Kinetics of Disulfide Exchange Reactions of Monomer and Dimer Loops of Cysteine–Valine–Cysteine Peptides†

Rumin Zhang and Grayson H. Snyder*

Department of Biological Sciences, State University of New York, Buffalo, New York 14260

Received September 8, 1987; Revised Manuscript Received December 11, 1987

ABSTRACT: Rate constants have been determined in 3 M guanidine hydrochloride for disulfide exchange reactions between glutathione and two synthetic peptides containing a cysteine–valine–cysteine region. Equilibrium experiments demonstrate the absence of noncovalent peptide aggregation in this solvent. Procedures are given for separating seven different components in quenched reactions, including the fully reduced cysteine cluster, the monomeric disulfide loop, parallel and antiparallel dimer loops, and the three monomers containing one or two mixed disulfides with glutathione. Intramolecular rate constants for (1) formation of a sterically strained monomer loop, (2) transfer of glutathione between the two cysteines on the same peptide chain, and (3) formation of unstrained dimer loops correspond to a series of processes forming rings of increasing size. In one sequence, these rate constants are 3, 6, and about 21 s^{-1}, respectively. The larger loops are formed more easily. In the other sequence, rate constants for formation and opening of monomer loops are accelerated 180- and 1300-fold, respectively, relative to analogous reactions in a peptide containing eight residues between the two cysteines. This gives a 7-fold smaller equilibrium constant for ring closure in the cysteine cluster. Dimer formation occurs by a mechanism utilizing the accelerated opening of monomer loops. Results provide information assisting efforts to develop strategies for directing disulfide pairing in novel protein structures. Results also help define factors contributing to formation of undesired oligomers during efforts to refold cysteine-containing proteins obtained by bacterial expression of mammalian genes.

Quantitative information about factors influencing the formation of small disulfide-containing loops in proteins would be useful in optimizing laboratory refolding of protein aggregates produced in *Escherichia coli* and in understanding natural processes such as activaton of membrane receptors, catalysis by thioredoxin reductase, and folding of small protease inhibitors. Disulfide oxidation and reduction may play a role in activation of a variety of membrane receptors (Malbon et al., 1987). An example is acetylcholine receptor, a protein containing a rarely observed disulfide loop formed by two cysteines that are adjacent to each other in the primary sequence. In this protein, the stability of the Cys-Cys loop is coupled to an agonist-induced conformational change at the receptor's binding site (Kao & Karlin, 1986). Small loops of the type Cys-X-X-Cys, where X is a non-cysteine residue, play a functionally important role in proteins involved in oxidation/reduction reactions. Examples include thioredoxin, glutaredoxin, thioredoxin reductase, and disulfide-exchange isomerase (O'Donnell & Williams, 1985; Edman et al., 1985; Klintrot et al., 1984). The redox potentials of these proteins may be dependent on details of the local sequences near the cysteines. Closely spaced cysteines play an important role in the folding of small disulfide-rich proteins. An example is the peanut double-headed trypsin inhibitor (Suzuki et al., 1987). Its sequence includes one, four, two, and one examples, respectively, of Cys-X_0-Cys, -X_1-, -X_2-, and -X_3- regions. In none of these eight regions are the two cysteines bonded to each other in the native protein. Pairing of closely spaced cysteines is sterically hindered. Cysteine clusters therefore assist folding of small proteins by maximizing the total number of final disulfides while simultaneously minimizing the number of energetically favorable combinations of cysteine pairs. When mammalian proteins are expressed in high concentrations in *E. coli*, those proteins frequently are isolated as aggregates that first must be dissociated into reduced monomers in a denaturing solvent (Schoemaker et al., 1985; Winkler & Blaber, 1986). Optimization of conditions for subsequent refolding of those monomers in an oxidizing buffer is a difficult problem. Results presented below will demonstrate that the presence of a Cys-Val-Cys region would significantly exacerbate the problem of dimer formation during that refolding. Greater knowledge of the sequence dependence of factors affecting disulfide exchange might enable genetic engineers to alter the local regions of cysteine clusters in a way that retains native tertiary structure but minimizes contributions to nonnative aggregate formation.

Systematic investigations of factors influencing loop formation are best performed by using a peptide containing two cysteines, such as YSRCVC,[1] and glutathione. Figure 1 depicts some of the disulfide exchange reactions between these

† This work was supported by the National Institutes of Health, Grant GM 26715, and by the Center for Applied Molecular Biology and Immunology at the State University of New York at Buffalo.

[1] Abbreviations: G and GG, reduced and oxidized forms of glutathione; SSG, a glutathione-containing disulfide; IAm, iodoacetamide; IAc, iodoacetate; SCam, carboxamidomethylated sulfur group; obsd, observed; int, intrinsic; subscript 0, initial; eq, equilibrium; concn, concentration; stat, statistical; HPLC, high-pressure liquid chromatography; Ig, immunoglobulin fragment; STI, soybean trypsin inhibitor fragment. Buffer components: guan, guanidine hydrochloride; Tris, tris(hydroxymethyl)aminomethane; TFA, trifluoroacetic acid; EDTA, ethylenediaminetetraacetic acid. The abbreviations for the synthetic peptide Tyr-Ser-Arg-Cys-Val-Cys-Ala and an analogue lacking alanine are YSRCVCA and YSRCVC, respectively. Peptide derivatives: P, the *p*eptide itself; R, *r*educed peptide; L, the monomeric form with an intramolecular disulfide *l*oop; S and D, monomer derivatives with *s*ingle or *d*ouble mixed disulfides; SD, *s*ingle mixed-disulfide *d*imer; LD, *l*oop *d*imer.

0006-2960/88/0427-3785$01.50/0

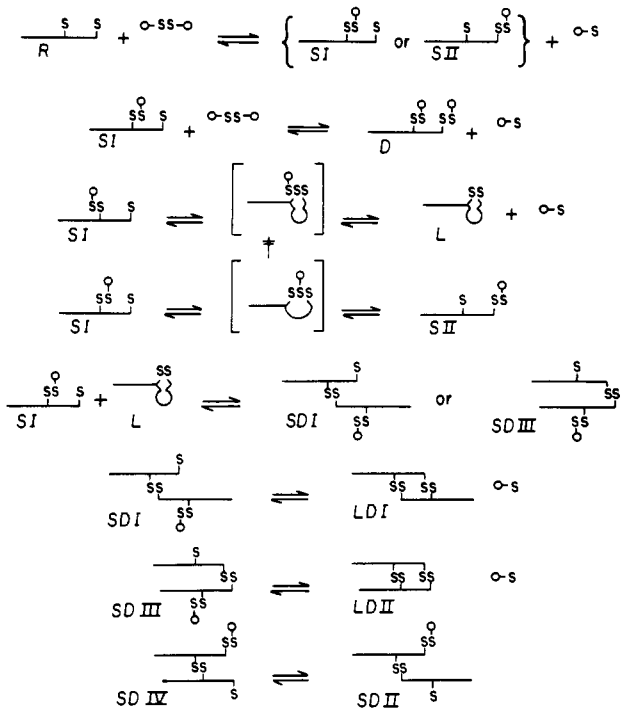


FIGURE 1: Examples of disulfide exchange reactions between glutathione and a peptide containing two cysteines.

molecules. The first reaction illustrates attack of GG by completely *r*educed (R) peptide to produce either SI or SII, containing a *s*ingle mixed (peptide/glutathione) disulfide. Reaction of GG with the internal cysteine of YSRCVC produces SI, whereas reaction with the carboxy-terminal cysteine produces SII. After SI is formed, it may participate in a variety of processes, including those illustrated by the next four reactions in Figure 1. It may react with another GG to form the species with *d*ouble (D) mixed disulfides. Alternatively, the free cysteine of a molecule of SI may attack the mixed disulfide on the same molecule to produce either the *l*oop L or SII. The choice of products depends on which sulfur of the mixed disulfide occupies the central position of the transition state [indicated by a double dagger (‡)]. Finally, the free cysteine of SI may attack one of the sulfurs in a molecule of L to form a *s*ingle mixed-disulfide *d*imer (SD). If it attacks the internal cysteine of L, it will produce the structure SDI by the reaction illustrated in Figure 1. Alternative attack of the carboxy-terminal cysteine of L would produce SDIII. Analogous reactions of SII with L would produce SDII and SDIV, whose structures are included in the figure. SD forms can undergo intramolecular rearrangements to give *l*oop *d*imer (LD) forms. Examples of such reactions are presented in Figure 1 for formation of LDI and LDII structures having peptides aligned in the antiparallel and parallel manners, respectively. The final illustrated reaction depicts direct intramolecular interconversion of SD forms, analogous to the SI/SII interconversion for monomers.

At present there is no direct kinetic data for loop formation and opening in systems with closely spaced cysteines. Analysis of the complete air oxidation of peptides of the type Cys-Gly_m-Cys (Heaton et al., 1956; Jarvis et al., 1961) revealed large amounts of oligomer formation for small values of m and increasing percentages of monomer loop formation as m is increased from 0 to 6. Cysteines separated by only a few residues thus are unlikely to pair with each other, possibly because steric constraints are imposed in a small ring. Relative effective rate constants for reduction of the same series of monomers by G (Weber & Hartter, 1974) exhibit monotonically decreasing values as m is increased from 1 to 4, possibly reflecting bond strain in unstable disulfides when m is small. However, these effective rate constants are difficult to interpret since they actually monitor spectroscopic changes associated with enzyme-catalyzed production of $NADP^+$ in the final step of the series of reactions:

$$L + G \underset{k_{S-L}}{\overset{k_{L-S}}{\rightleftharpoons}} S$$

$$S + G \rightleftharpoons R + GG$$

$$GG + NADPH + H^+ \xrightarrow{\text{glutathione reductase}} 2G + NADP^+$$

Depending on which step is rate limiting, the effective k_{obsd} may be (1) proportional to k_{L-S}, (2) proportional to the equilibrium constant given by the ratio k_{L-S}/k_{S-L}, or (3) neither of the above. Results presented below for peptides containing the Cys-Val-Cys group represent the first determination of microscopic rate constants for disulfide loop formation and opening in a sterically strained sequence.

Experimental Procedures

Iodoacetamide, iodoacetate, GSH, and GSSG were obtained from Sigma Chemical Co. The two forms of glutathione were further purified by isocratic reverse-phase HPLC on a Whatman C-18 preparative column (22 mm × 50 cm), using 0.01% TFA as the solvent.

Reduced synthetic peptides YSRCVC and YSRCVCA were prepared on a Biosearch Model 9500 automated peptide synthesizer, using standard Merrifield chemistry and hydrogen fluoride cleavage and deprotection. Crude peptides were purified by HPLC using the above preparative column and a 5-h, 10 mL/min linear gradient from 6% to 20% CH_3CN in 0.1% TFA. The resultant products eluted as single peaks during rechromatography on an analytical scale. Spectrophotometric assays using Ellman's reagent (Riddles et al., 1979) gave 1.9–2.1 SH/tyrosine by use of a value of 1197 M^{-1} cm^{-1} for the ϵ of tyrosine at 280 nm (Mihalyi, 1969). Amino acid analysis using Waters HPLC protocols (Bidlingmeyer et al., 1984) gave expected results. Products of different repeated syntheses for YSRCVC exhibited identical properties.

Four types of derivatives of these peptides were prepared for use as analytical HPLC standards for purposes of assignment of peaks to particular structures. A mixture of completely oxidized monomeric and oligomeric loops was prepared by overnight oxidation of 0.1 mM R in an oxygen-saturated buffer containing 3 M guan, 5 mM Tris, 5×10^{-7} M cupric sulfate, and 0.5 mM sodium azide at pH 8.3 and 23 °C. Ellman's reagent was used to confirm complete loss of all free SH groups. Oxidized peptides were desalted by acidifying the reaction with TFA, loading the reaction mixture on Millipore Sep-Pak C-18 cartridges, rinsing with 0.1% TFA to remove buffer components, and eluting peptide with 0.1% TFA and 70% CH_3CN. D was prepared by reaction of 2 mM R with 50 mM GG for 1 h at 23 °C in a buffer containing 3 M guan, 30 mM phosphate, and 4 mM EDTA at pH 7.1. The desalted product lacked SH groups and contained mostly D plus minor amounts of monomeric and dimeric loops. $R(SCam)_2$ containing two alkylated cysteines was prepared by reacting 3 mM R with 100 mM IAm for 5 min at 23 °C in the dark, using the above phosphate buffer. IAm-blocked S was prepared by a two-step process. First, 3 mM R containing 6 mM SH was combined with 3 mM IAm for 1 h in the above buffer. This produces a mixture of molecules containing 0, 1, or 2 SCam groups per chain. Then in a second step, remaining SH groups were reacted with 50 mM GG for a further hour. Both steps were performed in the dark.

Subsequent acidification and desalting gave a mixture of $R(SCam)_2$, $S(SCam)_1(SSG)_1$, and D.

L was prepared by the air-oxidation protocol described above, followed by the same preparative HPLC protocol for purifying R. S1 and S2 derivatives of YSRCVC were prepared after preliminary data had given estimates for disulfide exchange rate constants. Computer simulations then were utilized for predicting conditions giving optimal formation of S1 and S2 with minimal formation of R and LD forms. These predicted conditions subsequently were used in experimental preparations and consisted of reaction of 7.5 mM L with 50 mM G for 9 s at pH 6.6 followed by immediate addition of concentrated TFA with vortexing to rapidly lower the pH to a value of 2.0. The above protocols for preparative HPLC then yielded purified S1 and S2.

Peptide concentrations were determined by absorbance at 280 nm. The ϵ's of synthetic peptide derivatives at 280 nm were estimated by using the above value of ϵ for tyrosine and a value of 110 M^{-1} cm^{-1} for a cysteine disulfide (Bailey, 1966). G concentrations were measured with Ellman's reagent. GG concentrations were determined at 248 nm by using $\epsilon = 486 \pm 7$ M^{-1} cm^{-1}. This value was obtained from duplicate experiments consisting of complete air oxidation of a solution containing a known starting concentration of G.

Disulfide exchange reactions were performed in buffers saturated with low-oxygen N_2 (Airco, Inc.). Equilibrium studies and measurements of relatively slow kinetic processes were performed at pH 7.1–7.2. Relatively fast kinetic processes were studied at pH 6.4–6.5. For equilibrium experiments and studies of slow kinetic phases, stock solutions of sulfhydryl- and disulfide-containing molecules were prepared in vials sealed with silicone septums (Pierce Chemical Co.). The disulfide stock solution prepared in a buffer at pH 7 was injected into an unbuffered thiol-containing solution at pH 3 with a gastight Hamilton syringe to initiate the exchange reaction. Final reaction conditions were 3 M guan, 30 mM phosphate, 30 mM Tris, and 2 mM EDTA. For each experiment the pH value, peptide concentration, and glutathione concentration are given in the figure caption. Aliquots subsequently were withdrawn, and disulfide exchange was quenched by 5-fold dilution into TFA giving final conditions of 0.1% TFA at pH 2. These aliquots were immediately frozen at −80 °C. Some aliquots were thawed and assayed for total SH concentration by using Ellman's reagent. Results verified the absence of any measurable net oxidation by residual oxygen during the course of the disulfide exchange reactions. Other frozen aliquots were saved for HPLC analysis, at which time they were thawed and injected directly without desalting. For studies of fast kinetic processes, each time point constituted a separate reaction. Small amounts of G and L were mixed at time zero in a vortexed tube exposed to air. Concentrated TFA was added directly to the vortexed tube a short time later to rapidly quench the entire reaction.

Quenched reacton aliquots were separated into different components by analytical HPLC on a tandem arrangement of two Waters C-18 columns (4 mm × 30 cm followed by 8 mm × 10 cm). Chromatography conditions are given in Figure 2. Data were stored and integrated on an Apple computer using equipment and software supplied by Interactive Microware, Inc. Peak areas were corrected for differences in ϵ values at 280 nm to obtain calculated concentrations of each peptide species.

Procedures for purification and gel electrophoresis of a trypsin inhibitor fragment, STI, were described previously (Snyder, 1987). Experiments described below substitute

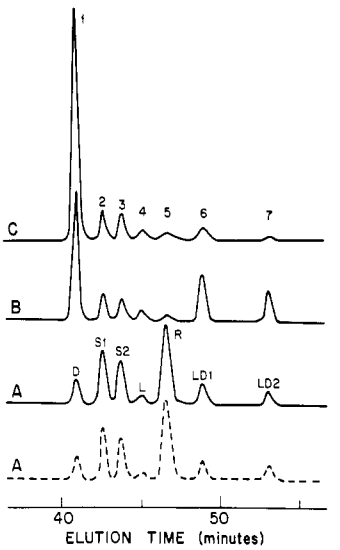


FIGURE 2: HPLC profiles of equilibrium mixtures of YSRCVC, G, and GG. Linear 60-min gradient from 0.1% TFA in H_2O to 0.1% TFA in 25% CH_3CN, detected from 13- to 73-min elution time. Flow rate = 1 mL/min; wavelength = 280 nm. Aliquots were quenched with acid after equilibrium has been reached in reactions where pH 7.16 and T = 23 °C. (A) Total peptide chains = 8.3×10^{-4} M; G = 1.9×10^{-2} M; GG = 3.0×10^{-3} M. (B) Chains = 8.3×10^{-4} M; G = 4.3×10^{-3} M; GG = 1.21×10^{-2} M. (C) Chains = 8.3×10^{-5} M; G and GG as in reaction B. Solid lines are aliquots sampled at 2 h; dashed lines is aliquot sampled at 1 h. Curves are normalized by total integrated area.

negative glutathione for the neutral *N*-acetylcysteine methyl ester used previously and substitute 1.0 M neutral IAm for the 0.1 M negative IAc quenching reagent employed before. The previously described procedures for purifying IAm and IAc also were employed here.

RESULTS

Assignment of HPLC Peaks of YSRCVC Derivatives. Figure 2 presents data for equilibrium mixtures of YRSCVC and glutathione. To ensure that equilibrium has been reached, aliquots were taken at half-hour intervals after the reaction was initiated. For reaction A, the 1-h data (dashed line) and 2-h data (solid line) are identical, indicating that equilibrium was reached before 1 h. Comparison of reaction A with reaction B reveals the effects of altering the G/GG ratio while a constant peptide concentration is maintained. Comparison of reaction B with reaction C reveals the effects of lowering total peptide chain concentration by a factor of 10 while the same G/GG ratio is maintained.

The standards D and R elute at positions of peaks 1 and 5, respectively, thereby identifying these peaks. The air-oxidized standard is a mixture of three species eluting at positions of peaks 4, 6, and 7. These peaks therefore correspond to completely oxidized loops lacking free SH or SSG groups. Derivatives of S have a side-chain composition that is intermediate between D and derivatives of R. If Y represents the SSG group and X represents the SCam or SH group, the structures of cysteine side chains of D, S, and R are given by Y_2, Y_1X_1, and X_2, respectively. This causes S derivatives to elute between D and R. On this basis, peaks 2 and 3 are initially assigned to S structures. S appears as two peaks, since SI and SII structures illustrated in Figure 1 interact slightly differently with the column. It is not known which structure (I or II) corresponds to which peak (1 or 2). This initial assignment is confirmed by examining HPLC profiles of aliquots quenched with 1.0 M IAm instead of TFA. The SCam group elutes earlier than the SH group in the TFA/CH_3CN

Table I: Determination of Monomer/Dimer/Trimer Nature of Air-Oxidized Species

peaks	test	ratio	equilibrium values[a] reaction B	reaction C
L	monomer	(G)(L)/(S1) × 10^3 (M)	2.4 ± 0.1	2.0 ± 0.1
LD1	monomer	(G)(LD1)/(S1) × 10^3 (M)	5.9 ● 0.3	0.6 ± 0.1
LD1	dimer	(G)(LD1)/(L)(S1) × 10^{-2}	2.1 ± 0.1	1.8 ± 0.3
LD1	dimer	(G)(G)(LD1)/(S1)(S1) × 10 (M)	4.3 ± 0.3	3.2 ± 0.6
LD1	trimer	(G)(G)(LD1)/(S1)(S1)(L) × 10^{-4}	1.6 ± 0.2	9 ± 2
LD1, LD2	similarity	(LD2)/(LD1)	0.64 ± 0.01	0.57 ± 0.01

[a] Reaction conditions are given in Figure 2.

gradient, causing peaks 2, 3, and 5 to elute 2, 2, and 6 min earlier, respectively, in the IAm-quenched aliquots. Shifted peaks 2 and 3 elute at positions exhibited by the $S(SCam)_1$-$(SSG)_1$ standard sample. The R peak is shifted more than the two S peaks, since alkylated R has twice as many SCam groups. Assignment of peaks to D, S, and R forms are further confirmed by comparisons of reactions B and A. Reaction B corresponds to a higher GG/G ratio promoting disulfide formation. This gives D > S > R since these species have 2, 1, and 0 disulfides, respectively. In contrast, reaction A with a higher G/GG ratio promotes sulfhydryl formation, thereby giving R > S > D since these species have 2, 1, and 0 SH groups, respectively.

Formation of oligomers depends on collisions between monomers, with such collisions decreasing at lower peptide concentrations in reaction C. The last two peaks in profile B decrease in profile C but maintain the same ratio of areas relative to each other. This indicates that peaks 6 and 7 correspond to oligomers with identical numbers of chains per molecule. The first five peaks increase in reaction C while maintaining the same areas relative to each other. Thus, they share a common but different number of chains per molecule. Peaks 1, 2, 3, and 5 have been assigned to monomers D, S1, S2, and R, respectively, by comparison of their elution times with peptide standards. Thus, peak 4 represents a monomer also, the loop L.

Quantitative confirmation of the assignment of peak L to a monomeric loop is obtained by determining whether the peak satisifies the equilibrium constant expression for intramolecular loop formation. As seen in Table I, the ratio (G × L)/S1 has the constant values 2.2 × 10^{-3} M at equilibrium. Peak LD1 fails the same monomer test.

Table I includes two tests of the possibility that peak 6 (LD1) corresponds to a dimer loop. The first test considers collisions of L and S, whereas the second test considers a mechanism involving collisions between two S molecules. At equilibrium, there are no further net changes in concentrations of any species. All possible reactions are simultaneously at equilibrium. Constant values should be obtained for an equilibrium ratio test corresponding to any possible pathway for generating the correct structure, independent of whether that pathway contributes significantly to the kinetic approach to final equilibrium.

Suppose that peak 2 corresponds to structure SI and that peak 6 is the antiparallel dimer structure, LDI. LDI could be formed by the following reactions that are illustrated in Figure 1:

$$L + S1 = SDI$$

$$SDI = LDI + G$$

$$L + SI = LDI + G \text{ (overall)}$$

Then at equilibrium, the ratios (LDI × G)/(L × SI) and (peak 6 × G)/(L × peak 2) would be constants. If SI elutes in peak 3 instead of in peak 2, then peak 2 still is proportional to SI since peak 3_{eq}/peak 2_{eq} = 0.9 in all reactions. Thus, it still would be true that the latter ratio above is a constant. Similarly, if LDI elutes in peak 7 insted of in peak 6, peak 6 still would be proportional to LDI. A reaction series relating loop dimers is

$$LDII + G = SDIII$$

$$SDIII = SDI$$

$$SDI = LDI + G$$

$$LDII + G = LDI + G \text{ (overall)}$$

Thus, the ratio (LDII × G)/(LDI × G) = LDII/LDI is a constant at equilibrium. As seen in Table I, peak 7/peak 6 = 0.6 in all equilibrium samples. Thus, if peak LD1 corresponds to a loop dimer, one should obtain a constant ratio for (LD1 × G)/(L × S1) at equilibrium, independent of the assignment of S1 or LD1 peaks to particular type I or type II structures. This test was passed, as was an alternative dimer test for collisions of two S molecules. One of many alternative trimer tests is given in Table I and was failed by peak LD1.

The LD2 peak passed the same tests as the LD1 peak. These two peaks probably correspond to the two structural types I and II for antiparallel and parallel chain arrangement, although it is not known which structural type elutes first. An ability to resolve parallel and antiparallel loop dimers by HPLC has been reported previously for a fibrinogen-related peptide (Hoeprich & Doolittle, 1983).

The purpose of 3 M guanidine in these experiments is to provide a denaturing solvent that prevents formation of noncovalent aggregates of YSRCVC peptides. This should generate the simplest possible behavior and thereby provide a frame of reference for consideration of dimer formation in more complicated cases. As an example, suppose that dimeric aggregates of loops exist prior to glutathione addition. Then the kinetics of covalent dimer formation might be described as follows, where L..L and L..SI are noncovalent aggregates:

$$L + L = L..L$$

$$L..L + G = L..SI$$

$$L..SI = SDI$$

$$SDI = LDI + G$$

The forward rate constant for the third step in this series corresponds to intraaggregate disulfide exchange. It would exhibit the properties of a unimolecular process that does not require collisions of L and SI. If initial concentrations of L were 100-fold greater than the equilibrium constant for L..L dissociation, 10-fold dilution of starting L would not significantly diminish the high percentage of L in the aggregated form. In this case the percentage of total peptide that exists as LDI dimers at equilibrium also would not be significantly diminished. In contrast, formation of LDI dimers of YSRCVC in Figure 2 is dependent on total peptide concentration in the quantitative manner expected for a mechanism requiring bi-

Table II: Equilibrium and Rate Constants for YSRCVC Reactions[a]

reaction	reactants	products	stat[b] factor	rate constants[c] forward	rate constants[c] reverse	equilibrium constants: equilibrium concn and stat factors	equilibrium constants: k_f/k_r[e]
1	R, GG	S1, G	2	14	*d*	1.4 ± 0.4	
2	R, GG	S2, G	2	11	*d*	1.2 ± 0.3	
3	S1	L, G	1	3.1 s^{-1}	2.1×10^3	$(2.1 \pm 0.3) \times 10^{-3}$ M	1.5
4	S2	L, G	1	2.7 s^{-1}	1.5×10^3	$(2.3 \pm 0.4) \times 10^{-3}$ M	1.8
5	S1	S2	1	6.0 s^{-1}	*d*	0.90 ± 0.01	
6	S1, GG	D, G	2			1.3 ± 0.1	
7	S2, GG	D, G	2			1.5 ± 0.1	
8	S1, L	LD1, G		4.6×10^2	2.6	$(1.9 \pm 0.2) \times 10^2$	1.8
9	S2, L	LD1, G		4.6×10^2	2.3	$(2.1 \pm 0.2) \times 10^2$	2.0
10	S1, L	LD2, G		6.5×10^2	*d*	$(1.1 \pm 0.1) \times 10^2$	
11	S2, L	LD2, G		6.5×10^2	*d*	$(1.2 \pm 0.1) \times 10^2$	
12	LD1, G	LD2, G		3.1	4.7	0.60 ± 0.04	0.67

[a] First seven reactions, single equilibria; last five, multiple equilibria. [b] Tabulated values are for forward direction; reverse statistical factors = 1 in all cases. [c] Units are s^{-1} M^{-1}, except where otherwise noted; errors are ±25%. [d] The reverse rate constant was calculated from the equilibrium constant. [e] Units and magnitudes are the same as in the concentration ratios; errors are ±25%.

molecular collisions of nonaggregated chains. Thus, 3 M guanidine achieves the simplifying effects desired in these studies.

Equilibrium Studies of Reactions of YSRCVC with G. Table II presents equilibrium constants for reactions of YRSCVC. Values are averages of results calculated from reactions A–C of Figure 2. For the first seven reactions, which are single-step reactions, tabulated values are K_{int}. For the last five reactions, tabulated values are statistically *un*corrected K_{obsd} for overall equilibria of multistep processes. An example of calculation of an intrinsic equilibrium constant is given by consideration of reaction 1. At equilibrium, the forward and reverse velocities are equal. R can attack either of the two equivalent sulfurs in GG. This gives

$$2k_{R\text{-}S1}[R]_{eq}[GG]_{eq}f_{S^-} = k_{S1\text{-}R}[S1]_{eq}[G]_{eq}f'_{S^-}$$

The terms f_{S^-} and f'_{S^-} give the fraction of R and G cysteine groups, respectively, in the chemically reactive thiolate anion form. For a simple isolated titrating group, this fraction is given by

$$f_{S^-} = (1 + 10^{pK-pH})^{-1}$$

Table II employs the simplifying assumption that all cysteines exhibit identifical pK_a values and thus identical f_{S^-}, a simplification that will be considered in greater detail below. This gives

$$K_{obsd} = \frac{[S1]_{eq}[G]_{eq}}{[R]_{eq}[GG]_{eq}} = \frac{2k_{R\text{-}S1}}{k_{S1\text{-}R}} = 2K_{int}$$

The simple disulfide exchange reactions converting R to S or converting S to D exhibit K_{int} values ranging from 1.2 to 1.5. Since these values are close to 1, there is no tendency for these reactions to lie to the right or left. Both disulfide types G–SS–G and P–SS–G therefore exhibit approximately equal stabilities. This is expected for the nonterminal cysteines represented by glutathione and cysteine 4 of YSRCVC. Equilibrium constants for reactions 8–11 describing formation of LD species from L and S species have values ranging from 100 to 200. This indicates that the combined stabilities of the two disulfides in L and S are less than the combined stabilities of the two disulfides in LD. One possible explanation is that the small loop in L is sterically strained.

Kinetics of Reaction of YSRCVC with Glutathione. Figure 3 exhibits data for reduction of the monomer loop of YSRCVC by a 10-fold excess of glutathione. There are three distinct kinetic phases. During the first minute, there is a fast kinetic phase, where L decreases from 100% to 26% and where S1 and S2 increase from 0% to 33% and 27%, respectively. At the end of this minute the L, S1, and S2 species are in approximate equilibrium with each other and establish a cyclic triangular system of three coupled reactions.

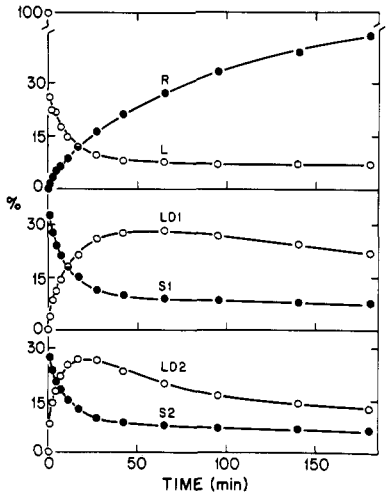
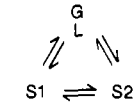


FIGURE 3: Overall kinetics of reduction of monomer loop of YSRCVC. $L_0 = 1.74 \times 10^{-4}$ M; $G_0 = 2.0 \times 10^{-3}$ M, pH 7.11, T = 23 °C. No D was observed. Ordinate gives percentage of total peptide chains in each species. Solid lines are results of numerical integrations employing the complete set of rate constants in Table II. Error of ±1% of total peptide chains is given by circle diameters.



During the next 70 min there is an intermediate kinetic phase leading to a buildup of LD2 and LD1 to maximum values. During this buildup L, S1, and S2 decline in parallel, maintaining approximately the same relative ratios. At the end of 70 min, L, S1, S2, LD1, and LD2 are in approximate equilibrium with each other. During the final slow phase they decline in parallel as fully reduced R continues to build up to its equilibrium value. No D is observed in this reducing reaction.

The observation of three distinct kinetic phases has implications for the relative rank order of particular bimolecular disulfide exchange rate constants. A typical value for disulfide exchange between small cysteine-containing molecules such as glutathione is 10 s^{-1} M^{-1}. The slow kinetic phase, reflecting the reaction of S with G to produce R, is expected to exhibit such a value and will be shown below to do so. Since the velocity of the fast-phase reaction of L with excess G is much greater than the velocity of the slow-phase reaction of S with

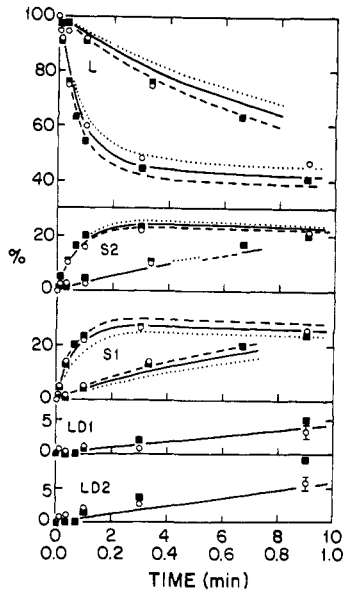


FIGURE 4: Fast phase of reduction of monomer loop of YSRCVC. $L_0 = 8.7 \times 10^{-5}$ M; $G_0 = 1.01 \times 10^{-3}$ M, pH 6.43, $T = 23$ °C. No R or D was observed. LD1 and LD2 data are plotted by using the 10-min scale; other species are plotted on both 10- and 1-min scales. Circles and squares correspond to two independent repetitions of the experiment. Error of ±1% is given by error bars in expanded plots for LD1 and LD2 and by symbol sizes in other plots. Solid lines are as in Figure 3. Dashed lines, k_{L-S1} increased by 25%; dotted lines, k_{L-S1} decreased by 25%.

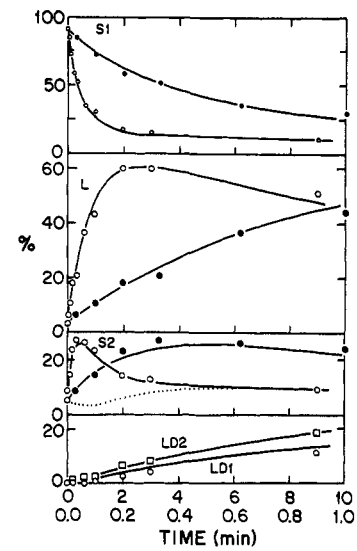


FIGURE 5: Kinetics of transfer of the glutathione group between the two cysteines of YSRCVC. $S1_0 = 4.2 \times 10^{-4}$ M; $S2_0 = 2.3 \times 10^{-5}$ M; $L_0 = 1.6 \times 10^{-5}$ M, pH 6.51, $T = 23$ °C. Open symbols correspond to the 10-min scale; filled symbols duplicate early time points on the expanded 1-min plot. Dotted line is the simulation for first 10 min if the S1/S2 exchange reaction is deleted. Solid lines as in Figure 3.

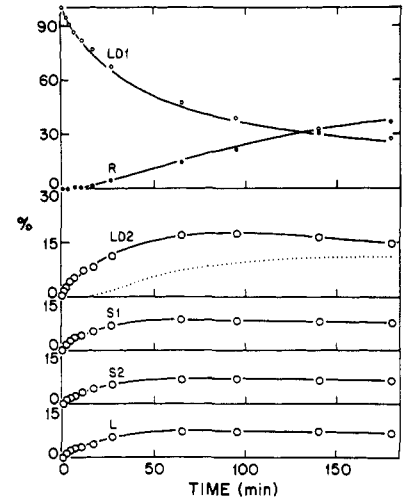


FIGURE 6: Kinetics of reduction of one of the dimer loop forms of YSRCVC. $LD1_0 = 9.5 \times 10^{-5}$ M; $G_0 = 2.0 \times 10^{-3}$ M, pH 7.16, $T = 23$ °C. Dotted line is the simulation if the G-catalyzed LD1/LD2 exchange reaction is deleted. Other features as in Figure 3.

the same excess G, the disulfide of L is reduced at greatly accelerated values relative to typical disulfide exchange reactions. This suggests that the disulfide of this small loop is strained and consequently unstable. The intermediate kinetic phase responsible for forming dimers requires productive collisions of a peptide with a peptide. This occurs faster than the slow-phase productive collisions of a peptide (S) with G, in spite of the fact that G is 10 times more concentrated than any peptide species. This indicates an accelerated disulfide exchange rate constant for dimer formation. Given the above suggestion of strain in L, it therefore is reasonable to propose that accelerated dimer formation is dominated by the mechanism whereby the L disulfide is reduced by the SH group of S1 or S2.

The fast phase in Figure 3 is not well-defined. The experiment was repeated using lower pH and lower starting concentrations to decrease the velocity of this phase. Lowering the pH from 7.1 to 6.4 decreases f_S- by a factor of 5. Twofold lower concentrations of both loop and glutathione decrease collisions by a factor of 4. Data are presented in Figure 4. Two independent repetitions of the experiment employing different synthetic preparations of YSRCVC demonstrate the reproducibility of the data. The fast phase is well-defined and complete in 3 min. This means that the fast phase in Figure 3 would have been complete in $^1/_{20}$ of 3 min, which is about 10 s. This is well before the 1-min sample was quenched with TFA in that experiment. The efficiency of the acid-quenching protocol will be discussed below, comparing time frames for the fastest disulfide exchange processes with the time frame for acid protonation of thiolate groups.

Figure 5 presents data for a pH-jump procedure capable of monitoring intramolecular conversion of S1 to S2. In this experiment, S1 purified at pH 2 and dissolved at pH 3.0 is rapidly raised to pH 6.5 at time zero to increase the f_S- 3000-fold and thereby facilitate S1 conversion to S2 and L. At time zero, some of the peptide is present as S2 and L because some disulfide exchange occurred at pH 3 during preparation and assays of the starting stock solution of S1. After the pH jump, both S2 and L build up to maximum concentrations, with maximum S2 occurring first. This indicates that one is observing at least some direct conversion of S1 to S2, without requiring that all S2 be made indirectly by the S1 to L to S2 route. It also indicates that S1/S2 conversion is a pH-dependent process, consistent with assignment of these two peaks to the two different structural types, SI and SII.

Figure 6 presents data for reduction of LD1. During the first 90 min LD2, L, S1, and S2 build up to a maximum value as dimer breaks down into monomers. At longer times, LD1 and these species decline in parallel during the final increase of R to its final equilibrium value. In effect, Figure 6 thus incorporates both the intermediate and slow phases of Figure 3, with the intermediate phase operating in the reverse direction. Data for reduction of LD2 are not shown but are qualitatively similar to the data of Figure 6 and are fit equally well as LD1 data by the final Runga–Kutta simulation.

Computer Modeling of Kinetic Data. Quantitative data analysis requires consideration of sulfhydryl pK_a values. In 3 M guanidine, the sulfhydryl groups in STI and G exhibit pK_a values of 8.9 ± 0.1 (Snyder, 1984). By use of similar procedures, the pK_a of *N*-acetylcysteine was determined to be 9.36 ± 0.02. *N*-Acetylcysteine and cysteine 6 of YSRCVC represent carboxy-terminal cysteines. STI, G, and cysteine 4 of YSRCVC represent nonterminal cysteines. Since the exact correspondence between the peaks S1 and S2 and the structures SI and SII is not known, the model used for numerical integrations temporarily assigns a pK_a of 8.9 to all cysteines. The carboxy-terminal cysteine with a higher pK_a is expected to exhibit slower effective kinetics at pH 7 resulting from lower f_{S^-} but also should exhibit faster effective kinetics resulting from greater nucleophilicity (Szajewski & Whitesides, 1980). To some extent these competing effects should cancel out. Final rate constants for S1 and S2 differ by no greater than a factor of 1.4 for YSRCVC. Since one of the forms represents reactions of a nonterminal cysteine and therefore correctly employs the pK_a value of 8.9, comparisons and conclusions based on the general order of magnitude of calculated *k*'s will be valid.

Kinetic results were simulated by Runga–Kutta–Fehlberg numerical integration (Danby, 1985) of differential equations representing the reactions in Table II. The initial step was to fit the fast kinetic phase data for establishing the L/S1/S2 cyclic equilibrium. Ratios of forward and reverse rate constants for the three contributing reactions are strongly determined by the need to fit the pseudoequilibrium levels of L, S1, and S2 observed at the end of the fast phase in Figure 4. Accurate determinations also are helped by the fact that the L/S equilibria are approached from both possible directions, starting either with L in Figure 4 or with S forms in Figure 5. In these simulations, the value of k_{S2-S1} was calculated from the relationship

$$k_{S1-S2}/k_{S2-S1} = K_{eq} = S2_{eq}/S1_{eq}$$

by using the parameter k_{S1-S2} and the equilibrium constant determined in separate equilbrium experiments. Two different people independently attempted to arrive at a set of values for k_{S1-L}, k_{L-S1}, k_{S2-L}, k_{L-S2}, and k_{S1-S2} giving an optimum fit for all experiments. These independent efforts gave values within a factor of 1.3 of each other for each of the five pairs of rate constants. Figure 4 indicates the results of increasing or decreasing k_{L-S1} by 25%. Most data points lie between curves for the two extremes. Each rate constant was considered in a similar manner and suggested that ±25% is a reasonable error estimate for each rate constant.

Results are given in Table II. Loop reduction exhibits an average rate constant of 1800 s^{-1} M^{-1}, which is 180-fold faster than a typical disulfide exchange rate constant of 10 s^{-1} M^{-1} for nonterminal cysteines. Intramolecular passing of the glutathione group from one cysteine to the other occurs with an average *k* equal to 6 s^{-1}, which is twice the size of the average rate constant for the loop-closing reactions (3 s^{-1}). Values of equilibrium constants calculated from rate constants agree within a factor of 1.4 of values calculated from statistical factors and ratios of equilibrium concentrations. Simulations also were attempted by using a model that deletes the fast S1/S2 reaction from consideration. As seen by the dotted line in Figure 5, there would be a significant lag in formation of S2 if one required that the S2 species only be formed by the long route which first converts S1 to L and subsequently converts L to S2.

The next step was to stimulate the medium and slow kinetic phases. The model was made as simple as possible (1) by employing overall equilibria for multistep processes, (2) by occasionally assigning a single value to S1 and S2 processes, and (3) by trying to match ratios of rate constants with values of independently determined equilibrium constants. Results of these fits show that the average overall rate constant for dimer formation is 550 s^{-1} M^{-1}, about 55-fold accelerated relative to rate constants of typical disulfide exchange reactions. This reaction has been assigned to a mechanism whereby an S species collides with a strained monomer loop. When an alternative mechanism was added that permitted collisions between two S forms with a maximum overall forward rate constant of 10 s^{-1} M^{-1}, this addition had no significant influence on the simulated curves. Simulations also were constructed by using an abbreviated model that deletes the fast LD1/LD2 interconversion reaction from consideration. This abbreviated model requires the complete breakdown of LD1 into monomers before LD2 can be re-formed. As seen by the dotted line in Figure 6, this would give a significant lag in production of LD2. Thus, it is necessary to propose that single-dimer forms (SD) can undergo fast intramolecular rearrangements just like single-monomer forms (S). Such a mechanism for LDI/LDII interconversion was described earlier.

Table II includes 13 rate constant parameters. There are three reasons why this is *not* too large a number to determine in this system. These reasons are as follows: (1) Each of seven different peptide species is fully resolved in HPLC chromatograms; (2) kinetic experiments were initiated from four different starting points (L, S1, LD1, and LD2); and (3) there are two or three well-separated kinetic phases in each experiment.

Disulfide Exchange Reactions of YRSCVCA and STI. The fast kinetic phase experiment illustrated in Figure 4 for YSRCVC was repeated using YSRCVCA. HPLC peaks eluted 8–11 min later in this sequence containing the nonpolar alanine (A) residue. Data from reproduced experiments are not shown but were similar in quality to the data of Figure 4. The ratio of S2/S1 after equilibrium was reached is 0.9, as before. The principal difference was observation of faster kinetics for all reactions. In YSRCVCA, k_{S1-L} and k_{S2-L} are 14 and 9.0 s^{-1}, respectively. Rate constants for loop opening, k_{L-S1} and k_{L-S2}, are 1.3×10^4 and 7.1×10^3 s^{-1} M^{-1}, respectively. Errors are ±25% as before. This gives equilibrium constants of $(1.1 \pm 0.2) \times 10^{-3}$ and $(1.3 \pm 0.3) \times 10^{-3}$ M for formation of L from S1 or S2. The equilibrium constants for YSRCVCA and YSRCVC derived from ratios of rate constants differ by a factor of 1.3 or less. In contrast, each of the four rate constants in YSRCVCA is 3–6 times faster than the corresponding value for the YSRCVC sequence given in Table II. Thus, the presence of a backbone carboxyl group adjacent to the Cys-Val-Cys sequence slows the kinetics of loop formation and opening but does not significantly affect the loop stability in 3 M guanidine at pH 7. Rate constants for S1 and S2 in YSRCVCA differ by a maximum factor of 1.8, demonstrating some inequivalence for the two nonterminal cysteines of this peptide.

The two cysteines in STI are nonterminal and are separated by eight intervening residues in the primary sequence. Loop formation in this fragment has been studied previously (Snyder, 1987) by using the negative IAc quenching reagent and the neutral cysteine analogue, *N*-acetylcysteine methyl ester. Experiments were repeated here employing the neutral IAm quenching reagent and the negative cysteine analogue, G. The effective charge of the G group, which contains two carboxyls and one amine, was observed to be −1 in the same

gel system used previously. Thus, it was possible to duplicate the previous ability to resolve L, D, S, and R species as four peaks on densitometer scans of stained slab gels, although D, S, and R migrated in the opposite order relative to the previous experiments. Data are not shown but gave $k_{S-L} = 0.065 \pm 0.009\ s^{-1}$, $k_{L-S} = 7.5 \pm 0.5\ s^{-1}\ M^{-1}$, and $k_{S-R} = 10 \pm 1\ s^{-1}\ M^{-1}$. The ratio of intrinsic rate constants for loop formation [$k_{S-L}/k_{L-S} = (1.1 \pm 0.2) \times 10^{-2}$ M] agreed with the equilibrium constant [$K_{S-L} = (0.9 \pm 0.2) \times 10^{-2}$ M] obtained from ratios of equilibrium concentrations in separate experiments. The average of the one k_{L-S} value for STI and three total k_{S-R} values for both STI and YSRCVC is $11 \pm 3\ s^{-1}\ M^{-1}$. The loop-opening rate constant in STI therefore exhibits a value typical of simple cysteine disulfide exchange reactions. Comparison of STI with the small nonterminal loop of YSRCVCA reveals that the smaller loop exhibits 180-fold faster loop formation and 1300-fold faster loop opening, giving an equilibrium constant K_{S-L} for loop formation that is 7-fold smaller than the value observed in STI.

Discussion

Choice of the Sequence YSRCVC. The sequence YSRCVC was selected as part of a program to develop strategies for directing disulfide pairing in novel protein structures. It imitates the carboxy-terminal end of a hypothetical four-cysteine sequence given by $^{+}$Cys-X^{+}...$^{-}$X-Cys-X^{-}...$^{+}$X-Cys-Cys^{-}. Such a sequence in principle can form three different combinations of disulfides, represented as ($1^{+}2^{-}$, $3^{+}4^{-}$), ($1^{+}3^{+}$, $2^{-}4^{-}$), and ($1^{+}4^{-}$, $2^{-}3^{+}$). The numbers give the rank order of cysteine positions in the primary sequence. Local electrostatic environments of each cysteine include charged side chains and the two backbone terminal ends. X is any amino acid containing the appropriate charge. For a small Cys-X-Cys loop, the $3^{+}4^{-}$ disulfide in the first combination above would be disfavored by steric constraints. The latter two combinations would have similar topologies consisting of two loops joined at the cysteine cluster. The ($1^{+}3^{+}$, $2^{-}4^{-}$) combination that juxtaposes like charges would be disfavored electrostatically. The ($1^{+}4^{-}$, $2^{-}3^{+}$) structure bringing together regions of opposite electrostatic charge would be favored. Thus, a combination of steric and electrostatic effects in this hypothetical sequence may operate to direct disulfide formation toward one specific combination out of the three possible forms. The results of the above experiments demonstrated a diminished equilibrium constant for loop formation in a Cys-Val-Cys region, consistent with the direction of effects desired for protein engineering applications. The selection of ...$^{+}$Arg-Cys-Val-Cys^{-} was based on the general absence of Cys-X-Cys loops in native proteins (Thornton, 1981) and a desire to simplify analysis by limiting available conformational space. The branching on the β carbon of valine's side chain limits available choices of backbone dihedral angles at the adjacent α carbon. An example of a protein containing two Cys-Val-Cys regions is the peanut double-headed trypsin inhibitor (Suzuki et al., 1987). The $^{+}$Tyr-Ser... soluble leader provides a convenient method for detecting and assaying peptide concentrations, taking advantage of the known value and dominant contribution of tyrosine to absorbance at 280 nm.

Efficiency of the Acid Quenching Protocol. Accurate determination of rate constants and equilibrium constants requires that the relative concentrations of species that are separated by HPLC or gel electrophoresis directly correspond to the relative concentrations of related species in reaction aliquots prior to addition of the quenching reagent. Quenching by TFA slows down disulfide exchange by lowering the pH and thereby decreasing f_{S^-}. The time frame for conversion of S^- to SH is very fast. Reaction of a cysteine S^- group with H_3O^+ is a diffusion-limited process exhibiting a rate constant $k \geq 10^9\ s^{-1}\ M^{-1}$ (Eigen, 1964). This recombination has a $\tau = (k[H_3O^+])^{-1}$, which therefore gives $\tau \leq 10^{-7}$ s at pH 2. This is the pH of acid-quenched aliquots and HPLC gradients in the above experiments. Thus, addition of a reaction aliquot to concentrated TFA drops the pH to a value of 2.0 within the time required for mixing in the vortexed tube. This mixing time is estimated to be ≤ 1 s.

For YSRCVCA, the fastest observed disulfide exchange velocities are the intramolecular formation of monomer loop and the reduction of that strained loop by excess G. Consider first the quenching of aliquots of reactions where S1 and L are in actual or approximate equilibrium with each other. Prior to addition of TFA the loop-closing and -opening velocities are identical. Time frames for these processes are given by $\tau_{closing} = (k_{S1-L}f_{S^-})^{-1}$ and $\tau_{opening} = (k_{L-S1}[G]f_{S^-})^{-1}$, respectively. During the 5-fold dilution of the aliquot by addition of excess TFA, bimolecular collisions between L and G will decrease. This dilution process therefore could give a possible net shift of S1 to L. The fastest possible time for this shift occurs at the highest pH value (7.1) prior to the pH drop and is given by $\tau_{closing} = 6$ s for a cysteine with a $pK_a = 8.9$. Since the pH drop occurs within 1 s, the pH will equal 2.0 before there is any significant shift of S1 to L. At pH 2, f_{S^-} values are approximately 10^5 smaller than at pH 7. Calculations of τ's for disulfide exchange at pH 2 give $\tau \geq 2$ days for all YSRCVCA processes, which is much longer than the 1 h required for HPLC separations. Finally, consider reaction aliquots where the S1/L reaction is not at equilibrium. Such aliquots are generated in experiments at pH 6.4 examining details of the initial fast kinetic phase of reduction of L by G. Here, $\tau_{closing} = 23$ s and $\tau_{opening} = 24$ s, both slow relative to the pH drop. Thus, the acid-quenching protocol meets all necessary requirements for studies of YSRCVC and YSRCVCA.

Possible Microscopic Model for Dimer Formation. There are four overall reactions for formation of dimers from L and S species, given by reactions 8–11 of Table II. Each of these overall reactions actually consists of two consecutive steps. For example, reaction 8 for formation of LD1 from S1 is given by

$$L + S1 \underset{1}{\overset{1}{\rightleftharpoons}} SD1 \underset{2}{\overset{1}{\rightleftharpoons}} LD1 + G$$

where numerals 1 and 2 are statistical factors. It is useful to try to estimate magnitudes of the four microscopic rate constants contributing to this reaction in YSRCVC. The reverse rate constant k_{SD1-L} corresponds to the same type of loop formation as in the k_{S-L} reaction and thus probably has a value of about $3\ s^{-1}$. The reverse rate constant $k_{LD1-SD1}$ probably has a value near 10 or $11\ s^{-1}\ M^{-1}$, the average of all measured simple bimolecular exchange rate constants for YSRCVC and STI. The overall rate constant for converting LD1 to SD1 to L is $2.6\ s^{-1}\ M^{-1} = 2(1.3\ s^{-1}\ M^{-1})$. This is $^1/_8$ as fast as expected for the microscopic step converting LD1 to SD1, given by $2(10\ s^{-1}\ M^{-1})$. One possible explanation is that an SD1 intermediate produced from LD1 proceeds to L $^1/_8$ of the time but reverts to LD1 $^7/_8$ of the time, thereby decreasing the effective velocity for overall LD1 to L conversion by a factor of 8. This possibility may be represented as

$$S1 \underset{1 \times FAST}{\overset{k_{S1-SD1}[L]}{\rightleftharpoons}} SD1 \underset{2k_{LD1-SD1}[G]}{\overset{7 \times FAST}{\rightleftharpoons}} LD1$$

Using the above estimates for $k_{LD1-SD1}$ and k_{SD1-L}, letting $k_{S1-SD1} = 500\ s^{-1}\ M^{-1}$, and using experimental concentrations

for L_0 and G_0 found in Figure 3, one may calculate approximations of pseudo-first-order rate constants for the bimolecular steps to obtain a mechanism in the common units of s^{-1}. This gives

$$L + S1 \underset{3\ s^{-1}}{\overset{0.09\ s^{-1}}{\rightleftharpoons}} SD1 \underset{0.04\ s^{-1}}{\overset{21\ s^{-1}}{\rightleftharpoons}} LD1 + G$$

Thus, the bimolecular steps are rate-limiting in both the forward and reverse directions. SD1 would not be populated to any significant extent since a molecule of SD1 disappears with $k = 21 + 3 = 24\ s^{-1}$, 270 times faster than disappearance of a molecule of L, S1, LD1, or G. Observed overall rate constants would be given by

$$k_{forward} = 500\ s^{-1}\ M^{-1}[21/(21 + 3)] = k_{S1\text{-}SD1}(7/8) = 440\ s^{-1}\ M^{-1}$$

and

$$k_{reverse} = 2(10\ s^{-1}\ M^{-1})[3/(21 + 3)] = 2k_{LD1\text{-}SD1}(1/8) = 2.5\ s^{-1}\ M^{-1}$$

Although experimental data are consistent with this microscopic model, one should not conclude that the four microscopic rate constants actually have the above values. However, since it is likely that reduction of LD1 by G exhibits a value close to $10\ s^{-1}\ M^{-1}$, one may reasonably conclude that $k_{SD1\text{-}LD1} > k_{SD1\text{-}L}$. In other words, a molecule of SD1 is more likely to undergo the intramolecular reaction forming the unstrained six-membered loop of LD1 than to form the sterically strained three-membered loop of L.

Comparison of Absolute Magnitudes of Disulfide Exchange Rate Constants. Rate constants for loop-opening processes suggest that the disulfide of the Cys-Val-Cys grouping is strained. Reactions of L with G or with S occur much faster than bimolecular disulfide exchange reactions in systems lacking loops or in systems where the loop is large, as in STI. The absence of trimers in the Cys-Val-Cys peptides suggests that there is no accelerated reduction of loop dimers by SI or SII to generate trimer intermediates. The six-membered $(Cys\text{-}Val\text{-}Cys)_2$ dimer loop therefore appears to be unstrained.

Accelerated rate constants for reduction of L by S would generate problems for efforts to express mammalian proteins containing the Cys-Val-Cys sequence in *E. coli*. In this bacterial system, proteins expressed at high concentrations sometimes are isolated as aggregates cross-linked by intermolecular disulfides. Examples include calf prochymosin (Schoemaker et al., 1985) and human urokinase (Winkler & Blaber, 1986). An example of the opposite case where the protein is isolated as noncovalent aggregates containing fully reduced cysteines is bovine growth hormone (Langley et al., 1987), which contains two disulfides in the native form. Protocols for obtaining refolded monomers begin by dissolving the aggregates in a denaturing solvent such as 6 M guanidine and adding a disulfide reducing agent if necessary. In principle, the reduced proteins then can be refolded into monomers by large-fold dilution into nondenaturing oxidizing buffers. However, for such dilution protocols, the gain achieved by minimization of dimer formation is offset by the increase in financial expense of recovery of monomers from large volumes of dilute solutions. Optimization of refolding conditions is a difficult problem in general. For sequences like carboxy-terminal Cys-Val-Cys, dilution protocols have an extra burden to overcome. Here, 7-fold dilution would give 49-fold slower monomer/monomer collisions, but this gain would be offset by a 50-fold acceleration of reaction of L with S ($k_{S1\text{-}SD1} = 500\ s^{-1}\ M^{-1}$).

Enzyme-catalyzed loop formation in DNA is influenced both by the probability that the two ends of the polymer will collide with each other and by a need to twist DNA to align the matching ends of the two helical strands (Shore & Baldwin, 1983). Rate constants for forming looped transition states in proteins also reflect contributions from several processes. As seen in Figure 1, the loop of the transition state for forming SII is slightly larger than the corresponding transition state for forming L. The rate constants $k_{S\text{-}L}$, $k_{S1\text{-}S2}$, and $k_{SD\text{-}LD}$ therefore correspond to formation of transition states of increasing size. For YSRCVC, they exhibit increasing values of 3, 6, and possibly $21\ s^{-1}$, respectively. Thus, one contributing factor is steric inhibition of formation of small loops, giving *more* frequent collisions between cysteines separated by a larger number of atoms. A second contributing factor operating in the opposite direction was demonstrated previously in a comparison of $k_{S\text{-}L}$ values for STI and a fragment of an immunoglobulin molecule, Ig (Snyder, 1987; Goto & Hamaguchi, 1981). The difference (n) in sequence positions of the two free cysteines in STI and Ig is 9 and 60, respectively. Observed values of $k_{S\text{-}L}$ for these species are generally consistent (within a factor of 2) with the expression $k_{S\text{-}L} = cn^{-3/2}$. This trend for unstrained random-coil chains in denaturing solvents indicates that ends separated by a larger number of links in the primary sequence will collide *less* frequently with each other. That observation remains true when the directly determined $k_{S\text{-}L}$ for the STI-SSG derivative reported here is used and c, the constant of proportionality, equals $3\ s^{-1}$ in 3 M guanidine. However, the smaller loops represented by $(Cys\text{-}Val\text{-}Cys)_2$ (where $n = 5$) and YSRCVCA ($n = 2$) exhibit loop formation rates that are 2 orders of magnitude faster than the time frame provided by fits of the above equation to the STI/Ig data. At present it is difficult to explain this difference, given the small number of two-cysteine species for which detailed kinetic data are available.

Explanations of the dynamics of formation of turns or small loops during protein folding will require experiments measuring microscopic rate constants of disulfide formation and reduction in many more sequences. The results presented here demonstrate that application of modern instrumentation for automated synthesis of peptides and HPLC resolution of peptide derivatives can provide detailed information about disulfide exchange reactions in sequences containing closely spaced cysteines.

ACKNOWLEDGMENTS

We appreciate the technical assistance of Yiping Sun.

Registry No. YSRCVCA, 113598-02-0; YSRCVCA disulfide, 113598-03-1; YSRCVC, 113598-04-2; YSRCVC disulfide, 113598-05-3; GSH, 70-18-8; GSSG, 27025-41-8; STI, 9078-38-0; cysteine, 52-90-4; cystine, 56-89-3.

REFERENCES

Bailey, J. E. (1966) Ph.D. Thesis, London University.

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93–104.

Danby, J. M. A. (1985) in *Computing Applications to Differential Equations*, pp 1–52 and 245–246, Reston Publishing, Reston, VA.

Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A., & Rutter, W. J. (1985) *Nature (London) 317*, 267–270.

Eigen, M. (1964) *Angew. Chem., Int. Ed. Engl. 3*, 1–72.

Goto, Y., & Hamaguchi, K. (1981) *J. Mol. Biol. 146*, 321–340.

Heaton, G. S., Rydon, H. N., & Schofield, J. A. (1956) *J. Chem. Soc.*, 3157–3168.

Hoeprich, P. D., Jr., & Doolittle, R. F. (1983) *Biochemistry 22*, 2049–2055.
Jarvis, D., Rydon, H. N., & Schofield, J. A. (1961) *J. Chem. Soc.*, 1752–1756.
Kao, P. N., & Karlin, A. (1986) *J. Biol. Chem. 261*, 8085–8088.
Klintrot, I. M., Hoog, J. O., Jornvall, H., Holmgren, A., & Luthman, M. (1984) *Eur. J. Biochem. 144*, 417–423.
Langley, K. E., Berg, T. F., Strickland, T. W., Fenton, D. M., Boone, T. C., & Wypych, J. (1987) *Eur. J. Biochem. 163*, 313–321.
Malbon, C. C., George, S. T., & Moxham, C. P. (1987) *Trends Biochem. Sci. (Pers. Ed.) 12*, 172–175.
Mihalyi, E. (1968) *J. Chem. Eng. Data 13*, 179–182.
Moore, J. W., & Pearson, R. G. (1981) in *Kinetics and Mechanism*, 3rd ed., p 313, Wiley, New York.
O'Donnell, M. E., & Williams, C. H., Jr. (1985) *Biochemistry 24*, 7617–7621.
Riddles, P. W., Blakeley, R. L., & Zerner, B. (1979) *Anal. Biochem. 94*, 75–81.
Schoemaker, J. M., Brasnett, A. H., & Marston, F. A. O. (1985) *EMBO J. 4*, 775–780.
Shore, D., & Baldwin, R. L. (1983) *J. Mol. Biol. 170*, 957–981.
Snyder, G. H. (1987) *Biochemistry 26*, 688–694.
Suzuki, A., Tsunogae, Y., Tanaka, I., Yamane, T., Ashida, T., Norioka, S., Hara, S., & Ikenaka, T. (1987) *J. Biochem. (Tokyo) 101*, 267–274.
Szajewski, R. P., & Whitesides, G. M. (1980) *J. Am. Chem. Soc. 102*, 2011–2026.
Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261–287.
Weber, U., & Hartter, P. (1974) *Hoppe-Seyler's Z. Physiol. Chem. 355*, 189–199.
Winkler, M. E., & Blaber, M. (1986) *Biochemistry 25*, 4041–4045.

Amino Acid Sequence of κ-Flavitoxin: Establishment of a New Family of Snake Venom Neurotoxins†

Gregory A. Grant*,‡,§ and Mark W. Frazier‡

Departments of Biological Chemistry and Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

Vincent A. Chiappinelli

Department of Pharmacology, St. Louis University School of Medicine, St. Louis, Missouri 63104

Received September 29, 1987; Revised Manuscript Received December 16, 1987

ABSTRACT: The complete amino acid sequence of κ-flavitoxin, a neurotoxin isolated from the venom of *Bungarus flaviceps*, has been determined by automated Edman analysis of the intact protein and of peptides derived from digests with trypsin and chymotrypsin. κ-Flavitoxin consists of a single 66-residue polypeptide chain which is completely devoid of methionine. The amino acid sequence of κ-flavitoxin demonstrates that although the toxin is related to the α-neurotoxin family, it displays a much higher degree of homology with κ-bungarotoxin. The conserved structural features of the κ-neurotoxins and their pharmacological profiles, which are distinct from those of all known α-neurotoxins, provide evidence for a new, structurally and functionally unique family of snake venom neurotoxins.

Poisonous snakes produce a number of toxic proteins, many of which are directed against sites in nervous tissue. Several different families of such neurotoxins have been identified in the venom of elapid snakes. β-Neurotoxins interfere with neurotransmitter release by binding to sites on presynaptic nerve terminals (Othman et al., 1982). α-Neurotoxins bind with high affinity to postsynaptic nicotinic acetylcholine receptors found in vertebrate skeletal muscle and in the muscle-derived electric tissue of electric fish (Changeux et al., 1984).

We have previously reported the complete amino acid sequence of κ-bungarotoxin, purified from the venom of *Bungarus multicinctus* (Grant & Chiappinelli, 1985). This toxin (molecular weight 7313) displays an average of 47% sequence identity with the long α-neurotoxins, but its physiological effects are quite different. κ-Bungarotoxin is a potent antagonist at a variety of neuronal nicotinic receptors which are insensitive to blockade by α-neurotoxins (Chiappinelli, 1983; Chiappinelli & Dryer, 1984). Conversely, κ-bungarotoxin binds to muscle nicotinic receptors with much lower affinity than do the α-neurotoxins (Wolf et al., 1987).

κ-Flavitoxin, a protein recently purified from the venom of *Bungarus flaviceps* (Chiappinelli et al., 1987), exhibits a number of properties which are similar to those of κ-bungarotoxin. Both toxins produce a complete blockade of nicotinic transmission in chick autonomic ganglia at 50–75 nM. This blockade is only slowly reversible and can be protected against by short-acting competitive nicotinic antagonists such as D-tubocurarine, but not by α-bungarotoxin or noncompetitive nicotinic antagonists (Chiappinelli, 1983; Sorenson et al., 1987). Radiolabeled κ-bungarotoxin and κ-flavitoxin bind with high affinity ($K_d \sim 5$ nM) to a nicotinic site in chick

† This work was supported by National Institutes of Health Grant NS17574 to V.A.C.
‡ Department of Biological Chemistry, Washington University School of Medicine.
§ Department of Medicine, Washington University School of Medicine.

0006-2960/88/0427-3794$01.50/0