# Kinetic Studies of the Regeneration of Recombinant Hirudin Variant 1 with Oxidized and Reduced Dithiothreitol†

Theodore W. Thannhauser, David M. Rothwarf, and Harold A. Scheraga*

*Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301*

*Received September 17, 1996; Revised Manuscript Received January 7, 1997*⊗

ABSTRACT: The regeneration of native recombinant hirudin variant 1 (rHV1) from the reduced unfolded form to the fully oxidized native state has been carried out with mixtures of oxidized and reduced dithiothreitol at pH 8.3 and 12 °C. The regeneration reaction was quenched at various times by the addition of 2-aminoethyl methanethiosulfonate to block unreacted sulfhydryl groups. The quenched protein-folding intermediates were fractionated by both capillary electrophoresis and a combination of anion exchange and reverse phase HPLC and characterized by mass spectrometry, amino acid analysis, and disulfide analysis. These intermediates (before quenching) were found to interconvert rapidly so as to achieve a steady-state distribution early in the regeneration process. The experimental data were fitted to a steady-state kinetic scheme. The analysis reveals that the rate-determining step in the regeneration of rHV1 with oxidized and reduced dithiothreitol involves the oxidation of one or more two-disulfide-containing species, most likely those already containing two native disulfide bonds. This regeneration mechanism is different from one that has been proposed by Chatrenet and Chang [(1993) *J. Biol. Chem. 268*, 20988]. The differences are discussed, and possible explanations for the differences are presented.

For the past 20 years, considerable effort has been devoted to trying to understand the process by which a reduced and denatured protein folds to the proper, biologically active native form (Anfinsen, 1973). Much of the experimental work has focused on identifying a detailed protein folding pathway. This work has been motivated by the assertion that such knowledge will allow for the generation of a set of empirical rules by which accurate predictions can be made concerning protein folding and structure.

Traditionally, two types of protein folding studies have been carried out: disulfide-intact folding and disulfide-coupled folding. Disulfide-intact folding involves the process by which a protein refolds to the native structure from a denatured state in which all elements of primary structure (i.e., all native covalent interactions) are present. Disulfide-coupled folding is concerned with the process by which the native structure is regenerated from a denatured state in which one element of the primary structure has been eliminated through the reduction of all disulfide bonds. Although the disulfide-coupled process is more complex, involving many chemically distinct species, it is also much easier to control experimentally through the use of an appropriate thiol/disulfide redox couple and thiol quenching agent.

Under anaerobic conditions, the distribution of intermediate disulfide-containing species and the rate of formation of native protein can be controlled by the concentration and type of thiol and disulfide redox couple present during the regeneration (Saxena & Wetlaufer, 1970; Hantgan et al., 1974; Konishi et al., 1981, 1982; Creighton and Goldenberg, 1984; Wetlaufer et al., 1987; Goldenberg, 1988; Rothwarf and Scheraga, 1993a−d). Furthermore, with rapid quenching of the thiol/disulfide interchange reaction (by blocking unreacted sulfhydryl groups), the regeneration process can be halted and the distribution of intermediate species trapped. The intermediate species present during the refolding reaction can then be isolated and characterized.

Despite these advantages, the problems associated with the determination of a disulfide-coupled protein folding pathway are complex, and few pathways have been characterized completely (Creighton, 1985; Pace & Creighton, 1986; Weissman & Kim, 1991, 1992, 1995). Even in these cases, the interpretation of the results is not always universally accepted (Creighton, 1992; Weissman & Kim, 1992). Despite the accumulation of large amounts of data, a general solution to the protein folding problem has yet to be found (Kim & Baldwin, 1990; Baldwin, 1995). Additional experiments carried out on alternative model protein systems would be helpful in realizing this goal.

Otto and Seckler (1991) have suggested that hirudin could be such an alternative model system with which to study the protein folding problem. Hirudin is the name given to any member of the family of thrombin-specific protease inhibitors (containing about 65 residues) isolated from the medicinal leech, *Hirudo medicinalis*. More then 20 different homologous (ca. 85%) forms of hirudin have been isolated and characterized (Stone & Maraganore, 1993). All variants of hirudin contain three highly conserved disulfide bonds: $Cys^{6}-Cys^{14}$, $Cys^{16}-Cys^{28}$, and $Cys^{22}-Cys^{39}$. Although hirudin has been traditionally isolated from the leech, recombinant hirudin is now available. In the natural hirudin, $Tyr^{63}$ is present as sulfotyrosine. This post-translational modification is not found in the recombinant material.

In their preliminary study, Otto and Seckler (1991) employed recombinant hirudin variant 3 (rHV3)[1] and successfully demonstrated that the native inhibitor could be fully

† This work was supported by grants from the National Institutes of Health (HL-30616 and GM-24893). Support was also received from the National Foundation for Cancer Research.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

S0006-2960(96)02340-9 CCC: $14.00 

regenerated from the reduced protein, when dissolved in a solution (pH 6−11) containing various concentrations of oxidized and reduced glutathione (GSSG and GSH). The folding kinetics, as observed spectroscopically or by the restoration of inhibitory activity, were complex, and the presence of well-populated intermediates with anti-thrombin activity was suggested.

Chatrenet and Chang (1992) have studied the folding mechanism of hirudin variant 1 (rHV1). These experiments were carried out at pH 8.3, initially in the absence of any exogenous redox couple, such as GSH or GSSG. Oxidation was achieved by means of dissolved atmospheric $O_2$. In order to achieve full regeneration of the native material, it was necessary for them to add a reductant, such as $\beta$-mercaptoethanol or GSH, to the fully oxidized mixture of non-native three-disulfide intermediates and native hirudin. On the basis of these experiments, they proposed that the folding of reduced hirudin proceeds by "trial and error", but they were unable to identify any preferred pathway.

In similar studies carried out subsequently on a *fragment* of rHV1, residues 1−49 (rHV1$^{1-49}$; Chang, 1990), Chatrenet and Chang (1993) proposed that the pathway by which reduced rHV1 folds is sequential and consists of two distinct phases. In the first phase, the reduced protein (R) is oxidized irreversibly to form the set of intermediates that contain one intramolecular disulfide bond (1S), which then is oxidized irreversibly to form the set of intermediates that contain two intramolecular disulfide bonds (2S), which finally is oxidized irreversibly to the set of intermediates containing three intramolecular disulfide bonds (3S) as indicated in eq 1.

$$R \rightarrow 1S \rightarrow 2S \rightarrow 3S \quad (1)$$

Chatrenet and Chang (1993) suggested that this entire process, from the reduced protein to the set of 3S intermediates, is driven by hydrophobic collapse. In the second phase, the native protein (N) is formed directly from the 3S intermediates, as indicated in eq 2 by a process that is said to be "guided" by specific tertiary interactions.

$$3S \rightarrow N \quad (2)$$

Here, we present a study of the kinetics of the folding of rHV1 carried out under anaerobic conditions in the presence of oxidized and reduced dithiothreitol (DTT$^{ox}$ and DTT$^{red}$, respectively). The data have been interpreted according to the general kinetic scheme of protein folding proposed by Konishi et al. (1982), which will hereafter be referred to as the KOS approach. In this general approach, all possible pathways through which a protein may fold are considered. A pathway is defined by the type of intermediate and type of reaction involved in the rate-determining step. The purpose of this work is to determine both the type of intermediate(s) and the nature of the reaction(s) involved in the rate-determining step in the folding of reduced rHV1 to the native oxidized state.

The experimental methodology used here has been adapted from that which has been used recently to elucidate the folding pathway of RNase A (Rothwarf and Scheraga, 1993a−d). Briefly, reduced rHV1 was allowed to fold at 12 °C and pH 8.3 in the presence of various concentrations of DTT$^{ox}$ and DTT$^{red}$. The choice of temperatures was not arbitrary. Preliminary experiments had demonstrated that measurable concentrations of chemically modified forms of rHV1 accumulated during regeneration experiments carried out at temperatures of ≥20 °C. The nature of these modifications will be discussed in detail in the Discussion. The presence of these chemically modified species would have complicated both the identification of the disulfide-bond-containing regeneration intermediates and the overall kinetic analysis. Therefore, it was decided that all of these experiments would be carried out at 12 °C. At this temperature, no measurable quantities of these chemically modified species were detectable at regeneration times up to 500 min.

At various times after the initiation of folding, aliquots were removed and the folding was halted by the addition of 2-aminoethyl methanethiosulfonate (AEMTS) to block the unreacted sulfhydryl groups. This reagent reacts rapidly and specifically with thiols, thereby arresting the regeneration process (Bruice & Kenyon, 1982; Rothwarf & Scheraga, 1991, 1993a,b). It is important to note that AEMTS introduces a positive charge at each site at which it reacts with a protein thiol. This provides the means by which the blocked intermediates can be predictably fractionated. The speed and specificity with which this reagent reacts with thiols are significant advantages when this reagent is compared to more traditional thiol-blocking groups such as iodoacetamide and iodoacetate (Rothwarf & Scheraga, 1993a). Furthermore, the reaction is reversible which would allow the regeneration process to be restarted from isolated intermediates.

The folding intermediates that were trapped by this procedure were desalted rapidly and analyzed by capillary electrophoresis. This technique was able to resolve the mixture of folding intermediates into five electrophoretically distinct bands (R, 1S, 2S, 3S, and N). With determination of the relative concentrations of these electrophoretically distinct intermediate species as a function of time and concentration of DTT$^{ox}$ and DTT$^{red}$, and employing the kinetic approach of Konishi et al. (1982), it is possible to propose a folding pathway that quantitatively accounts for the experimentally observed kinetics.

## MATERIALS AND METHODS

The rHV1 was a gift from Ciba Geigy Pharmaceuticals (Horsham, West Sussex). It was found to be >98% pure as judged by RP-HPLC and was used without further purification to produce reduced rHV1 which was purified as described in the next subsection. DTT$^{red}$ was obtained from Sigma Chemical Co., and was found to be >99% pure by HPLC, and was used without purification. DTT$^{ox}$ was obtained from Sigma and was purified by HPLC as described

[1] Abbreviations: rHV1, recombinant hirudin variant 1; rHV3, recombinant hirudin variant 3; rHV1$^{1-49}$, recombinant hirudin variant 1, fragment 1−49; DTT$^{red}$, DL-dithiothreitol; DTT$^{ox}$, *trans*-4,5-dihydroxy-1,2-dithiane (oxidized DL-dithiothreitol); GSH, reduced glutathione; GSSG, oxidized glutathione; RNase A, bovine pancreatic ribonuclease A; BPTI, bovine pancreatic trypsin inhibitor; AEMTS, 2-aminoethyl methanethiosulfonate; NTSB, 2-nitro-5-thiosulfobenzoic acid; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; TFA, trifluoroacetic acid; TEAA, triethylammonium acetate; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); Bis-Tris, [bis(2-hydroxyethyl)imino]tris(hydroxymethyl)methane; PITC, phenyl isothiocyanate; CE, capillary electrophoresis; TMAPS, trimethylammonium propanesulfonate; ESI, electrospray ionization; MALDI, matrix-assisted laser desorption ionization; HPLC, high-performance liquid chromatography; RP-HPLC, reverse phase HPLC.

below. The AEMTS was synthesized by the method of Bruice and Kenyon (1982). The NTSB was synthesized as described by Thannhauser et al. (1984).

The reduction buffer was 2 M guanidine thiocyanate, 100 mM Tris-HCl, and 3 mM EDTA at pH 8.0. The regeneration buffer was 100 mM Tris-HCl at pH 8.3 and 3 mM EDTA. The HPLC eluants A and B were 0.1% TFA in water and 0.1% TFA in acetonitrile, respectively. The TEAA buffer was 100 mM acetic acid adjusted to pH 5.0 with triethylamine. All other reagents were of the highest grade commercially available.

*Reduction of rHV1*. rHV1 (1−5 mg) was reduced with a large excess of $DTT^{red}$ in 1−2 mL of reduction buffer at 25 °C for 2 h. The pH of the reaction mixture was subsequently reduced to 5 by the addition of acetic acid. The reduced protein was then purified on a Waters 600E HPLC instrument equipped with a model 486 variable wavelength detector using a 10 × 250 mm YMC ODS-AQ column, equilibrated with 90% eluant A and 10% eluant B. The reduced protein was then eluted, employing a linear gradient of 10 to 50% eluant B in 60 min at 4.0 mL/min. Fractions corresponding to the reduced protein were pooled. The concentration of this solution was then determined by UV absorbance using the extinction coefficient of reduced rHV1 ($\epsilon^{275} = 3050\ cm^{-1}\ M^{-1}$), and measured aliquots containing the appropriate amount of reduced rHV1 were pipetted directly into the regeneration vessels and lyophilized. The regeneration vessels were sealed under Ar with a self-sealing Teflon-coated septum and stored desiccated at −20 °C.

The reduced rHV1 produced by this procedure was found to have a mass of 6971 ± 5 Da (expected, 6970 Da) when analyzed by MALDI MS and a single symmetric peak when analyzed by CE. Furthermore, when subjected to thiol analysis with DTNB (Ellman, 1959), it was found to contain 6.02 ± 0.18 thiols per mole of protein in accordance with expectation for fully reduced protein.

*Purification of $DTT^{ox}$*. Oxidized DTT was dissolved in water (ca. 15 mg/mL) and purified by RP-HPLC using a 3.5 × 50 cm (10 μm) ODS AQ column (YMC), employing a flow rate of 20 mL/min and a linear gradient from the starting condition of 5% $CH_3CN$ in $H_2O$ to 30% $CH_3CN$ in $H_2O$ in 60 min. The fraction corresponding to $DTT^{ox}$ was lyophilized. After purification, the $DTT^{ox}$ was reanalyzed by RP-HPLC and found to contain no detectable impurities.

*Regeneration of rHV1 at 12 °C*. The regeneration process was started by adding (to the lyophilized reduced protein) a measured volume of deoxygenated regeneration buffer that had been equilibrated at 12 °C containing the desired concentration of $DTT^{ox}$. The buffer was introduced into the reaction vessel with a syringe through the self-sealing septum. Dissolution of the lyophilized protein was rapid because of the high velocity of the buffer stream generated by the syringe, aided by magnetic stirring. For these experiments, the concentration of reduced rHV1 ranged from 27 to 172 μM. The concentration of $DTT^{ox}$ ranged from 50 to 150 mM. The *starting* concentration of $DTT^{red}$ was always zero, but it increased rapidly early in the regeneration run to a steady-state concentration. The regeneration reaction mixture was purged continuously with a flow of humidified Ar. The temperature was regulated to ±0.1 °C by using an insulated jacketed outer vessel connected to a Fisher Scientific model 9101 refrigerated circulating bath.

*Thiol Quenching*. Aliquots (50−200 μL) of the regeneration mixture were removed with a syringe through the self-sealing septum at various times during the regeneration process. The free thiols were blocked by the addition of 200 μL of 21 mM AEMTS (a 75−100-fold excess). After 2 min, the pH of the sample was adjusted to 5 with 100 mM acetic acid, and the sample was desalted as described below.

*Desalting Blocked Samples*. The samples generated by the thiol quenching reaction were desalted as follows. A YMC ODS column (4.6 × 50 mm, 3 μm particle size) was equilibrated with 90% TEAA buffer containing 10% acetonitrile at 2.0 mL/min. The sample, together with three rinses of the sample vial with the TEAA buffer, was injected. Under these conditions, only the protein is retained on the column. Once the salt, excess reagents, and reaction by products had eluted from the column (<1 min), and the baseline was returned to its initial position, the column was washed with 10% acetonitrile in water and then equilibrated with 90% eluant A and 10% eluant B. The protein was then eluted with a linear gradient of 10 to 50% eluant B in 5 min. All the protein eluted within a single 2.0 mL fraction. Recovery, as judged by UV absorbance, was >95%. This material was lyophilized and characterized as described below.

*Identification of Intermediates*. A Mono Q HR 5/5 column (Pharmacia) was equilibrated with 50 mM Bis-Tris at pH 6.5 at a flow rate of 1.0 mL/min. A lyophilized preparation containing 1 mg of blocked and desalted material (see above) was dissolved in this buffer and injected. The intermediates were eluted from the column using a linear NaCl gradient from 0 to 200 mM in 30 min. Each fraction that contained protein was rechromatographed by RP-HPLC on a YMC ODS-AQ column (4 × 50 mm, 3 μm) equilibrated with 90% eluant A and 10% eluant B. The protein was eluted using a linear gradient of 10 to 40% B in 30 min. The fractions containing protein were then lyophilized and dissolved in an appropriate volume of 5% acetic acid. This material was then characterized by MALDI and ESI MS, amino acid analysis, disulfide analysis (Thannhauser et al., 1984, 1987), and CE as described below. These procedures permitted the determination of the composition of each intermediate as to its number of intramolecular disulfide bonds and cysteamine blocking groups. The solutions of identified intermediate groupings were used as standards to identify, characterize, and quantitate the five electrophoretically distinct bands revealed by the CE analysis described below.

*MS Analysis of the Intermediates*. Both ESI and MALDI mass spectrometry were used to characterize the intermediates. The MALDI MS experiments were carried out in the positive ion mode on a Lasermat 2000 (FinniganMAT) mass spectrometer using a matrix of α-cyano-4-hydroxycinnamic acid. ESI MS was carried out on a modified Millipore Extrel FTMS 2000 FT/ICR mass spectrometer equipped with a 6.2 T magnet and fitted with an external ESI source (Beu et al., 1993).

*Amino Acid Analysis*. Quantitative amino acid analysis was used to determine the concentrations of stock solutions of rHV1, reduced rHV1, and each of the intermediates. Triplicate samples were hydrolyzed in constant-boiling HCl at 150 °C for 90 min. The amino acids were derivatized with PITC and analyzed by RP-HPLC (Heinrikson & Meredith, 1984; Bidlingmeyer et al., 1984) using a Waters Pico-Tag amino acid analysis system.

In order to determine the number of intramolecular disulfide bonds and cysteamine blocking groups in each intermediate, the samples were first oxidized with performic acid (Hirs, 1956). This procedure quantitatively converts the inter- and intramolecular disulfides into taurine and cysteic acid, respectively (Thannhauser et al., 1997). The samples were then subjected to amino acid analysis as described above.

*Disulfide Analysis*. As an additional check on the accuracy of the MS and amino acid analysis, the standard solutions of N, R, and each intermediate grouping (1S, 2S, and 3S) were subjected to disulfide analysis with NTSB (Thannhauser et al., 1984, 1987).

*Fractionation of Intermediates*. CE was used to determine the distribution of intermediates at each time during the folding reaction. The capillary was fused silica (72 cm, 50 $\mu$m inside diameter) (Applied Biosystems). The carrier electrolyte solution was 25 mM citric acid and 1.0 M TMAPS. Hydrodynamic injections were made (1−3 s). Electrophoresis was carried out in the positive mode, on an Applied Biosystems model 270A capillary electrophoresis system, at 30 °C, with a potential of 208 V/cm. The electropherogram was monitored at 200 nm. The choice of 200 nm, as opposed to a more typical longer wavelength, e.g., 275 nm, was necessary because of the short path length available to measure absorbance in this instrument, 50 $\mu$m. In all cases, an internal standard (BPTI) was used to adjust for variations in electrophoretic mobility.

The standard solutions of R, N, and each ensemble of intermediate species (1S, 2S, and 3S) were subjected to analysis by CE. The electrophoretic mobility of each was determined and used to confirm the identity of each of the bands detected by the CE analysis of the quenched folding reaction mixture.

*Relative UV Absorbance of the Intermediates at 200 nm*. Since the electropherograms used to obtain the kinetic data were monitored at 200 nm, it was necessary to determine the relative absorbance of each intermediate species at 200 nm. Because of the lack of precision in hydrodynamic sample loading procedures used in CE, this was accomplished by RP-HPLC. A YMC ODS-AQ column (2.0 × 150 mm) was equilibrated with 90% eluant A and 10% eluant B at 200 $\mu$L/min. Injections of known amounts of each separate species (R, 1S, 2S, 3S, and N) were made from the standard solutions described above. The protein was eluted with a linear gradient of 10 to 50% eluant B in 30 min at a flow rate of 200 $\mu$L/min. The chromatogram was recorded at 200 nm and integrated electronically. A response factor, in area units per picomole, was determined for each species. These were then averaged to obtain an average response factor for all species. The relative absorbance of each separate species was obtained as the ratio of the individual response factor of that species to the average response factor for all species. These values are reported in Table 1.

*Fitting the Kinetic Data*. The kinetic data collected by the procedures described above were interpreted by application of the general KOS approach as described in the Results. The experimental data for the concentration of each species as a function of time were fit to this scheme by application of a simplex minimization procedure which had been modified to permit simultaneous fitting of all data sets (Rothwarf & Scheraga, 1993b; Li et al., 1995). Error estimates were determined by a Monte Carlo procedure (Straume & Johnson, 1992) and are given at the 95% confidence limit.

Table 1: Characterization of the Species Present during the Regeneration Process

| species[a] | $M_w$[b] | number of disulfides[c] | number of blocking groups[d] | relative absorbance at 200 nm[e] |
|---|---|---|---|---|
| R | 7420 (±1) | 6.0 (±0.2) | 6.0 (±0.18) | 0.95 (±0.06) |
| 1S | 7268 (±1) | 5.0 (±0.2) | 4.0 (±0.12) | 1.03 (±0.07) |
| 2S | 7115 (±1) | 4.1 (±0.1) | 2.1 (±0.06) | 1.06 (±0.07) |
| 3S | 6964 (±4) | 3.0 (±0.1) | 0.0 (±0.05) | 0.99 (±0.04) |
| N | 6964 (±4) | 3.0 (±0.1) | 0.0 (±0.05) | 0.98 (±0.05) |

[a] The unreacted sulfhydryl groups of R, 1S, and 2S are blocked with cysteamine. [b] Molecular mass in daltons measured by MALDI and ESI MS as described in text. The expected $M_w$ for the species listed are as follows: R = 7420, 1S = 7268, 2S = 7116, 3S = 6964, and N = 6964. [c] Moles of intra- and intermolecular disulfide per mole of protein as determined by NTSB assay. [d] Number of blocking groups per molecule of intermediate as determined by amino acid analysis after performic acid oxidation which converts the cysteamine blocking groups into taurine (Thannhauser et al., 1997). [e] The relative absorbance at 200 nm, determined as described in the text. The errors in all measured quantities are given at the 95% confidence limit.

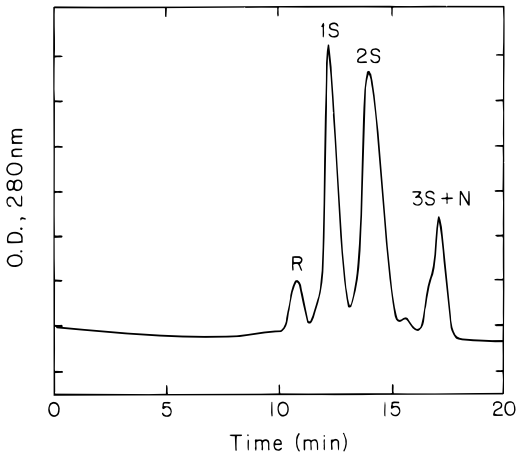


FIGURE 1: Anion exchange (Mono Q HR 5/5) separation of the intermediates present in the regeneration reaction at 120 min. The chromatographic conditions used are described in the text. The starting conditions for the regeneration reaction were pH 8.0, 12 °C, 72 $\mu$M rHV1, 100 mM $DTT^{ox}$, and 0 mM $DTT^{red}$.

## RESULTS

*Identification of the Intermediates*. The mixture of blocked intermediates was fractionated by anion exchange chromatography as described above. Figure 1 shows a representative chromatogram. The most positively charged intermediate, R, eluted first, followed by 1S, then by 2S, and finally by 3S + N, which eluted together because they have the same net charge. The ensemble of 3S intermediates and N was subsequently resolved by the RP-HPLC method which was used to desalt the pooled fractions obtained from the anion exchange separation. The identifications of the intermediates were confirmed, after desalting, by disulfide analysis, amino acid analysis, and MS. The results are summarized in Table 1.

Although MALDI MS was used to determine the mass of 3S and N, it was not useful in the characterization of the *blocked* materials (R, 1S, and 2S). The MALDI spectra of each of these intermediates showed significant amounts of

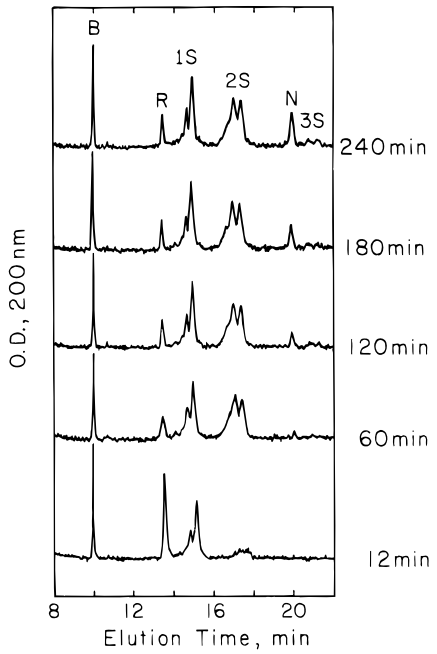


FIGURE 2: Series of electropherograms showing the appearance of intermediates with regeneration time (the right-hand ordinate). The starting regeneration conditions were pH 8.0, 12 °C, 72 $\mu$M rHV1, 100 mM $DTT^{ox}$, and 0 mM $DTT^{red}$.

material with a molecular mass lower than expected. In every case, the difference between the expected mass of the intermediate and the mass of the lower-molecular mass species was found to be a whole number multiple of the mass of the blocking group (76 Da), indicating the loss of increasing numbers of blocking groups during the MALDI ionization process. Although MALDI is generally considered a soft ionization process and usually results in little or no fragmentation of the molecular ion (Hillenkamp & Karas, 1990), the lability of disulfide bonds in MALDI MS analysis has been reported before (Beavis & Chait, 1989); in fact, this forms the basis of recently published methods of determining the disulfide bond pairing patterns in proteins (Patterson & Katta, 1994: Crimmins et al., 1995). Therefore, intermediates R, 1S, and 2S were analyzed by electrospray MS. Using this method of ionization, no evidence of disulfide bond cleavage was detected, and it is these values that are reported for R, 1S, and 2S in Table 1.

*Separation of Intermediates by CE*. Figure 2 shows a series of electropherograms demonstrating the appearance of intermediates as a function of regeneration time. The bands elute in the order expected on the basis of their charge. The identity of each was confirmed by CE analysis of the standard solutions of each intermediate that was isolated chromatographically and characterized as described above.

*Analysis of Kinetic Data*. The experimental data reported here were interpreted using the KOS approach to the analysis of protein regeneration (Konishi et al., 1982; Scheraga et al., 1987; Rothwarf & Scheraga, 1993b). According to this approach, a reduced protein and the various intramolecular disulfide-bonded intermediates interconvert rapidly, leading to a pre-equilibrium steady state, as shown in eq 3.

$$R \rightleftarrows 1S \rightleftarrows 2S \rightleftarrows 3S \qquad (3)$$

The intermediates are assumed to be separated from the native protein by a high energy barrier. The rate-determining

**A)**

| STEP | RATE EQUATION |
|---|---|
| $R \rightarrow R^*$ | $k_1[R]$ |
| $R + DTT^{ox} \rightarrow 1S^* + DTT^{red}$ | $k_2[R][DTT^{ox}]$ |
| $1S + DTT^{red} \rightarrow R^* + DTT^{ox}$ | $k_3[1S][DTT^{red}]$ |
| $1S \rightarrow 1S^*$ | $k_4[1S]$ |
| $1S + DTT^{ox} \rightarrow 2S^* + DTT^{red}$ | $k_5[1S][DTT^{ox}]$ |
| $2S + DTT^{red} \rightarrow 1S^* + DTT^{ox}$ | $k_6[2S][DTT^{red}]$ |
| $2S \rightarrow 2S^*$ | $k_7[2S]$ |
| $2S + DTT^{ox} \rightarrow 3S^* + DTT^{red}$ | $k_8[2S][DTT^{ox}]$ |
| $3S + DTT^{red} \rightarrow 2S^* + DTT^{ox}$ | $k_9[3S][DTT^{red}]$ |
| $3S \rightarrow 3S^*$ | $k_{10}[3S]$ |

**B)**

$$dN/dt = \sum_{i=1}^{3} k_i f_i (1-N)[DTT^{ox}] + \sum_{i=1}^{3} k_i f_i (1-N)[DTT^{red}] + \sum_{i=1}^{4} k_i f_i (1-N)$$

FIGURE 3: (A) The ten possible rate-determining steps in the regeneration of a three-disulfide-containing protein. The species marked with an asterisk are those that subsequently fold rapidly to form the native protein. (B) The general form of the rate equation describing the appearance of native protein. The quantity $k_i$ is the apparent rate constant for the $i$th pathway; $f_i$ is the fractional concentration of the $i$th grouping of intermediates in the regeneration reaction mixture.

step(s) in the regeneration process is then the step(s) that crosses this barrier. The rate of regeneration of the native protein depends on the concentration(s) of the intermediate-(s) that directly precedes the rate-determining step(s) and, as will be shown below, the type of reaction(s) involved.

There are three different types of reactions that can be involved in a rate-determining step: rearrangement, reduction, and oxidation. Consider the following examples. A 2S species can undergo a conformational transition or an intramolecular disulfide rearrangement to form another species $2S^*$ that then folds rapidly to the native protein. In this case, the rate of formation of the native protein would depend on the concentration of 2S in the steady state. The rate equation for this step would have no direct dependence on the concentration of the redox couple. Alternatively, a 2S species can undergo a reduction to a species $1S^*$ which would then fold rapidly to the native protein. This reaction would depend on both the concentration of 2S and the concentration of $DTT^{red}$. Finally, a 2S species can undergo an oxidation to a species $3S^*$ which then would fold rapidly to the native protein. The appearance of native protein would then depend on the concentration of 2S and $DTT^{ox}$.

The general KOS approach makes no assumptions concerning the intermediates that are involved in the rate-determining step. Rather, it considers that all species present in the steady state (including the reduced protein, R) may undergo a rate-determining transition (i.e., a rearrangement, reduction, or oxidation) that will convert it into an intermediate that will fold rapidly to the native protein.

In the case of a three-disulfide protein such as rHV1, there are 10 possible rate-determining steps. These are listed together with the corresponding rate equations in Figure 3A. The general form of the rate equation describing the appearance of the native protein is given in Figure 3B.

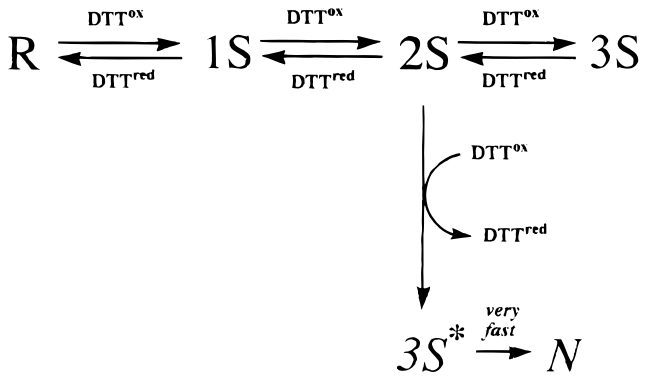


FIGURE 4: Best-fit regeneration model. The rate constants are given in Table 2. The bimolecular rate constant for the rate-determining step, 2S → *3S**, is $k_8$ which equals $(8.4 \pm 0.2) \times 10^{-3}$ $min^{-1}$ $M^{-1}$.

The data from all kinetic experiments were fit simultaneously to each of the single-pathway models given in Figure 3A. A good fit to the experimental data was found only for the reaction in which the rate-determining step was presumed to be of the type 2S → *3S** ($k_8$ in Figure 3A). The second-order rate constant for this reaction was found to be $k_8$ which equals $8.4 \times 10^{-3}$ $min^{-1}$ $M^{-1}$. This pathway alone can account for >90% of the regenerated protein under all of the redox conditions used. Furthermore, when the data were fit to a model in which all possible rate-determining steps indicated in Figure 3A were included, the 2S → *3S** pathway still accounted for over 80% of the regenerated protein, under all of the redox conditions used. The proposed regeneration model is given in Figure 4. The rate constants for the various reactions indicated in Figure 4 are given in Table 2.

The species *3S** occurs after the rate-determining step in the regeneration process. It is kinetically indistinguishable from the native protein and proceeds rapidly to N. It presumably has correct disulfide bonds but may possess a conformation that is different from that of N.

Figure 5 shows the curves for the time dependence of the relative populations of the various species calculated using the rate constants given in Table 2. The symbols represent the experimental data, and the solid lines represent the calculated curves. Figure 5 demonstrates that the intermediates rapidly achieve a steady-state condition (20−100 min). After this time, the only change in the distribution of intermediates is due to the generation of native protein and the consequent formation of $DTT^{red}$. The $DTT^{red}$ produced by the formation of native protein results in a rapid redistribution of the intermediates which converts most of it back to $DTT^{ox}$. Therefore, the increase in $DTT^{red}$ is slow after the steady state has been achieved, as shown in Figure 6.

## DISCUSSION

*Origin of Chemical Modifications.* The source of the chemical modifications of rHV1 that were observed at temperatures of >20 °C and times of >400 min can be determined by consideration of the known behavior of other homologous variants of hirudin, specifically variant 2 (rHV2). Treatment of rHV2 at pH 9 and 37 °C results in the deamidation of both Asn33 and Asn53 (Toung et al., 1992). In the amino acid sequence of rHV2, Asn33 and Asn53 are both followed by Gly. Asn-Gly sequences are known to deamidate at appreciable rates under the conditions used here (pH 8.3 and 12 °C) for regeneration of native rHV1 (Thannhauser & Scheraga, 1985; Bischoff & Kolbe, 1994). The mechanism of this base-catalyzed reaction is such that it results in two products, one with an α-linked and one with a β-linked aspartic acid (Meinwald et al., 1986). In the sequence of rHV1, the two labile asparagine residues have been replaced by aspartic acid. These Asp-Gly sequences of rHV1 cannot deamidate; however, they are still prone to rearrangement to β-linked aspartic acid through a similar mechanism, although at a slower rate (Bischoff & Kolbe, 1994). It is these species, containing β-linked aspartic acid at positions 33 and 53, that account for the accumulation of the chemically modified forms of rHV1 that was discussed above. In order to minimize the effects of these chemical modifications on the isolation and characterization of the regeneration intermediates and on the analysis of the kinetic data, 12 °C was selected as an appropriate temperature for these experiments. At this temperature, none of these chemically modified forms of rHV1 was detectable at regeneration times up to 500 min.

*Application of the KOS Approach to the Regeneration of rHV1; Validity of the Steady-State Assumption.* The analysis of the data presented above demonstrates that the rate-determining step in the regeneration of rHV1 is the oxidation of a two-disulfide intermediate to a three-disulfide species (*3S**) which then folds rapidly to and is kinetically indistinguishable from the native protein. It has already been demonstrated that this type of analysis can be used to determine the regeneration pathways of RNase A (Konishi et al., 1981, 1982; Rothwarf & Scheraga, 1993b). We now show that such an analysis is applicable to the regeneration of rHV1 as well.

Figures 5 and 6 demonstrate that the curves generated from the kinetic analysis described above accurately fit the experimentally observed concentrations of the groupings of disulfide-bonded intermediates and $DTT^{red}$. The ability of the model to fit the data over the full range of experimental conditions used suggests that the kinetic scheme shown in Figure 3A is valid, at least to the extent that a kinetic analysis can rule out alternative mechanisms.

It is important to stress that the electrophoretically distinct groupings of isomeric intermediates, 1S, 2S, and 3S, are not homogeneous, chemically distinct species. Each of these intermediates is actually a mixture of chemically distinct species that have the same number of disulfide bonds and cysteamine blocking groups. The KOS approach presumes that each grouping of isomeric intermediates can be treated as a single kinetic species. This can be done if the distribution of species within the groupings of isomeric intermediates remains constant over the time course of the regeneration process and under the range of redox conditions used (Scheraga et al., 1987; Rothwarf & Scheraga, 1993b).

Inspection of the relevant electropherograms indicates (roughly) that the distribution of intermediates within each grouping does not vary with time (see Figure 2). Furthermore, comparisons between the distributions of species found within groupings of isomeric intermediates isolated from regeneration experiments carried out under different redox conditions (data not shown) also reveal no obvious differences. However, because of our limited ability to resolve the individual components of the various intermediate groupings, this type of analysis would be unable to detect small changes in these distributions. The possibility of the

Table 2: Rate Constants[a] at 12 °C and pH 8.0

| reaction | $(k^{obs})_f$ ($\times 10^2$) ($min^{-1}$ $M^{-1}$) | $(k^{ave})_f$ ($\times 10^2$) ($min^{-1}$ $M^{-1}$) | $(k^{ave})_{intra}$ ($min^{-1}$) | $(k^{obs})_r$ ($\times 10^{-2}$) ($min^{-1}$ $M^{-1}$) | $(k^{ave})_r$ ($\times 10^{-2}$) ($min^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| R ⇄ 1S | 123 ± 2 | 8.2 | 11.6 | 2.20 ± 0.01 | 2.20 |
| 1S ⇄ 2S | 42 ± 1 | 7.1 | 10.1 | 2.79 ± 0.08 | 1.40 |
| 2S ⇄ 3S | 2.3 ± 0.2 | 2.3 | 3.3 | 2.14 ± 0.20 | 0.71 |

[a] These rate constants were determined from the data for the whole time course of the reaction under all experimental conditions. $(k^{obs})_f$ and $(k^{obs})_r$ are the observed forward and reverse rate constants for the formation or reduction of a disulfide bond using $DTT^{ox}$ and $DTT^{red}$, respectively. $(k^{ave})_f$ and $(k^{ave})_r$ are the observed forward and reverse rate constants, respectively, corrected for statistical factors, as described in the text. $(k^{ave})_{intra}$ is the average rate constant for the formation of an intramolecular disulfide bond. The rate constant for the 2S → *3S** step is $k_8$ which equals 8.4 × $10^{-3}$ $min^{-1}$ $M^{-1}$. Errors are expressed at the 95% confidence limit.

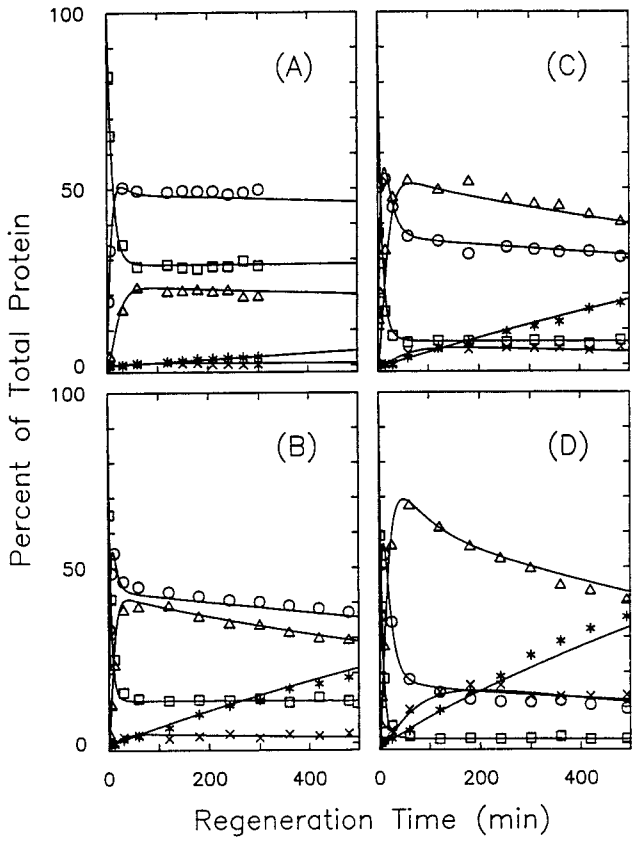


FIGURE 5: Concentration of native protein and the intermediates as a function of regeneration time at pH 8.0 and 12 °C: (□) reduced protein; (○) one-disulfides; (△) two-disulfides; (×) three-disulfides, and (*) native rHV1. The starting concentrations were as follows: (A) 172 $\mu$M rHV1 and 50 mM $DTT^{ox}$, (B) 172 $\mu$M rHV1 and 150 mM $DTT^{ox}$, (C) 66 $\mu$M rHV1 and 100 mM $DTT^{ox}$, and (D) 27.8 $\mu$M rHV1 and 150 mM $DTT^{ox}$. The symbols represent the experimental data; the curves are drawn using the rate constants given in Table 2.

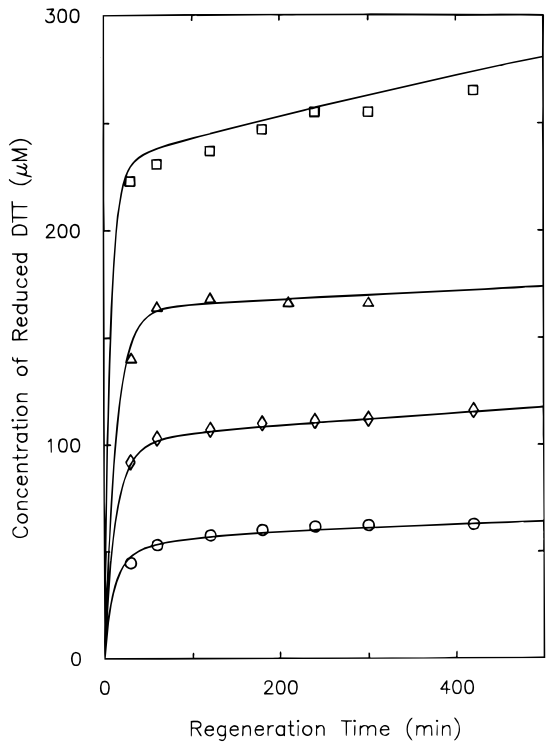


FIGURE 6: Appearance of $DTT^{red}$ as a function of regeneration time. The symbols represent the experimental data; the curves are generated by fitting the data to the regeneration model shown in Figure 4 using the rate constants given in Table 2. The symbols △, □, ◇, and ○ correspond to the conditions described in A−D, respectively, in Figure 5.

accumulation of small concentrations of slowly interconverting intermediates must be considered, but this possibility does not invalidate the kinetic model (Rothwarf & Scheraga, 1993b).

*Factors Affecting the Rate Constants*. There are several factors that are known to contribute to the observed forward and reverse rate constants of disulfide interchange reactions. These include electrostatic and other sequence-specific local interactions, entropic considerations, and steric factors. A brief discussion of these factors is relevant because they provide insight that will aid in the subsequent analysis of the rate constants given in Table 2.

For the purposes of this discussion, the electrostatic effects are considered to be local interactions, at the ionic strength of 50 mM used here. It has been determined that, at ionic strengths of >20 mM, the effects of charges other than those of neighboring amino acids are negligible (Snyder et al., 1981). Electrostatic effects have been studied in a variety of disulfide exchange reactions (Hird, 1962; Creighton, 1975; Szajewski & Whitesides, 1980; Snyder et al., 1981). The electrostatic factors include two major effects.

The first is the effect that charged residues have on the p*K* of neighboring thiols. The presence of one or two positively charged neighbors will tend to decrease the p*K* of the thiol, thereby increasing the concentration of active thiolate species, which will in turn increase the rate of thiol/disulfide interchange. Conversely, the presence of one or two negatively charged neighbors will raise the p*K* of the thiol, decrease the concentration of the active thiolate ion, and decrease the rates of thiol disulfide interchange.

The second of these electrostatic effects arises from the influence of neighboring charged residues on the rate of nucleophilic attack of the negatively charged thiolate ion on a target disulfide. The most favorable case would be that of a disulfide flanked by four positively charged neighbors; the least favorable case would involve a disulfide flanked by four negatively charged neighbors.

In combination, the effects of these two factors can be considerable. It has been estimated that the Coulombic forces due to the presence of neighboring charged residues in the local environment of a cysteine residue provide the means of achieving a $10^6$-fold range in rate constants in disulfide exchange reactions (Snyder et al., 1981). Furthermore, other sequence-specific local interactions (Morgan et al., 1978; Bodner et al., 1980; Némethy & Scheraga, 1981) can also influence the observed average rate constants. Given the magnitude of these sequence-specific interactions, it is difficult to gain insight into the regeneration mechanism through a direct comparison of the rates of disulfide bond formation and reduction obtained from the regeneration of different proteins. However, it is useful to compare the trends observed in the rates of disulfide bond oxidation and reduction as a *given* protein regenerates from a reduced to a more oxidized form.

Entropic factors also contribute to the observed values of the average rate constants (Poland & Scheraga, 1965; Creighton, 1979; Lin et al., 1984; Xu et al., 1996). As has been discussed by Rothwarf and Scheraga (1993b), some of the entropic factors would be expected to increase the values of $(k^{ave})_f$ with the formation of each successive disulfide bond, while other entropic factors would tend to decrease these values. The opposing nature of these effects complicates the task of explaining the observed changes in the average forward rate constants in terms of entropy. The tendencies observed in any particular protein will be due to a unique combination of all entropic factors.

Steric factors are also important. The accumulation of structure, particularly native structure, can restrict the accessibility of disulfide bonds, rendering them resistant to nucleophilic attack. Furthermore, thiols can be effectively buried and inhibited from further participation in a disulfide exchange reaction (Weissman & Kim, 1995).

When using the KOS approach, statistical factors must also be considered. Since the groupings of isomeric intermediates are not chemically distinct species, the observed rate constants $(k^{obs})_f$ and $(k^{obs})_r$[2] given in Table 2 do not correspond to any single process. Every step in the regeneration process involves an electrophoretically distinct mixture of species, each member of which is capable of being converted into a member of another electrophoretically distinct mixture of species through the formation or reduction of a disulfide bond. The observed rate constants are, therefore, weight-averaged sums of the rate constants for all species involved in the particular conversion (Scheraga et al., 1987; Rothwarf & Scheraga, 1993b). Using the data presented here, it is not possible to determine the rate constants for the interconversion of any specific pair of chemically distinct species. However, it is possible to use these data to determine average values for the rate constants of each conversion.

Consider the R $\rightarrow$ 1S step. The reduced protein can form any one of 15 possible 1S species. The observed rate constant would be the weight-averaged sum of the 15 individual rate constants. The average rate constant for this step would then be $(k^{obs})_{f1}/15$. Similarly, any 1S species can be oxidized to form six possible 2S species. A 2S species can be oxidized to form only a single 3S species. The average forward rate constants for these reactions would be $(k^{obs})_{f2}/6$ and $(k^{obs})_{f3}$, respectively. Analogous arguments can be made for the reduction rate constants. The average rate constants for these steps would be $(k^{obs})_{r1}$, $(k^{obs})_{r2}/2$, and $(k^{obs})_{r3}/3$, respectively.

*Disulfide Bond Formation*. A comparison between the values of $(k^{ave})_f$ in Table 2 is of interest because it can yield information about the average conformational state of the protein at the various stages of regeneration. While the average rate constants for the formation of an intramolecular disulfide bond from the reduced protein and from 1S intermediates are approximately equal, the value of $(k^{ave})_f$ for the formation of a 3S from a 2S species is smaller. The rough agreement between the values of $(k^{ave})_f$ for the oxidation of R and 1S suggests that the mobility of the remaining protein thiols in these species is about equivalent and not significantly restricted by the formation of the first intramolecular disulfide bond.

The lower value for $(k^{ave})_f$ found for the formation of 3S suggests that the mobility of the remaining thiols in 2S is more restricted than those of either R or 1S. The formation of the second disulfide appears to be accompanied by the development of partially stable structures (not necessarily native-like) which reduce the conformational freedom of the protein, and thereby reduce the rate of formation of the third disulfide bond.

It is appropriate to compare the values of $(k^{ave})_f$ reported here for rHV1 with those reported for RNase A (Rothwarf & Scheraga, 1993b). When adjusted for differences in temperature and pH, the average forward rate constants $(k^{ave})_{f1}$ and $(k^{ave})_{f2}$ for rHV1 are approximately twice those found for all steps in the regeneration of RNase A. Since the sequences of these two proteins are so dis-similar, a 2-fold difference in the values of $(k^{ave})_f$ is not surprising.

What is revealing is that the value of $(k^{ave})_{f3}$, the average rate of formation of a 3S species from a 2S intermediate in rHV1, is much less than the values of $(k^{ave})_{f1}$ and $(k^{ave})_{f2}$. This is not the same trend that was found in the regeneration of RNase A. In the latter protein, the various values of $(k^{ave})_{fn}$ remained relatively constant for each step in the regeneration. This was interpreted as further evidence that the intermediates populated during the regeneration of RNase A are largely unstructured and disordered (Rothwarf & Scheraga, 1993b). For rHV1, on the other hand, the average rate constant for the formation of the third disulfide bond, $(k^{ave})_{f3}$, is significantly less than the average rate constants for the formation of the first and second disulfide bond, $(k^{ave})_{f1}$ and $(k^{ave})_{f2}$. This suggests that the formation of the second disulfide bond restricts the conformational freedom of the protein chain by a combination of the entropic and steric factors described above, rendering the remaining thiols relatively less reactive.

*Disulfide Bond Reduction*. The value of $(k^{ave})_r$ for the reduction of a 3S intermediate to a 2S intermediate is 1/2 of that for the reduction of a 2S to a 1S intermediate and is 1/3 of that for the reduction of a 1S intermediate. This implies that the disulfides of the grouping of 3S intermediates are not as readily accessible to the reducing agent and suggests the presence of some sort of "protective structure".

Although it appears that while, on average, the disulfide bonds of the ensemble of 3S intermediates are somewhat protected from reduction through the formation of the third disulfide bond, the "structure" responsible for the reduced rates of reduction is much less stable than the native structure.

[2] When the number 1, 2, or 3 follows the letters f or r in the subscript of these symbols, they refer to the reactions R $\rightleftarrows$ 1S, 1S $\rightleftarrows$ 2S, and 2S $\rightleftarrows$ 3S, respectively.

Preliminary experiments have demonstrated that the rate constant for reduction of the native protein is more than 3 orders of magnitude smaller than that for reduction of the ensemble of 3S species (T. W. Thannhauser and H. A. Scheraga, unpublished results).

The larger value of $(k^{ave})_r$ for the grouping of 2S intermediates, compared to that for the 3S grouping, indicates that, on average, the disulfide bonds in the 2S species are more accessible to reductant than those of the grouping of 3S species. However, the 2S grouping of intermediates does appear to retain some of this protective structure because the value of $(k^{ave})_r$ for this grouping of intermediates is less than that for the grouping of 1S species. The various constituents of the grouping of 1S intermediates appear to be largely unstructured, as indicated by the magnitude of $(k^{ave})_r$ (220 $min^{-1}$ $M^{-1}$). When adjusted for differences in temperature and pH, this value is roughly equivalent to that obtained for the reduction of the grouping of 1S intermediates of RNase A (Rothwarf & Scheraga, 1993b). It should be noted that the intermediates isolated from the regeneration pathway of RNase A have been shown to be disordered and unstructured (Konishi & Scheraga, 1980; Rothwarf & Scheraga, 1993b).

*$(k^{ave})_{intra}$.* Although it is possible to discuss the formation of disulfide bonds in a protein in terms of the values of $(k^{ave})_f$, it must be kept in mind that, in addition to their dependence on the intramolecular processes involved in the regeneration of rHV1, the values of $(k^{ave})_f$ depend strongly on the redox couple used. In order to focus attention on the intramolecular processes involved in the disulfide-coupled regeneration mechanism, it is helpful to use a quantity that is independent of the redox reagent and reflects only the conformational properties of rHV1.

For this purpose, consider the reaction shown schematically in Figure 7 which depicts the processes by which disulfide bonds are formed in proteins through oxidation with $DTT^{ox}$. The intermediate mixed disulfide species generated by the reduction of $DTT^{ox}$ by a protein thiol either can be converted rapidly to a species containing an intramolecular disulfide bond and a molecule of $DTT^{red}$ or can be converted back to the fully reduced protein and a molecule of $DTT^{ox}$ by the recyclization of the dithiothreitol. The observed forward rate constant for this process can be written in terms of an equilibrium constant $K_{DTT}$ and a rate constant $k_{intra}$ as

$$(k^{obs})_f = K_{DTT} k_{intra} \quad (4)$$

The rate constant $k_{intra}$ may be obtained by dividing the observed forward rate constant by $K_{DTT}$. The quantity $k_{intra}$ is the rate constant for the formation of an intramolecular disulfide bond. It pertains to an intramolecular process involving only the protein and, therefore, should be independent of the redox couple used (Creighton & Goldenberg, 1984; Rothwarf & Scheraga, 1993b). However, as was the case with the observed forward and reverse rate constants, the observed values of $k_{intra}$ are not representative of any single process since each step involves a mixture of many species. The observed values of $k_{intra}$ were adjusted for statistical factors (15, 6, and 1 for 1S, 2S, and 3S, respectively), as described above, to yield an average quantity, $(k^{ave})_{intra}$. The values for $(k^{ave})_{intra}$ reported in Table 2 were determined by using the value of $7.05 \times 10^{-3}$ $M^{-1}$ for $K_{DTT}$ (Chau & Nelson, 1991; Rothwarf & Scheraga, 1992, 1993c). These values of $(k^{ave})_{intra}$ are introduced here because they can be used to compare data obtained using different redox reagents, and they can be used to estimate the rates of disulfide bond exchange within each grouping of intermediates (see below).

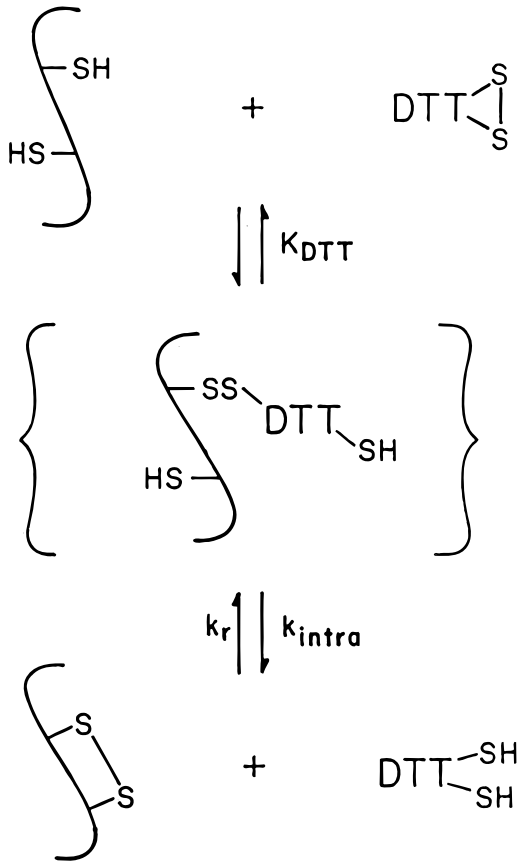


FIGURE 7: Formation of disulfide bonds by oxidation with $DTT^{ox}$/$DTT^{red}$.

*Rate of Thiol/Disulfide Interchange within the Groupings of Isomeric Intermediates*. In order to understand the regeneration process more fully, it is important to obtain some measure of the rate of thiol/disulfide interchange within the groupings of isomeric intermediate species. To do this directly would require the isolation of unique intermediates with specific disulfide pairs within each grouping. This has been possible in the case of certain intramolecular rearrangements in the regeneration of BPTI (Creighton & Goldenberg, 1984; Weissman & Kim, 1991, 1992, 1995); however, the regeneration pathway of BPTI is populated by relatively few stable native-like intermediates. The regeneration pathway of rHV1 appears to be characterized by many, mostly disordered intermediates (Chatrenet & Chang, 1992, 1993; Chang, 1995). This makes it difficult to obtain a direct determination of interchange rates for the regeneration of rHV1 by isolation of individual intermediates. Fortunately, it has been shown that $(k^{ave})_{intra}$ provides a reasonable estimate of the thiol/disulfide interchange rates within the groupings of isomeric intermediates (Rothwarf & Scheraga, 1993b).

The values of $(k^{ave})_{intra}$ for the formation of the 1S group from R and for the formation of the 2S group from 1S are approximately equal, averaging 10.9 $min^{-1}$. While that for the formation of the 3S group from the 2S group is less, it is still on the same order, 3.3 $min^{-1}$. These values can be compared to the average rates of oxidation and reduction of these groupings of intermediates by multiplying the values

of $(k^{ave})_f$ and $(k^{ave})_r$ by the concentrations of $DTT^{ox}$ and $DTT^{red}$, respectively. This comparison indicates that, under the conditions used in these experiments, the *average* rate of intramolecular disulfide interchange within isomeric intermediate groupings is a factor of $10^2-10^4$ greater than the *average* rates of oxidation ($10^{-3}$ $min^{-1}$) or reduction ($10^{-3}$ $min^{-1}$). This result is significant because, if the rates of interconversion within the groupings of isomeric intermediates are much faster than the rates of oxidation or reduction, then the distribution of species within each grouping will remain constant with time and redox conditions as is required by the KOS approach.

*Regeneration Pathway*. Through both an examination of the experimental data and an analysis of the rate constants presented above, a picture of the regeneration process of rHV1 emerges. The reduced protein (R) and the ensembles of isomeric disulfide bond-containing intermediates (1S, 2S, and 3S) rapidly reach a pre-equilibrium steady state. The rate of regeneration of native protein was found to depend both on the concentration of the ensemble of 2S intermediates *and* on the concentration of $DTT^{ox}$. This finding indicates that the rate-determining step in the major regeneration pathway of rHV1 is the oxidation of one or more 2S species, i.e., 2S → *3S**.

Since the rate-determining step in the regeneration of rHV1 appears to be the oxidation of a 2S species, it is of interest to consider the disulfide composition of this 2S species. There are three possibilities. It can contain two non-native disulfide bonds, one native and one non-native disulfide bond, or two native disulfide bonds.

The first two possibilities would require that the rate-determining step in the regeneration process be the formation of a $3S^*$ species that contains either two or three non-native disulfide bonds. However, the $3S^*$ species occurs after the rate-determining step and is kinetically indistinguishable from the native protein. Therefore, it seems unlikely that any 3S species, containing two or three non-native disulfide bonds, could undergo the intermolecular reduction, multiple intramolecular disulfide rearrangements, and intermolecular oxidation steps necessary to produce the native protein at such a rapid rate and with no kinetically detectable intermediates. Such a situation is made even more unlikely by the use of $DTT^{ox}/DTT^{red}$ as the redox couple. Consequently, it seems more reasonable to postulate that the rate-determining step in the regeneration of rHV1 involves the formation of the final native disulfide bond. This conclusion suggests that the critical 2S intermediate(s) contains two native disulfide bonds. Since three such species are possible,

$2S^{6-14,16-28}$, $2S^{6-14,22-39}$, and $2S^{16-28,22-39}$, this raises the possibility that as many as three distinct regeneration pathways may exist. From the data presented here, we cannot determine which of the three possible critical intermediates is actually involved in the major regeneration pathway of rHV1.

While the analysis of the kinetic data has led to the conclusion that, under the experimental conditions used here, the rate-determining step involves an oxidation and is of the type 2S → $3S^*$, it is possible that, at sufficiently high oxidizing conditions, the rate of formation of the critical 2S species will become slower than its subsequent oxidation. Such a situation would result in a change in the kinetically observable pathway to the type 2S → *2S**. However, in experiments carried out under much higher oxidizing conditions then those used here, in which GSSG/GSH was used as the redox couple, the rate of regeneration of rHV1 was greatly accelerated (Chatrenet & Chang, 1993). This observation is consistent with a pathway in which the rate-determining step involves an oxidation process (see below).

*Comparison with a Previously Proposed Regeneration Pathway*. A comparison between the regeneration model proposed here for rHV1 and the one proposed by Chatrenet and Chang (1993), based on their experiments on $rHV1^{1-49}$, reveals differences (see the introductory section). The origin of these differences lies in the different model proteins and techniques used to obtain the experimental data.

Chatrenet and Chang (1993) proposed their pathway for the regeneration of hirudin (rHV1), on the basis of experiments carried out on the amino-terminal fragment of hirudin, $rHV1^{1-49}$, even though their earlier experiments had demonstrated that there was a significant difference in the rates at which the native structures of these two proteins are regenerated (Chatrenet & Chang, 1992). The presence of the C-terminal tail of hirudin modifies its regeneration kinetics and is probably responsible for some of the differences in the proposed regeneration pathways.

In the experiments of Chatrenet and Chang (1993) used to develop their model, dissolved atmospheric oxygen was the oxidizing agent. These studies were carried out in the absence of any small molecule thiol/disulfide redox couple or in the presence of a small amount (60 $\mu$M) of $\beta$-mercaptoethanol. Under these types of conditions, the rate of reduction of disulfides is very slow (Rothwarf & Scheraga, 1993c).

The exaggerated importance of the scrambled 3S species in their model would be a consequence of the use of such highly oxidizing conditions that caused the rate-determining step in the regeneration process to become the reduction of the non-native 3S intermediates. Clearly, a scrambled 3S species (which must contain at least two non-native disulfide bonds) must proceed to the native protein through a 2S intermediate. However, because the concentration of reducing agent was too low to produce measurable amounts of 1S or 2S, they were unable to detect meaningful concentrations of these intermediates; hence, Chatrenet and Chang (1993) concluded that these intermediates are not important in the regeneration of the native protein.

Another potential difficulty that would tend to exacerbate the problem of determining the significance of the 1S and 2S species under the conditions used by Chatrenet and Chang was their reliance on acid quenching as a method of terminating the refolding process. Although reduction of the pH does decrease the rate of disulfide reshuffling, it does not provide a means by which R, N, and the 1S, 2S, and 3S intermediates can be reliably and predictably separated. The inability to resolve the groupings of disulfide-bonded intermediates would complicate the task of estimating their concentrations. The fact that most of the chromatographic fractions used by Chatrenet and Chang to characterize the regeneration actually were mixtures of intermediates containing 1S, 2S, and 3S species probably contributed to their failure to recognize the significance of the 2S species as the ones involved in the rate-determining step (Chatrenet & Chang, 1993).

Chatrenet and Chang (1992, 1993) found it necessary to add small amounts of reductants such as $\beta$-mercaptoethanol or GSH to the mixture of scrambled 3S intermediates and N

in order to ensure full recovery of the native protein. However, in the absence of significant concentrations of reducing agent, the rate-determining step in the formation of native protein will still be the reduction of the scrambled 3S species. This is a consequence of the redox conditions and the type of redox reagent used and not a reflection of the conformational preferences of the protein. An acyclic thiol reagent such as $\beta$-mercaptoethanol or GSH reacts with the disulfides of the 3S species by a disulfide interchange mechanism that results in the formation of stable mixed disulfide bonds with the protein thiols (Rothwarf & Scheraga, 1993a,c). Consequently, acyclic thiol reagents are much weaker reducing agents than $DTT^{red}$, which has been used here. Therefore, it is very unlikely that any significant population of 1S or 2S species would form from the scrambled 3S species under the experimental conditions used by Chatrenet and Chang (1993).

While it was not incorporated into their model, Chatrenet and Chang did investigate the regeneration rate of $rHV1^{1-49}$ using various concentrations of GSSG/GSH. Their results demonstrate that, under some conditions, the regeneration of native protein was significantly more rapid than the rates observed starting with scrambled 3S species or than those that we observed for rHV1 using $DTT^{ox}/DTT^{red}$ as the redox couple. GSSG is a much more potent oxidizing agent than $DTT^{ox}$, and its use would therefore be expected to accelerate the rate of the rHV1 regeneration greatly if the rate-determining step involved an oxidation process. Therefore, the results of Chatrenet and Chang (1993), using GSSG/GSH, are consistent with the conclusion reported here, that the rate-determining step in the major regeneration pathway(s) is of the type 2S → *3S**. Such consistency, when comparing the effects of different types of redox reagents, has also been observed in regeneration studies of RNase A (a four-disulfide-containing protein) in which the rate-determining step in the major pathway at 25 °C and pH 8.0 is of the type 3S → *3S**. Consequently, it would be expected that the optimal rates of regeneration of native RNase A using $DTT^{ox}/DTT^{red}$ or GSSG/GSH would be very similar (Rothwarf & Scheraga, 1993b,c), and this has been verified experimentally (Rothwarf & Scheraga, 1993a).

## CONCLUSIONS

The experiments reported here were carried out under conditions such that a pre-equilibrium steady state was rapidly established among the reduced protein and the various disulfide-containing intermediates. The regeneration process was arrested at various times, and the distribution of intermediate species was frozen by the addition of AEMTS. The positive charge contributed by each blocking group provided the means by which the intermediates were separated into isomeric groupings (R, 1S, 2S, and 3S), each of which contained the same number of intramolecular disulfides and blocking groups. The concentrations of these species were then determined as a function of time and redox conditions. Fitting the experimental results, using the general KOS approach to protein regeneration, reveals that the rate of recovery of the native protein can best be described by a mechanism involving a pathway in which the oxidation to form the final disulfide represents the rate-determining step. This finding makes it reasonable to propose that the specific intermediate(s) directly involved in the regeneration of the native protein are the ones that already contain two native disulfide bonds. At this stage of the investigation, it cannot be determined which of the three possible native 2S intermediates is/are involved in the rate-determining reaction, or which structural features are associated with the 2S → *3S** transition. Experiments designed to identify these structural features are currently in progress.

## ACKNOWLEDGMENT

We thank Dr. R. B. Wallis (Ciba Geigy, Horsham) for the recombinant hirudin, Mr. R. W. Sherwood for assistance with the amino acid analysis, Dr. D. P. Little, Dr. N. A. Kelleher, and Professor F. W. McLafferty for the ESI MS results, and Dr. R. W. Dodge for helpful discussions and critical reading of the manuscript.

## REFERENCES

Anfinsen, C. B. (1973) *Science 181*, 223.
Baldwin, R. L. (1995) *J. Biol. NMR 5*, 103.
Beavis, R. C., & Chait, B. T. (1989) *Rapid Commun. Mass Spectrom. 3*, 233.
Beu, S. C., Senko, M. W., Quinn, J. P., Wampler, F. M., III, & McLafferty, F. W. (1993) *J. Am. Soc. Mass Spectrom. 4*, 557.
Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93.
Bischoff, R., & Kolbe, H. V. J. (1994) *J. Chromatogr. B 662*, 261.
Bodner, B. L., Jackman, L. M., & Morgan, R. S. (1980) *Biochem. Biophys. Res. Commun. 94*, 807.
Bruice, T. W., & Kenyon, G. L. (1982) *J. Protein Chem. 1*, 47.
Chang, J.-Y. (1990) *J. Biol. Chem. 265*, 22159.
Chang, J.-Y. (1995) *J. Biol. Chem. 270*, 25661.
Chatrenet, B., & Chang, J.-Y. (1992) *J. Biol. Chem. 267*, 3038.
Chatrenet, B., & Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 20988.
Chau, M. H., & Nelson, J. W. (1991) *FEBS Lett. 291*, 296.
Creighton, T. E. (1975) *J. Mol. Biol. 96*, 767.
Creighton, T. E. (1979) *J. Mol. Biol. 129*, 411.
Creighton, T. E. (1985) *J. Phys. Chem. 89*, 2452.
Creighton, T. E. (1992) *Science 256*, 111.
Creighton, T. E., & Goldenberg, D. B. (1984) *J. Mol. Biol. 179*, 497.
Crimmins, D. L., Saylor, M., Rush, J., & Thoma, R. S. (1995) *Anal. Biochem. 226*, 355.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70.
Goldenberg, D. P. (1988) *Biochemistry 27*, 2481.
Hantgan, R. R., Hammes, G. G., & Scheraga, H. A. (1974) *Biochemistry 13*, 3421.
Heinrikson, R. L., & Meredith, S. C. (1984) *Anal. Biochem. 136*, 65.
Hillenkamp, F., & Karas, M. (1990) *Methods Enzymol. 193*, 280.
Hird, F. J. R. (1962) *Biochem. J. 85*, 320.
Hirs, C. H. W. (1956) *J. Biol. Chem. 219*, 611.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631.
Konishi, Y., & Scheraga, H. A. (1980) *Biochemistry 19*, 1308.
Konishi, Y., Ooi, T., & Scheraga, H. A. (1981) *Biochemistry 20*, 3945.
Konishi, Y., Ooi, T., & Scheraga, H. A. (1982) *Biochemistry 21*, 4734.
Li, Y.-J., Rothwarf, D. M., & Scheraga, H. A. (1995) *Nat. Struct. Biol. 2*, 489.
Lin, S. H., Konishi, Y., Denton, M. E., & Scheraga, H. A. (1984) *Biochemistry 23*, 5504.
Meinwald, Y. C., Stimson, E. R., & Scheraga, H. A. (1986) *Int. J. Pept. Protein Res. 28*, 79.
Morgan, R. S., Tatsch, C. E., Gushard, R. H., McAdon, J. M., & Warme, P. K. (1978) *Int. J. Pept. Protein Res. 11*, 209.
Némethy, G., & Scheraga, H. A. (1981) *Biochem. Biophys. Res. Commun. 98*, 482.
Otto, A., & Seckler, R. (1991) *Eur. J. Biochem. 202*, 67.
Pace, C. N., & Creighton, T. E. (1986) *J. Mol. Biol. 188*, 477.
Patterson, S. D., & Katta, V. (1994) *Anal. Chem. 66*, 3727.
Poland, D. C., & Scheraga, H. A. (1965) *Biopolymers 3*, 379.
Rothwarf, D. M., & Scheraga, H. A. (1991) *J. Am. Chem. Soc. 113*, 6293.

Rothwarf, D. M., & Scheraga, H. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7944.
Rothwarf, D. M., & Scheraga, H. A. (1993a) *Biochemistry 32*, 2671.
Rothwarf, D. M., & Scheraga, H. A. (1993b) *Biochemistry 32*, 2680.
Rothwarf, D. M., & Scheraga, H. A. (1993c) *Biochemistry 32*, 2690.
Rothwarf, D. M., & Scheraga, H. A. (1993d) *Biochemistry 32*, 2698.
Saxena, V. P., & Wetlaufer, D. B. (1970) *Biochemistry 9*, 5015.
Scheraga, H. A., Konishi, Y., Rothwarf, D. M., & Mui, P. W. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5740.
Snyder, G. H., Cennerazzo, M. J., Karalis, A. J., & Field, D. (1981) *Biochemistry 20*, 6509.
Stone, S. R., & Maraganore, J. M. (1993) *Methods Enzymol. 223*, 312.
Straume, M., & Johnson, M. L. (1992) *Methods Enzymol. 210*, 117.
Szajewski, R. P., & Whitesides, G. M. (1980) *J. Am. Chem. Soc. 102*, 2011.
Thannhauser, T. W., & Scheraga, H. A. (1985) *Biochemistry 24*, 7681.
Thannhauser, T. W., Konishi, Y., & Scheraga, H. A. (1984) *Anal. Biochem. 138*, 181.
Thannhauser, T. W., Konishi, Y., & Scheraga, H. A. (1987) *Methods Enzymol. 143*, 115.
Thannhauser, T. W., Sherwood, R. W., & Scheraga, H. A. (1997) *Anal. Biochem.* (submitted for publication).
Tuong, A., Maftouh, M., Ponthus, C., Whitechruch, O., Roitsch, C., & Picard, C. (1992) *Biochemistry 31*, 8291.
Weissman, J. S., & Kim, P. S. (1991) *Science 253*, 1386.
Weissman, J. S., & Kim, P. S. (1992) *Science 256*, 112.
Weissman, J. S., & Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 1123.
Wetlaufer, D. B., Branca, P. A., & Chen, G.-X. (1987) *Protein Eng. 1*, 141.
Xu, X., Rothwarf, D. M., & Scheraga, H. A. (1996) *Biochemistry 35*, 6406.

BI962340W