Size of Hirudin Sequence Required To Fold into an Active Core Domain

Jui-Yoa Chang

Pharmaceuticals Research Laboratories, Ciba-Geigy Ltd., Basel CH-4002, Switzerland

Received May 9, 1995; Revised Manuscript Received June 19, 1995[⊗]

ABSTRACT: The active core domain of hirudin contains three native disulfides (Cys^{6}−Cys^{14}, Cys^{16}−Cys^{28}, and Cys^{22}−Cys^{39}) and 49 amino acid residues (Hir^{1-49}). This compact structure folds spontaneously, and its folding pathway has been elucidated recently [Chatrenet and Chang (1993) *J. Biol. Chem. 268,* 20988−20996]. The folding mechanisms of Hir^{1-35} and Hir^{1-43} were investigated in order to determine the minimum structural elements required to fold into this active core structure. Hir^{1-35} includes two native disulfides (Cys^{6}−Cys^{14} and Cys^{16}−Cys^{28}) and an extra Cys^{22}. When reduced/denatured Hir^{1-35} was allowed to fold, it folded into a collection of equilibrated 2-disulfide isomers. At least eight fractions of the 2-disulfide species have been observed. Structural analysis revealed that out of the 10 possible disulfide pairings, only five were detected to exist in the 2-disulfide isomers, and all have their half-cystines separated by less than 8−10 amino acid residues. One of the native disulfides, Cys^{16}−Cys^{28}, has not been found in any of those 2-disulfide species. On the other hand, the C-terminal extension of an octapeptide permitted reduced/denatured Hir^{1-43} to fold into a defined active structure possessing the three native disulfides. These results also demonstrate that the folding mechanism of Hir^{1-35} resembles what occurs during the early stage of Hir^{1-43} (and Hir^{1-49}) folding, which involves a process of nonspecific packing of unfolded polypeptide chains.

Hirudin is a thrombin specific inhibitor obtained from the leech *Hirudo medicinalis* (Markwardt & Walsmann, 1958). The inhibitor contains a disordered carboxyl-terminal tail (Hir^{50-65})[1] and a compact amino-terminal core (Hir^{1-49}) which is stabilized by three native disulfides (Dodt et al., 1985; Haruyama & Wuethrich, 1989; Folkers et al., 1989; Chang, 1990). The core domain requires proper folding in order to maintain function, and its disulfide folding pathway has been recently elucidated (Chatrenet & Chang, 1992, 1993; Chang, 1994). The results unveil a number of important features of the folding mechanism, and some of them indeed had not been fully anticipated by us. (1) The folding begins with a nonspecific packing of the polypeptide, followed by the consolidation of partially packed intermediates to attain the active structure. The course of packing involves a sequential and irreversible flow of unfolded, 0-disulfide hirudin (R) through equilibrated 1-disulfide isomers and 2-disulfide isomers to equilibrated 3-disulfide (scrambled) species. The process of consolidation is characterized as scrambled hirudins reorganize to reach the native structure. (2) The non-covalent specific interactions—forces that stabilize the folded structure—apparently do not actively participate in guiding the folding of Hir^{1-49} until the unfolded protein condenses to a certain degree of compactness. In this case, the effect of these forces becomes obvious at the final stage of folding as scrambled (mispaired 3-disulfide) species of Hir^{1-49} reorganize to adopt the active conformation. As a consequence, the inclusion of denaturant in the folding buffer exerts no influence on the compositions of 1- and 2-disulfide intermediates. (3) Scrambled hirudins are not aborted species. Thermodynamically, they are relevant intermediates and passage to the active hirudin. (4) The folding intermediates exist in dynamic equilibrium and are enormously heterogeneous. (5) The relative rates of packing and consolidation, and consequently the level of accumulation of scrambled hirudins along the folding pathway, can be selectively regulated by a number of external factors. For example, in the presence of GSSG or cystine, the speed of packing is greatly accelerated. Under these conditions, the process of consolidation becomes the major rate-limiting step, and scrambled species become the most dominant folding intermediates.

Hirudin is not alone in displaying these properties. Two other single-domain cystine-containing proteins, namely, potato carboxypeptidase inhibitor (39 amino acids and 3 disulfides) and human epidermal growth factor (53 amino acids and 3 disulfides), have also been shown to refold through a similar mechanism (Chang et al., 1994, and 1995). This mechanism of protein folding is distinguished from the well-documented BPTI model (Creighton, 1978, 1990; Weissman & Kim, 1991) in two crucial aspects: (1) The apparent absence of non-covalent specific interactions in guiding the initial stage of folding. In the case of BPTI, well-populated 1- and 2-disulfide intermediates were shown to contain exclusively native disulfides. This has led to the conclusion that specific interactions play an essential role during the entire process of folding (Weissman & Kim, 1991). (2) The presence of non-native 3-disulfide (scrambled) species as folding intermediates. Scrambled species were found in significant concentration along the folding pathways of hirudin, PCI, and EGF but have not been observed in the case of BPTI (Creighton, 1978, 1990). Those discrepancies remain to be clarified, but they may possibly demonstrate the vast diversity of mechanisms engaged in the process of protein folding. On the other hand, the folding mechanism of hirudin appears to bear a crucial resemblance to that of

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1995.

[1] Abbreviations: BPTI, bovive pancreatic trypsin inhibitor, CI2, chymotrypsin inhibitor 2; GdmCl, guanidinium chloride; GSH, reduced glutathione; GSSG, oxidized glutathione; Hir, hirudin; HPLC, high-performance liquid chromatography.

0006-2960/95/0434-11160$09.00/0

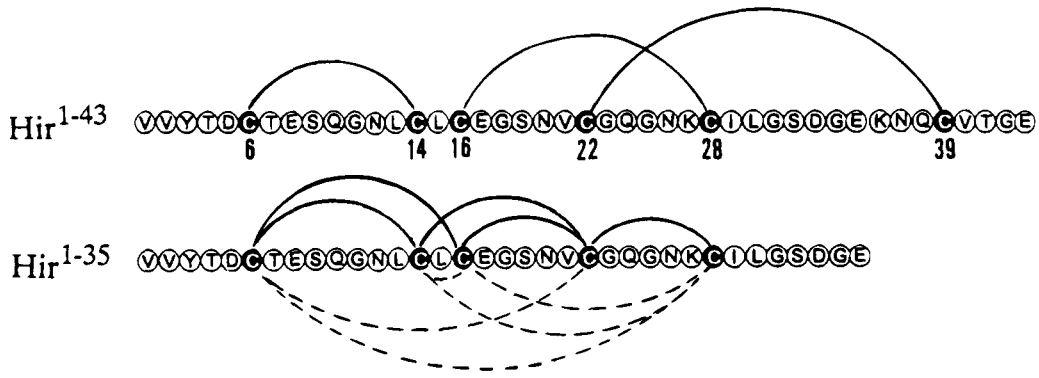

FIGURE 1: Primary structures of Hir^{1-43} and Hir^{1-35}. Reduced/denatured Hir^{1-43} folds into an active structure (see Figure 8) stabilized by three native disulfides (Cys^6-Cys^{14}, $Cys^{16}-Cys^{28}$, and $Cys^{22}-Cys^{39}$). For Hir^{1-35}, the folding leads to a mixture of equilibrated 2-disulfide isomers. Only 5 out of 10 possible disulfide bonds (solid lines) were detected to exist within these 2-disulfide isomers. The remaining 5 disulfide bridges (broken lines) were not observed in any significant amount.

pro-BPTI (Weissman & Kim, 1992). When reduced/denatured pro-BPTI is allowed to refold in the presence of GSH/GSSG, a group of scrambled pro-BPTIs are shown to appear in the folding pathway as well-populated intermediates. Furthermore, the folding of hirudin is characteristically similar to that found in the case of ribonuclease A (Anfinsen et al., 1961; Scheraga et al., 1984; Wetlaufer et al., 1987). A two-stage mechanism has also been observed during the renaturation of ribonuclease A (Anfinsen et al., 1961; Hantgan et al., 1974; Pigiet & Schuster, 1986), in which an initial stage of random disulfide formation is promoted by oxidized glutathione, followed by disulfide reshuffling catalyzed by reduced thioredoxin.

In this report, the folding mechanisms of two hirudin amino-terminal fragments, Hir^{1-35} and Hir^{1-43}, have been investigated. The goal of this study is two-fold: (1) to identify the minimum structural elements of hirudin required to fold into an active core domain and (2) to distinguish between the folding mechanisms of polypeptide fragments derived from the same protein which are able and unable to reach the functional structure.

EXPERIMENTAL PROCEDURES

Preparation of Hir^{1-35} and Hir^{1-43}. Hir^{1-43} was prepared from the intact recombinant hirudin (Hir^{1-65}) by selective removal of its C-terminal tail using V8 protease (Chang et al., 1990). Hir^{1-35} was chemically synthesized by the solid-phase method using Fmoc chemistry and was kindly supplied by Hans Rink (Ciba-Geigy, Basel). The purity and structure of both fragments were confirmed by HPLC profile, composition analysis, amino acid sequencing, and mass spectrometry. Amino acid sequences of the two fragments are shown in Figure 1.

Folding Experiments and Trapping of Folding Intermediates. Folding of Hir^{1-35} was performed in either sodium bicarbonate buffer (50 mM, pH 8.3) or Tris-HCl buffer (0.1 M, pH 8.5). Hir^{1-35} (4 mg) was first reduced and denatured by incubating for 90 min in 0.8 mL of Tris-HCl buffer (0.5 M, pH 8.4) in the presence of 5 M GdmCl and 30 mM dithiothreitol. To initiate the folding, the sample was passed through a PD-10 column (Pharmacia) equilibrated in the $NaHCO_3$ buffer (or Tris-HCl buffer). The desalting process took 1−2 min. Reduced/denatured Hir^{1-35} was recovered in 1.4 mL and was immediately diluted with the same $NaHCO_3$ buffer (or Tris-HCl buffer) to a final concentration of 140 μM. Selected concentrations of glutathione or β-mercaptoethanol were introduced simultaneously. Folding was carried out at 22 °C. The folding intermediates were trapped in a time-course manner by acidifying aliquots of the folding sample with 2 vol of 4% trifluoroacetic acid. Those acid-trapped intermediates were analyzed by HPLC (Figures 2 and 3), and purified fractions of intermediates were subsequently used for stop/go folding experiments (Figure 4) and structural analysis following carboxymethylation (Figures 5 and 6).

Hir^{1-43} was reduced/denatured and allowed to refold under similar conditions in the sodium bicarbonate buffer, in both the absence and the presence of β-mercaptoethanol (100 μM) (Chatrenet & Chang, 1993). Acid-trapped folding intermediates of Hir^{1-43} were analyzed under the HPLC conditions described in Figure 7.

Stop/Go Folding Experiments with Acid-Trapped Intermediates. Acid-trapped intermediates were further purified by HPLC (see Figure 2), freeze-dried, and reconstituted (10 μM) in a $NaHCO_3$ solution (50 mM, pH 8.3) to resume the folding (Chang, 1993). The folding process was again trapped at different time intervals by mixing with 2 vol of 4% trifluoroacetic acid.

Carboxymethylation of Acid-Trapped Folding Intermediates and Products. For structural analysis, acid-trapped intermediates and products (fractions purified by HPLC) which contain free cysteines must first be carboxymethylated. This was achieved by using the following conditions specifically designed to suppress the disulfide shuffling (Chang, 1993): Isolated 1-disulfide intermediate or 2-disulfide product of Hir^{1-35} (5−10 nmol, freshly dried from the acid solution) was dissolved in 200 μL of a mixed solution containing 40% (by volume) dimethyl formamide, 60% Tris-HCl buffer (0.5 M, pH 6.5), and 1 M iodoacetic acid. Carboxymethylation was carried out at 22 °C for 20 min. The sample was then passed through a NAP-5 column (Pharmacia) equilibrated in ammonium bicarbonate solution (50 mM) and was recovered in a total volume of 0.55 mL.

Structural Analysis. The disulfide content was determined by the dabsyl chloride method (Chang & Knecht, 1991). The amino acid sequence was analyzed by using a 477A sequencer from Applied Biosystems Inc. (Hunkapiller et al., 1981). Molecular mass was determined by laser desorption ionization mass spectrometry using a home-built instrument (Boernsen et al., 1990).

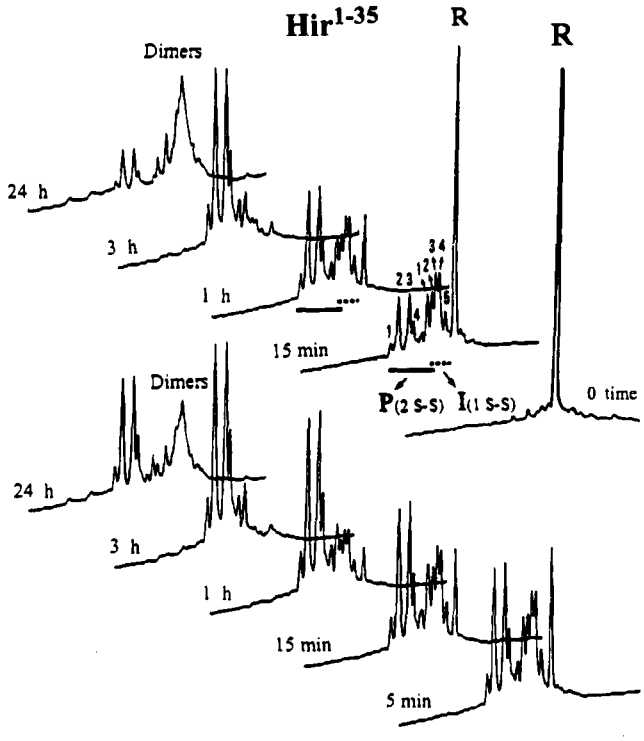

FIGURE 2: HPLC chromatograms of acid-trapped folding intermediates of Hir^{1-35}. The folding was carried out either in the $NaHCO_3$ buffer (50 mM, pH 8.3) (top panel) or in the Tris-HCl buffer (0.1 M, pH 8.5) (bottom panel). R is fully reduced/denatured Hir^{1-35}. Fractions within group I are 1-disulfide intermediates, and those within group P are 2-disulfide products. Upon extended incubation in the folding buffer (24 h), all 2-disulfide products proceed to form heterogeneous dimers. The HPLC column was Vydac C-18 for peptides and proteins. Solvent A was 0.1% trifluoroacetic acid in water. Solvent B was 0.1% trifluoroacetic acid in acetonitrile/water (9:1, v/v). The gradient was 22−40% solvent B, linear, for 25 min. Detector wavelength was 214 nm. P-2 and R elute at 16.9 and 20.5 min, respectively.

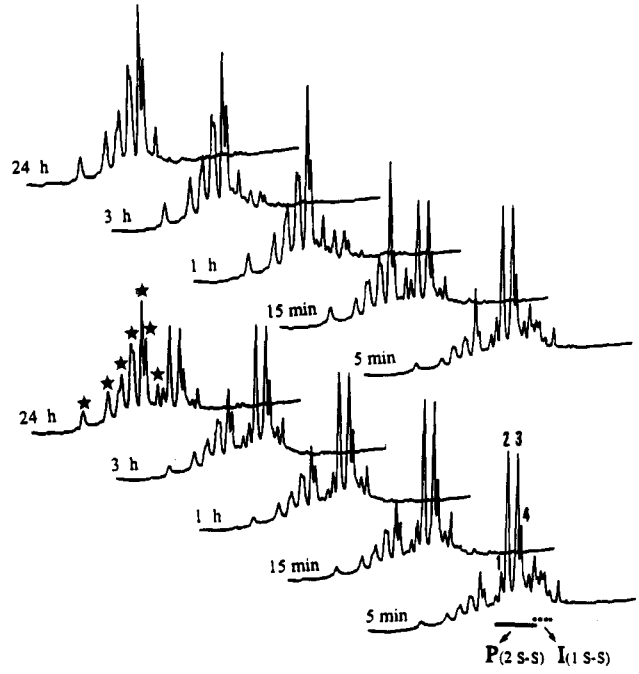

FIGURE 3: Folding of Hir^{1-35} performed in the presence of GSSG (1 mM) (top panel) and a GSH/GSSG mixture (2/1 mM) (bottom panel). Folding intermediates were trapped by acid and analyzed under the same HPLC conditions described in Figure 2. The formation of 2-disulfide products (P) was significantly accelerated under these folding conditions. Upon prolonged folding, all 2-disulfide products proceeded to form mixed disulfides with reduced glutathione (peaks marked by stars). This was confirmed by mass spectrometry. The heterogeneity of mixed disulfide species indicates that there are at least eight isomers of 2-disulfide Hir^{1-35}.

RESULTS

Folding of Hir^{1-35}. Folding of Hir^{1-35} was first carried out in the buffer alone (sodium bicarbonate buffer or Tris-HCl buffer), and acid-trapped folding intermediates were analyzed by HPLC (Figure 2). As the folding progressed, the decrease of R was accompanied by the recovery of two groups of equilibrated isomers, designated as I for intermediate and P for product, which were ascending and descending along the folding pathway in a sequential and irreversible manner. More than 95% of the starting material converted to P after 180 min. In the Tris-HCl buffer, the initial collapse of R was significantly accelerated as compared to that performed in the sodium bicarbonate buffer; however, the patterns and compositions of I and P remained indistinguishable (Figure 2, bottom panel). The mechanism of the effect of Tris buffer is not clear, but similar results have been observed during the refolding experiments of Hir^{1-49} (Chatrenet & Chang, 1993) and EGF. These two groups of isomers partially overlap (P-5 overlaps with I-1). Fractions of I and P were subsequently isolated. They were either (1) carboxymethylated (see Experimental Procedures) and then subjected to structural analysis or (2) directly used for stop/go folding experiments (Figure 4). Group I was found to contain only 1-disulfide isomers, and group P consisted of 2-disulfide isomers. This was confirmed by both amino acid analysis and mass determination of carboxymethylated samples. Carboxymethylated 1-disulfide and 2-disulfide Hir^{1-35} have molecular weights of 3771 and 3655, respectively. Upon prolonged incubation (24 h) in the folding buffer, the 2-disulfide products (P) slowly oxidized to form heterogeneous dimers. The precise molecular weight of these dimers (7184) was also confirmed by mass spectrometry. It is important to notice that the pairing of inter-disulfide bonds which leads to the formation of dimers is significantly slower than the formation of intra-disulfides. This has also been found during the preparation of active hirudin dimer by *in vitro* folding (Chang et al., 1993). The presence of an extra unpaired cysteine serves to catalyze disulfide exchange, facilitates the process of folding, and apparently has no effect on the pathway of folding (Weissmann & Kim, 1992; Chang et al., 1993).

When folding of Hir^{1-35} was performed in the sodium bicarbonate buffer containing GSSG (1 mM) or a mixture of GSH/GSSG (2/1 mM), the flow from R *via* I to P was complete within 5 min (Figure 3). The acceleration of the flow is dependent on the concentration of the oxidized glutathione which provides a redox potential that favors the disulfide formation of reduced polypeptides (Saxena & Wetlaufer, 1970; Light et al., 1974; Hantgan et al., 1974; Orsini et al., 1975; Pigiet & Shuster, 1986; Lyles & Gilbert, 1991). This finding is consistent with what has been observed in the folding of Hir^{1-49} (Chang, 1994) in which oxidized glutathione alone greatly enhances the efficiency of packing by promoting the disulfide formation. Again, the pattern of I and P (Figure 3) remains indistinguishable from that performed in the absence of redox agents (Figure 2). Mixed disulfide species of I may exist in such minute

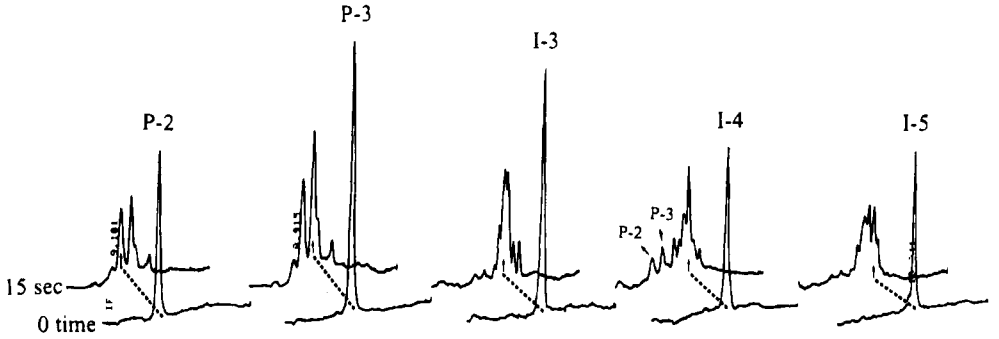

FIGURE 4: Stop/go folding experiments initiated with purified 1-disulfide intermediates and 2-disulfide products. Purified P-2, P-3, I-3, I-4, and I-5 were freeze-dried and reconstituted (10 μM) in a $NaHCO_3$ solution (50 mM, pH 8.3) to resume the folding. The samples were trapped after 15 s by mixing with 2 vol of 4% trifluoroacetic acid and were directly analyzed by the HPLC systems described in the caption of Figure 2.

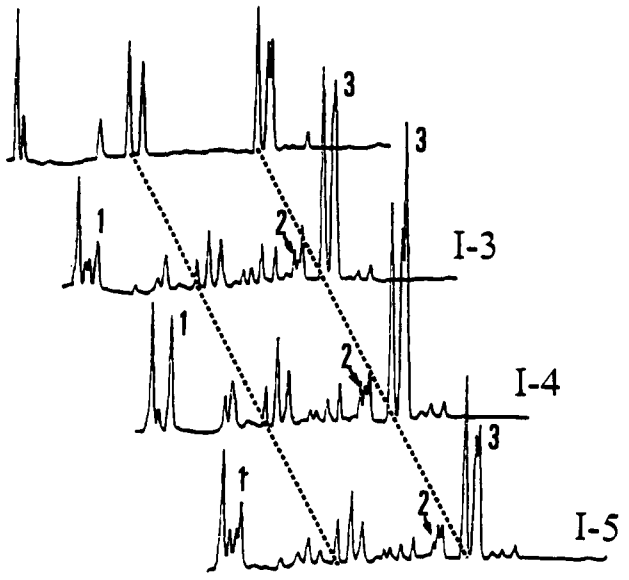

FIGURE 5: Peptide mapping of thermolysin-digested I-3, I-4, and I-5. Acid-trapped I-3, I-4, and I-5 were carboxymethylated, digested by thermolysin, and separated by HPLC using the gradient described in the caption of Figure 6. Nine fractions of peptides were sequenced, and the results are given in Table 1. A control sample similar to that of Figure 6 is also presented.

concentrations that they are not noticeable on the chromatogram. The major difference is that, instead of generating the dimers, the 2-disulfide products (P) proceed to form heterogeneous mixed disulfides with glutathione (see the 24-h samples of Figure 3, peaks marked by stars). At least eight fractions of the mix disulfide species were observed. Analysis by mass spectrometry has shown that these end-products all have a molecular weight (3899) equivalent to the combination of Hir^{1-35} (3593) and a reduced glutathione (306). It is also interesting to point out that, in the system containing GSH, the formation of the GSH-P mixed disulfide species is significantly retarded (compare the top and bottom panels of Figure 3).

Stop/Go Folding Experiments with Acid-Trapped 1-Disulfide Species of Hir^{1-35} and Identification of Productive Folding Intermediates. The isolated 1-disulfide intermediates (fractions I-2, I-3, I-4, and I-5) were reconstituted in the alkaline buffer ($NaHCO_3$, 50 mM) to resume the folding. Their ability to generate the 2-disulfide products was taken to measure their productivity (Chang, 1993) (Figure 4). All 1-disulfide intermediates equilibrated swiftly following reconstitution in the folding buffer but only fraction I-4 contained highly productive species that were ready to burst forth to form the 2-disulfide Hir^{1-35}. Within 15 s of resumed folding, about 17% of I-4 converted to the 2-disulfide products. Other 1-disulfide fractions achieved less than 1−2% conversion under the same conditions. The small amount of productivity observed in fraction I-3 was most likely due to the overlapping of I-3 and I-4. To further identify the structure of the productive 1-disulfide intermediate, I-3, I-4, and I-5 were carboxymethylated and digested by thermolysin. Peptides were isolated by HPLC (Figure 5) and characterized by sequence analysis. Three disulfide-containing peptides which were present in I-4 but absent (or present in relatively lower concentration) in I-3 and I-5 were identified. They were shown to contain Cys^{22}−Cys^{28}, Cys^{14}−Cys^{22}, and Cys^{6}−Cys^{16}, respectively (Table 1).

Analysis of the 2-Disulfide Products of Hir^{1-35}. Reduced/denatured Hir^{1-35} was allowed to refold in the $NaHCO_3$ buffer for 3 h and trapped by acid (see Figure 2, top panel). Four fractions of the 2-disulfide products (P-1, P-2, P-3, and

Table 1: Sequences of Thermolytic Peptides Derived from the 1-Disulfide Intermediates of Hir^{1-35} (See Figure 5)

peak	sequence (position)	Cys−Cys	recovery[a] (pmol) from I-3	I-4	I-5
1	V-C-G-Q-G-N-K-C (21−28)	Cys22−Cys28	142	1160	132
2	V-C-G-Q-G-N-K-C[b] (21−28)	Cys14−Cys22	87	153	36
	L-C (14−22)				
3[c]	V-V-Y-T-D-C-T-E-S-G-Q-N (1−12)	Cys6−Cys16	0	25	0
	L-C-E-G-S-N (15−20)				
	I-L-G-S-D-G-E (29−35)		450	1250	522
	V-V-Y-T-D-C[b]-T-E-S-G-Q-N (1−12)		111	445	103

[a] Calculated from the recoveries of PTH-Val, PTH-Leu, or PTH-Ile and calibrated by percentages of the total samples applied for HPLC separation and sequencing. [b] Identified as (carboxymethyl)cysteine. [c] Analysis is complicated by the presence of excess amounts of two non-cystine-containing peptides, residues 1−12 and 29−35. The conclusion that peak 3 of I-4 contains an additional cystine peptide is solely based upon the detection of a minor sequence, L-?-E-G-S-N. This minor sequence has not been detected in either I-3 or I-5.

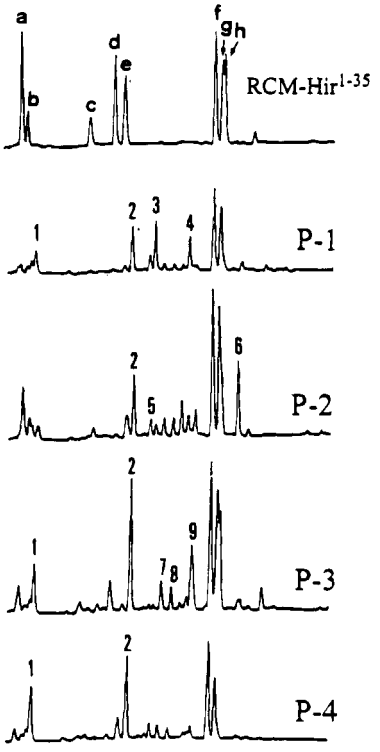

FIGURE 6: Peptide mapping of thermolysin-digested P-1, P-2, P-3, and P-4. Acid-trapped 2-disulfide product (~30 μg) was carboxymethylated and treated with 2 μg of thermolysin (Sigma P-1512) in 100 μL of *N*-ethylmorpholine/acetate buffer (50 mM, pH 6.5). Digestion was carried out at 23 °C for 14 h. Peptides were separated under the HPLC conditions described in Figure 2, except for using a different gradient consisting of 5−22% B, linear, over 30 min, 22−60% B from 30 to 31 min, 60% B for 5 min, and then, a returned to 5% B within 2 min. A control sample derived from the digestion of fully reduced, carboxymethylated Hir$^{1-35}$ is given in the top panel. Sequences of peptides denoted by a−h or numbered 1−9 are presented in Table 2.

P-4) were isolated for further analysis. When a 2-disulfide fraction was reconstituted in the same $NaHCO_3$ buffer (pH 8.3), full equilibrium among the 2-disulfide isomers occurred within 20 s (Figure 4). The disulfide shuffling was catalyzed by the extra free cysteine, and no additional thiol reagent, such as β-mercaptoethanol, was required. 2-Disulfide Hir$^{1-35}$ isomers do not equilibrate with any detectable amount of 1-disulfide species. Upon prolonged incubation in the folding buffer, all 2-disulfide species proceed to form disulfide-linked dimers (see Figure 2, 24-h sample). For structural characterization, they were carboxymethylated and treated with thermolysin. Disulfide-containing peptides were purified and characterized. The results, presented in Figure 6 and Table 2, demonstrate that many of them comprise more than one isomer. Only P-4 contains a relatively pure species, Hir$^{1-35}$[Cys^6−Cys^{14}, Cys^{22}−Cys^{28}]. However, among the ten possible disulfide bonds (Figure 1), only five were found to exist within these 2-disulfide products. One of the native disulfides, Cys^{16}−Cys^{28}, has not been found in any of those analyzed fractions. However, we cannot rule out the possibility that Cys^{16}−Cys^{28} may exist within the two minor 2-disulfide fractions that overlap with 1-disulfide intermediates.

Characterization of Acid-Trapped Folding Intermediates of Hir$^{1-43}$. The folding intermediates of Hir$^{1-43}$ were trapped by acidification and separated by HPLC. The patterns of folding intermediates are presented in Figure 7. Unfolded Hir$^{1-43}$ proceeded to form the active structure quantitatively only in the presence of supplementing free thiols. This phenomenon is indistinguishable from that observed in the case of Hir$^{1-49}$ (Chatrenet & Chang, 1993). The pathway can be dissected into two stages. An initial stage of nonspecific packing of the polypeptide chain is characterized by the irreversible and sequential flow of R to equilibrated 1-disulfide isomers to equilibrated 2-disulfide isomers to equilibrated 3-disulfide (scrambled) isomers. The final stage of folding involves consolidation and refinement of the partially packed intermediates, in this case scrambled 3-disulfide Hir$^{1-43}$, to form the native structure (N). These two

Table 2: Sequences of Thermolytic Peptides Derived from Reduced Carboxymethylated (RCM-) Hir$^{1-35}$ and the 2-Disulfide Folding Products of Hir$^{1-35}$ (See Figure 6)

species	peak	sequence (position)	Cys−Cys (Cys)
RCM-Hir$^{1-35}$	a	T-D-C^a-T-E-S-Q-G-N (4−12)	(Cys6)
RCM-Hir$^{1-35}$	b	V-C^a-G-Q (21−24)	(Cys22)
RCM-Hir$^{1-35}$	c	L-C^a (13−14)	(Cys14)
RCM-Hir$^{1-35}$	d	V-C^a-G-Q-G-N-K-C^a (21−28)	(Cys22 + Cys28)
RCM-Hir$^{1-35}$	e	L-C^a-E-G (15−18)	(Cys16)
RCM-Hir$^{1-35}$	f	V-V-Y (1−3)	none
RCM-Hir$^{1-35}$	g	I-L-G-S-D-G-E (29−35)	none
RCM-Hir$^{1-35}$	h	V-V-Y-T-D-C^a-T-E-S-Q-G-N (1−12)	(Cys6)
P-1	3	T-D-C-T-E L-C-E-G	Cys6−Cys16
P-1	4	L-C V-C-G-Q-G-N-K-C^a	Cys14−Cys22
P-2	5	L-C V-C-G-Q	Cys14−Cys22
P-2	6	L-C-E-G V-C-G-Q-G-N-K-C^a	Cys16−Cys22
P-3	1	V-C-G-Q-G-N-K-C	Cys22−Cys28
P-3	2	T-D-C-T-E-S-G-Q-N L-C	Cys6−Cys14
P-3	7	L-C-E-G V-C	Cys16−Cys22
P-3	8	L-C-E-G-S-N V-C-G-Q-G-N-K-C^a	Cys16−Cys22
P-3	9	L-C-E-G-G-S-N-V-C	Cys16−Cys22

[a] Identified as (carboxymethyl)cysteine.

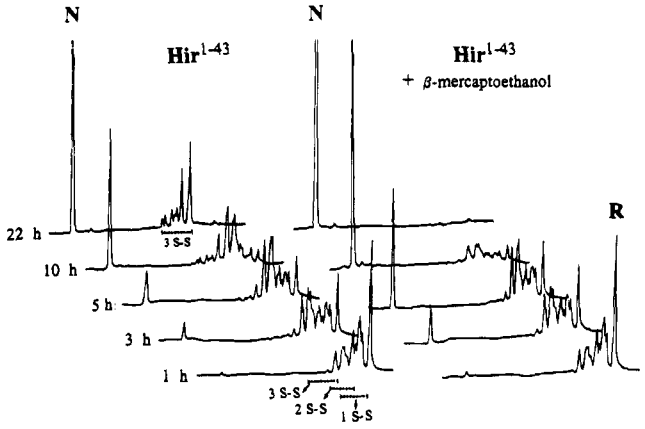

FIGURE 7: HPLC analysis of acid-trapped folding intermediates of Hir^{1-43}. The folding was performed in both the absence (left) and the presence (right) of β-mercaptoethanol (100 μM). About 15–20 μg of each time-course trapped sample was injected. N and R stand for native and fully unfolded Hir^{1-43}, respectively. Three groups of intermediates advance along the folding pathway in a sequential and irreversible manner. They are denoted as 1 S-S (equilibrated 1-disulfide isomers), 2 S-S (equilibrated 2-disulfide isomers), and 3 S-S (equilibrated 3-disulfide (scrambled) isomers). In the absence of β-mercaptoethanol, about 40% of Hir^{1-43} was stuck at scrambled 3-disulfide species. The column was a Vydac C-18 for peptides and proteins. Solvent A was water containing 0.1% trifluoroacetic acid. Solvent B was acetonitrile/water (9:1, v/v) containing 0.1% trifluoroacetic acid. The gradient was 16–30% solvent B over 50 min. The flow rate was 1 mL/min. N and R elute at 18.7 and 47.2 min, respectively.

stages of folding can be selectively regulated by a number of external factors (Chang, 1994). The efficiency of packing is mainly enhanced by oxidized glutathione, whereas the speed of reorganization is catalyzed and assisted by reagents containing free thiols. In the absence of supplementing thiols (β-mercaptoethanol or reduced glutathione), free cysteines of R or 1- or 2-disulfide intermediates act as effective thiol catalysts in promoting the reorganization during the early stage of folding. As the folding progresses, more free cysteines are engaged in the disulfides, and fewer are available as thiol catalysts; therefore, scrambled 3-disulfide intermediates become trapped, accumulate, and stay (Figure 7, left column). This problem is overcome by including β-mercaptoethanol or reduced glutathione in the folding sample (Figure 7, right column). The refolded Hir^{1-43} (N) cannot be distinguished from its native structure in terms of anticoagulant activity and HPLC behavior.

DISCUSSION

The folding mechanism of two hirudin N-terminal fragments, Hir^{1-35} and Hir^{1-43}, has been investigated. The results demonstrate that, in the absence of sufficient sequence information, the polypeptide Hir^{1-35} folds into a collection of equilibrated isomers (conformations). The assembly of complete structural elements eventually favors the polypeptide Hir^{1-43} to adopt a defined native structure. The structure of active Hir^{1-43} is shown in Figure 8 (Haruyama & Wuethrich, 1989; Folkers et al., 1989). Hir^{1-43} differs from Hir^{1-35} by the C-terminal extension of an octapeptide. This additional structural element, which adopts β-strand (residues 36–41), contributes to stabilize the hirudin core domain in two major ways: (1) It forms an antiparallel β-sheet with residues extending from residue 26 to 31. Four hydrogen bonds are involved in this tertiary structure. (2) Cys^{39} forms

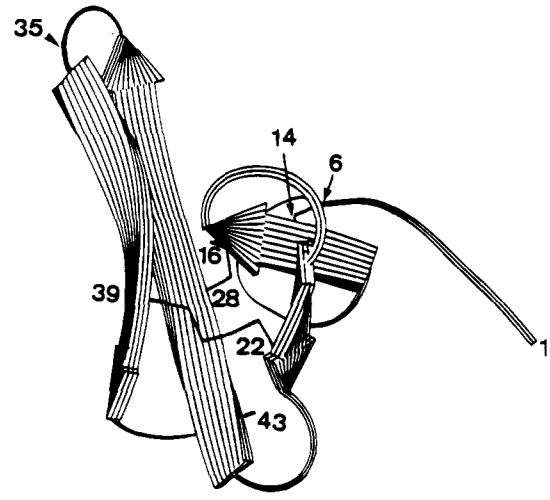

FIGURE 8: Structure of hirudin N-terminal core domain Hir^{1-43}. The core domain contains four stretches of β-strand but no α-helix. One double-stranded β-sheet consists of residues Asn^{26} to Glu^{31} and Lys^{36} to Thr^{41}. Another short one is made of residues Leu^{13} to Cys^{16} and Val^{21} to Gly^{23}.

a disulfide bridge with Cys^{22}. This disulfide helps to fix the relative spatial locations of the two β-sheets. Nonetheless, Hir^{1-35} alone does contain a substantial portion of the native structures. Aside from the two native disulfides, 18 out of the 24 identified hydrogen bonds of the core domain are formed between amino acid residues within Hir^{1-35} (Haruyama & Wuethrich, 1989). We had indeed expected that the isomer containing the two native disulfides, namely, $Hir^{1-35}[Cys^{6}-Cys^{14}, Cys^{16}-Cys^{28}]$, should be formed at least as one of the predominant 2-disulfide products. The finding that this isomer does not exist in significant quantity is intriguing.

Despite the difference in their ability to form the active structure, the folding mechanisms of Hir^{1-35} and Hir^{1-43} bear some important similarities. Indeed, the folding mechanism of Hir^{1-35} closely resembles what occurs during the early stage of Hir^{1-49} (and Hir^{1-43}) folding, in which both involve a process of nonspecific packing (disulfide formation) and are accelerated by oxidized glutathione (Chatrenet & Chang, 1993). They are characterized as R flows spontaneously and sequentially through equilibrated 1-disulfide isomers to equilibrated 2-disulfide isomers. In the case of Hir^{1-43}, the equilibrated 2-disulfide isomer continues the journey of folding *via* the equilibrated 3-disulfide (scrambled) isomer to settle at the structure of the native disulfide (N). The intermediates presented in these folding processes are enormously heterogeneous.

Two aspects of the data derived from the structural analysis of Hir^{1-35} need to be further elaborated. One concerns the identification of productive intermediates (Chang, 1993), and the other relates to the implication of the native-like structure found in the folding intermediates (Kim & Baldwin, 1990). By definition, productive intermediates represent the folding pathway. They are responsible for the direct conversion of progressive folding structures and account for the flow of folding intermediates. We have maintained (Chang, 1993) that those species cannot be simply deduced from the kinetic analysis of well-populated intermediates when isomers of intermediates exist in a state of dynamic equilibrium. For instance, in the proposed BPTI model (Weissman & Kim, 1991), apart from their presence in the pathway, there is no

Table 3: Sizes of the Hirudin N-Terminal Sequences and Their Folding Products

sequence (% of the core)	containing		folding products, found/possible isomers
	Cys	native Cys−Cys	
Hir^{1-27} (55%)	6, 14, 16, 22	6−14	2/3[a]
Hir^{1-35} (72%)	6, 14, 16, 22, 28	6−14, 16−28	9/15
Hir^{1-43} (88%)	6, 14, 16, 22, 28, 39	6−14, 16−28, 22−39	1/15
Hir^{1-49} (100%)	6, 14, 16, 22, 28, 39	6−14, 16−28, 22−39	1/15[b]

[a] Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 4043−4049. [b] Chatrenet, B., & Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 20988−20996.

direct evidence that well-populated 1-disulfide intermediates [Cys^{30}−Cys^{51}] and [Cys^{5}−Cys^{55}] are productive species and convert directly to [Cys^{30}−Cys^{51}; Cys^{14}−Cys^{38}] and [Cys^{5}−Cys^{55}; Cys^{14}−Cys^{38}] without undergoing additional rearrangement. We suggest that one way to identify those productive species is through the kinetic analysis of purified, structurally defined intermediates using stop/go folding experiments. By this approach, we have located the fraction of 1-disulfide intermediate which contains the productive species. Three potential candidates, Hir^{1-35}[Cys^{22}−Cys^{28}], Hir^{1-35}[Cys^{6}−Cys^{16}], and Hir^{1-35}[Cys^{14}−Cys^{22}], were identified, but none of them are native disulfide. Another unexpected outcome is the absence of the native disulfide Cys^{16}−Cys^{28} in the folded 2-disulfide products. Among the ten possible disulfide bridges, only five were detected to exist in the 2-disulfide products, and all have their half-cystines separated by fewer than 8−10 residues (Figure 1). The other disulfide bridges, including Cys^{16}−Cys^{28}, which are disengaged by either more than 12 or fewer than 2 residues, were not found in any significant amount. These findings also raise a crucial question as to whether the formation of Cys^{6}−Cys^{14} along the folding pathway of Hir^{1-35} and Hir^{1-27} (Chang, 1993) is dictated by the preference of local native-like structure or is merely a consequence of the favored size of its disulfide loop. Our doubt is whether one can take it for granted that the formation of native disulfide implies the existence of native-like local structure, and *vice versa*. It is surprising indeed to find out that a peptide fragment (Hir^{1-35}) containing 80% of the native sequence could not even adopt a favored structure, let alone the native (disulfide) structure.

Together, the folding mechanisms observed with Hir^{1-27} (Chang, 1993), Hir^{1-35}, Hir^{1-43}, and Hir^{1-49} (Chatrenet & Chang, 1993) may provide a useful *in vitro* model for the renaturation of partially synthesized polypeptide chains (Table 3). There has been intensive interest in elucidating the *in vivo* folding mechanisms of nascent proteins (Freedman, 1989; Ellis, 1991; Gething & Sambrook, 1992). Finally, the folding mechanism of nascent hirudins appears to be compatible with those observed with chymotrypsin inhibitor 2 (CI2), a single-module protein with 64 amino acids and no disulfide (De Prat Gay et al., 1995). The folding behavior of a family of polypeptides consisting of increasing lengths of the N-terminal sequence of CI2 has been investigated. The growing polypeptide chains do not form a stable native-like structure until more than 83% of the N-terminal sequence (residues 1−53) has been assembled.

ACKNOWLEDGMENT

The author is deeply grateful to Mr. Hans Rink for supplying the synthetic hirudin fragment (residues 1−35) and Mr. Ueli Ramseier for the analysis of peptide sequences.

REFERENCES

Anfinsen, C. B., Haber, E., Sela, M., & White, F. H., Jr. (1961) *Proc. Natl. Acad. Sci. U.S.A. 47*, 1309−1515.
Boernsen, K. O., Schaer, M., & Widmer, M. (1990) *Chimia 44*, 412−416.
Chang, J.-Y. (1990) *J. Biol. Chem. 265*, 22159−22166.
Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 4043−4049.
Chang, J.-Y. (1994) *Biochem. J. 300*, 643−640.
Chang, J.-Y., & Knecht, R. (1991) *Anal. Biochem. 197*, 52−58.
Chang, J.-Y., Grossenbacher, H., Meyhack, B., & Märki, W. (1993) *FEBS Lett. 336*, 53−56.
Chang, J.-Y., Cannals, F., Schindler, P., Querol, E., & Aviles, F. X. (1994) *J. Biol. Chem. 269*, 22087−22094.
Chang, J.-Y., Schindler, P., Ramseier, U., & Lai, P.-S. (1995) *J. Biol. Chem. 270*, 9207−9216.
Chatrenet, B., & Chang, J.-Y. (1992) *J. Biol. Chem. 267*, 3038−3043.
Chatrenet, B., & Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 20988−20996.
Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231−297.
Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.
De Prat Gay, G., Ruiz-Sanz, J., Neira, J. L., Itzhhaki, L. S., & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3683−3686.
Dodt, J., Seemuller, U., Maschler, R., & Fritz, H. (1985) *Hoppe-Seyler's Z. Physiol. Chem. 366*, 379−385.
Ellis, R. J. (1991) *Annu. Rev. Biochem. 60*, 321−347.
Folkers, P. J. M., Clore, G. M., Driscoll, P. C., Dodt, J., Koehler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601−2617.
Freedman, R. (1989) *Cell 57*, 1069−1072.
Gething, M.-J., & Sambrook, J. (1992) *Nature 355*, 33−45.
Hantgan, R. R., Hammer, G. G., & Scheraga, H. A. (1974) *Biochemistry 13*, 3421−3431.
Haruyama, H., & Wuethrich, K. (1989) *Biochemistry 28*, 4301−4312.
Hunkapiller, M. W., Hewick, R. M., Dreyer, W. J., & Hood, L. E. (1983) *Methods Enzymol. 91*, 399−412.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Light, A., Taniuchi, H., & Chen, R. F. (1974) *J. Biol. Chem. 249*, 2285−2293.
Lyles, M. M., & Gilbert, H. F. (1991) *Biochemistry 30*, 613−619.
Markwardt, F., & Walsmann, P. (1958) *Hoppe-Seyler's Z. Physiol. Chem. 312*, 85−98.
Orsini, G., Skrzynia, C., & Goldberg, M. E. (1975) *Eur. J. Biochem. 59*, 433−440.
Pigiet, V. P., & Schuster, B. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7643−7647.
Scheraga, H. A., Konishi, Y., & Ooi, T. (1984) *Adv. Biophys. 18*, 21−41.
Sexena, P., & Wetlaufer, D. B. (1970) *Biochemistry 9*, 5015−5022.
Weissman, J. S., & Kim, P. S. (1991) *Science 253*, 1386−1393.
Weissman, J. S., & Kim, P. S. (1992) *Cell 71*, 841−851.
Wetlaufer, D. B., Branca, P. A., & Chen, G.-X. (1987) *Protein Eng. 1*, 141−146.

BI951050X