conformations were calculated with a subroutine of the program AMBER (4). Modeling indicated that for C24/C87,[4] a right-handed disulfide ($X3 \sim 90°$) could be built requiring very little main-chain movement. The disulfide in C22/C87 could not be modeled with any combination of staggered dihedral angles without also requiring significant main-chain rearrangements. The model requiring minimal main-chain movement for Cys-22—Cys-87 was a left-handed disulfide ($X3 = -85°$) with two high energy dihedrals ($X2$ = 112°, $X2'$ = 130°).

Although some examples of high energy dihedral ($X2$ or $X2'$) angles in disulfides in proteins have been reported (*e.g.* BPTI (Cys-30—Cys-51, $X2 = -121°$, Cys-14—Cys-38, $X2' = -119°$), crambin (Cys-4—Cys-32, $X2' = -118°$), ribonuclease (Cys-58—Cys-110, $X2' = -125°$)) (5), such cases are rare. Therefore, no reliable prediction could be made as to what the most likely geometry of the Cys-22—Cys-87 disulfide would be or as to how its incorporation might affect the local environment.

Purification and Crystallization—A culture of the C24/C87-producing bacteria was obtained from J. Wells of Genentech and grown in

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by Grant GM29616 from the National Institutes of Health.

[1] J. E. Villafranca, personal communication.

[2] The Cys-22—Cys-87 disulfide-containing subtilisin mutant studied here was actually a triple mutant Tyr-21→Ala/Thr-22→Cys/Ser-87→Cys. The substitution of Tyr-21 for Ala was made for a purpose independent of this disulfide study. The triple mutant will be called C22/C87 throughout this paper.

[3] The initial modeling was done by J. Wells and R. Bott of Genentech.

[4] The abbreviations used are: C24/C87 and C22/C87, the Cys-24—Cys-87 and Cys-22—Cys-87 disulfide-containing mutants, respectively; Cys-24—Cys-87 and Cys-22—Cys-87, the disulfide cross-links; WT, wild type; BPTI, basic pancreatic trypsin inhibitor; MES, 2-(*N*-morpholino)ethanesulfonic acid; HEPES, *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; SAS, saturated ammonium sulfate; F_o and F_c, observed and calculated structure factor; F^+ and F^-, structure factor amplitude and its Friedel mate; I_{hkl}, intensity of the *hkl* reflection; ϕ_c, calculated phase; ϕ_{WT}, wild type phases; F(SS), F(disulfide); conventional least squares R factor (or residual), $R = \sum_{i=1}^{i=N}||F_o| - |F_c||/\sum_{i=1}^{i=N}|F_o|$, where N = the number of reflections.

TABLE I

Data collection parameters and statistics

Variant	Crystal dimensions	Resolution	Decay at highest resolution	Total no. of observed measurements collected	No. of unique reflections collected	No. of Friedel pairs collected	R_{sym}[a]	RI_{iso}[b]
	mm	*Å*	%				%	%
Wild type	0.35 × 0.35 × 0.50	22.0–1.80	5	15,461	15,461		4.0	
C24/C87	0.20 × 0.20 × 0.50	22.0–2.15	<1	16,844	12,132	3,583	8.8	7.8
		22.0–4.45				1,317	4.2	
		2.31–2.15				2,266	9.2	
C22/C87	0.25 × 0.30 × 0.60	22.0–2.05	7	17,461	13,864	2,690		7.8
		22.0–3.30				2,690	2.4	

[a] $R_{sym} = \sum_{i=1}^{i=N} |I_i^+ - \langle I \rangle_i| / \sum_{i=1}^{i=N} \langle I \rangle_i$, where $\langle I \rangle_i = (I_i^+ + I_i^-)/2$ and N = no. of Friedel pairs.

[b] $RI_{iso} = \sum_{i=1}^{i=N} |I_i(\text{mutant}) - I_i(\text{WT})| / \sum_{i=1}^{i=N} \langle I \rangle_i$, where $\langle I \rangle_i = (I_i\ (\text{mutant}) + I_i\ (\text{WT}))/2$ and N = no. of reflections common to both data sets.

4 shake flasks, each containing 1 liter of medium, for 24 h under previously described conditions (3). The subtilisin was purified in the presence of 1 mM PMSF (6), giving a yield of 10 mg/4 liters. The C22/C87 mutant protein (2 mg) was a gift of J. Wells.

Large single rod-shaped crystals of phenylmethanesulfonyl fluoride-inhibited C24/C87 and C22/C87 were grown by vapor diffusion in 20-μl hanging drops seeded with small blocky pieces of orthorhombic wild type crystals. The best crystals of C24/C87 grew from drops initially containing 2.85 mg/ml protein, 23% (by volume) SAS solution, 0.8% (by volume) 2-methyl-2,4-pentanediol, 0.15 M MES, pH 5.9. The drops were equilibrated against 1-ml wells containing 38–43% SAS, 6% 2-methyl-2,4-pentanediol, 0.3 M MES, pH 5.9. (The optimal crystallization conditions for wild type or for C22/C87 varied slightly from the above conditions.)

WT and C24/C87 crystals were stabilized by soaking in a synthetic mother liquor of 65% SAS, 0.2 M MES, pH 5.90, containing millimolar concentrations of $CaCl_2$. The soaks also included 1% styrene saturated in 50% SAS to prolong the crystal lifetime in the x-ray beam. The crystal of C22/C87, grown at pH 5.90, was transferred through a series of soaks of increasing pH and of increasing $CaCl_2$ concentration to a final soak of 65% SAS, 0.3 M HEPES, pH 7.0, 30 mM $CaCl_2$, and 1% saturated styrene.

Data Collection and Structure Refinement—Data were collected on an ENRAF NONIUS CAD4 diffractometer by the step scan method (7) with monochromated copper Kα radiation (40 kV, 27 mA). Data collection was done in 10 shells, starting from the high resolution limit. Typically, raw intensities were obtained by summing the highest five consecutive steps (0.02° (in ω)/step). The data were corrected for absorption by the method of North *et al.* (8). Table I contains data collection parameters and statistics.

C24/C87—For the C24/C87 mutant a complete data set to 2.15 Å resolution was collected in 12 days with no detectable decay. Friedel pairs collected at low and high resolution (see Table I) were locally scaled[5] to correct for any systematic errors. R_{sym} (defined in Footnote *a* of Table I) for 1317 Friedel pairs from 22.0 to 4.45 Å was 4.2% (on intensity); for 2266 pairs from 2.31 to 2.15 Å, R_{sym} was 9.2%. Local scaling[6] of the intensities of C24/C87 to those of WT gave an RI_{iso} value (defined in Footnote *b* of Table I), reflecting the difference between the WT and the C24/C87 structure, of 7.8%.

C22/C87—For C22/C87 a complete data set to 2.05 Å resolution was collected in 14 days. No decay was observed for the low resolution data; the average decay at higher resolution (2.15–2.05 Å) after 14 days was 7%, but because these data were collected first, their measured intensities were not affected by decay. Therefore, no decay correction was applied. Local scaling of F^+ to F^- yielded an R_{sym} value of 2.4% for the data to 3.3 Å. Locally scaling the C22/C87 intensities to those of WT gave an RI_{iso} value of 7.8%.

Structure Refinement—Initial models of the disulfide structures were built from difference electron density maps calculated with terms $F(SS) - F(WT)$ or $2F(SS) - F(WT)$ and WT phases (ϕ_{WT}).[4] Steigemann's PROTEIN package (9) was used for these computations. The WT structure used for phases (1.8 Å, R = 14.8%) and for the structure comparisons was provided by R. Bott.[7]

The structures of C24/C87 (2.15 Å resolution) and of C22/C87 (2.05 Å) were refined through iterations of restrained least squares refinement (10) and refitting of the protein model where necessary. Bond lengths and angles were restrained to ideal values (root mean square deviation of bond lengths from ideality = 0.017 Å). Refitting was based on $(F_0 - F_c, \phi_c)$, and $(2F_0 - F_c, \phi_c)$[4] maps and was carried out with the program FRODO (11) supported on an Evans and Sutherland PS300 graphics system. The current residuals[4] for the C24/C87 and C22/C87 structures are 16.3 and 15.6%, respectively.

RESULTS

The C24/C87 difference (F(C24/C87) − F(WT), ϕ_{WT}) map at 2.15 Å resolution showed two resolved positive peaks corresponding to the scattering of the sulfur atoms and two negative peaks at the positions occupied by the Ser-24 Oγ and Ser-87 Oγ atoms in the WT structure (Fig. 1). A model of a C24/C87 double cysteine mutant was made by replacing Ser-24 and Ser-87 with cysteines and rotating the $X1$ dihedral of Cys-24 by −85° and of Cys-87 by −173°, with respect to the WT serine orientations, to yield a right-handed disulfide geometry ($X3$ = 85°). The fit was further optimized by a ψ rotation of −9° about the C_α87-C87 bond.

[5] For each reflection this local scaling program searches for a specified number (typically 50–125) of neighboring reflections and their Friedel mates, and from the resulting subset of Friedel pairs, calculates the local slope of F^- *versus* F^+ and the local intercept. Corrections are applied to both F^+ and F^- with the constraint that the resulting local slope of F^- *versus* F^+ be unity and that the intercept be zero. The regions of reciprocal space used in locally scaling a Friedel pair are continually updated along with the local slope and intercept, so that each pair receives its own correction. The maximum number of neighboring reflections used to scale a particular Friedel pair can be varied with resolution. Thus, at low resolution, where neighbors are farther apart in reciprocal space, the localness of the scaling can be maintained by using fewer neighbors to scale the pair. The program is advantageous in correcting data where absorption due to glass and solvent is significant and insufficiently corrected with a ϕ curve alone. (The original version of the program was written by A. Goldman at Yale.) Because collecting Friedel pairs for all unique data to 2 Å resolution is time consuming with a conventional diffractometer, we collected Friedel pairs at low resolution and at high resolution, deduced the ϕ, χ, and θ dependence of the corrections, and applied them to the data lacking Friedel mates.

[6] The program used to locally scale mutant data sets to the wild type data set determines, for each reflection, the local slope of I_{hkl} (mutant) *versus* I_{hkl} (WT) and applies the slope to I_{hkl} (mutant). The calculations can be weighted by $1/I_{hkl}$. The other algorithms of this program are similar to those in the program described in Footnote 5. With the data collection strategies used here, such a local scaling procedure can greatly minimize the effects of x-ray decay on difference Fouriers. The signal-to-noise ratio in the F(mutant) − F(WT) difference electron density maps was shown to be dramatically improved compared to that in the corresponding maps obtained when the mutant data are scaled to the WT data with only one scale factor and one temperature factor.

[7] R. Bott, B. A. Katz, M. Ultsch, and A. Kossiakoff, manuscript in preparation.

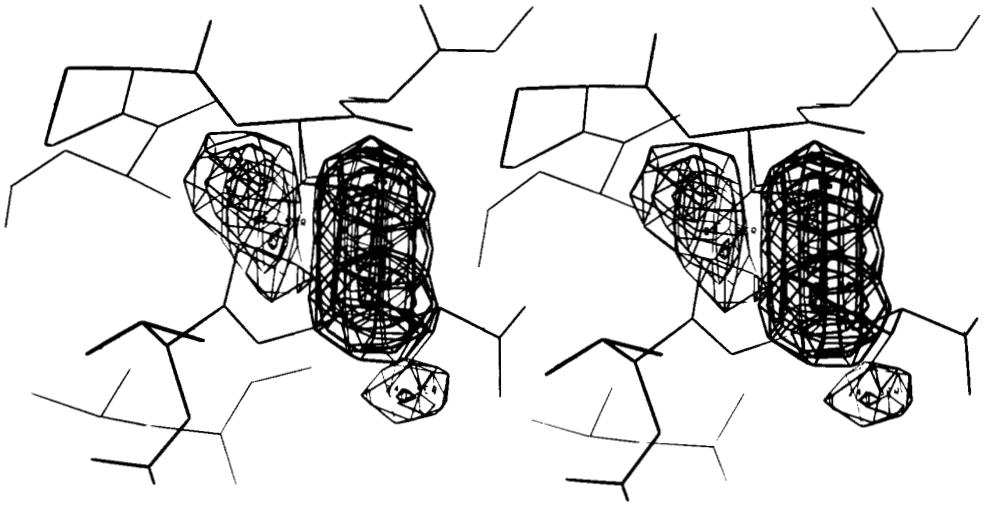


FIG. 1. **Difference (F(C24/C87) − F(WT)) electron density map superimposed on the wild type structure and on the initial crystallographic model of C24/C87.** The positive density contoured at 3, 4, 8, and 12 σ surrounds the sulfur atoms while the negative density at 3, 4.5, and 6 σ appears at the positions in the WT structure of the Ser-24 and Ser-87 Oγ atoms. The sulfur atoms of residues 24 and 87 are 2.2 and 2.7 Å, respectively, from the 24 and 87 Oγ positions in the WT structure. A negative ripple (an artifact of Fourier transform termination errors) around the disulfide bond was removed for clarity.

FIG. 2. **Difference (F(C22/C87) − F(WT)) electron density map superimposed on the WT structure.** The positive density (below Cβ87) is contoured at 3, 6, and 9σ, while the negative density (at Cγ22 and covering Tyr-21) is contoured at 3, 4.5, and 6σ).

A left-handed Cys-22—Cys-87 disulfide group was fit to the density in a 2.05 Å resolution ($2F$(C22/C87) − F(WT), ϕ_{WT}) map. The difference map is shown superimposed on the WT structure in Fig. 2. A good fit of the disulfide was accomplished by small ψ,ϕ rotations (less than 10°) of residues 85–90, along with a 40° rotation of the Cys-87 $X1$ dihedral. No rotation of the Cys-22 dihedral was needed. The need for ψ,ϕ rotations to fit the density in the disulfide region was later determined to result from a translation in the crystal lattice of the entire C22/C87 structure with respect to WT.[8]

[8] The finding here that in highly isomorphous crystals there can be subtle but real changes in the packing of protein molecules in the lattice identifies the need, when comparing two structures, to consider the possibility that long range systematic differences could sometimes be artifacts of crystallization.

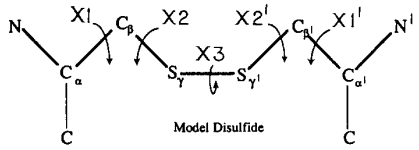
The resulting disulfide dihedral angles for the initial unrefined crystal structures of C22/C87 and of C24/C87 are given in Table II. In most respects, the experimentally determined disulfide structures agree well with modeling predictions. The Cys-22—Cys-87 structure showed that the disulfide modeled with high energy $X2$ and $X2'$ dihedral angles was the correct one. Formation of the Cys-24—Cys-87 disulfide requires much larger side-chain rotations than formation of Cys-22—Cys-87. Disulfide incorporation in both cases results in minor subtle local rearrangement in main-chain conformation.

No large changes in structures were observed between the initial and refined disulfide structures. As inferred from modeling, little main-chain alteration occurs in C24/C87 when compared to WT. In the crystal structure of C22/C87, the entire molecule is translated by about 0.3 Å, primarily in the

TABLE II

The most highly populated classes of left- and right-handed disulfide geometries, determined by inspection of disulfide dihedrals observed in protein crystal structures, along with the modeled and crystallographically observed geometries of Cys-24—Cys-87 and of Cys-22—Cys-87

The dihedral angles in a disulfide group are defined in the diagram below.



The structural data considered here were obtained from the Brookhaven Data Bank (5) or from Ref. 22. The directionality of the bridge, *i.e.* the assignment of the dihedrals as primed or unprimed, has been chosen to emphasize similarities in the various disulfide geometries. $X1$ is the dihedral angle about the Cα-Cβ bond; $X2$ is that about the Cβ-Sγ bond; $X3$ about Sγ-Sγ′; $X2'$ about Sγ′-Cβ′; and $X1'$ about Cβ′-Cα′. The dihedral energies were calculated with the equation obtained from a subroutine of AMBER (4): E (kcal/mol) = 2.0 (1 + cos (3$X1$)) + 2.0 (1 + cos (3$X1'$)) + 1.0 (1 + cos (3$X2$)) + 1.0 (1 + cos (3$X2'$)) + 3.5 (1 + cos (2$X3$)) + 0.6 (1 + cos (3$X3$)). Note that the $X3$ dihedral energy function is a sum of a 2-fold term whose minima occur at ±90° plus a smaller 3-fold term whose minima occur at ±60°, 180°; the minima in the overall $X3$ dihedral energy function, therefore, occur at ±83°, close to the values observed in some small molecule crystal structures. The Cα-Cα and other distances were calculated assuming ideal bond lengths and bond angles. Angles given are in degrees, distances are in Å, and energies are in kcal/mol. Standard deviations in dihedral angles are in parentheses.

Class	No. of examples	$X1$	$X2$	$X3$	$X2'$	$X1'$	Cα-Cα	E
Left-handed disulfides								
1	15	−59 (10)	−57 (14)	−85 (9)	−64 (21)	−62 (17)	5.49	0.54
2[a]	4	−168 (12)	173 (4)	−84 (13)	−171 (5)	173 (11)	6.60	1.18
3[b]	5	−165 (3)	171 (6)	−82 (7)	−171 (6)	−55 (5)	6.57	1.36
4	6	−66 (10)	−65 (7)	−80 (11)	−93 (5)	−172 (13)	5.82	1.97
5[c]	7	−76 (4)	−87 (6)	−89 (5)	−148 (9)	−85 (4)	6.32	4.66
All[d]	39			−85.1 (8.7)			5.88	1.68
							(±0.49)	(±1.51)
Cys-22—Cys-87 (modeled)		64	112	−85	130	−57		
Cys-22—Cys-87 (initial build)		49	109	−102	131	−45		
Cys-22—Cys-87 (refined)		53	121	−98	143	−49	5.37	4.79
Right-handed disulfides								
1	4	−67 (2)	−61 (4)	98 (7)	63 (7)	179 (3)	4.66	1.13
2	4	−60 (2)	−56 (6)	103 (2)	74 (7)	65 (8)	4.95	1.68
3	4	−62 (8)	−78 (5)	103 (2)	−166 (13)	71 (9)	5.00	2.29
4	6	−56 (8)	170 (18)	92 (9)	43 (25)	−154 (17)	5.48	2.83
5	4	−63 (9)	−71 (8)	103 (7)	84 (17)	−77 (13)	5.13	2.95
6	8	−53 (7)	−83 (4)	101 (3)	−121 (11)	−55 (3)	4.18	4.02
All[e]	45			99 (11)			5.07	3.19
							(±0.73)	(±1.43)
Cys-24—Cys-87 (modeled)		−70	−35	92	−152	−174		
Cys-24—Cys-87 (initial build)		−71	−60	116	−149	176		
Cys-24—Cys-87 (refined)		−65	−50	96	−171	−157	4.59	2.46

[a] Immunoglobulins.
[b] Cys-168—Cys-82 in 5 homologous serine proteases.
[c] Cys-58—Cys-42 in 7 homologous serine proteases.
[d] Left-handed disulfide geometries in the protein crystal structures tabulated here.
[e] Right-handed disulfide geometries in the protein crystal structures tabulated here as well as some "miscellaneous" right-handed geometries occurring in proteins (data not shown).

direction of the *b* axis, with respect to the WT molecule. After a rigid body rotation and translation matrix is applied to superimpose C22/C87 on WT (resulting root mean square deviation between the two structures = 0.18 Å), the main-chain in the disulfide region shows only minor differences. The structures of the two mutants compared to WT are shown in Fig. 3. Fig. 4 shows the region of the subtilisin molecule where the disulfides were introduced.

DISCUSSION

Geometric Trends of Naturally Occurring Disulfides in Proteins—Available protein structure coordinates indicate that two general disulfide families can be defined by the chirality of the dihedral angle about the sulfur-sulfur bond: right-handed ($X3 \sim 90°$) and left-handed ($X3 \sim -90°$). The families can be subdivided according to their other side-chain dihedral angles into several distinct classes (1, 12). Because the data base of protein crystal structures has expanded since Richardson's compilation (12), we have re-examined the geometric trends in the disulfide families.

The dihedral angles for the most highly populated geometric types of right- and left-handed disulfides in the protein data bank (5) are given in Table II. Right-handed and left-handed disulfide groups in proteins differ in several geometric characteristics. Whereas many (15 out of 39) left-handed disulfide groups in proteins adopt (approximately) a preferred geometry (called a left-handed spiral ($X1 = -60°$, $X2 = -60°$, $X3 = -85°$, $X2' = -60°$, $X1' = -60°$)), the right-handed groups are more variable in their dihedral angle patterns. There are many more examples of right-handed disulfides with $X1$ (or $X1'$) = 60° than there are left-handed ones. Finally, Richardson (12) pointed out that while the average Cα-Cα distance of left-handed disulfides in proteins (5.88 ± 0.49 Å in our survey) is very close to that observed in small molecule disulfide crystal

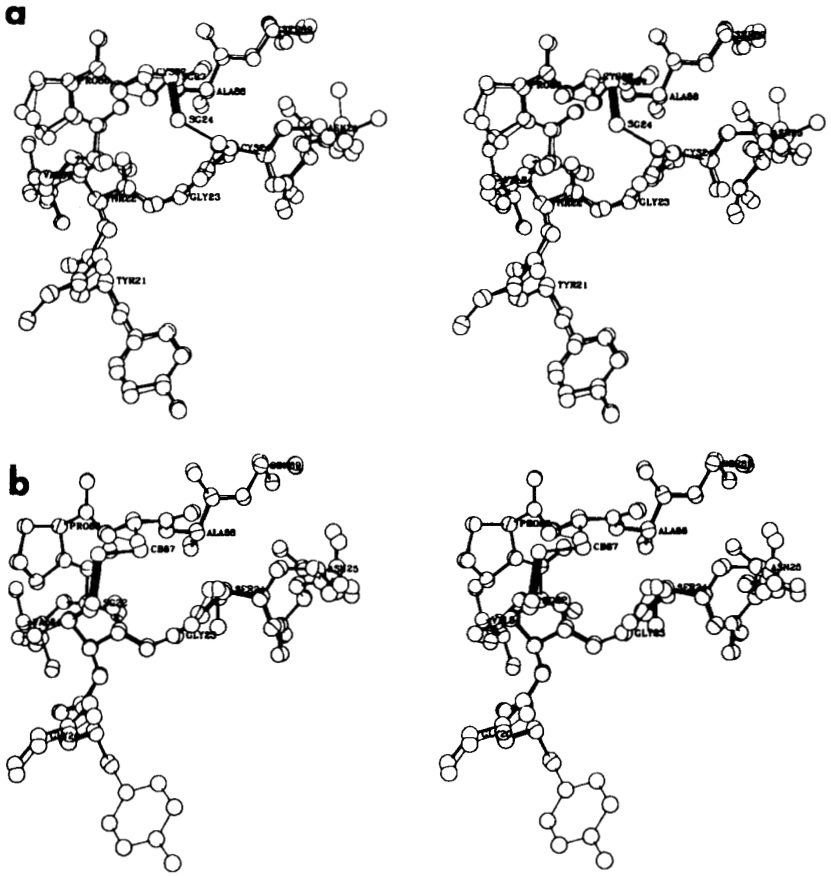


FIG. 3. *a*, stereo plot of C24/C87 superimposed on WT in the region of the disulfide. *b*, stereo plot of C22/C87 superimposed on WT in the region of the disulfide.

FIG. 4. **Backbone of C24/C87 showing the location on the surface of the protein of the Cys-24—Cys-87 cross-link (marked by an *arrow*) with respect to the active site residues (Asp-32, His-64, and Ser-221).** Ser-221 is in the foreground at the end of the long helix running through the middle of the molecule.

structures, that of right-handed ones is somewhat smaller (5.07 ± 0.73 Å in our survey).

Included in Table II are calculated values of the dihedral energies for each disulfide class. On average, the left-handed types have lower calculated dihedral energy (1.68 ± 1.51 kcal/mol *versus* 3.19 ± 1.43 kcal/mol). Crystallographic (13–16) and theoretical (17) studies on small molecules have established that the electronically and sterically preferred values of $X3$ are close to ±83°. In proteins, the average $X3$ dihedral angle for left-handed groups agrees well with these findings ($X3 = -85° \pm 9°$); however, for right-handed groups the average is 99° ± 11°.

Comparison of C24/C87 and C22/C87 Disulfide Geometries with Naturally Occurring Disulfide Geometries—From the current structures in the Protein Data Bank (5) we found that neither Cys-24—Cys-87 nor Cys-22—Cys-87 fit well into any geometric category with respect to all five dihedral angles. The Cys-24—Cys-87 disulfide resembles one family (right-handed family 3), differing significantly in only one dihedral ($X1' = -157°$ *versus* 71°). C24/C87 has a short Sγ87-□87 contact of 3.12 Å.

The Cys-22—Cys-87 disulfide adopts a novel set of dihedral angles. Although $X1$ (53°) is of low dihedral energy, the occurrence in proteins of disulfides with similar $X1$s (or $X1'$s) is rare, presumably because of unfavorable non-bonded interactions between Sγ and the main-chain (12). C22/C87 has a short Sγ22-N22 contact of 3.02 Å and a short Sγ87-N87 contact of 2.92 Å. Short 1,4 sulfur-carbon (methylene or methyl) interactions have been observed in a significant number of small molecule organosulfur compounds (18), and semiempirical molecular orbital calculations suggest that favorable but weak 1,4 sulfur-carbon interactions can occur (18).

The unique geometry of Cys-22—Cys-87 is also reflected in its $X2$ (121°) and $X2'$ (143°) values. The former dihedral is

of essentially maximal $X2$ dihedral energy (the maxima in the $X2$ dihedral energy function occur at 0°, ±120°). (Similar dihedral strain also occurs in some rare cases in natural proteins (*vide infra*).) The dihedral strain energy in disulfide groups can also be partitioned nearly equally among all five dihedrals. The dihedral energy barrier to rotation about a RCH_2—SR bond (defining $X2$ or $X2'$) is lower than the other dihedral energy barriers in a disulfide group, and thus $X2$ and $X2'$ are expected to be the most variable.[9]

Insight into the amount of strain in a disulfide due to its geometry and its local environment can be obtained by measuring the disulfide's redox potential. For example, we find a correlation between the redox potentials of the three disulfide groups in BPTI (19) and their calculated dihedral energies (data not shown). Since the disulfides in C24/C87 and in C22/C87 are in the same region of the molecule and share a common residue, these mutants comprise an ideal system to relate changes in disulfide structures to differences in redox potentials. The disulfide bond in C22/C87 has a higher calculated dihedral energy than that in C24/C87 (4.8 kcal/mol *versus* 2.5 kcal/mol). Wells and Powers (3) experimentally showed that the C22/C87 disulfide is the more strained by determining the redox potentials of C24/C87 and of C22/C87.

Attempts to quantify the thermal stabilities of C24/C87 and of C22/C87 have been complicated by aggregation and autolysis.[10] Although Wells and Powers (3) have shown that neither mutant is more stable toward autolysis than WT, a definitive evaluation of the stabilities of these disulfides must await a full thermodynamic treatment.

With regard to developing strategies for engineering disulfides, it is important to recognize that a disulfide with a non-optimal high energy geometry may still markedly increase thermodynamic stability; disulfides with high energy dihedrals are found in some very stable proteins. For example, reduction of the high energy Cys-14—Cys-38 disulfide in BPTI causes a 21° decrease in the reversible thermal denaturation midpoint temperature at pH 2.1 (20). Removal, by site-directed mutagenesis, of either the Cys-14—Cys-38 cross-link or of another high energy disulfide in BPTI, Cys-30—Cys-51, in each case effects a similar pronounced decrease in the thermal stability of the protein.[11] The decrease in entropy in the denatured state of BPTI presumably outweighs the strain that these disulfides add to the native state. Ueda *et al.* (21) have described the effect of the strain of intramolecular cross-links of various lengths, introduced by chemical modification into lysozyme, on the stability of this protein.

CONCLUSIONS

The primary goals of this structural study were 1) to evaluate the nature and extent of the structural changes that accompany the introduction of disulfide bonds into proteins and 2) to identify the structural and energetic criteria and constraints most important in influencing conformations of disulfide cross-links. Summarizing our results, we come to the following conclusions.

1. Disulfides with atypical dihedral geometries can be inserted into proteins by site-directed mutagenesis. Note, however, that their description as atypical should be qualified by the limited size and accuracy of the data base available.

2. Both engineered and natural disulfides can manifest high energy dihedral angles as well as unfavorable non-bonded interactions. Dihedral strain can be reflected in $X2$ (or $X2'$) dihedral angles near ±120°.

3. Modeling can give reasonable insight into locations in a protein structure where introduction of a disulfide cross-link is possible and into the resulting disulfide's geometry.

4. Insertion of the disulfide is accommodated by the protein's ability to relax around the structural constraints of the cysteine substitutions. The substitutions are accommodated by small subtle structural rearrangements and energy compromises localized about the mutagenesis site.

5. From the point of view of engineering disulfides into structures to increase their stabilities, one need not always be restricted to disulfides of optimal low energy geometries. Indeed, disulfides of relatively high dihedral energy can still be quite effective at imparting to the protein molecule increased stability toward unfolding.

Acknowledgments—We are grateful to J. A. Wells for providing strains and/or proteins used in this study. The initial wild type structure and phases were generously provided by R. Bott. We appreciate the AMBER energy minimization experiments done by S. Rao at the University of California, San Francisco. Critical reviews of this manuscript by J. A. Wells, R. Wetzel, G. Rose, and R. M. Stroud were very helpful.

[9] In subtilisin, we have observed that the dihedrals about the Cγ-Sδ bonds of 3 out of the 5 methionines are near ±120°.

[10] When phenylmethanesulfonyl subtilisin is heated, the Ser-221 Oγ-sulfur bond is hydrolyzed regenerating active enzyme that autolyzes.

[11] S. Anderson and I. D. Kuntz, private communication.

REFERENCES

1. Thornton, J. M. (1981) *J. Mol. Biol.* **151,** 261–287
2. Perry, L. J., and Wetzel, R. (1984) *Science* **226,** 555–557
3. Wells, J. A., and Powers, D. B. (1986) *J. Biol. Chem.* **261,** 6564–6570
4. Weiner, S. J., Kollman, P. A., Case, D., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984) *J. Am. Chem. Soc.* **106,** 765–784
5. Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY
6. Estell, D. A., Graycar, T. P., and Wells, J. A. (1985) *J. Biol. Chem.* **260,** 6518–6521
7. Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allewell, N. M., Kelly, D. M., and Richards, F. M. (1967) *J. Mol. Biol.* **27,** 563–578
8. North, A. C. T., Phillips, D. C., and Mathews, F. S. (1968) *Acta Crystallogr. Sect. A Cryst. Phys. Diffr. Theor. Gen. Crystallogr.* **24,** 351–359
9. Steigemann, W. (1974) Ph.D. thesis, TU Muenchen
10. Hendrickson, W. A., and Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., ed) Vol. 1, pp. 43–57, Pergamon Press, Ltd. Oxford
11. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11,** 268–272
12. Richardson, J. (1981) *Adv. Protein Chem.* **34,** 167–339
13. Beagley, B., and McAloon, K. T. (1971) *Trans. Faraday Soc.* **67,** 3216–3222
14. Rosenfield, R. E., Jr., and Parthasarathy, R. (1974) *J. Am. Chem. Soc.* **196,** 1925–1930
15. Rajewaran, M., and Parthasarathy, R. (1985) *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **C41,** 726–728
16. Jones, D., Bernal, I., Frey, M., and Koetzle, T. (1974) *Acta Crystallogr.* **B30,** 1220–1226
17. Rauk, A. (1984) *J. Am. Chem. Soc.* **106,** 6517–6528
18. Van Wart, H. E., Shipman, L. L., and Scheraga, H. A. (1975) *J. Phys. Chem.* **79,** 1436–1447
19. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol.* **179,** 497–526
20. Vincent, J. P., Chicheportiche, R., and Ladzunski, M. (1971) *Eur. J. Biochem.* **23,** 401–411
21. Ueda, T., Yamada, H., Hirata, M., and Imoto, T. (1985) *Biochemistry* **24,** 6316–6322
22. Wang, D., Bode, W., and Huber, R. (1985) *J. Mol. Biol.* **185,** 595–624