THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 19, Issue of May 10, pp. 11301–11308, 1996

Printed in U.S.A.

Refolding and Oxidation of Recombinant Human Stem Cell Factor Produced in *Escherichia coli**

(Received for publication, October 1, 1995, and in revised form, February 15, 1996)

Michael D. Jones, Linda O. Narhi, Wen-Chang Chang‡, and Hsieng S. Lu§

From Amgen Inc., Thousand Oaks, California 91320

Oxidative folding of recombinant human stem cell factor (rhSCF) produced in *Escherichia coli* was investigated *in vitro*. Folding of denatured and reduced rhSCF involves at least five intermediate forms, I-1 to I-5, detectable by their differences in hydrophobicity using reverse-phase high performance liquid chromatography. Both I-1 and I-2 contain a native-like disulfide bond, Cys^{4}–Cys^{89} and Cys^{43}–Cys^{138}, respectively, and I-3 forms a mispaired disulfide, Cys^{43}–Cys^{89}. These forms appear to reach steady state equilibrium and are important folding intermediates. I-1 was found to be the prominent intermediate that directly folds into native rhSCF (N); and the thermodynamically less stable I-2 favors rearrangment into I-1. I-3 may serve as an intermediate for disulfide rearrangement between I-1 and I-2. I-4 and I-5, which are disulfide-linked dimers, are in equilibrium with reduced rhSCF and other intermediates and may not play an important role in rhSCF folding. Both trifluoroacetic acid-trapped I-1 and I-2, after isolation by high performance liquid chromatography, proceed with the remaining oxidative folding process after reconstitution. Iodoacetate-trapped I-1 and I-2 contain low α-helical content and some tertiary structure, while I-3 and reduced rhSCF have little ordered structure. Gel filtration/light-scattering experiments indicate that reduced rhSCF and iodoacetate-trapped I-1, I-2, and I-3 exist as dimeric forms, indicating that rhSCF dimerization precedes formation of disulfide bonds. I-1, I-2, I-3, and the C43,138A analog lacking Cys^{43}–Cys^{138} bond are not biologically active or exhibit significantly lower activity. The two disulfide bonds in rhSCF seem to be essential for the molecule to maintain an active conformation required for its receptor binding and biological activities.

The mechanistic folding of proteins is a critical area of study in understanding their structural and conformational behaviors in solution. Much experimental work has been done on the thermodynamics and kinetics of folding pathways of proteins (1–4). Disulfide-containing proteins have the advantages that the intra- or interchain disulfide bond is a natural covalent cross-link which, by thermodynamic requirements, stabilizes the cross-linked conformation (1). Oxidative refolding of a disulfide-reduced protein allows one to study the rates of folding and to chemically trap the folding intermediates. Many such folding intermediates are sufficiently long-lived during folding and can be isolated and characterized. In several cases, detailed folding pathways have been defined in terms of occurrence of disulfide-containing intermediates at various stages of folding (4–7). Through the analysis of disulfide intermediates, it is apparent that disulfide formation does not necessarily occur through a simple sequential pathway. As a result, specific issues regarding protein folding pathways have been brought to attention (3, 6–10). These include the significance of non-native disulfides, the kinetic importance of well populated and thermodynamically stable intermediates, and the clarification of the consequence and role of kinetically trapped species that accumulate in the pathway.

Human stem cell factor (hSCF)[1] is a hematopoietic cytokine that plays an important role in the stimulation, proliferation, differentiation, and functional activation of blood cells (11, 12). The molecules obtained naturally or expressed in recombinant CHO cells are heavily glycosylated (13–15). Recombinant hSCF (rhSCF) can also be produced in genetically engineered *Escherichia coli* (16). The glycosylated or nonglycosylated molecule exists as a noncovalently linked homodimer in solution (17). Dissociation-reassociation of rhSCF dimer spontaneously occurs at a fairly rapid rate (18). There are two intramolecular disulfide bonds present in the molecule (Cys^{4}–Cys^{89} and Cys^{43}–Cys^{138}), and the production of active rhSCF from *E. coli* requires an oxidative folding procedure to recover its biological activity (16). In this study, the refolding pathway of reduced rhSCF is examined by monitoring the distribution of HPLC-separable disulfide-bonded intermediates during refolding and by characterizing physicochemical and biological properties of the isolated intermediates. The folding of rhSCF involves intermediates containing native and non-native disulfides, thermodynamically stable and unstable intermediates, as well as disulfide-bonded dimers. These studies allow us better to understand the contribution of native and mispaired disulfide-containing intermediates to the *in vitro* folding pathway of rhSCF.

EXPERIMENTAL PROCEDURES

Materials—*E. coli*-produced rhSCF was purified according to methods described previously (16). The recombinant molecule contains 165 amino acids plus an N-terminal methionine at position −1. Ultrapure DTT^{red} was purchased from Boehringer-Mannheim, while DTT^{ox} was from Sigma. These two reagents were verified to be highly pure by reverse-phase HPLC using a C4 column. Expression, isolation and some

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Amgen's visiting scientist from the Institute of Biological Chemistry, Academia Sinica and Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan.

§ To whom correspondence should be addressed: Amgen Inc., M/S 14-2-E, Amgen Center, Thousand Oaks, CA 91320-1789. Tel.: 805-447-3092; Fax: 805-499-7464.

[1] The abbreviations used are: hSCF, human stem cell factor; rhSCF, recombinant hSCF; CHO, Chinese hamster ovary; DTT, dithiothreitol; DTT^{ox}, oxdized DTT; DTT^{red}, reduced DTT; TFA, trifluoroacetic acid; IAA, iodoacetic acid; RP, reversed phase; HPLC, high performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; Cys(Cm), *S*-carboxymethylated Cys; GdnHCl, guanidine HCl; FWHM, full width at half emission maximum; N, fully oxidized rhSCF; R, fully reduced rhSCF; and P, SDS-nondissociable but DTT-reducible rhSCF dimer.

characterization of rhSCF C43,138A analog have been described previously (19). Crystallized iodoacetic acid was purchased from Sigma. HPLC solvents and water were purchased from Burdick and Jackson. Sequencing reagents and solvents were supplied by Applied Biosystems and Hewlett Packard. All other reagents were of the highest quality available.

Denaturation and Reduction of rhSCF—rhSCF (3 mg/ml) was reduced with a 100-fold molar excess of DTT^{red} in 0.3 M Tris containing 6 M GdnHCl (pH 8.5) for 4 h at 25 °C. The sample was quickly buffer-exchanged into refolding buffer using an HR10/10 desalting column (Pharmacia Biotech Inc.) run in a fast protein liquid chromatography system. The refolding buffer is 50 mM Tris-HCl containing 2.5 M urea, 2 mM EDTA, and 60 mM NaCl (pH 8.5). Oxygen in the buffer was completely removed by a continuous nitrogen flushing of the buffer before use. Urea is introduced in the buffer to keep denatured and reduced rhSCF soluble. The concentration of protein was spectrophotometrically determined at 280 nm (absorbance = 0.62 for rhSCF at 1 mg/ml) (17). Refolding studies were carried out immediately after buffer exchange.

Refolding and Oxidation of Reduced rhSCF with DTT^{ox}—Fully reduced and denatured rhSCF was diluted to 0.25–1.0 mg/ml in oxygen-free refolding buffer as described. Samples were allowed to refold in the presence of DTT^{ox} at 66.6 mM concentration. At various time intervals, aliquots of folding mixtures were taken, which were then reacted immediately with iodoacetic acid (IAA) or acidified with TFA (see below).

The refolding rhSCF mixtures, after IAA or acid treatment, were injected onto C-4 reverse phase columns (Syn-Chro Pak, 0.46 × 25 cm, 300 Å). Solvent A is 0.1% TFA and solvent B is 0.1% TFA in 90% acetonitrile. The column was initially equilibrated with 55% B and eluted by linear gradients from 55 to 65% B over 80 min and from 65 to 95% B over 5 min on an HP-1090 liquid chromatograph equipped with an autosampler and a photodiode array detector. Elution of protein peaks was detected at 215 nm, and protein fractions obtained from HPLC were collected for further analysis.

IAA Modification and Acid Quenching of Intermediates—Aliquots of folding mixture were removed at selected time intervals and unreacted thiols blocked by the addition of 0.1 volume of 1 M iodoacetic acid in 1.0 M Tris-HCl (pH 8.0) for 2 min at room temperature. Samples were then immediately analyzed by HPLC or desalted as described above. TFA-quenched samples were prepared by adding a 0.1 volume of 25% TFA into the folding mixture.

Partial Reduction of rhSCF—Native rhSCF at 1 mg/ml was incubated in the folding buffer (see above) containing 4 mM DTT^{red} at 25 °C. Aliquots were removed at various times and acidified by the addition of 0.1 volume of 25% TFA for subsequent RP-HPLC analysis.

Preparation and Refolding of Acid-quenched Intermediates—Intermediate forms obtained from acid-quenching of the folding samples or from partial reduction of native rhSCF as described can be prepared by RP-HPLC (see "Results"). The samples obtained were partially dried in a vacuum centrifuge and reconstituted in 100 μl of 10 mM HCl containing 2.5 M urea (pH 2.0). Folding of each isolated intermediate was resumed by adding 5 volumes of the above-mentioned folding buffer containing DTT^{ox}.

Protease Digestion and Peptide Mapping—IAA-alkylated intermediates isolated from HPLC were dried and reconstituted in 50 mM Tris buffer containing 2.5 M urea, 2 mM EDTA, and 60 mM NaCl (pH 8.5) for endoproteinase Glu-C digestion. Peptide mapping was performed as described previously (20). Samples prepared for Asp-N peptide mapping were reconstituted using the same buffer described above but without addition of 2 mM EDTA, and detailed procedures were also as reported previously (16).

Sequence Analysis and Mass Spectrometry—Sequence analysis of peptides was performed on an automatic protein sequencer (Applied Biosystems models 477A or 470 or Hewlett Packard G1000A). Mass spectroscopic analysis of peptides was performed on a Sciex API III electrospray mass spectrometer as described elsewhere (20).

Circular Dichroism (CD) and Fluorescence Spectroscopic Analyses—CD spectra were obtained using a Jasco J-720 spectropolarimeter and cuvettes with a 0.02 cm (for far UV CD) or a 1 cm (for near UV CD) path length. Protein samples were clarified through sterile filtration using a 0.2- μm membrane, and concentrations were determined by their absorbance at 280 nm. All proteins were analyzed in 50 mM Tris-HCl containing 2.5 M urea, 2 mM EDTA, and 60 mM NaCl (pH 8.5). The α-helical content of proteins was estimated according to Greenfield and Fasman (21). The thermal stability was measured by the temperature-induced change in ellipticity at 222 nm, using a JTC-345 Peltier thermal control unit, and a rectangular cuvette with a path length of 0.1 cm. The proteins were heated at 20 °C/h, and the spectra from 240 to 200 nm were collected every 2 ± 0.5 °C. The ellipticity at 222 nm alone was simultaneously monitored every 0.5 °C. The curve of ellipticity at 222 nm *versus* temperature was fitted by a computer program which fits the base lines before and after heating individually, and then fits the transition region compensating for the base lines.

The fluorescence spectra were determined at ambient temperature on an SLM Aminco 500C spectrofluorimeter, using a cuvette with a path length of 0.5 cm and protein concentrations of 0.22 mg/ml. The solutions were excited at 280 nm, and the fluorescence spectra from 280 to 420 nm were recorded, using slit widths of 5 nm.

rhSCF Biological Assay and SDS-PAGE—[^{3}H]Thymidine incorporation by the human megakaryoblastic leukemia cell line UT-7 was monitored in cell proliferation assays as described previously (22). Nonreducing and reducing SDS-PAGE using 16% precast gels (Novex, San Diego, CA) was performed using a Laemmli (23) gel system.

Size-exclusion Chromatography and Light-scattering Detection—Gel filtration was performed in Superdex 75 column (Pharmacia) preequilibrated with 0.1 M phosphate buffer (pH 7.0) containing 0.5 M NaCl. Samples were analyzed by an HP-1050 Ti system (Hewlett Packard) equipped with an autosampler, a diode array detector, and a ChemStation PC-based software package for data acquisition and analysis using conditions as described elsewhere (17). Molecular weights were determined using two detectors in series: an on-line laser light-scattering detector (Mini-Dawn, Wyatt Technologies) and a refractive index detector (Hewlett Packard HP1047A). Chromatograms were collected and analyzed on the Hewlett Packard ChemStation. Molecular weights were obtained from the ratio of light-scattering and refractive index signals, using procedures similar to those described previously (24, 25), after calibration using ribonuclease, ovalbumin, and bovine serum albumin monomer/dimer. The molecular weight of samples determined by this procedure is independent of their respective elution times.

RESULTS

Refolding of Reduced and Denatured rhSCF—In its native state, folded and disulfide-bonded rhSCF forms a noncovalently linked dimer in solution (17). In RP-HPLC, native rhSCF has been dissociated into monomers under acidic elution conditions and behaved as a sharp symmetrical peak. Under conditions using a shallow acetonitrile gradient elution in TFA, the reduced rhSCF form was well separated from the oxidized form and eluted approximately 10 min later. As shown in Fig. 1*A*, this HPLC procedure allows the separation of various folding intermediates for studying oxidative folding of denatured and reduced rhSCF.

Fig. 1*A* shows RP-HPLC analysis of IAA-treated sample aliquots taken at different time intervals. Folding after TFA acidification also exhibited a similar HPLC profile containing the same relative distribution of intermediate forms. RP-HPLC can resolve several IAA-trapped intermediates in addition to the native (N) and fully reduced (R) forms in rhSCF folding mixtures. In the initial stages of the reaction (<2 h), R and N species were readily identified along with five intermediate forms, designated as I-1, I-2, I-3, I-4, and I-5. I-4 and I-5 seemed to form very early (<20 min) and decreased to a very low level later. Intermediate I-1, I-2, and I-3 still remain to some extent throughout the whole folding process. At later stages in the folding (>20 h) a particular form P that elutes more hydrophobically than R also appeared. P was identified to be an SDS-nondissociable but DTT-reducible rhSCF dimer as described in a companion study (26).

The folding kinetics were investigated further by plotting folding times *versus* rate of generation for each intermediate estimated from HPLC peak integration. Fig. 1*B* shows the kinetic profiles for the five intermediate forms as well as R and N, while Fig. 1*C* depicts generation rates for intermediates within the first 20 min of folding. Initially R disappeared rapidly with concomitant formation of various intermediate forms and the finally folded form N. The data clearly show that the I-1 concentration increased rapidly and persisted for a longer period of time during folding. After 48 h of folding, 72.9% of the total protein has completely folded into N, with 12.1% remaining in I-1 and 12.9% in P. Only 1.3% of the

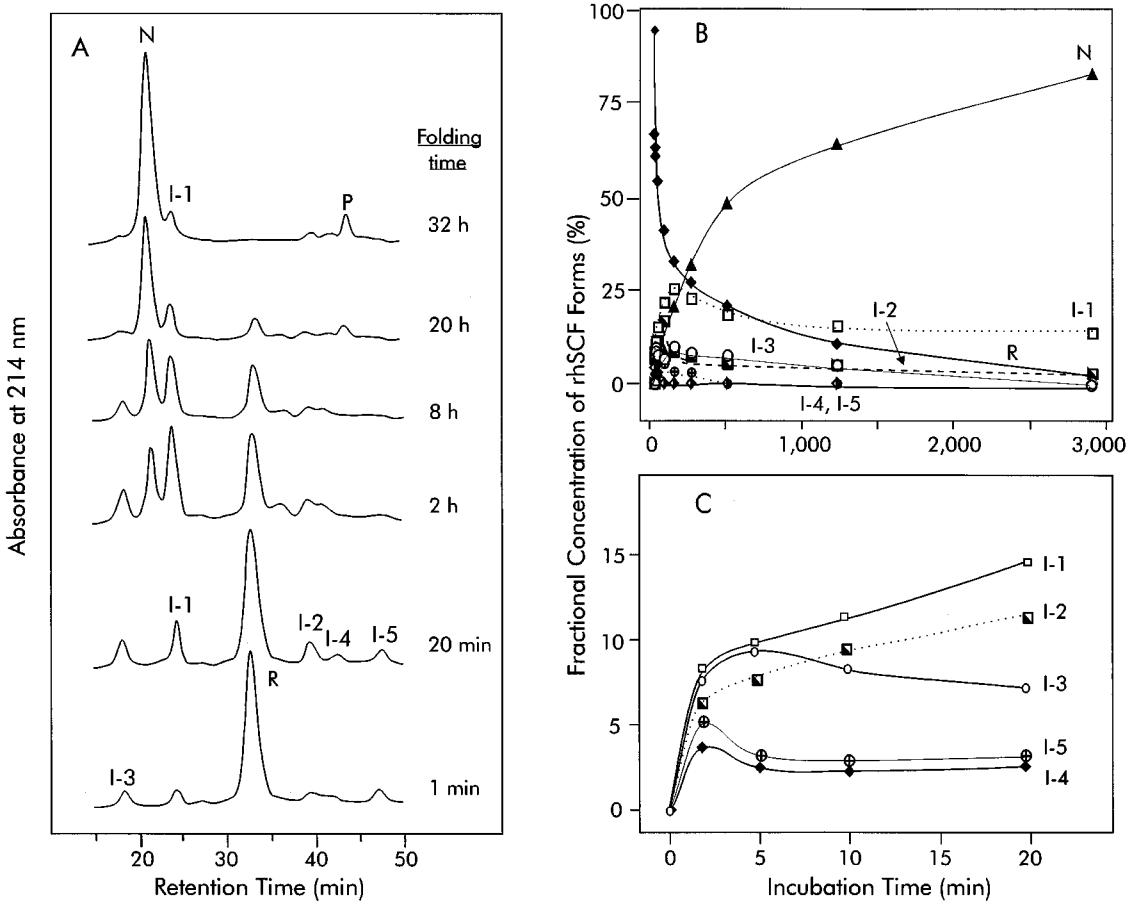

FIG. 1. **HPLC analysis and kinetics of rhSCF folding products.** RP-HPLC of rhSCF intermediates after iodoacetic acid reaction prepared at 1 mg/ml (*panel A*) and folding kinetics for the disappearance of reduced rhSCF (*R*) and the appearance of the fully folded rhSCF (*N*) and partially folded intermediates (*I-1* to *I-5*) over a period of 50 h and 20 min (*panels B* and *C*, respectively). The percent generation of rhSCF forms represents peak area of each intermediate divided by the sum of the peak area of all intermediates at the given time.

TABLE I
Initial folding rates and times required to reach maximal concentrations for rhSCF folding intermediates

Folding products[a]	Initial rate		Time reaching maximal concentration		Disulfide formed
	1 mg/ml	0.43 mg/ml	1 mg/ml	0.43 mg/ml	
A. Folding of denatured and reduced rhSCF:	*percent/s*		*min*		
I-1	1.24	0.24	120	240	Cys^{4}–Cys^{89}
I-2	1.07	0.12	60	240	Cys^{43}–Cys^{138}
I-3	0.94	0.10	5	240	Cys^{43}–Cys^{89}
I-4	ND[b]	ND	2	ND	Cys^{43}–Cys^{43}
I-5	ND	ND	2	ND	Cys^{89}–Cys^{89}
B. Folding of reduced and denatured Cys(Cm) I-2 and C43,138A analog at 0.43 mg/ml:					
Cys(Cm) I-2	ND	0.281		ND	Cys^{43}–Cys^{138}
C43,138A analog	ND	0.298		ND	Cys^{4}–Cys^{89}

[a] Cys(Cm) I-2 was prepared by HPLC isolation from the IAA-trapped folding mixture.
[b] ND, not determined.

protein remained fully reduced.

The initial first order folding rates for conversion of R into each of the intermediates and the time required to reach maximal concentration derived from the results shown in Fig. 1 are listed in Table I. Initial rates of I-1, I-2, and I-3 formation are similar. However, the time required for I-1 to reach its higher maximal concentration is much longer than those for I-2 and I-3. I-4 and I-5 reached maximal concentration by two min after which their rates of formation decreased. As a result, their initial rate constants could not be determined. In separate folding experiments, decreasing rhSCF concentration (*i.e.* 0.43 mg/ml) has dramatically reduced the initial folding rate and increased the times for intermediates to reach their maximal concentrations (Table I). For, example, the initial folding rate has decreased 5 to 10-fold for I-1, I-2, and I-3. In the folding at lower rhSCF concentrations (0.2–0.5 mg/ml), the disulfide linked dimer, I-4, and I-5, have diminished.

Different rhSCF intermediates were subjected to nonreducing SDS-PAGE. As listed in Table II, R migrates as an extended, larger molecule than the native (18 and 15 kDa, respectively). IAA-trapped I-1 and I-3 also have expanded molecular sizes between N and R. Both I-4 and I-5 migrated as a dimer having a molecular mass of 38 kDa. Upon reduction, their molecular masses reduced to 18 kDa.

Structural Characterization of IAA-trapped Intermediates—For structural characterization, the above-mentioned IAA-modified R and intermediates were isolated by RP-HPLC as described (Fig. 1*A*) and subjected to peptide mapping after endoproteinase Glu-C digestion. Fig. 2 compares peptide maps derived from the digestion of IAA-trapped I-1, I-2, I-3, I-4, I-5, as well as Cys(Cm)-rhSCF. The obtained peptides were characterized by sequence analysis and electrospray mass spectrometry. R after IAA modification had a profile identical to that of Cys(Cm)-rhSCF (*chromatogram 1*). I-1 contains a single Cys^{4}–Cys^{89} disulfide bond with Cys^{43} and Cys^{138} remaining unpaired, and I-2 has a Cys^{43}–Cys^{138} disulfide bond with unpaired Cys^{4} and Cys^{89} as well (*chromatograms 3* and *4*, respectively). I-3 forms a non-native Cys^{43}–Cys^{89} disulfide bond and leaves Cys^{4} and Cys^{138} unpaired (*chromatogram 2*). Both I-4 and I-5 also form an incorrect disulfide bond, *i.e.* Cys^{43}–Cys^{43} and Cys^{89}–Cys^{89}, respectively, to become disulfide-linked dimers (*chromatograms 6* and *7*). I-4 and I-5 have three free cysteines as seen in their respective peptide maps. The disulfide structures of form N was confirmed to be identical to rhSCF standard by direct comparison of peptide maps obtained from Asp-N endoproteinase digestion (data not shown, also see Ref. 16).

TABLE II
Molecular properties of rhSCF standard, folded forms, and related intermediates

rhSCF species	Tm[a]	Molecular mass	
		Nonreducing SDS-PAGE[b]	Gel filtration/ light-scattering[c]
	°C	kDa	
Standard	61	15.0	34.3
N	60	15.0	34.4
R	ND[d]	18.0	— (dimer)[e]
I-1	48	17.0	38.1
I-2	36	16.0	34.9
I-3	ND	17.0	33.3
I-4	ND	38.0	ND
I-5	ND	38.0	ND
C43, 138A analog	48	15.0	35.4

[a] Tm, the melting temperature, is defined as the temperature required to induce 50% denaturation, and determined by CD analysis at 222 nm shown in Fig. 5*C*.
[b] Determined by nonreducing SDS-PAGE using 14% gels according to Laemmli (23).
[c] Determined by gel filtration/laser light scattering as indicated in Fig. 6.
[d] ND, not determined.
[e] Determined by gel filtration using regular calibration standards according to their retention times.

Folding of Acid-quenched Intermediates and Partial Reduction of Native rhSCF—Following HPLC isolation of TFA-trapped intermediates, folding of individual intermediates was subsequently investigated. Fig. 3, *A* and *B*, shows the folding of I-1 and I-2 by HPLC analysis. Folding of I-1 appears to generate only N. In contrast, I-2 folds into I-1 and I-3, which then convert to N. A small amount of R and an uncharacterized form X near I-1 were detected as well.

Folding of rhSCF C43,138A Analog and Alkylated I-2—Both rhSCF C43,138A analog and alkylated I-2 were isolated, denatured, reduced, and refolded. Fig. 4*A* and Table I shows that the initial rates for formation of the Cys^{4}–Cys^{89} disulfide for reduced rhSCF C43,138A analog and wild type rhSCF are similar. Their overall folding rates over 20 h are also similar. This observation appears to indicate that I-1 may directly fold into N in reduced rhSCF. Fig. 4*B* compares the folding kinetics between alkylated I-2 and reduced rhSCF after denaturation and reduction. The initial oxidation rate for Cys^{43}–Cys^{138} from the reduced form of IAA-trapped I-2 was also similar to that of the formation of the same disulfide in the wild type protein. However, the overall folding rate for formation of I-2 in the wild type rhSCF is significantly slower.

Physicochemical Analyses—Fig. 5*A* shows the far UV CD spectra of various folding intermediates. Both N and rhSCF standard in 2.5 M urea have approximately identical spectra.

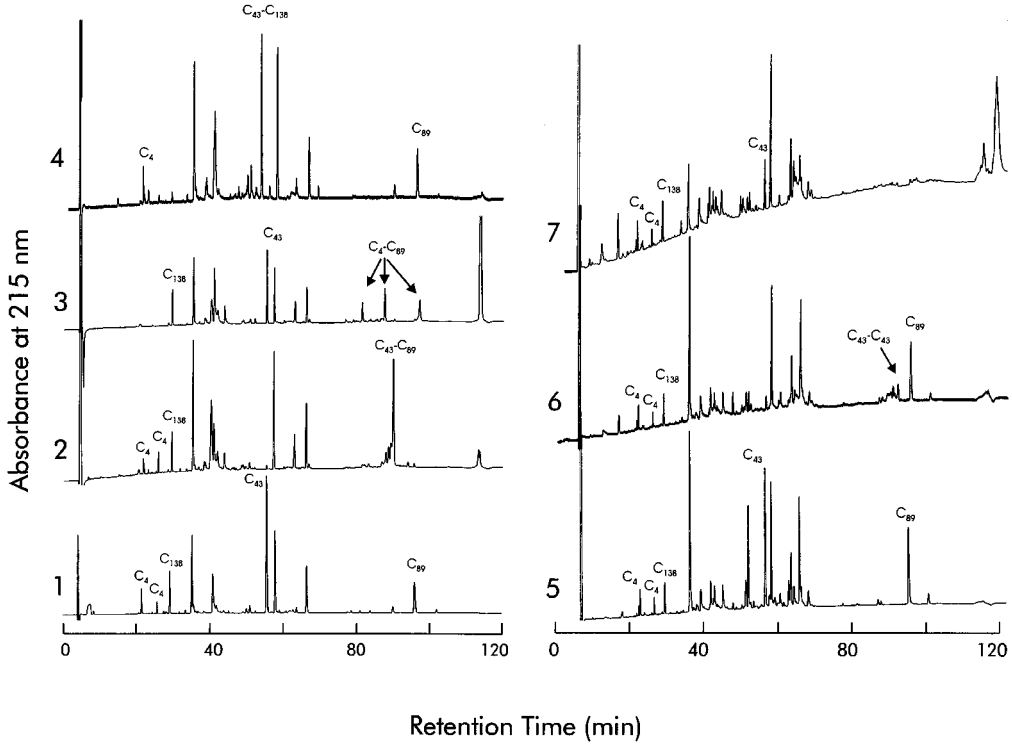

FIG. 2. **Peptide mapping of alkylated intermediates after endoproteinase Glu-C digestion.** *Chromatograms 1–4*, R, I-3, I-1, and I-2, respectively, separated on a C4 column (inside diameter = 0.46 × 25 cm). *Chromatograms 5–7*, R, I-4, and I-5, respectively, separated on a narrow bore C4 column (0.21 × 25 cm). The assigned Cys(Cm)- and disulfide-containing peptides are identified.

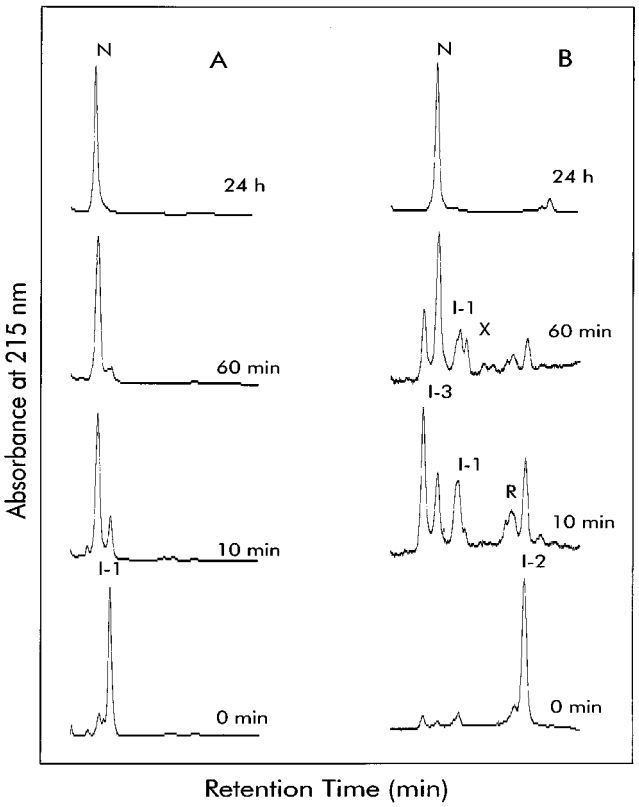

FIG. 3. **Analysis of folding and oxidation of isolated acid-trapped intermediates by RP-HPLC.** *Panels A* and *B*, folding of I-1 and I-2, respectively. Generation of fully fully folded (*N*), reduced (*R*), and intermediates are identified. *X* is not characterized.

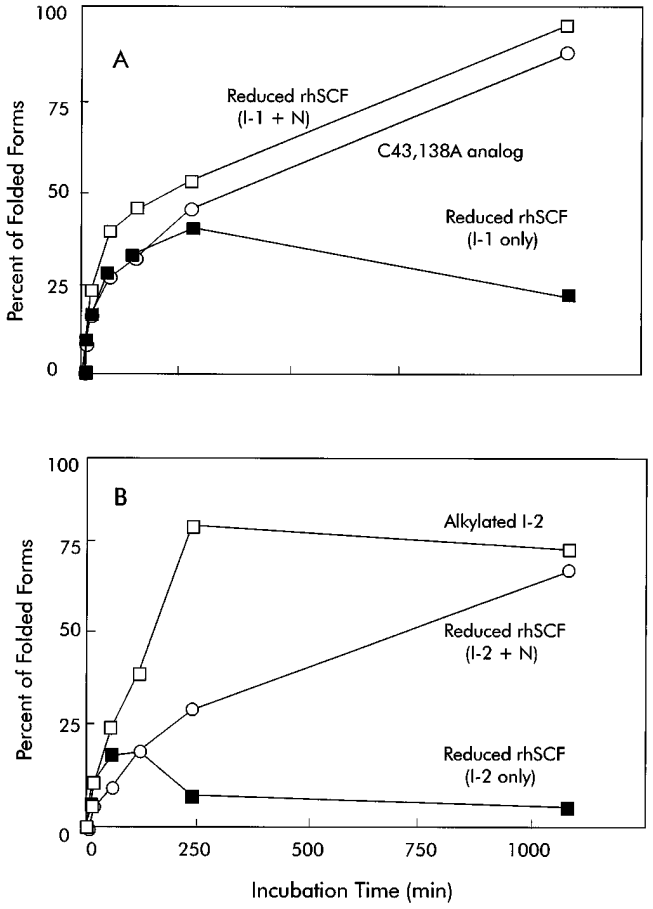

FIG. 4. **Refolding kinetics for rhSCF modified forms.** *Panel A*, percent regeneration represents the formation of Cys^4–Cys^{89} from folding of rhSCF C43,138A analog (○——○) which is compared to the sum of the formation for both I-1 and N (□——□) as well as I-1 alone (■——■) from reduced rhSCF. *Panel B*, percent regeneration represents formation of Cys^{43}–Cys^{138} bond (□——□) from refolding of iodoacetate-modified I-2, which is compared to the formation of both I-2 (■——■) and the sum of both I-2 and N (○——○) from R are also indicated.

Both samples contain about 45% α-helix, consistent with results observed in the absence of urea (17). The reduced rhSCF has a spectrum reflecting primarily a random-coiled structure that contains only 3–5% α-helix. Secondary structures of I-1 and I-2 are similar, and both contain approximately 18% α-helix. I-3 appears to have very little structure, containing only trace amounts of α-helix (<5%).

Fig. 5*B* shows the near UV CD spectra. N also has a spectrum identical in shape to that of standard rhSCF, with a slight decrease in magnitude. The spectrum of R indicates that the reduced protein is structurally disordered. I-1 and I-2 have identical spectra that are also similar in shape to that of N, but with signals reduced at levels between N and R. Both of the single disulfide intermediates appear to partially display a native-like tertiary structure.

The changes in ellipticity at 222 nm *versus* temperature for various rhSCF forms are depicted in Fig. 5*C*. Increase of temperature causes loss of α-helical structure as a result of thermal denaturation. Table II lists the transition midpoint of thermal denaturation of various rhSCF intermediates and analog. SCF standard and fully refolded rhSCF (N) are very stable molecules, with identical transition temperature of 60 °C. I-1 and C43,138A analog with a temperature of 48 °C are of intermediate stability. I-2 with a temperature of 36 °C is the least stable, consistent with increased flexibility as indicated by its fluorescence spectrum (described below).

The fluorescence spectra of rhSCF species are shown in Fig. 5*D*. The spectra for both refolded rhSCF and rhSCF standard consist of a single peak at 322 nm, consistent with the single Trp (at position 44) being in a hydrophobic environment. The FWHM is about 50 nm. As is usually the case for Trp-containing proteins, there is no Tyr signal, indicating energy transfer is occurring between Tyr and Trp. R has a single peak at 350 nm with a wider FWHM of 56 nm, typical of a Trp spectrum in solution; and its Tyr fluorescence is not apparent. I-1 has a spectrum slightly red shifted to 329 nm, indicating that the Trp is more exposed to the solvent than N. However, Trp in I-1 appears to be in a native-like environment. I-2 has a fluorescence maximum at 333 nm with a FWHM of 57 nm, indicating that the Trp in this molecule is slightly more exposed to solvent than I-1. Like R, the spectrum of I-3 has a maximum at 350 nm, apparently due to Trp being fully exposed. I-3 also contains a clear shoulder representing the presence of substantial Tyr fluorescence, indicating that the Tyr is no longer transferring energy to the Trp. This suggests that the distance between the Tyr and Trp has increased. C43,138A analog lacking the Cys^{43}–Cys^{138} disulfide bond, exhibits both CD and fluorescence spectra identical to those for the rhSCF standard. Its thermostability, however, is similar to I-1 (see above).

Hydrodynamic properties of several intermediates and analog were analyzed by size exclusion chromatography in conjunction with light-scattering detection. As expected, the determined molecular weights for rhSCF standard, N and intermediates are all around 33,000–35,000, indicative of being a dimer in solution (Table II). Fig. 6*A* illustrates a typical chromatographic separation profile for I-2 detected by UV absorption, refractive index, and light scattering. Fig. 6*B* shows the calibration plot for the molecular weight determination of

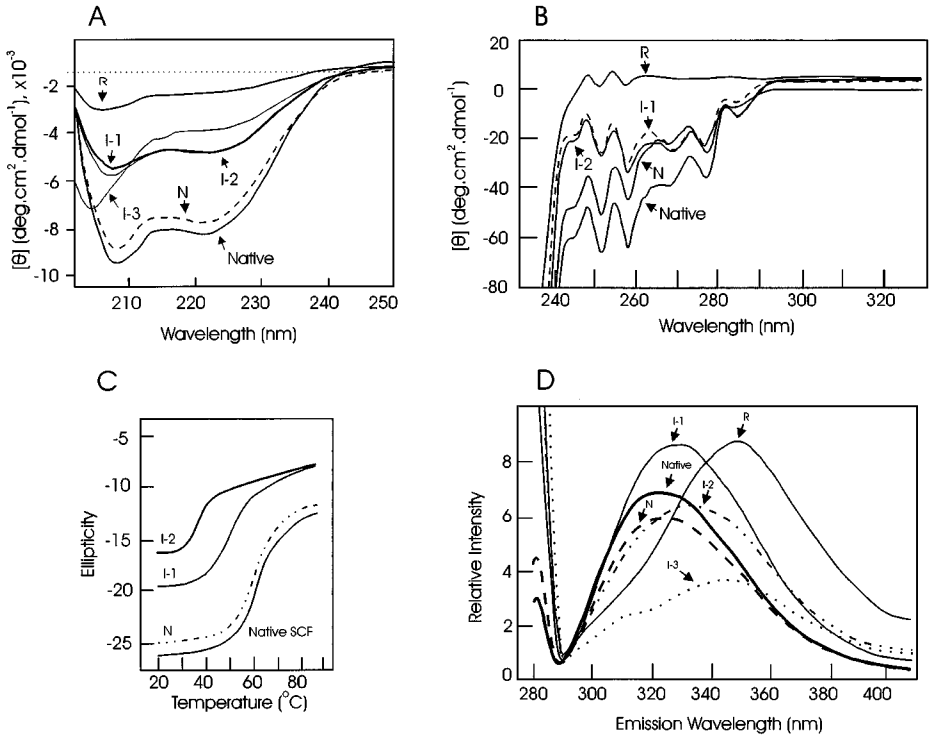

FIG. 5. **Spectroscopic properties of rhSCF and intermediates.** Far UV CD spectra of rhSCF species (*panel A*), near UV CD spectra of rhSCF species (*panel B*), temperature-dependent changes in ellipticity at 222 nm for rhSCF species (*panel C*), and fluorescence spectra of rhSCF species (*panel D*).

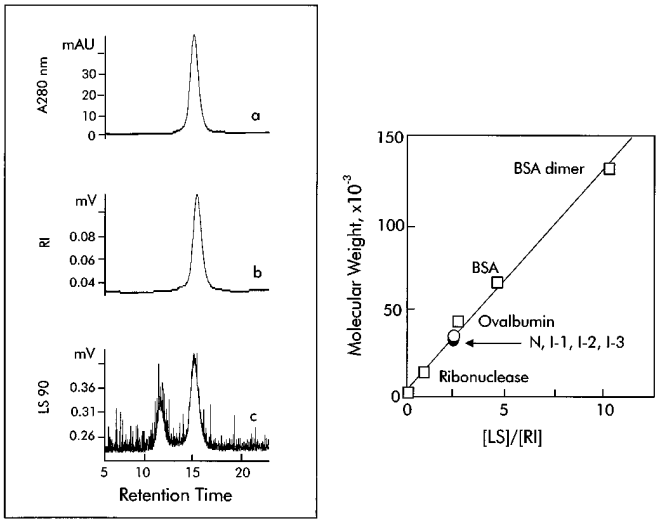

FIG. 6. **Gel filtration of rhSCF and intermediates.** *Left panel*, elution profiles of I-2 detected by UV absorption at 280 nm, refractive index [*RI*], and the 90° light scattering [*LS*] (*chromatograms a–c*, respectively). The column was calibrated with standard proteins as described in the *right panel*. The main peak eluting at 15 min was detected all methods and identified to be the dimeric SCF. The early eluting peak in chromatogram c, possibly the aggregated tetrameric I-2, can only be detected by light scattering. *Right panel*, molecular weight determination of rhSCF forms. [LS]/[RI] values were plotted *versus* molecular weight of standard proteins as described under "Experimental Procedures." *BSA*, bovine serum albumin.

various species. N was determined to have a molecular weight close to that of the rhSCF standard. IAA-trapped I-1, I-3, and I-2 are also dimeric molecules. With the exception of standard rhSCF and N, every isolated intermediate and R contain small amounts of high molecular weight aggregates (only detectable by light scattering; see Fig. 6*A* for I-2), indicating that intermediates and R are less stable in solution (in an order of N $\ll$ I-1 $<$ I-2, I-3 $<$ R).

Biological Activity—Table III summarizes the data obtained from the cell proliferation and radioreceptor binding assays. N, as expected, elicits full mitogenic activity on UT-7 assay but is about half active in radioreceptor assay. The lower activity in radioreceptor assay may be due to loss of activity during sample handling in HPLC step as rhSCF standard proceeded through an HPLC step also yielded a reduced affinity to the receptor. In contrast to N, iodoacetate-modified R, I-2, and I-3 show no or very low rhSCF biological activity and receptor binding affinity. Iodoacetate-modified I-1 and the C43,138A analog (see Ref. 19 also) retain only some residual biological activity in the UT-7 assay, but display a significant amount of binding to the kit receptor (15).

TABLE III
Biological activities of rhSCF and iodoacetate-modified folded forms analyzed by UT-7 and radioreceptor assays

In general, statistical error for both UT-7 assay and radio receptor binding analysis is approximately 10% and is not listed.

rhSCF Forms	Biological activity[a]	Receptor binding
	%	
Standard rhSCF	100	100
Standard rhSCF after HPLC	100	50
N	100	49
R	0	0
I-1	5	20
I-2	<1	0
I-3	0	0
C43,138A analog	10	50[b]

[a] Determined by the ratio of concentration (ng/ml) of rhSCF standard required for half-maximal mitogenic activity to UT-7 megakaryocytic cells to that of an rhSCF standard.
[b] Also see Ref. 16.

DISCUSSION

In vitro folding of rhSCF proceeds via a number of intermediate forms containing native and non-native disulfide bonds. There are five partially oxidized intermediates together with reduced rhSCF. At the completion, the folding mixture contained a major folded and oxidized form N along with small fractions of I-1 and P. I-1, I-2, and I-3 appear to reach steady state equilibrium during folding (Fig. 1), implying that they are important intermediates in the rhSCF folding pathway. From

refolding studies of the isolated intermediates and analog (Figs. 3 and 4), it appears that I-1 is an essential intermediate directly folding into N, while the majority of I-2 needs to reconvert to I-1 via I-3 by S–S rearrangement. By estimation of equilibrium concentrations detected by HPLC (Table I), I-1 is the most stable intermediate while I-2 and I-3 are less stable.

Several observations indicated that the conformation of I-1 is mostly native-like. For example, I-1 is the only intermediate still retaining residual biological activity (5%) (Table III). Upon SV-8 proteolytic digestion, an appreciable amount of I-1 remains as large, undigested fragments, whereas other intermediates can be easily digested (Fig. 2). By fluorescence spectroscopic analysis, Trp^{44} in I-1 or I-2 appears to be in a hydrophobic environment similar to that found for rhSCF standard (17). By partial reduction of native rhSCF, I-2 is the major species accumulated; and therefore stability of Cys^{43}–Cys^{138} bond is higher than Cys^{4}–Cys^{89} bond (see Ref. 26 for details). Taken together with the single Trp being in a hydrophobic core structure, it is highly possible that Cys^{43}–Cys^{138} is stabilized by being in this environment. This hydrophobic environment should also stabilize I-1 by sequestering the free Cys^{43} and Cys^{138} away from the aqueous surrounding needed for the formation of Cys^{43}–Cys^{138} bond. As a result, I-1 appears to be stable and remains as the major intermediate form.

Our data indicated that the disulfide-mispaired I-3 may serve as an intermediate for the conversion between I-1 and I-2. This is supported by refolding of isolated I-2 which converts to I-1 through I-3 by disulfide rearrangement (Figs. 3*B* and 4). Formation of Cys^{4}–Cys^{89} disulfide bond does not necessarily require the rearrangement since rhSCF C43,138A analog, a homolog of I-1, forms Cys^{4}–Cys^{89} bond at the same rate as that in the wild type molecule. However, during folding of the reduced rhSCF, disulfide rearrangement between I-1 and I-2 via I-3 actually decreased the folding rate for I-2 (Fig. 4*B*). The folding of isolated I-2 produced I-1 and I-3 before N, indicating that the formation of N from I-2 is preferentially directed to I-1 via disulfide rearrangement (Fig. 3*B*). Since a small fraction of N was detected in addition to I-1 and I-3 at the early folding stage of isolated I-2, direct folding of I-2 into N may also occur at a slow rate.

The generation of intermolecular disulfide bonds found in I-4 and I-5 is interesting. They are relatively unstable as they only transiently exist at an early stage of folding. In the refolding of I-4 and I-5, it appears that both intermediates refold back into R which then fold into N via I-1, I-2, and I-3. Since dimerization of either I-4 or I-5 uses the same cysteine residue (Cys^{43} or Cys^{89}, respectively), their structures may be different from any of the isolated intermediates as well as product P found at the end of folding (see subsequent paper, ref. 26). These observations suggest that I-4 and I-5 may not be as important as other intermediates in rhSCF folding. In summary, properties of all characterized intermediates appear to indicate that I-1, I-2, and I-3 are on-pathway folding intermediates important in the folding of rhSCF, while I-4 and I-5 may be off-pathway intermediates.

In comparison with native rhSCF, each trapped intermediate and R appears to be unstable in solution as they form a large amount of high molecular weight aggregates detectable in gel permeation chromatography using light-scattering detection. This observation implies that rhSCF disulfide bonds and/or the fully folded conformation play a role in the stability of rhSCF molecule. Characterization of partially oxidized intermediates seems to indicate that each of the two rhSCF disulfide bonds may contribute equally in maintaining proper tertiary structural folding of rhSCF.

All rhSCF folding species exist as noncovalently linked dimers (Table III). We observe that the completely reduced rhSCF (R) is also dimeric. However, the completely denatured and reduced R requires 1–2 h to become dimeric in oxygen-free folding conditions, and R prepared from partial reduction experiments in the absence of 6 M GdnHCl is present in dimeric form as well.[2] Refolding rate of this reduced rhSCF preparation is similar to that for the reduced rhSCF prepared in 6 M GdnHCl. Therefore, the formation of dimer may occur at an earlier folding time preceding oxidation of disulfide bonds and may not affect the overall folding.

To elicit its full biological function, rhSCF has to be fully oxidized and folded. None of the isolated intermediates and analog show significant biological activity, indicating that they do not have native rhSCF conformations. Cys^{4}–Cys^{89} and Cys^{43}–Cys^{138} are essential to maintain a proper structural folding of rhSCF for binding to its receptor, c-*kit*, and exerting full biological functions. Although the completely folded and purified C43,138A analog has a gross conformation indistinguishable from the wild type rhSCF and displays significant receptor binding activity, this molecule is biologically less active (Table III).

In summary, a putative *in vitro* folding pathway was postulated. I-1 is identified as an important and productive intermediate for formation of fully oxidized and folded rhSCF. I-2 seems to favor rearrangment through I-3 to I-1 and N. I-2 may also directly fold into N at a much slower rate. Our preliminary data indicated that the folding of rhSCF can be affected by protein disulfide isomerase.[2] It would be interesting to further study the effect of the isomerase or chaperone proteins (2, 22) in the catalysis of rhSCF folding. I-4 and I-5, which are disulfide-linked dimers, are in equilibrium with other intermediates including R. These off-pathway intermediates are kinetically identifiable and play less important roles in the *in vitro* folding of rhSCF. The observation that R, I-1, I-2, and I-3 exist as dimeric molecules appears to indicate that dimer formation together with a small extent of secondary and tertiary structural folding occur at a very earlier folding stage.

Acknowledgments—We are indebted to Analytical Resource Group at Amgen Inc. in performing SCF biological assays, to Dr. Michael Mann in designing rhSCF analog, and to Joan Bennett for help in typing the manuscript.

REFERENCES

1. Creighton, T. E. (1990) *Biochem. J.* **270,** 1–16
2. Creighton, T. E., Bagley, C. J., Cooper, L., Darby, N. J., Freedman, R. B., Kemmink, J., and Sheikh, A. (1993) *J. Mol. Biol.* **232,** 1176–1196
3. Chang, J.-Y. (1993) *J. Biol. Chem.* **268,** 4043–4049
4. Mui, P. W., Konishi, Y., and Scheraga, H. A. (1985) *Biochemistry* **24,** 4481–4489
5. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol.* **179,** 497–526
6. Weissman, J. S., and Kim, P. S. (1991) *Science* **253,** 1386–1393
7. Lu, H. S., Clogston, C. L., Narhi, L. O., Merewether, L. A., Pearl, W. R and Boone, T. C. (1992) *J. Biol. Chem.* **267,** 8770–8777
8. Creighton, T. E. (1991) *Science* **156,** 111–112
9. Zhang, J.-X., and Goldenberg, D. P. (1993) *Biochemistry* **32,** 14075–14081
10. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem.* **59,** 631–660
11. Zsebo, K. M., Wypych, J., McNiece, I. K., Lu, H. S., Smith, K. A., Karkare, S. B., Sachdev, R. K., Yuschenkoff, V. N., Birkett, N. C., Williams, L. R., Satyagal, V. N., Tung, W., Bosselman, R. A., Mendiaz, E. A., and Langley, K. E. (1990) *Cell* **63,** 195–201
12. Martin, F. H., Suggs, S. V., Langley, K. E., Lu, H. S., Ting, J., Okino, K. H., Morris, C. F., McNiece, I. K., Jacobsen, F. W., Mendiaz, E. A., Birkett, N. C., Smith, K. A., Johnson, M. J., Parker, V. P., Flores, J. C., Patel, A. C., Fisher, E. F., Erjavec, H. O., Herrera, C. J., Wypych, J., Sachdev, R. K., Pope, J. A., Leslie, I., Wen, D., Lin, C. W., Cupples, R. L., and Zsebo, K. M. (1990) *Cell* **63,** 203–211
13. Lu, H. S., Clogston, C. L., Wypych, J., Fausset, P. R., Lauren, S., Mendiaz, E. A., Zsebo, K. M., and Langley, K. E. (1991) *J. Biol. Chem.* **266,** 8102–8107
14. Lu, H. S., Clogston, C. L., Wypych, J., Parker, V., Lee, T. D., Swiderek, K., Baltera, R. F., Patel, A. C., Chang, D. C., Brankow, D. W., Liu, X.-D., Ogden, S. G., Karkare, S. B., Hu, S. S., Zsebo, K. M., and Langley, K. E. (1992) *Arch. Biochem. Biophys.* **298,** 150–158

[2] M. D. Jones, L. O. Narhi, W.-C. Chang, and H. S. Lu, unpublished data.

15. Langley, K. E., Mendiaz, E. A., Liu, N., Narhi, L. O., Zeni, L., Perseghian, C., Clogston, C. L., Leslie, I., Pope, J. A., Lu, H. S., Zsebo, K. M., and Boone, T. C. (1994) *Arch. Biochem. Biophys.* **311,** 55–61
16. Langley, K. E., Wypych, J., Mendiaz, E. A., Clogston, C. L., Parker, V. P., Farrar, D. H., Brothers, M. O., Satygal, V. N., Leslie, I., Birkett, N. C., Smith, K. A., Baltera, R. J., Jr., Lyons, D. E., Hogan, J. M., Crandall, C., Boone, T. C., Pope, J. A., Karkare, S. B., Zsebo, K. M., Sachdev, R. K., and Lu, H. S. (1992) *Arch. Biochem. Biophys.* **295,** 21–28
17. Arakawa, T., Yphantis, D. A., Lary, J. M., Narhi, L. O., Lu, H. S., Prestrelski, S. J., Clogston, C. L., Zsebo, K. M., Mendiaz, E. A., Wypych, J., and Langley, K. E. (1991) *J. Biol. Chem.* **266,** 18942–18948
18. Lu, H. S., Chang, W. C., Mendiaz, E. A., Mann, M. B., Langley, K. E., and Hsu, Y.-R. (1995) *Biochem. J.* **305,** 563–568
19. Langley, K. E., Mendiaz, E. A., Liu, N., Narhi, L. O., Zeni, L., Perseghian, C., Clogston, C. L., Leslie, I., Pope, J. A., Lu, H. S., Zsebo, K. M., and Boone, T. C. (1994) *Arch. Biochem. Biophys.* **311,** 55–61
20. Jones, M. D., Merewether, L. A., Clogston, C. L., and Lu, H. S. (1994) *Anal. Biochem.* **216,** 135–146
21. Greenfield, M., and Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4116
22. Smith, K. A., and Zsebo, K. M. (1992) in *Current Protocols in Immunology* (Coligan, J. E., Kruisbeek, A. M., Margulies, D. H., Shevach, E. M., and Strober, W., eds.) pp 6.17.1–6.17.11, John Wiley & Sons, New York
23. Laemmli, U. K. (1970) *Nature* **227,** 680–685
24. Takagi, T. (1990) *J. Chromatogr.* **280,** 409–416
25. Kato, A., Kameyama, K., and Takagi, T. (1992) *Biochim. Biophys. Acta* **1159,** 22–28
26. Lu, H. S., Jones, M. D., Shieh, J.-H., Mendiaz, E., Feng, D., Watler, P., Narhi, L. O., and Langley, K. E. (1996) *J. Biol. Chem.* **271,** 11309–11316