Dependence of Formation of Small Disulfide Loops in Two-Cysteine Peptides on the Number and Types of Intervening Amino Acids*

Rumin Zhang and Grayson H. Snyder
From the Department of Biological Sciences, State University of New York, Buffalo, New York 14260

Microscopic disulfide-exchange rate constants have been measured for the formation and opening of small disulfide loops in reactions between glutathione and peptides containing 2 cysteines. Twelve cysteine-X-cysteine peptides have been studied, where X is an amino acid and m is the number of amino acids between the cysteines. Homopolymers of alanine for m equaling 0–5 are evaluated, as well as X1 and X2 series employing glycine, valine, or proline. Equilibrium constants Ke for loop closing are only slightly dependent on the nature of X. Loops with even values of m generally are favored relative to loops with odd values. Ke increases in the rank order X1, X3, X0, X4, and X2. Formation of a disulfide between sequentially adjacent cysteines therefore is not especially difficult. The dependence of Ke on the odd-even nature of m is compared with similar patterns observed both in statistics of disulfide formation in naturally occurring proteins and in theoretical studies of peptide cyclization. The relative equilibrium populations of intramolecular disulfides in peptides containing cysteine-cysteine-cysteine and cysteine-serine-cysteine-serine-cysteine clusters are consistent with predictions based on the values of Ke in the two-cysteine peptides.

Quantitative information about factors influencing the formation of small disulfide loops is important for understanding naturally occurring proteins and for designing new engineered proteins. Examples of native proteins with small loops include acetylcholine receptor and protein disulfide isomerase. A -Cys-Cys-loop formed by 2 cysteines which are adjacent in the primary sequence affects agonist-induced conformational changes at the acetylcholine receptor’s binding site (1). Two -Cys-X-X-Cys-loops in the isomerase form the active site of this redox enzyme (2). Proteins with short sequences often employ high densities of disulfides to generate a folded topology. An example is the double-headed trypsin inhibitor from a Chinese medical herb (3), having 41 amino acids including 6 cysteines.

Early systematic studies of air oxidation of peptides of the type Cys-Gly-Cys (4–6) at low ionic strength revealed no intramolecular loop formation for m = 0 or 1, 15 and 40% monomeric loops for m = 2 and 3, respectively, and 90% or more monomeric loops for m = 4–6. Reduction of monomer loops for the same series by reduced glutathione exhibited monotonically decreasing values for the overall macroscopic rate constant as m increases from 1 to 4 (7). However, these overall rate constants are difficult to interpret since they were generated by data requiring a series of several consecutive chemical reactions in an enzyme-coupled spectrophotometric assay system. There have been no published data on details of the microscopic rate constants for loop formation as a function of the number and types of residues in the sequence between the paired cysteines.

Fig. 1 depicts several examples of disulfide exchange reactions pertinent to the data discussed below. Reaction 1 illustrates attack of the completely reduced peptide (R) by oxidized glutathione to produce either of two species (SI/SII) containing a single peptide-glutathione mixed disulfide. Those species can go on to produce a monomeric loop (L) by intramolecular attack of their mixed disulfide by the free cysteine on the same peptide chain, as shown in Reaction 2. The two rate constants are given by km describing loop closing, and koc describing loop opening. Alternatively, SII may react again with another GSSG to produce double mixed-disulfides (D), as in Reaction 3. It also is possible for one of the mixed disulfide forms, such as SI of Reaction 4, to convert directly to the other, SII, by an intramolecular rearrangement. Looped dimers (LD) may be formed by a variety of processes. One such process given by Reactions 5 and 6 begins with reduction of L by the free sulfhydryl group on SI. The last process in Fig. 1 depicts intramolecular exchange in a 3-cysteine species (LR) containing one loop and 1 reduced cysteine.

We recently presented a detailed study of such reactions in a peptide containing a -Cys-Val-Cys region (8), in which 12 rate constants were determined by kinetic studies beginning with four different monomeric and dimeric forms of the peptide. Data presented below extend those studies to a series of 11 additional synthetic peptides. Here however we focus only on the rate constants km and koc of Reaction 2 of Fig. 1, determined by a single type of kinetic experiment using L as the starting reagent. Rate constants are evaluated as a function of m for -Cys-Ala-Cys sequences where m varies from 0 to 5. Rate constants are evaluated as a function of the type of residue between the cysteines in -X1- and -X2-series employing Ala, Gly, Val, and Pro variations. These four types of residues have very different effects on formation of reverse turns in polypeptide chains (9). Alanine is the prototype L-amino acid. Glycine provides additional conformations not available to L-amino acids, since its tiny side chain consists of a single hydrogen atom. In contrast, the side chain of valine is branched at the β position adjacent to the backbone, leading to constraints on allowed backbone dihedral angles. Proline provides unique effects by virtue of attachment of its side chain to both the α carbon and the nitrogen atoms of the backbone.

Experiments described below also consider intramolecular disulfide exchange in peptides containing the three-cysteine...
Formation of Small Disulfide Loops in Two-Cysteine Peptides

1. \[ \text{R} + \text{S-S} \rightarrow \text{S-S-R} + \text{S} \]
2. \[ \text{S-S} \rightarrow \frac{k_1}{k_2} \text{S-L} \]
3. \[ \text{S-S} + \text{D} \rightarrow \text{S-D} + \text{S} \]
4. \[ \text{S-L} + \text{D} \rightarrow \text{S-L-D} + \text{S} \]
5. \[ \text{S-L-D} + \text{S} \rightarrow \text{S-L-D-S} + \text{S} \]
6. \[ \text{S-L-D-S} + \text{D} \rightarrow \text{S-L-D-S-D} + \text{S} \]
7. \[ \text{S-L-D-S-D} + \text{S} \rightarrow \text{S-L-D-S-D-S} + \text{S} \]

Fig. 1. Examples of disulfide exchange reactions in mixtures of glutathione and peptides containing 2 or 3 cysteines. Oxidized and reduced glutathione: O-SS-O and O-S, respectively.

MATERIALS AND METHODS

Iodoacetate, GSH, and GSSG were obtained from Sigma. The two forms of glutathione were further purified by reverse phase HPLC. Synthetic peptides were prepared on a Biosearch model 9500 peptide synthesizer using standard Merrifield chemistry. Crude reduced peptides were purified by preparative HPLC. Their amino acid composition was confirmed by analysis of acid hydrolysates. Their fully reduced nature was confirmed by assaying peptide concentration by Ellman's reagent (11). Oxidized forms at equilibrium were distinguished by holding peptide chain concentration constant while varying the GSH/GSSG ratio. Final peak assignments may be confirmed quantitatively by verifying that ratios of equilibrium concentrations which should correspond to equilibrium cysteine peptides are in fact constant under different experimental conditions. Details of these methods were presented previously for Val-Ala (8).

A typical kinetics experiment begins by mixing 0.1 mM fully oxidized loop monomer with 5 mM reduced GSH at pH 7.0 in the same buffers employed above. At different time points, aliquots are removed and trifluoroacetic acid is added to lower the pH. Disulfide exchange requires the nucleophile protonated thiolate \(-S^2-\) form of cysteines. Protonation of sulphydryl groups at low pH is capable of slowing disulfide exchange to the point where it is negligible during the time scale of HPLC separations present in this work. Quenched aliquots were frozen and stored at \(-80^\circ C\) until needed for HPLC analysis. Fig. 2 exhibits kinetic data for such an experiment. The starting reagent is monomer loop. This is reduced by excess GSH to produce approximately equal amounts of the SI and SII species. It is not known which peak, SI or SII, correlates with structure SI or SII. At long times, SI and SII are further reduced to generate R with two free sulphydryl groups. For some peptides, transient population of oligomeric species is observed. Such oligomers in principle may be formed by processes such as Reactions 5 and 6 of Fig. 1. No peptide was observed to produce any D under the strongly reducing conditions of these kinetics experiments.

A different protocol was used to monitor intramolecular exchange between LR forms of a 3-cysteine peptide, given by Reaction 7 of Fig. 1. The peptide ending in -Cys-Cys-Ala-amide has two stable LR forms at equilibrium. One of these was purified at pH 2 by reverse phase HPLC. In the kinetics experiment, the pH was then jumped up to 7.0, without addition of any GSH or GSSG, to accelerate intramolecular disulfide exchange. As time proceeds, the LRI form exchanged almost completely with the SI form, confirming that the LRI form exists.

Kinetic data were analyzed by Runga-Kutta-Fehlberg numerical integration of the set of differential rate equations describing the contributing reactions (12). Values of rate constants used as input parameters to the simulation were altered to given minimum deviations between observed and simulated concentrations of species. Plots of concentration versus elapsed time were constructed. Observed data were fit by the least-squares method. Observed data were found to lie between simulated plots given by the final values for the \(k\) values, \(\pm 20\%\) variation. This is the margin of error for all rate constants determined below. Equilibrium constants have a \(\pm 20\%\) margin of error. Microscopic rate constants were determined for the two separate equilibria describing formation of LI by SI or SII. In all cases, rate constants for the SI species were similar to the analogous rate constants for the SII species, differing by no more than \(30\%\). Values reported below are the average of the separate values for the two exchanges. Depending on the calculations, values reported below have been normalized to the high pH limit of 100% thiolate form. Calculations
Formation of Small Disulfide Loops in Two-Cysteine Peptides

assume a typical value of 8.9 for the pK\textsubscript{a} of each cysteine (8). Thus, for example, the Henderson-Hasselbalch relationship permits one to calculate that a typical cysteine studied at pH 7.0 would spend only one-eighth of its time in the reactive deprotonated state. As such, effective values for observed rate constants at pH 7 are multiplied by 80 to generate the normalized magnitudes.

Fully reduced 3-cysteine peptides were slowly oxidized at pH 8.3 by addition of oxygen to a buffer containing 0.3 mM peptide. Buffers were similar to those above, but lacked supplementary Cu\textsuperscript{2+} which would have accelerated the slow oxidation process. No glutathione forms were present in the solution. Ellman’s reagent was used to determine the time point at which 50% of the total sulfhydryl groups had been oxidized to disulfides. This generally occurred about 3 h after starting the reaction. At that time, the pH was lowered by addition of trifluoroacetic acid. This quenched both net oxidation by oxygen and also disulfide exchange reactions. Different oxidized forms of the 3-cysteine peptide then were separated by preparative HPLC at pH 2.

Some of the partially oxidized forms were LR monomers containing one free cysteine and one intramolecular disulfide. To identify which of the 3 cysteines was free, each LR species purified at pH 2 was combined with 1.0 mM iodoacetate at pH 7 to irreversibly alkylate the free sulphydryl groups. The alkylated derivative then was loaded on an Applied Biosystems model 470A protein sequencer synchronized with an on-line model 120A microbore HPLC analyzer. For each species studied, carboxymethylated cysteine was detected primarily in only 1 of the 3 cysteine positions. Cysteines joined in a disulfide did not give a CmCys signal in the sequencing chromatogram.

**RESULTS**

**Two-Cysteine Peptides**—The effect of the negative carboxyl group on disulfide exchange reactions involving carboxy-terminal cysteines is seen by comparing -Cys-Cys with -Cys-Cys-Ala and by comparing -Cys-Val-Cys with -Cys-Val-Cys-Ala. For -Cys-Cys, addition of an alanine at the carboxyl terminus increases \( k_1 \), 4.0-fold from 0.30 to 1.2 s\(^{-1}\) and increases \( k_2 \), 5.0-fold from 11 to 55 s\(^{-1}\) M\(^{-1}\). Since both rate constants are increased by approximately similar amounts, the equilibrium constant for loop formation given by the ratio \( K_c = k_1/k_2 \) is essentially unaffected. Similarly, addition of an alanine at the carboxyl end of -Cys-Val-Cys increases \( k_1 \) and \( k_2 \) by factors of 4.8 and 5.6, respectively, again not affecting \( K_c \). Thus the negative carboxyl group slows disulfide exchange by a factor of 5, presumably by inhibiting the presence of a neighboring negative -S\(^-\) group, both when the carboxy-terminal cysteine is attacking a disulfide and when it is part of the disulfide being attacked. The carboxyl group does not, however, have a significant impact on the equilibrium constant for forming an adjacent disulfide loop.

Rate constants for carboxyl-terminal loops are arranged in rank order of increasing magnitude of \( K_c \), in Table I. The rate constant for loop closing, \( k_c \), varies over a 370-fold range from 0.10 s\(^{-1}\) in Pro\(_2\) to 37 s\(^{-1}\) in Gly\(_1\). The three slowest loop closing reactions occur in the largest loop, Ala\(_6\), and in the two species -Cys-Cys- and Pro\(_2\). These latter two species have strong limitations on allowed values of backbone dihedral angles. Formation of a disulfide in Cys-Cys requires prior formation of an unfavorable cis peptide bond between the cysteines (13). Prolines have only one favorable value for the dihedral angle \( \phi \) describing rotation about the N-C\(_\alpha\) bond. Loop opening processes, given by \( k_o \), vary over a 5000-fold range from 5.0 s\(^{-1}\) M\(^{-1}\) in Pro\(_2\) to 25 \( \times 10^5 \) s\(^{-1}\) M\(^{-1}\) in Gly\(_1\). The five peptides with the smallest values for \( k_o \) also have the smallest values for \( k_c \). Gly\(_1\) exhibits the fastest example of both kinetic processes. Peptides with the structure Cys-X-X-Cys have the fastest values for \( k_o \) regardless of the nature of \( X \).

The values of equilibrium constants determined by calculating the ratio of rate constants obtained in kinetics experiments are given in Table I. They are identical, within experimental error of ±30%, with values calculated from the ratios of molecular concentrations obtained in equilibrium experiments. Observed numbers vary over a 50-fold range, from 1.5 mM for the most forbidden loop Gly\(_1\) to more than 71 mM for the most favored loop Ala\(_6\). The four smallest \( K_c \) values correspond to the four X\(_1\) species, each of which has a very fast \( k_o \). It is important to note that the case of two adjacent cysteines is not the most difficult disulfide to form. In fact, -Cys-Cys- has an intermediate value for \( K_c \). For a given value of \( m \), the value of \( K_c \) is not strongly dependent on the nature of \( X \). Thus in the X\(_1\) series, \( K_c \) exhibits a maximum 5-fold difference from the 1.5 mM in Gly\(_1\) to 7.3 mM in Pro\(_2\). For X\(_2\) peptides, \( K_c \) again exhibits a maximum 5-fold range, in this case from 16 mM in Pro\(_2\) to 86 mM in Ala\(_6\).

Table II aids in consideration of the effect of varying \( m \).

| Peptide | \( k_1 \) | \( k_2 \) | \( K_c = k_1/k_2 \) |
|---------|--------|--------|-------------|
| Pro\(_2\) | 0.10   | 5.0    | 20          |
| Cys-Cys | 0.30   | 11     | 37          |
| Ala\(_6\) | 0.90   | 16     | 55          |
| Ala\(_5\) | 0.55   | 18     | 30          |
| Ala\(_3\) | 0.90   | 50     | 18          |
| Gly\(_1\) | 4.0    | 0.13 \( \times 10^3 \) | 31          |
| Val\(_2\) | 10     | 0.14 \( \times 10^3 \) | 71          |
| Ala\(_2\) | 10     | 0.14 \( \times 10^3 \) | 71          |
| Pro\(_1\) | 1.5    | 0.21 \( \times 10^3 \) | 7.3         |
| Val\(_1\) | 2.9    | 1.8 \( \times 10^3 \) | 1.6         |
| Ala\(_1\) | 5.0    | 1.6 \( \times 10^3 \) | 3.1         |
| Gly\(_1\) | 37     | 25 \( \times 10^3 \) | 1.5         |
For all types of X examined, increasing m from 1 to 2 has a greater effect on \( k_2 \) than on \( k_1 \). For example, changing Ala\(_1\) to Ala\(_2\) increases \( k_2 \) by a factor of 2.0 but decreases \( k_1 \) by a factor of 12. The larger effect, 12, is indicated in the table by bold print. Extensive variation in \( m \) is given only for the Ala\(_m\) series, whose data appear in the lower part of the table. Focusing only on the bold print, one notes that the dominant effect alternates back and forth between \( k_1 \) and \( k_2 \). In other words, the larger effect occurs with opening for the case of changing \( X_1 \) to \( X_0 \), closing for \( X_2 \) versus \( X_0 \), opening for \( X_0 \) versus \( X_1 \), and closing for \( X_1 \) versus \( X_0 \). All of these dominant effects are decreases in the corresponding rate constant. Furthermore, their absolute magnitudes (12, 11, 3.1, 1.6) decrease monotonically as \( m \) increases. Each time a dominant decrease occurs in the forward rate constant, \( k_1 \), the equilibrium constant for loop formation exhibits a decrease. Alternatively each time a dominant decrease occurs in the reverse rate constant, \( k_2 \), the equilibrium constant increases. The net result on the equilibrium constant, given by the right column in Table II, therefore is an alternating pattern of increases and decreases in \( K \), with a monotonically declining value in absolute magnitude of the effect as \( m \) gets larger. This means that a loop with an even number \((m)\) of alanines between the cysteines is more stable than a loop with \( m \pm 1 \) intervening alanines, where \( m \pm 1 \) is an odd number. For example, \( K \) for Ala\(_2\) is greater than \( K \) for Ala\(_1\) or Ala\(_3\). Although it also is true that \( K \) for Cys-Cys \((m = 0)\) is greater than \( K \) for Cys-Ala-Cys \((m = 1)\), the case of two sequentially adjacent cysteines requires special considerations to be discussed below.

### Table II

| Comparison | \( k_1 \) (forward) | \( k_2 \) (reverse) | Equilibrium ratio \( K = k_2/k_1 \) |
|------------|---------------------|---------------------|-----------------------------------|
| Pro\(_1\)  | 15 dec\(^a\)        | 42 dec              |                                  |
| Gly\(_1\)  | 9.3 dec             | 190 dec             |                                  |
| Val\(_1\)  | 3.5 inc             | 13 dec              |                                  |
| Ala\(_1\)  | 2.0 inc             | 12 dec              | 23 inc                            |
| Ala\(_2\)  | 11 dec              | 2.8 dec             | 3.9 dec                           |
| Ala\(_3\)  | No change           | 3.1 dec             | 3.1 inc                           |
| Ala\(_4\)  | 1.6 dec             | 1.1 inc             | 1.8 dec                           |

\(^a\) dec, decrease; inc, increase.

Three-Cysteine Peptides—Peptides in the 50% oxidized sample of Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide eluted in two HPLC peaks at pH 2. One of these was fully reduced monomeric R, and the other had an average of one sulphydryl group and one disulfide per peptide chain. This latter peak tentatively was assigned to the monomeric LR form. To determine if it really was monomer, equilibrium experiments were performed where the concentration of total peptide was varied 10-fold in solutions containing a constant ratio of excess GSH/GSSG. Ten peaks were observed, including the monomeric R peak and the tentatively assigned LR peak observed in the air-oxidation reaction. The relative magnitudes of these 10 peaks were independent of peptide concentration, thereby indicating that all species were monomeric like R and that the tentative LR assignment was correct. A complete analysis of the glutathione-containing equilibrium mixture was not attempted. In reactions between a 3-cysteine species and glutathione, there are 14 possible monomeric forms which could contribute to the HPLC data.

The monomeric LR peak in the slow air oxidation solution was purified by preparative HPLC at pH 2 and then alkylated with 1.0 M iodoacetate at pH 7. Sequencing data for the carboxymethylated derivative are presented in Fig. 3. PTH-derivatives of tyrosine, serine, CmCys, and arginine all elute between 17 and 27 min. Standards containing equimolar amounts of these amino acid derivatives gave peak heights in the rank order Tyr > CmCys > Arg > Ser. The tyrosine ring itself contributes to absorption of 270 nm light in PTH-Tyr. PTH-Arg has a broadened and therefore relatively shorter peak. PTH-Ser is partially degraded by the temperature and timing parameters of the sequencing program used in this particular experiment. Cycle 1 is dominated by the PTH-Tyr peak, and cycle 2 is dominated by PTH-Arg. Similar analysis of subsequent steps gives the complete sequence Tyr-Arg-X-Ser-CmCys-Ser-X, where X corresponds to cycles in which only a small quantity of CmCys is observed. Thus the LR peak predominantly corresponds to the form containing a free cysteine at position 5 and an intramolecular disulfide joining cysteines 3 and 7. The partially oxidized sample at pH 7 therefore contains no significant amount of either of the two possible species of the Cys-X-Cys type joining cysteines 3 and 5 or joining cysteines 5 and 7.

A similar set of experiments was performed for the peptide Tyr-Arg-Cys-Cys-Cys-amide. In this case, there are two stable monomeric LR forms giving separate HPLC peaks midway thru the slow O\(_2\)-induced oxidation process. Sequencing analysis indicated that they correspond to the LRI and LRHI structures given in Reaction 7 of Fig. 1, where the central cysteine is joined by a disulfide to either of its immediately adjacent neighbors. Thus at pH 7, there is no significant amount of the one possible disulfide of the Cys-X-Cys type joining cysteines 3 and 5.
Formation of Small Disulfide Loops in Two-Cysteine Peptides

**DISCUSSION**

The dependence of the rate constants $k_c$ and $k_s$ on the number of residues between the cysteines reflects several contributing factors. For 2 cysteines separated by many residues in the sequence, the probability of those cysteines colliding with each other in a denaturing solvent decreases as $m$ increases (14, 15). This configurational entropy factor probably is responsible for much of the general trend for non-proline residues in Table I. For such residues, $k_c$ values for different $m$ occur in the rank order $(m = 5) < (m = 4) < (m = 2) < (m = 1)$. As $m$ gets smaller and smaller, the tendency for cysteines to find each other more rapidly will be offset by unfavorable nonbonding interactions. The need to pack backbone groups more tightly together in small loops will lead to distortion of bond lengths and bond angles, preventing formation of a very stable conformation for the looped molecule. This strain factor may contribute to the increase in $k_s$ as $m$ decreases. For non-proline molecules, $k_s$ exhibits the rank order $(m = 5) < (m = 3) < (m = 2) < (m = 1)$.

The case of two sequentially adjacent cysteines requires special consideration. For such cysteines, theoretical studies (13) and crystallographic data (16) indicate that formation of a disulfide requires a cis peptide bond between the cysteines. Formation of a cis peptide bond is very unfavorable unless it precedes proline, in which case it is slightly unfavorable data for these coupled equilibria could be complicated. Since $K_c$, $K_{SOM}$, pH-independent conformational change prior to the HPLC configuration. The overall reaction for forming an intramolecular disulfide is

$$K_{SOM} \frac{[H_2SOMCys-Cys{SOM}]}{[Cys-Cys]} = K_{SOM} \frac{k_1}{k_2} \text{Cis-SS + GSH}$$

Cis-SS is the looped L form. In principle, analysis of kinetic data for these coupled equilibria could be complicated. Since velocities of isomerization are independent of pH but velocities of disulfide exchange depend on the percentage of molecules in the -S- form, different steps might be rate-limiting at the extremes of very low and very high pH. For the data observed here at pH 7, mathematical analysis is relatively uncomplicated. The low pH quench method will trap $H^2SOMCys-Cys{SOM}$ and allow it to convert to the more favored trans form via a pH-independent conformational change prior to the HPLC analysis. The HPLC peak for a single mixed-disulfide species (SI or SII) therefore will correspond to the sum of cis and trans S forms at the moment of the pH quench. For the species eluting in the SI peak, this gives concentration ratio:

$$K_c = \frac{[\text{Cis-SS}] [\text{GSH}]}{[\text{Trans}^{SSG}]} = \frac{[\text{Trans}^{SSG}]}{[\text{Cis-SS}][\text{GSH}]} = \frac{[\text{Cis-SS}] [\text{GSH}]}{[\text{Trans}^{SSG}]}$$

where concentrations and peak areas are those observed in quenched aliquots of equilibrium mixtures. The symbol representing an approximate equality above is allowed since $H^2SOMCys-Cys{SOM}$ at equilibrium. The final ratio is the exact expression for the overall equilibrium constant for the exact consecutive reactions, such that

$$K_c = \frac{[\text{Cis-SS}][\text{GSH}]}{[\text{Trans}^{SSG}]} = \frac{K_{SOM} k_1}{k_2}$$

The observed rate constant $k_c$ corresponds to the microscopic rate constant $k_{-1}$ describing reduction of the disulfide loop by GSH. Thus $k_c = K_{SOM} \times k_i$ at pH 7. Since $K_{SOM} < 1/20$, $k_i > 20 k_c$. The value of $k_c$ is $0.30 \text{ s}^{-1}$. Thus once the cis configuration is achieved, the microscopic rate constants for closing the loop, $k_i$, is $> 6.0 \text{ s}^{-1}$. This is at the fast end of the rank order of observed loop closing rate constants in Table I. As mentioned earlier, the equilibrium constant for $K_c$ is generally favorable in -Cys-Cys- sequences. Therefore the unfavorability of having to form a cis peptide bond between the cysteines is more than compensated by 1) the very high probability of collision between the two closely spaced sulfur groups in the cis configuration and 2) the absence of significantly unfavorable nonbonded interactions.

Previous studies of air oxidation of Cys-Cys at low ionic strength (4, 6) detected no monomeric loop. Dimer formation was dominant, in contrast to results observed above. In the previous studies, Cys-Cys was present at a concentration of 45 mM (10 g/liter, cited in Ref. 6), which is significantly higher than the 1.6 mM concentration of Tyr-Ser-Arg-Cys-Cys used here. That higher concentration would favor peptide/peptide collisions and corresponding dimer formation. Moreover at low ionic strength, electrostatic factors may have enhanced formation of antiparallel dimers of (‘H$_2$N-Cys-Cys-(COO)-), which would juxtapose the negative carboxyl terminus of one chain with the positive amino terminus of the other chain.

Short polypeptides consist of a chain of rigid links, with each link having a fixed distance between the backbone nitrogen and carboxyl carbon atoms of an amino acid. Given restricted choices of allowed backbone dihedral angles, a chain of such rigid links has a tendency to persist in travelling in a given direction instead of readily reversing its orientation. Monte-Carlo theoretical methods have been applied to the problem of joining the amino- and carboxyl-terminal ends of short polyanaline chains to form a small loop in a denaturing solvent (19). This simulation required not only that the two termini be brought near to each other to create a new peptide bond, but also that they be mutually oriented at the proper angle for peptide bond formation. The probability of forming a cyclic peptide was calculated as a function of $m$. For small $m$, that calculated probability alternated between high and low values as $m$ varied from even to odd numbers, with the amplitude of the oscillation decreasing as $m$ increased. This is the same qualitative behavior observed in Table II for equilibrium constants of loop formation in -Cys- Ala-...-Cys-peptides in 3 M guanidine. It is possible that the pattern observed for this experimental series originates in the theoretically described persistence length effect for short polymers. Although one might postulate that hydrogen bonds in $\beta$ turns could enhance loop stability for even-numbered loops such as $m = 2$ or 4, such bonds are not very stable in the 3 M guanidine conditions of these experiments.

This alternating odd-even pattern also appears in statistics
of disulfide loop formation in naturally occurring proteins. A survey of known disulfide positions is presented in the Mini
print. Fig. 4 summarizes data in that survey. The value at the top of each bar gives the absolute number of currently known
examples of sequentially unrelated cysteine pairs for each value of \( m \). The vertical axis plots the percentage of occasions
where the 2 cysteines of those pairs are joined to each other
in an intramolecular disulfide. Cysteine pairs with 2 or 4
intervening residues have a greater probability of loop for-

mation than loops with \( m = 1, 3, \) or 5. This pattern ceases for
\( m > 5 \). Competition between configurational entropy, loop
strain, and persistence length effects appears to give the same
net dominant influence in the biological and synthetic peptide
data. A possible exception is the absence of many cases of
Cys-Cys loops in naturally occurring proteins, perhaps reflect-
ing the lack of much biological utility for this tiny but ener-
getically permitted loop.

Given the very different effects of glycine, valine, proline,
and alanine on turn formation as mentioned in the intro-
duction, it is not possible to give a definitive explanation of the
general lack of dependence of loop stability on the nature of
\( X \) in \( X_1 \) and \( X_2 \) series. Each peptide is probably a special case of its own, with different factors compensating to give roughly
equivalent values of \( K \), for a particular value of \( m \). In studies of
several \( X_1 \) species, the positioning of valine, glycine, and
proline residues had a more noticeable effect (20). Loop
formation correlated with predicted potentials for \( \beta \) turn
formation, yielding no monomer for \( \text{Val} \), and an 8-fold max-
umum range in \( K \), for the heteropolymers examined.

The determination of the predominant disulfide in a given
3-cysteine \( \text{LR} \) form relies on alkylation of the free cysteine
not involved in the disulfide. For a peptide having two or
more stable \( \text{LR} \) forms, it therefore is necessary that alkylation of
each purified stable form be faster than intramolecular disulfide
exchange between those forms. In the Tyr-Arg-Cys-
Cys-Cys-amide species described above, the interconversion
of \( \text{LRI} \) and \( \text{LRII} \) was depicted in Fig. 1 and is characterized by forward and reverse rate constants which both have the value of \( 0.024 \text{ s}^{-1} \). Iodoacetate carboxymethylates free cysteines with a rate constant of about \( 16 \text{ s}^{-1} \text{ M}^{-1} \) (21). Thus in the experiments described above, alkylation by \( 1.0 \text{ M} \) iodoac-
etate occurs with a pseudo-first order rate constant of \( 16 \text{ s}^{-1} \), considerably faster than the disulfide exchange. In contrast,
for a peptide having only one stable \( \text{LR} \) form, such as Ty-
Arg-Cys-Ser-Cys-Ser-Cys-amide, it is irrelevant whether al-
kylation is faster than intramolecular disulfide exchange. If
the intramolecular exchange were to be faster, then the solu-
tion would establish and maintain an equilibrium population of
different \( \text{LR} \) forms during the slower alkylation. The most
stable \( \text{LR} \) form would predominate among \( \text{LR} \) forms at all
time points. Given equal probabilities of reaction of iodoac-
etate with any cysteine in the solution, one therefore would obtain the expected single major alkylated product in this
fast-exchange limit.

A disulfide loop between adjacent cysteines in -Cys-Cys-
Cys-amide will have the same bonding interactions and ther-

modynamic stability as a disulfide loop between adjacent cysteines in -Cys-Cys-Ala. The difference in free energy (\( \Delta G \))
between the reduced chain and the loop-containing chain should have the same value in both sequences. Values of \( K_s \),
which are equal to exp \( (-\Delta G/RT) \), therefore should also be inde-
pendent of the loop’s environment. Thus it should be possible to use the relative values of the equilibrium constants
\( K \) in 2-cysteine peptides to predict the relative populations
of \( \text{LR} \) species at equilibrium in a 3-cysteine species. For
example, \( K_s \) for forming \( \text{Ala} \) and \( \text{Ala}_3 \) loops are 3.1 and 18
\( \text{mm} \), respectively. It is reasonable to estimate that equilibrium
constants for small serine-containing loops will be similar.
Given the lack of significant dependence of \( K \) on the type of
\( X \) observed above for very dissimilar types of amino acids.
Thus relative populations of the 3-5, 3-7 and 5-7 disulfides in
Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide should be 3:18:3, re-
spectively. The sequencing data in Fig. 3 exhibit predomin-
nance of the 3-7 disulfide, as expected. Similarly, the relative
populations of the 3-4, 3-5, and 4-5 disulfides in Tyr-Arg-
Cys-Cys-Cys-amide should be 27:3:27, respectively. Experi-
mental observation of equimolar amounts of the 3-4 and 4-5
disulfides, with no significant 3-5 loops, is consistent with
the prediction.

Although equilibrium constants are expected to be similar
in the 2- and 3-cysteine peptides for the reasons given above,
this is not true for the rate constants. Fig. 1 depicts two types of
reactions which are unimolecular in both directions. These
are interconversion of \( S \) forms in Reaction 4, and intercon-
version between \( \text{LR} \) forms in Reaction 7. Interconversion of the
\( S \) forms was directly measured only for the \( \text{Val} \) peptide
examined previously (6). Those experiments demonstrated
that the intramolecular interconversion rate constants and
the intramolecular loop-closing rate constant \( k \) differ only by
a factor of 2. It therefore is reasonable to expect that the SI
and SII forms of -Cys-Cys-Ala will interconvert with a rate
constant similar to the value of \( 1.2 \text{ s}^{-1} \) observed for \( k \),
in this peptide. This estimated value for the 2-cysteine peptide is 50
times faster than the \( 0.024 \text{ s}^{-1} \) interconversion of LRI
and LRII forms of the 3-cysteine peptide, -Cys-Cys-Cys-amide. In
the former case, a free cysteine is trying to attach to an
adjacent cysteine participating in a mobile mixed disulfide
with glutathione. In the latter case, a free cysteine is trying to
attach to an adjacent cysteine participating in a rigidly
constrained Cis-SS loop. These are very different processes,
exhibiting rate constants which are estimated to differ by
more than a factor of 10.

Acknowledgments—We appreciate the contributions by Wenhan
Chang, who obtained the data for the Cys-Ala-Cys peptide during a
graduate laboratory rotation. We also thank Robert L. Baldwin for
directing our attention to the literature on alternating odd-even
effects in polymers of amino acids. We are indebted to Yiping Sun in
our university’s core facility for dependable synthesis of HF-cleaved
peptides, development of HPLC separation protocols, and assistance
with peptide sequencing.

REFERENCES
1. Kao, P. N., and Karlin, A. (1986) J. Biol. Chem. 261, 8085-8086
2. Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A., and Rutter,
Formation of Small Disulfide Loops in Two-Cysteine Peptides

W. J. (1985) *Nature* **317**, 267–270
3. Tan, F., Zhang, G., Mu, J., Lin, N., and Chi, C. (1984) *Hoppe-Seyler’s Z. Physiol. Chem.* **365**, 1211–1217
4. Heaton, G. S., Rydon, H. N., and Schofield, J. A. (1966) *J. Chem. Soc.* 3157–3168
5. Jarvis, D., Rydon, H. N., and Schofield, J. A. (1961) *J. Chem. Soc.* 1752–1756
6. Hardy, P. M., Ridge, B., Rydon, H. N., and Serrao, F. O. S. P. (1971) *J. Chem. Soc. Sect. C Org. Chem.* 1722–1731
7. Weber, U., and Hartter, P. (1974) *Hoppe-Seyler’s Z. Physiol. Chem.* **355**, 189–199
8. Zhang, R., and Snyder, G. H. (1988) *Biochemistry* **27**, 3785–3794
9. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1986) *Adv. Protein Chem.* **37**, 1–109
10. Mihalyi, E. (1968) *J. Chem. Eng. Data* **13**, 179–182
11. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) *Anal. Biochem.* **94**, 75–81
12. Danby, J. M. A. (1985) in *Computing Applications to Differential Equations*, pp. 1–52 and 245–246, Reston Publishing, Reston, VA
13. Chandrasekaran, R., and Balasubramanian, R. (1969) *Biochim. Biophys. Acta* **188**, 1–9
14. Jacobson, H., and Stockmayer, W. H. (1950) *J. Chem. Phys.* **16**, 1600–1606
15. Kauzmann, W. (1939) in *Sulfur in Proteins* (Benesch, R., Benesch, R. E., Beyer, P. D., Klots, I. M., Middlebrook, W. R., Szent-Gyorgyi, A. G., and Schwarz, D. R., eds) pp. 93–108, Academic Press, New York
16. Capusso, S., Mattia, C., Mazzarella, L., and Puliti, R. (1977) *Acta Crystallogr. Sect. B Struct. Sci.* **33**, 2080–2083
17. Creighton, T. E. (1983) *Proteins: Structures and Molecular Properties*, pp. 162–165, W. H. Freeman & Co., New York
18. LaPlante, L. A., and Rogers, M. T. (1964) *J. Am. Chem. Soc.* **86**, 347–351
19. Mutter, M. (1977) *J. Am. Chem. Soc.* **99**, 8307–8314
20. Milburn, P. J., Meinwald, Y. C., Takahashi, S., Ooi, T., and Scheraga, H. A. (1988) *Int. J. Pept. Protein Res.* **31**, 311–321
21. Snyder, G. H. (1987) *Biochemistry* **26**, 688–694

Additional references are found on p. 18479.
### Table 1: Data on distribution of cysteine peptides in various tissues

| Tissue         | peptide |
|----------------|---------|
| Human heart    | 0.87    |
| Porcine heart  | 0.12    |
| Ruminant heart | 0.05    |

**References**

1. Sallan, S. M. and Hwang, S. L. (1969) J. Biol. Chem. 242, 3432-3437.
2. Dayhoff, M. O. (1973) in Atlas of Protein Sequence and Structure Vol. 5, pp. 121. National Biomedical Research Foundation, Silver Spring, MD.
3. Dayhoff, M. O. (1978) in Atlas of Protein Sequence and Structure Vol. 5, pp. 68. National Biomedical Research Foundation, Silver Spring, MD.
4. Brown, N. N., Mallal, D. B., and Mallal, J. L. (1983) Proc. Natl. Acad. Sci. USA 80, 429-433.
5. Greenfield, K. M., Ryan, C. A., and James, M. N. G. (1985) J. Biol. Chem. 260, 201-209.
6. Drouin, F. C., Cronon, F. M., and Barr, J. (1978) Biochemistry 17, 2160-2167.
7. Heil, G. O., Le-Nguyen, D., and Costa, B. (1989) Biochemistry 28, 2930-2936.
8. Tschumper, J. M., Massier, T., Mes, S., and Salmeron, J. F. (1989) J. Biol. Chem. 264, 14835-14839.
9. Hohne, A., Samsa, A., Krut, J. G., Salmeron, J. F., and Muller, J. W. (1981) J. Biol. Chem. 256, 3531-3539.
10. Krug, T. L. (1981) J. Biol. Chem. 256, 1739-1742.
11. Wright, C. S. (1977) J. Biol. Chem. 252, 411-418.
12. Vanueckelen, J. N., Lee, W. D., and Comer, P. M. (1983) Nature 303, 24-30.
13. Bernaud, A., Guimard, A. A., Desnuelle, P., and Raftery, M. V. (1972) FEBS Lett. 30, 256-261.
14. Dostler, W. B., Dainton, J., Kahl, K. H., and Varmus, P. J. (1977) J. Biol. Chem. 262, 54-56.
15. Sladen, R. A. and Sambrook, J. (1974) J. Biol. Chem. 249, 971-977.
16. Allan, G. and Gunning, L. P. (1983) Biochem. J. 209, 481-487.
17. Mikes, D. G., Holley, L., and Sambrook, J. (1960) J. Biol. Chem. 235, 87-95.
18. Brown, H. H., Herrick, M. A., and Brandt, A. B. (1973) J. Physiol. 235, 593-601.
19. Mardon, G. and Sambrook, J. (1980) J. Biol. Chem. 255, 5079-5087.
20. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3873-3880.
21. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3881-3887.
22. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3888-3892.
23. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3893-3897.
24. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3898-3902.
25. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3903-3907.
26. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3908-3912.
27. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3913-3917.
28. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3918-3922.
29. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3923-3927.
30. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3928-3932.
31. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3933-3937.
32. Lennard, J. C., Lennard, J. C., and Sambrook, J. (1980) J. Biol. Chem. 255, 3938-3942.