Factors Governing Selective Formation of Specific Disulfides in Synthetic Variants of α-Conotoxin†

Rumin Zhang‡ and Grayson H. Snyder*
Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York 14260
Received February 19, 1991; Revised Manuscript Received July 31, 1991

ABSTRACT: α-Conotoxin GI is a snail toxin protein consisting of 13 amino acids cross-linked by 2 intramolecular disulfide bridges. This toxin is an antagonist of acetylcholine receptors. The native sequence has been synthesized, along with nine additional variants in which non-cysteine residues are replaced by alanine or the cysteine positions are altered. Each reduced peptide has been oxidized by reaction with oxygen or glutathione both in a folding buffer and in 6 M guanidine hydrochloride. Purified products of oxidation have been characterized with respect to molecular weights and the positions of disulfides. The four cysteines in conotoxin can form two intramolecular disulfides in three different combinations. Relative yields of each of the three isomers have been determined, thereby permitting evaluation of the roles of non-cysteine residues and cysteine placements in the folding of conotoxin. Cysteine positions dominate factors directing formation of the nativelike isomer in a manner that may be predicted from equilibrium constants for loop formation in model peptides containing two cysteines. Alanine substitutions at several positions which are conserved in naturally occurring conotoxins affect the discrimination between the two most favored disulfide arrangements. Substitutions at three nonconserved positions have no structural effect on isomer yields. It therefore is possible to vary these latter three positions in a manner which might help to generate a functional binding surface which is complementary to receptors in the specific prey of a particular species of snail, without affecting the toxin's folding.

Part of the challenge of understanding protein folding is to evaluate the role of intramolecular disulfide cross-links in small disulfide-rich proteins. Many naturally occurring neurotoxic peptides and small protease inhibitors employ two or more disulfides to achieve a highly stable, compact, and unique structural mainframe. This mainframe then is decorated with a protruding amino acid at a particular location, to generate an inhibitory function for the protein. For example, lysine-14 provides trypsin-binding specificity to bovine pancreatic trypsin inhibitor, a 58-residue protein containing 3 disulfide bonds. Modifications of that lysine have been made to evaluate alterations in specificity of inhibitory binding to proteases (Beckman et al., 1988). Determinations of the relative populations of isomers containing different disulfides have permitted evaluation of the energetics of folding of this protein (Creighton & Goldenberg, 1984).

Peptide toxins such as bee venom apamin, snail venom conotoxins, and *Escherichia coli* enterotoxin are even smaller than pancreatic trypsin inhibitor and contain a higher density of cysteines separated by only a few intervening residues in the sequence. α-Conotoxins are paralytic peptides which block nicotinic acetylcholine receptors in the prey of the snails (Gray et al., 1988). They represent the smallest two-disulfide proteins of known tertiary structure (Kobayashi et al., 1989; Pardi et al., 1989). The sequence of one of the naturally occurring variants (GI from *Conus geographus*) is included in Table I. The roles of individual positions in the sequence on protein activity have been investigated previously (Gray et al., 1984; Nishiuchi & Sakakibara, 1984; Florance et al., 1986; Zafaralla et al., 1988). Variations at positions 1, 4, 10, and 11 and the carboxy terminus have a small effect on activity, whereas changes at proline-5 and glycine-8 substantially decrease activity. The presence of a positively charged lysine or arginine residue in the vicinity of residues 9–10 also is important for activity.

There are three possible monomeric two-disulfide forms of conotoxin. The form having disulfides joining cysteines-2–3 and -7–13 resembles a string of two independent beads. The form with disulfides 2–7 and 3–13 is found in native conotoxins. In the discussion below, it is named the globular form since drawings of that isomer resemble a folded spherical protein. The third form with disulfides 2–13 and 3–7 is called the ribbon topology since for a sequence having equally spaced cysteines, one could draw a planar projection of the structure in the form of a two-stranded antiparallel sheet looking like a short piece of ribbon.

The experiments below define the roles of individual cysteine and non-cysteine residues on the relative stabilities of these three isomeric forms, both in a folding solvent and in a denaturing solvent. Results in concentrated guanidine hydrochloride are compared with predictions based on model studies of peptides containing two cysteines, which were studied in the same type of denaturing conditions (Zhang & Snyder, 1989).

Experimental Procedures

All peptides, in their amidated forms, were synthesized with *p*-methylbenzhydrylamine resins and *t*-Boc coupling chemistry on an automated peptide synthesizer (Biosearch Model 9500). The four Cys residues were differentially blocked with HF-labile *p*-methylbenzyl groups (for the first and third Cys) and HF-resistant acetamidomethyl (Acm)[1] groups (for the second

†This work was supported by Grant GM26715 from the National Institutes of Health.
‡Present address: Schering-Plough Research, 60 Orange St., Bloomfield, NJ 07003.

[1] Abbreviations: HPLC, high-performance liquid chromatography; DTT, dithiothreitol; SS, disulfide; Acm, acetamidomethyl; PTH, phenylthiohydantoin; FAB, fast atom bombardment; GSH and GSSG, reduced and oxidized forms of glutathione, respectively.

0006-2960/91/0430-11343$02.50/0

Table I: Relative Yields (%) of Disulfide Forms after Oxidative Folding[a]

Name[b]	Sequence[c]	%'s					
		Denaturing solvent[d,e]			Folding solvent[d,e]		
	2 3 7 13	Ribb	Bead	Glob	Ribb	Bead	Glob
Shf2	. . . . A C A	58	30	12	66	20	14
GI	E C C N P A C G R H Y S C	22	31	47	21	8	71
8	. . . . A	21	29	50	14	10	76
9	. . . . A . . A	14	23	63	11	12	77
10	. . . A . . . . A A . . .	13	42	45	11	15	74
11	. . . A . . . . A A A . .	15	40	45	18	28	54
12	- . . A . . . . A A . . .	13	45	42	13	30	57
13	- . . A . . . . A A A . .	15	40	45	18	28	54
14	. . . A . . . A A A . . .	14	42	44	15	26	59
15	. . . A . . . . A A . A .	15	48	37	14	24	62
pred	8 and 9 (pro absent)	18	26	56			
avg	8 and 9 (pro absent)	18±5	26±4	56±9	13±2	11±1	76±1
avg	11 thru 15 (# changes ≥ 4)	14±1	43±3	43±3	16±2	27±2	57±3
pred	Shf2 (modified loop sizes; pro absent)	56	30	14			

[a] avg = average, pred = predicted. [b] Peptides 8–15 are variants of naturally occurring GI. "Shf2" is a peptide which substitutes Ala for Pro at position 5, and switches the positions of Ala-6 and Cys-7. [c] A period indicates the same amino acid type as occurs in GI. A minus sign indicates that there is no residue preceding the cysteine at position 2. [d] Ribb, Bead, and Glob are abbreviations for the ribbon, string-of-beads, and globule topologies, respectively. [e] Errors are estimated to be within ±5 relative percentage units. Values are averages of essentially identical results obtained in air oxidation and glutathione oxidation experiments.

and fourth Cys). Crude peptides were cleaved off from the resin and partially deprotected by the standard HF method. The partially deprotected peptides containing two free Cys and two Acm-blocked Cys were purified on preparative reversed-phase HPLC. A portion of each of the purified peptides was converted to the fully reduced form having four free sulfhydryl groups by mercuric acetate removal of Acm groups (van Wandelen et al., 1989). The sequences of the partially deprotected and the fully reduced peptides were confirmed by analysis on an Applied Biosystems Model 470A protein sequencer equipped with a Model 120A on-line PTH analyzer. PTH-Cys (Acm) eluted before PTH-Ala. PTH-Cys in a dehydrated and DTT adduct form eluted at a characteristic position between PTH-Ala and PTH-Tyr.

Ten different synthetic variants were prepared. Their sequences are given in Table I. These variants included the native conotoxin GI molecules, as well as a derivative in which the proline was replaced by alanine and the cysteine at position 7 was switched with the alanine at position 6. This variant alters the size of the disulfide loops. Eight additional forms replace native non-cysteine amino acids with alanine at different positions in the sequence, or delete the N-terminal glutamate.

The purified and partially deprotected peptide having two Cys-Acm groups was used to prepare the standard having the native disulfide pairs, using a two-step oxidation process (Gray et al., 1984). First, air oxidation is used to form the disulfide between cysteines at positions 2 and 7 having free SH groups. Then iodine is used at low pH to remove Acm groups while forming the second native disulfide between cysteines-3 and -13 and simultaneously preventing any disulfide interchange reactions. A single major product was observed by analytical reversed-phase HPLC on a C-18 column. The elution position of this product, corresponding to the isomer with native disulfide pairings, was compared with chromatograms of samples of air-oxidized products of the completely reduced peptide. This permitted assignment of one of the air-oxidized products to the native isomer form.

A solution containing 0.1 mM fully reduced conotoxin was combined with 4 mM HPLC-purified oxidized glutathione at pH 8.5 under a blanket of low-oxygen nitrogen. Aliquots were removed after the first and second hours of reaction. Ellman's reagent, 5,5′-dithiobis(2-nitrobenzoic acid), was used to test aliquots for any net loss of sulfhydryl groups by residual oxygen present in the solution. These assays consistently demonstrated no net loss of the 0.4 mM total SH. In separate aliquots, disulfide exchange was quenched by adding glacial acetic acid to lower the pH to a value of 4.0. These quenched aliquots, representing the kinetic progress of the reaction, were compared with each other by analytical HPLC to access the approach to equilibrium. The validity and utility of this acid-quenching protocol have been discussed previously (Zhang & Snyder, 1988, 1989). For a given reaction, the HPLC chromatograms for 1-h and 2-h time points were consistently identical with each other.

Reduced conotoxin also was fully oxidized in separate reactions by oxygen present in an air-saturated buffer. Use of glutathione facilitates more controllable and more reproducible results than air oxidation, but has the disadvantage of generating some additional products. These extra products correspond to peptides containing intermolecular disulfides between the peptide and glutathione. Oxidation buffers consisted either of a folding solvent containing 20 mM tris(hydroxymethyl)aminomethane and 2 mM ethylenediaminetetraacetic acid at pH 8.5 or of a denaturing solvent made from the same buffer plus 6 M guanidine hydrochloride. Use of two types of oxidizing reagents and two types of buffers generated four sets of data for each synthetic variant. Products of air oxidation, when analyzed by reversed-phase HPLC, typically consisted of a mixture of three species. Relative yields of these species were calculated either by computer integration using an analog-to-digital converter from Interactive Microware Inc., or by estimates calculated from peak heights and peak widths at half-maximum height. Both methods agreed with each other within ±5%. The three oxidation products for each synthetic variant were purified by HPLC for further analysis of molecular weights and disulfide positions.

Molecular weights of purified components were determined by FAB mass spectroscopy. This permitted a determination of whether the molecular species was a monomer having two intramolecular disulfides or was instead a dimer or higher molecular weight oligomer formed by intermolecular disulfide cross-linking of peptide chains.

It is possible to distinguish the string-of-beads isomer by detecting the presence of the disulfide-containing di-PTH-cystine derivative in particular cycles of an automated microsequencing run using DTT-free solvents. When Edman degradation cleaves the peptide bond between an N-terminal cysteine and the rest of the chain to which it is still joined by a disulfide, that derivatized cysteine remains attached to the peptide deposited on the membrane inserted in the sequencer. At a later point, when Edman degradation cleaves off the second cysteine of the disulfide-bridged pair, di-PTH-cystine appears on the analyzer. It elutes either at the same position as PTH-Tyr or slightly earlier, depending on the gradiant and the age of the column. In principle, the string-of-beads isomer (disulfides joining cysteines-2–3 and -7–13) should generate di-PTH-cystine in cycles corresponding to positions 3 and 13

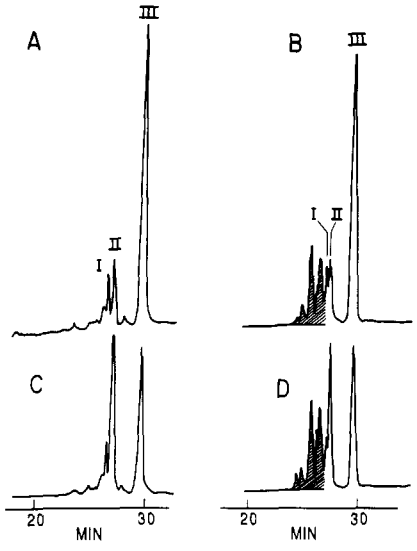

FIGURE 1: HPLC of products of oxidation of conotoxin GI analogue 10. Linear 40-min gradient from 0.1% trifluoroacetic acid in water to 0.1% TFA in 30% acetonitrile, flow rate = 1 mL/min, 23 °C, 3.9 mm × 15 cm Waters Nova-pak C-18 column, detection at 280 nm. Shaded areas represent products containing mixed disulfides between glutathione and the synthetic peptide. Top (A and B) = folding solvent, bottom (C and D) = 6 M guanidine. Left (A and C) = reaction with oxygen, right (B and D) = reaction with glutathione.

of the conotoxin GI sequence. The other two forms, the (2–7, 3–13) and (3–7, 2–13) topologies, both should generate the cystine derivative in cycles corresponding to positions 7 and 13 of the GI sequence. Thus, the string-of-beads isomer may be assigned to any monomeric species producing di-PTH-cystine in the cycle corresponding to position 3.

The overall reaction between completely reduced peptides and oxidized glutathione to generate a particular disulfide isomer is given as

$$\text{peptide-(SH)}_4 + 2\text{GSSG} \rightleftharpoons \text{peptide-(SS)}_2 + 4\text{GSH}$$

Absolute values of the overall equilibrium constant for this reaction for each of the three peptide isomers were determined using synthetic variant "9"; 50 μM reduced peptide, corresponding to 0.2 mM sulfhydryl groups, was equilibrated with mixtures of excess reduced and oxidized glutathione in a denaturing buffer under a blanket of low-oxygen nitrogen. The assays for net oxidation and removal of aliquots were as described above. In this case, however, three different sets of ratios of reduced/oxidized glutathione were used to provide a rigorous test of calculated equilibrium constants. These ratios were 11.4/3.2, 11.4/1.1, and 7.6/3.2, where numbers give concentrations of reduced and oxidized forms, respectively, in millimolar. In each case, the 1- and 2-h data for a given reaction were identical. Under these strongly reducing conditions, measurable amounts of both fully reduced and fully oxidized peptide forms were present in each experiment. Calculations of ratios of observed concentrations gave essentially identical values for equilibrium constants in all three experiments having different ratios of glutathione forms.

Results

HPLC of Aliquots Obtained from the Oxidation Reactions. Figure 1 illustrates HPLC data from the four sets of experiments investigating oxidation of synthetic variant 10. There are three products formed by air oxidation. The same three products are formed in the presence or absence of guanidine, although in very different relative amounts. As will be discussed below, each of the three peaks corresponds to one of the three monomeric isomers. Peak III elutes at the same position as the synthetically prepared standard having the native globular disulfide pairing. In guanidine, one of the nonnative isomers, corresponding to peak I, is strongly disfavored. In a folding solvent, the other nonnative isomer corresponding to peak II also is disfavored. Glutathione oxidation of the peptide, illustrated on the right, generates the same three isomers (peaks labeled I, II, and III). Moreover, those isomers are formed in the same relative ratios which were observed in air oxidation reactions, illustrated on the left. Additional products in glutathione oxidation reactions, indicated by shaded areas in Figure 1, correspond to mixed-disulfide forms incorporating glutathione. They elute before peak I.

Use of Ellman's reagent to analyze fractions corresponding to peaks I, II, and III revealed the absence of free sulfhydryl groups in those fractions. FAB mass spectroscopy of material purified from each peak indicated the presence of a monomer of molecular weight 1115 and the complete absence of dimer. Thus, the peaks correspond to the three possible oxidized monomers containing two intramolecular disulfides. Microsequencing of peak III gave a di-PTH-cystine peak in cycle 3, proving that peak III contains the string-of-beads isomer. Since peak II elutes at the position of the synthetically generated globular isomer standard, peak I therefore is assigned to the ribbon isomer.

Similar methods were used to assign peaks for each synthetic variant. For native conotoxin, the elution order determined by this procedure was ribbon–beads–globule, the same as reported by other investigators (Nishiuchi & Sakakibara, 1982) who used the two-step oxidation protocol to generate synthetic standards for each of the three isomeric forms.

Relative Yields of Different Isomers. The relative yields of different monomeric oxidized forms are given in Table I. In a denaturing solvent, each of the nine species preserving the conotoxin-like cysteine spacing exhibits a disfavoring of the ribbon form relative to the beads and globule topology. In the GI sequence, the nativelike globular disulfide pairing predominates (47%), even in the absence of electrostatic and hydrogen-bonded folding interactions which are unlikely in this solvent. Variants 8 and 9, in which proline has been replaced with alanine, are the best for comparisons with predictions of relative percentages on the basis of previously determined properties of alanine-containing disulfide loops in concentrated guanidine. A discussion of such predictions will be presented later. For the moment, one notices that the average behavior for variants 8 and 9 is identical to their predicted behavior presented in the table, namely 18, 26, and 56% relative population of the ribbon, bead, and globule topologies, respectively. Variants 11–15, having the largest number of changes from the native sequence, exhibit nearly identical data when compared with each other, as seen by the small standard deviations for their average behavior summarized at the bottom of the table.

In a folding solvent, the conotoxin-like variants still exhibit disfavoring of the ribbon form. Now, however, there is enhancement of formation of the nativelike globular disulfide pairing in each derivative. Variants 11–15 exhibit nearly identical behavior compared with each other, as also was true in the denaturing solvent.

The cysteine-shifted proline-substituted variant shf2 exhibits relative percentages which are in the opposite rank order from those observed for the conotoxin variants. As seen in Table I, shf2 exhibits disfavoring of the globule form (only 12–14% population) and favoring of the ribbon form (58–66% population). For this variant, the relative yield for any given isomer exhibits approximately the same value in both the denaturing

Table II: Overall Equilibrium Constant for Formation of Oxidized Monomers from Sequences Lacking Proline, Starting from Fully Reduced Peptide and Oxidized Glutathione[a]

	conotoxin variants		Shf2
isomer	predicted	measured[b]	predicted
ribbon	1.3	1.3 ± 0.4	5.1
beads	2.2	2.2 ± 0.8	2.7
globule	4.0	8.5 ± 3.3	1.3

[a] K equilibrium values given as $K \times 10^3$ (M^2). [b] Measured data are for variant 9, which lacks proline. The tabulated value is the average of determinations using three different ratios of GSH/GSSG in reaction mixtures.

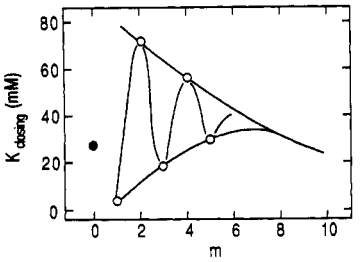

FIGURE 2: $K_{closing}$ in concentrated guanidine hydrochloride: extrapolation of data for m = 1–5 to obtain an estimate for the value at m = 6–9. Data are replotted from published data (Zhang & Snyder, 1989) for peptides containing a Cys-Ala_m-Cys region. The closed circle (●) represents the special case requiring a cis peptide bond when m = 0. m is the number of intervening residues between the two cysteines forming the disulfide loop. $K_{closing}$ is for the intramolecular loop-forming reaction between a free SH group and a mixed-peptide–glutathione disulfide in the same chain, such as the reactions given by the second and fourth steps in Figure 3.

and folding solvents. Apparently, the guanidine does not have much influence on this sequence.

Absolute values of the overall equilibrium constant for formation of the three isomers from the fully reduced form of variant 9 were determined experimentally in a denaturing solvent. These values were essentially independent of the ratio of excess reduced glutathione to excess oxidized glutathione in the reaction mixture, as required by chemical principles. Average measured values are given in Table II. Measured values are consistent, within experimental error, with predictions to be discussed below which are based on previous studies (Zhang & Synder, 1989) of alanine-containing disulfide loops in concentrated guanidine.

Discussion

Estimation of Predicted Yields for Proline-Absent Variants. The dependence of disulfide loop formation on the number and types of residues in the loop has been defined previously (Table I; Zhang & Snyder, 1989). Those studies indicated that the type of residues was less important than the number of residues for loops with 5 or fewer positions (m) between the bridged cysteines. The case of m = 0 is special since adjacent cysteine residues can form a disulfide only when the peptide bond between them is in the cis configuration. For loops using only alanines between the cysteines, there is a distinctive odd–even pattern in the value of loop stability. This pattern is given in Figure 2, which illustrates the favorable nature of small loops with even-numbered m and the unfavorable nature of small loops with odd-numbered m. The amplitude of the oscillating pattern diminishes at larger values of m, and is predicted to degenerate into a general dependence of the type where K is proportional to $m^{-3/2}$ (Kauzmann, 1959; Mutter, 1977). This latter dependence represents the general decrease in probability of random collisions between two cysteines as they are separated by longer distances in the linear sequence.

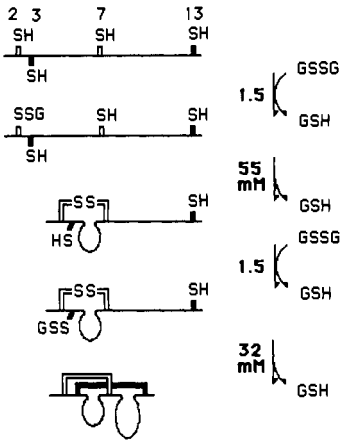

FIGURE 3: Estimated equilibrium constants for one pathway forming the globule topology. The dimensionless equilibrium constant with a value of 1.5 for bimolecular reactions with glutathione is based on previous experimental kinetic studies with cysteine-containing peptides (Zhang & Snyder, 1988), not on conotoxin data. The equilibrium constant for intramolecular loop formation having a value of 55 mM was observed not in conotoxin itself but rather in a peptide containing two cysteines separated by four alanines (Zhang & Snyder, 1989). The final loop-closing reaction step having an estimated equilibrium constant of 32 mM is based on extrapolated curves in Figure 2, for m having a value in the range of 7–9 effective residues.

Figure 3 illustrates one particular pathway for forming the globular isomer of conotoxin GI by reaction between the fully reduced peptide and oxidized glutathione. By the principles of coupled equilibria, all different pathways for converting the reduced form to the globule isomer must have the same numerical value for the equilibrium constant for the composite overall pathway. Estimation of that equilibrium constant, using the details for one given pathway, therefore suffices as an estimate for the value expected by any other pathway.

Reactions of oxidized glutathione with individual reduced cysteines to generate a mixed-peptide–glutathione disulfide typically exhibit a value of 1.5 (Zhang & Snyder, 1988). Figure 3 incorporates two such steps. Previous studies of the peptide containing the Cys-Ala-Ala-Ala-Cys sequence had a $K_{closing}$ = 55 mM. This value therefore is incorporated into Figure 3 as an estimate of the magnitude of the corresponding closing reaction in a conotoxin variant in which proline-5 is replaced by alanine. After formation of the m = 3 loop joining cysteines-3 to -7, cysteines-2 and -13 are brought closer to each other than in the original reduced chain. Thus, one should *not* use m = 11 for formation of the second loop. Instead, one may estimate that the situation most likely is equivalent to loop formation in a chain which is shortened by several residues, having m somewhere in the range of 7–9 residues. Thus, Figure 3 incorporates a value of 32 mM for the final step in the reaction, since this is the estimate generated in Figure 2 from hand-drawn extrapolation of the dampened oscillations in values of $K_{closing}$. A prediction for the overall equilibrium constant in a proline-absent variant therefore is 4.0×10^{-3} M^2, equal to the product 1.5 × 55 mM × 1.5 × 32 mM.

This estimate for the overall equilibrium constant is included in Table II. Values for similar predictions for the ribbon and bead forms also are included. One expects the estimate for the beads form to be most accurate, since formation of each of the two loops is independent of what happens in the other loop and therefore allows a direct comparison with model peptide data. As seen in Table II, the observed and predicted values of overall equilibrium constants are identical to each other for the bead and ribbon forms. Moreover, predictions and observations for the globule form are consistent with each

other within experimental error. This agreement between observed and predicted values supports the approach taken in Figure 3. At pH 8.5, the half-times for the individual steps in forming the disulfide bonds are no greater than 5 min, suggesting that the similar profiles obtained after 1 and 2 h are true equilibrium profiles, although this has not been demonstrated directly.

The relative values of K given in Table II should be directly proportional to the relative equilibrium concentrations of the three isomers. The bottom two rows of Table I give relative concentrations predicted with equations of the following type for the fraction of isomers in the globule form:

$$f_{\text{globule}} = K(\text{globule})/[K(\text{globule}) + K(\text{ribbon}) + K(\text{beads})]$$

Observed Relative Yields of Different Isomeric Forms. Data for conotoxin variants in a denaturing solvent demonstrate that disulfide loop sizes dominate the results. The ribbon form is disfavored for all sequences because it includes the small odd-numbered ring with $m = 3$, the second most unfavorable size. The globule topology is generally favored, because it includes the even-numbered $m = 4$ loop, the second most favored size. This dominance leads to the close agreement between observed and predicted values of relative yields and overall equilibrium constants. The agreement is particularly good in the proline-absent derivatives, since calculations are based on alanine-containing model two-cysteine peptides (Zhang & Snyder, 1989).

Under folding conditions, steric and configurational entropy factors mentioned above will continue to contribute and will be supplemented by additional factors such as hydrogen-bonding interactions. Variants 11–15, having three to four side-chain replacements by alanine and occasional loss of the N-terminal glutamate, all exhibit identical values for relative yields of the three possible isomeric forms. In these five variants, only one or two non-alanine non-cysteine side chains are present, and their choice does not affect the observed yields. It is possible that these yields, 16–27–57% for ribbon–beads–globule, represent the basal level of effects created by backbone interactions alone.

Synthetic variants 8–10 have fewer than four differences from native conotoxin. For these variants, as well as native conotoxin, the globule form occurs in more than 70% of the molecules in a folding solvent. This represents an increase relative to the percentage observed in the highly substituted variants in the same solvent, or to any variant in a denaturing solvent. Thus, side-chain substitutions at positions 4, 5, 8, 9, and 10 in variants 8–10 appear to have only minor effects on folding. This behavior is not surprising for positions 4, 9, and 10, which are not highly conserved among the naturally occurring conotoxin variants.

It appears, therefore, that the conotoxin folding is defined primarily by the relative placement of cysteines at positions 2, 3, 7, and 13. In proline- and glycine-containing variants with many alanine side chains, this predominance is at the level of 57% globule form. Placement of Glu, Tyr, and Ser at positions 1, 11, and 12 enhances formation of the globule form, increasing the yield from 57% to about 72%. This framework then may be decorated with a variety of residues at positions 4, 9, and 10 to confer functional binding to receptor sites which are unique to the particular prey attacked by the particular snail.

The behavior of the cysteine-shifted analogue shf2 also is dominated by the sizes of the loops in its isomers. In this case, the ribbon form is favored since it incorporates a loop with $m = 2$, the most favorable size. The globule form is unstable since it includes the case of $m = 3$, the second most disfavored size. Behavior in guanidine matches predictions based on the two-cysteine model peptides previously studied in concentrated guanidine. In contrast with conotoxin variants, switching from the denaturing solvent to folding conditions does not affect relative populations of isomers. This is reasonable, since changes in disulfide positions would prevent adoption of the conotoxin-like folding. The shf2 variant does not resemble a known naturally occurring protein and was not designed to employ a predicted set of side-chain and backbone folding interactions.

Comparison with Other Proteins. Conotoxin, apamin, and endothelin are 3 small proteins containing 21 or less amino acids, including 4 cysteines which form 2 intramolecular disulfides. Although it has been hypothesized that distant regions of precursor chains might play a role in the formation of disulfides in the cysteine-rich regions (Woodward et al., 1990), this has not been demonstrated experimentally. The final forms of conotoxin, endothelin, and apamin are fully competent to re-form at least 70% of the nativelike isomer after reduction and reoxidation of their disulfides. As demonstrated above, the conotoxin sequence produces slightly more than 70% native monomeric globule form in glutathione-catalyzed disulfide exchange experiments. The disulfide pairing in this protein is directed primarily by the relative positioning of the cysteines, since at least 54% globule form is produced in folding buffers for each of the variants examined. Apamin produces more than 90% of its native globule topology in similar experiments, with endothelin producing about 75% of its native ribbon topology (Nakajima et al., 1989). Since endothelin and apamin have identical relative positioning of cysteines but adopt different disulfide pairings in solution, at least one of them must employ a major contribution from side-chain interactions for purposes of stabilizing the native form and thereby directing formation of the correct corresponding disulfides. The bead forms of apamin and endothelin would require formation of loops of size $m = 1$ and $m = 3$, the two least favored sizes. Thus, as was the case in conotoxin, cysteine placements appear to be utilized to disfavor this topology.

Acknowledgments

We thank Dr. Alice Bergmann of the Department of Chemistry at the State University of New York at Buffalo for help in obtaining FAB mass spectroscopy data.

Registry No. **8**, 136706-37-1; **9**, 136706-38-2; **10**, 136706-39-3; **11**, 136706-40-6; **12**, 136706-41-7; **13**, 136735-98-3; **14**, 136706-42-8; **15**, 136735-99-4; GI, 78249-65-7; Shf2, 136706-36-0.

References

Beckmann, J., Mehlich, A., Schröder, W., Wenzel, H. R., & Tschesche, H. (1988) *Eur. J. Biochem. 176*, 675–682.

Creighton, T. E., & Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497–526.

Florance, J., Galdes, A., Konteatis, Z., Gold, B. & Lenz, G. (1986) *Int. Symp. HPLC Pept. Proteins Oligonucleotides 739*, 34.

Gray, W. R., Luque, F. A., Galyean, R., Atherton, E., Sheppard, R. C., Stone, B. L., Reyes, A., Alford, J., McIntosh, M., Olivera, B. M., Cruz, L. J., & Rivier, J. (1984) *Biochemistry 23*, 2796–2802.

Gray, W. R., Olivera, B. M., & Cruz, L. J. (1988) *Annu. Rev. Biochem. 57*, 665–700.

Kauzmann, W. (1959) in *Sulfur in Proteins* (Benesch, R., Benesch, R. E., Boyer, P. D., Klotz, I. M., Middlebrook, W. R., Szent-Gyorgyi, A. G., & Schwarz, D. R., Eds.) pp 93–108, Academic Press, New York.

Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., Braun, W., & Go, N. (1989) *Biochemistry 28*, 4853–4860.
Mutter, M. (1977) *J. Am. Chem. Soc. 99*, 8307–8314.
Nakajima, K., Kumagaye, S. I., Nishio, H., Kuroda, H., Watanabe, T. X., Kobayashi, Y., Tamaoki, H., Kimura, T., & Sakakibara, S. (1989) *J. Cardiovasc. Pharmacol. 13 (Suppl. 5)*, S8–S12.
Nishiuchi, Y., & Sakakibara, S. (1982) *FEBS Lett. 148*, 260–262.
Nishiuchi, Y., & Sakakibara, S. (1984) in *Peptide Chemistry 1983* (Sakakibara, S., Ed.) pp 191–196, Protein Research Foundation, Osaka, Japan.
Pardi, A., Galdes, A., Florance, J., & Maniconte, D. (1989) *Biochemistry 28*, 5494–5501.
van Wandelen, C., Zeikus, R., & Tsou, D. (1989) in *Chemistry Update*, pp 1–16, Milligen/Biosearch, Burlington, MA.
Woodward, S. R., Cruz, L. J., Olivera, B. M., & Hillyard, D. R. (1990) *EMBO J. 9*, 1015–1020.
Zafaralla, G. C., Ramilo, C., Gray, W. R., Karlstrom, R., Olivera, B. M., & Cruz, L. J. (1988) *Biochemistry 27*, 7102–7105.
Zhang, R., & Snyder, G. H. (1988) *Biochemistry 27*, 3785–3794.
Zhang, R., & Snyder, G. H. (1989) *J. Biol. Chem. 264*, 18472–18479.

Characterization of Phosphate Binding in the Active Site of Barnase by Site-Directed Mutagenesis and NMR†

Elizabeth M. Meiering, Mark Bycroft, and Alan R. Fersht*

MRC Unit for Protein Function and Design, Cambridge IRC for Protein Engineering, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

Received May 23, 1991; Revised Manuscript Received August 7, 1991

ABSTRACT: Phosphate is a competitive inhibitor of transesterification of GpC by the ribonuclease barnase. Barnase is significantly stabilized in the presence of phosphate against urea denaturation. The data are consistent with the existence of a single phosphate binding site in barnase with a dissociation constant, K_d, of 1.3 mM. The 2D ^{1}H NMR spectrum of wild-type barnase with bound phosphate is assigned. Changes in chemical shifts and NOEs for wild type with bound phosphate compared with free wild type indicate that phosphate binds in the active site and that only small conformational changes occur on binding. Site-directed mutagenesis of the active site residues His-102, Lys-27, and Arg-87 to Ala increases the magnitude of K_d for phosphate by more than 20-fold. The 2D ^{1}H NMR spectra of the mutants His-102 → Ala, Lys-27 → Ala, and Arg-87 → Ala are assigned. Comparison with the spectra of wild-type barnase reveals that His-102 → Ala and Lys-27 → Ala have essentially the same structure as wild type, while some structural changes occur in Arg-87 → Ala. It appears that phoshate binding by barnase is effected mainly by positively charged residues including His-102, Lys-27, and Arg-87. This may have applications for the design of phosphate binding sites in other proteins.

About half of all known proteins bind compounds containing phosphoryl groups (Schultz & Schirmer, 1979). The recognition and binding of phosphoryl groups are important in the interaction of proteins with a wide range of molecules including substrates, cofactors, allosteric effectors, and ligands. The structures of phosphoryl binding sites in a number of proteins are known from X-ray crystallography, for example, nucleotide binding proteins such as human c-H-*ras* oncogene protein complexed with GDP (de Vos et al., 1988), triosephosphate isomerase complexed with dihydroxyacetone phosphate (Lolis & Petsko, 1990), and bovine pancreatic ribonuclease A complexed with mono- and dinucleotide inhibitors [for a summary see Wlodawer (1985)]. Electrostatic interactions between the negatively charged phosphate and positive groups on proteins are important for the binding of phosphate by proteins. This includes charge–charge interactions between phosphate and positively charged side chains that are found in virtually all phosphate binding sites, charge–dipole interactions between phosphate and α-helix dipoles (Hol et al., 1978), and hydrogen bonding between phosphate and side-chain or main-chain peptide groups as, for example, in the Gly-X-X-X-X-Gly-Lys phosphate binding loop of purine nucleotide binding proteins (Campbell-Burk, 1989; Redfield & Papastavros, 1990). There is, however, little experimental data on the energetics of interactions between proteins and phosphate, and the specificity of phosphate binding sites is poorly understood. Here, we characterize a phosphate binding site in which phosphate is bound mainly by positively charged side chains in the ribonuclease barnase.

Barnase is a small endoribonuclease of 110 amino acids from *Bacillus amyloliquefaciens*. It is a member of a family of guanine-specific and guanine-preferential microbial endoribonucleases, including enzymes from both bacterial and fungal species (Hill et al., 1983). Hydrolysis of RNA catalyzed by the microbial ribonucleases occurs in a two-step

†E.M.M. was the recipient of a 1967 Science and Engineering Scholarship from the Natural Sciences and Engineering Research Council of Canada.
*Author to whom correspondence should be addressed.

0006-2960/91/0430-11348$02.50/0