Journal of Protein Chemistry, Vol. 12, No. 5, 1993

Cysteine 17 of Recombinant Human Granulocyte-Colony Stimulating Factor is Partially Solvent-Exposed

Tsutomu Arakawa,[1,2] Steven J. Prestrelski, [1] Linda O. Narhi, [1] Thomas C. Boone, [1] and William C. Kenney[1]

Received March 23, 1993

Oh-eda *et al.* have shown instability of granulocyte-colony stimulating factor (G-CSF) upon storage above *p*H 7.0 [*J. Biol. Chem.* (1990) **265**, 11,432–11,435]. To clarify the mechanism of this instability, the accessibility of a free cysteinyl residue at position 17 for disulfide exchange reaction was examined using a sulfhydryl reagent. The results show that the cysteine is partially solvent-exposed in both glycosylated and nonglycosylated forms, suggesting that the exposure of the cysteine plays a critical role in the instability of the protein. This is supported by the facts that at low *p*H where the cysteine is protonated, both proteins have much greater stability and that a Cys17 → Ser analog is extremely stable at neutral *p*H and 37°C. It was observed that the rate of sulfhydryl titration is slower for the glycosylated form than for the nonglycosylated form, suggesting that the cysteine residue is less solvent-exposed for the former protein or that the pK_a is somewhat more basic. In either case, the carbohydrate appears to affect the reactivity of the sulfhydryl group through steric hindrance or alteration in local conformation. Both the glycosylated and nonglycosylated proteins showed essentially identical conformation as determined by circular dichroism, fluorescence, and infrared spectroscopy. Unfolding of these two proteins, induced either by guanidine hydrochloride or by *p*H, showed an identical course, indicating comparable conformational stability. Contribution of conformational changes to the observed instability at higher *p*H is unlikely, since little difference in fluorescence spectrum occurs between *p*H 6.0 and 8.0. Based on these observations, G-CSF, whether glycosylated or not, should not be stored above pH 7.0 in solution. On the other hand, G-CSF is extremely stable in acidic solution as expected from the proposed mechanism.

KEY WORDS: Circular dichroism; FTIR; disulfide exchange; G-CSF; sulfhydryl titration.

1. INTRODUCTION

Granulocyte-colony stimulating factor (G-CSF[3]) has been identified as a specific growth factor for the granulocyte lineage (Metcalf, 1984, 1985; Kitagawa *et al.*, 1987; You *et al.*, 1987). Natural human G-CSF has one free cystein at residue 17 and two intramolecular disulfide bonds (Lu *et al.* 1989), and is glycosylated at Thr-133 (Oh-eda *et al.*, 1988, 1990; Kubota *et al.*, 1990). Recombinant G-CSF has been obtained using *E. coli* or CHO expression systems, and the corresponding protein purified and extensively characterized (Lu *et al.*, 1989; Souza *et al.*, 1986). The average carbohydrate content of the CHO-derived protein was estimated to be 8% by densimetry (Kolvenbach *et al.*, 1991). Both forms of the protein show similar secondary structure, with a high α-helical content (Lu *et al.*, 1989; Oh-eda *et al.*, 1990; Narhi *et al.*, 1991*a*). We have shown that the *E. coli*-derived protein undergoes a complex, reversible denaturation when the protein structure is perturbed by *p*H (Narhi *et al.*, 1991*a*).

[1] Amgen, Inc., Amgen Center, Thousand Oaks, California 91320.
[2] To whom all correspondence should be addressed.
[3] Abbreviations used: G-CSF, granulocyte-colony stimulating factor; *E. coli, Escherichia coli;* FTIR, Fourier transform infrared; CD, circular dichroism; DTGS, deuterated triglycine sulfate; GdnHCl, guanidine hydrochloride; CHO, Chinese hamster ovary; DTNB, 5,5′-dithiobis-(2-nitrobenzoic acid).

0277-8033/93/1000-0525$07.00/0

Formulation is extremely important for developing a protein pharmaceutical since many proteins are marginally stable in solution (Pace, 1970; Privalov and Khechinashvili, 1974). Although we find no stability problem of G-CSF in acidic solution, Oh-eda *et al.* (1990) have observed instability of glycosylated and deglycosylated forms of G-CSF above *p*H 7.0 and at elevated temperatures. These observations suggest involvement of either a free sulfhydryl group or of conformational change as a source of the observed instability above *p*H 7.0. Therefore, we examined both the state of sulfhydryl group using a sulfhydryl titration and the protein conformation using spectroscopic methods. It has been previously observed that glycosylated erythropoitein has greater conformational stability than its nonglycosylated counterpart when denatured by *p*H, heat and GdnHCl (Narhi *et al.*, 1991*b*). Therefore, we compared the two protein forms with regard to the above properties, as reported in this paper, and, unlike for erythropoietin, found that the conformations of glycosylated and unglycosylated G-CSF are comparable.

2. MATERIALS AND METHODS

2.1. Materials

CHO-derived G-CSF was purified as described previously (Kolvenbach *et al.*, 1991) and stored frozen before use. *E. coli*-derived G-CSF and the Cys17→Ser analog were purified from inclusion bodies by previously described procedures (Zsebo *et al.*, 1986) and stored at 4°C and low *p*H. The CHO and *E. coli*-derived G-CSF as well as the Cys17→Ser analog all show equivalent activity and no apparent instability under the conditions of the assay.

2.2. Methods

2.2.1. Infrared Spectroscopy

Both the *E. Coli*- and CHO-derived G-CSF samples were prepared for infrared (IR) spectroscopy by dialyzing against a 0.32 mM HCl solution and lyophilizing the dialyzed protein solutions. The lyophilized proteins were then dissolved in a 0.32 mM DCl, pD = 3.5 in D_2O (Sigma Chamical Co., 99.99% isotopic purity) or into a 20 mM imidazole buffer, pD = 7.5. The protein concentration was 1.7 mM. IR spectra were collected at room temperature using a Nicolet 800 FTIR system equipped with a Globar source, a Ge-coated KBr beamsplitter, and a DTGS detector. For each spectrum 512 double-sided interferograms were collected at 2 cm^{-1} resolution and Fourier transformed after application of a Happ-Genzel apodization function. Spectral features arising from residual H_2O vapor in the light path and from buffer components were subtracted using the Nicolet FTIR software. Second derivative spectra were calculated using the method of Susi and Byler (1983). Fourier self-deconvolution was also performed using the Nicolet software, which is based on the the method of Kauppinen *et al.* (1981). Curve-fitting was performed using the program PeakFit (Jandel Scientific Co.).

2.2.2. Denaturation

The GdnHCl stability of the G-CSF species were determined in 20 mM citrate, 100 mM NaCl, *p*H 7.1, at ambient temperature. Stock solutions of G-CSF were diluted into buffer containing the desired GdnHCl concentration to a final protein concentration of 1.5 mg/ml. The ellipticity was determined using a Jasco J-500C spectropolarimeter controlled by a Samtron computer. Cuvettes with a pathlength of 1 cm and 0.02 cm were used for the near and far UV CD spectra, respectively. After incubating in GdnHCl for 5 min, the far UV CD spectra from 260 nm to 205 nm were recorded. No dependence on incubation time was observed. For some samples, the near UV CD spectrum from 340 to 240 nm were also recorded. Stability was determined from the change in ellipticity at 221 nm.

2.2.3. pH Titration

CHO- and *E. Coli*-derived G-CSF were dialyzed into 0.3 mM HCl and concentrated to 7 mg/ml. Aliquots of these stock solutions were diluted into 20 mM citrate, 100 mM NaCl, at the appropriate *p*H to an absorbance at 280 nm of 0.2. The fluorescence spectra were determined on an SLM-Aminco SPF-500C spectrofluorimeter at 25°C, exciting at 280 nm, using slit widths of 5 nm, and a cuvette with a 0.5 cm pathlength.

2.2.4. Thiol Titration

Titration of the sulfhydryl group was carried out using DTNB (Sigma), essentially according to the method of Habeeb (1972). *E. coli*- and CHO-derived G-CSF at 1.1 mg/ml in 20 mM sodium phosphate, 0.05% EDTA, *p*H 7.5 were mixed with the DTNB

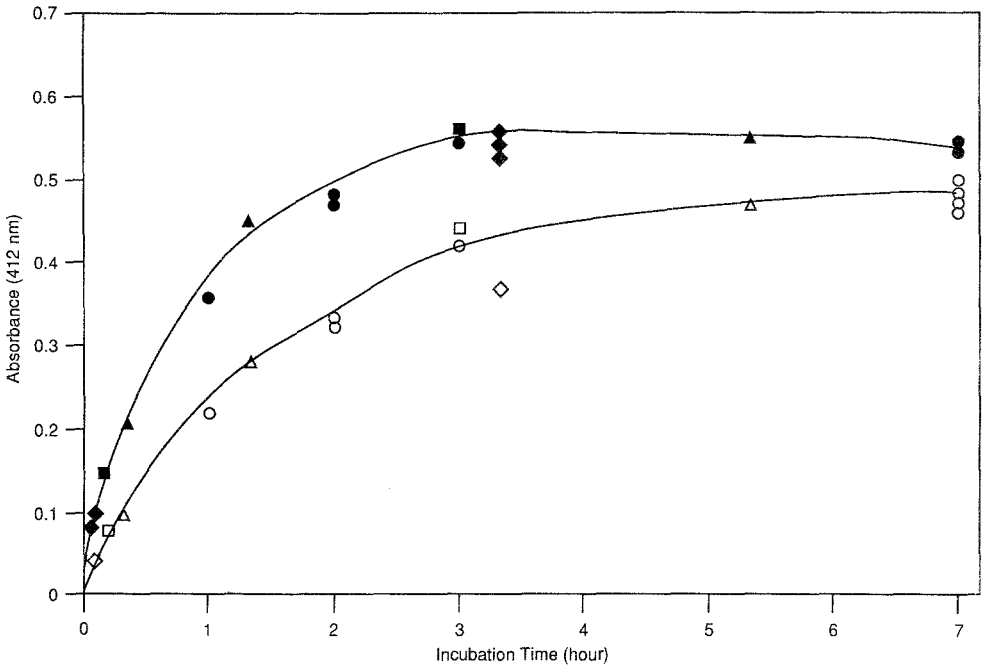

Fig. 1. Thiol titration of *E. coli*- and CHO-derived G-CSF at *p*H 7.5 and room temperature. Solid symbols, *E. coli*-G-CSF; open symbol, CHO-G-CSF. ○, no incubation at *p*H 7.5 and room temperature prior to addition of DTNB; △, 100 min incubation; □, 4 hr incubation; ◇, 8 hr incubation.

reagent solution and incubated at room temperature. The time course of the reaction was followed by the increase in absorbance at 412 nm. Sulfhydryl titration experiments were also carried out in the presence of 4 M GdnHCl. In this case, the absorbance was measured 10 and 20 min after mixing with DTNB. The incubation time had little effect on the absorbance.

3. RESULTS

3.1. Thiol Titration

Oh-eda *et al.* (1990) have shown that G-CSF, regardless of glycosylation, becomes more unstable as the *p*H is increased above 7.0. There are two possibilities for the observed instability: conformational change and reactivity of free sulfhydryl group. G-CSF has a free cysteine at position 17.

Therefore, we examined first the accessibility of the free thiol in the *E. coli*-derived G-CSF using DTNB titration. The result is shown in Fig. 1 (closed symbol). To demonstrate that the protein is stable during incubation with DTNB, some of the samples were incubated at room temperature before mixing with DTNB. It is evident that the results are identical, independent of the length of pre-incubation at room temperature. The results show that the DTNB titration is a slow process, the half-time being about 50 min. A solvent-exposed cysteine should react with the DTNB within several minutes (Habeeb, 1972). This result demonstrates that the free thiol is accessible, but not fully solvent-exposed. Since we observed little change in fluorescence spectra between *p*H 6.0 and 8.0, as described below, the observed instability above *p*H 7.0 could be ascribed to the solvent-exposed free thiol, but not to a conformation change. This hypothesis was confirmed by the observation that mutation of Cys 17 to Ser in the *E. coli*-derived protein abolishes the instability of the protein at *p*H 7.5 (data not shown). We also examined the state of the free thiol in the CHO cell-derived protein, as shown in Fig. 1 (open symbol). The result indicates that the free thiol in CHO cell-derived protein is also reactive with DTNB, the half time of reaction being about 2 hr. Thus, the free thiol appears somewhat less solvent-exposed in the CHO cell-derived protein than in the *E. coli*-derived one, or that the pK_a of the sulfhydryl group is somewhat higher.

3.2. Conformation of CHO and *E. coli*-Derived G-CSF

In order to see if conformational differences contribute to the observed difference in the thiol reactiv-

Table I. Peak Positions (cm^{-1}) and Relative Integrated Intensities for *E. coli* and CHO-Derived G-CSF at pD = 3.5 and pD = 7.5

pD = 3.5				pD = 7.5			
E. coli		CHO		*E. coli*		CHO	
υ	A	υ	A	υ	A	υ	A
1630	.17	1630	.15	1629	.18	1629	.17
1638	.10	1638	.12	1637	.10	1637	.09
1645	.13	1645	.12	1644	.20	1644	.23
1655	.49	1654	.50	1654	.35	1654	.38
1670	.07	1670	.07	1667	.11	1667	.08
1681	.04	1681	.04	1679	.05	1678	.04

ity between the two forms of G-CSF, we examined the conformation and conformational stability of the protein. Infrared spectroscopy was used to examine quantitatively the conformation of G-CSF. Table I lists the peak positions and relative integrated intensities of the amide 1 component bands for both *E. coli*- and CHO-derived G-CSF at neutral and low *p*H, as determined by curve-fitting of the spectra after application of Fourier self-deconvolution (see Fig. 2). The amide 1 region (C = O stretching, ~1700–1620 cm^{-1}) of protein infrared spectra is highly sensitive to the backbone conformation, and when combined with band narrowing algorithms, such as derivative spectrocopy and Fourier self-deconvolution, provides extensive information concerning secondary structure. The secondary structure assignment for the individual components is made based on previous reports (Byler and Susi, 1986; Surewicz and Mantsch, 1988; Prestrelski *et al.*, 1991): 1654 cm^{-1}, α-helix; 1629 and 1680 cm^{-1}, extended strands; 1645 cm^{-1}, irregular structures; 1667–1670 cm^{-1}, reverse turns. The assignment of the component bands near 1637–1638 cm^{-1} is somewhat less straightforward. Bands near this frequency have been attributed to 3_{10}-helices (Prestrelski, 1991) as well as extended structures (Byler and Susi, 1986). A short 3_{10}-helix is observed in the recently determined crystallographic structure of G-CSF (T. Osslund, personal communication).

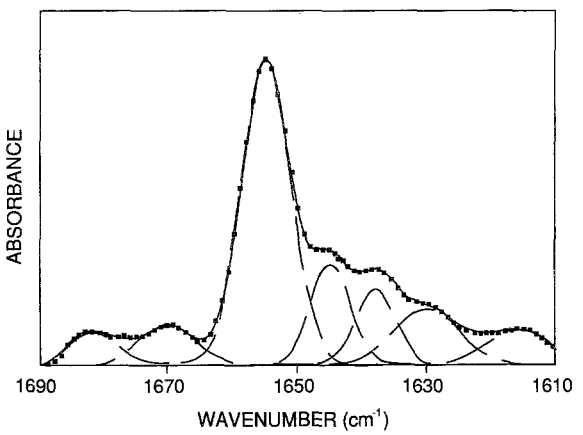

Fig. 2. Deconvoluted spectrum of *E. coli*-derived G-CSF in the amide 1 region at pD = 3.5. (Each square represents one data point.) The broken and solid lines represent individual Gaussian components and their sum, respectively, determined by an iterative curve-fitting procedure. Values of 18 cm^{-1} and 3 were used in the deconvolution algorithm for the undeconvoluted halfwidth and K value, respectively. Note: a peak is fitted below 1620 cm^{-1} which is not an amide 1 component but is included to avoid the approximation otherwise incurred by addition of a sloping baseline parameter to the analysis.

The spectra of G-CSF derived from *E. coli* and CHO at pD = 3.5 indicate that the two forms show little or no difference in their secondary structures at the resolution of this method. The predominant feature in these spectra is the intense band near 1655 cm^{-1}, which accounts for 50% of the amide 1 intensity in the spectra of both proteins. Thus, the IR spectra indicate that G-CSF is ~50% α-helical structure. This is somewhat low in comparison with the structure derived from CD spectroscopy which estimates ~75% (Lu *et al.*, 1989; Narhi *et al.*, 1991*a*). It has been shown previously, however, that in studies of highly α-helical proteins, IR spectrosocpy systematically underestimates the fraction of α-helix (Byler and Susi, 1986). It appears that such is the case here. In addition to the large amount of α-helical structure, the infrared spectra indicate that the G-CSF contains 12–13% irregular sedondary structure and 7% turns. The remainder exists as extended strands and a small amount of 3_{10}-helix ($\leq$10%).

The spectra of *E. coli*- and CHO-derived G-CSF at neutral *p*H again show little difference, but are clearly different from the spectra of these proteins at lower *p*H. The difference in the relative intensities between the *E. coli*- and CHO-derived G-CSF at pD = 7.5, while still small, are somewhat greater than that observed at the lower pD. The most significant difference between the structures at neutral *p*H and those at low *p*H is the loss of α-helical structure

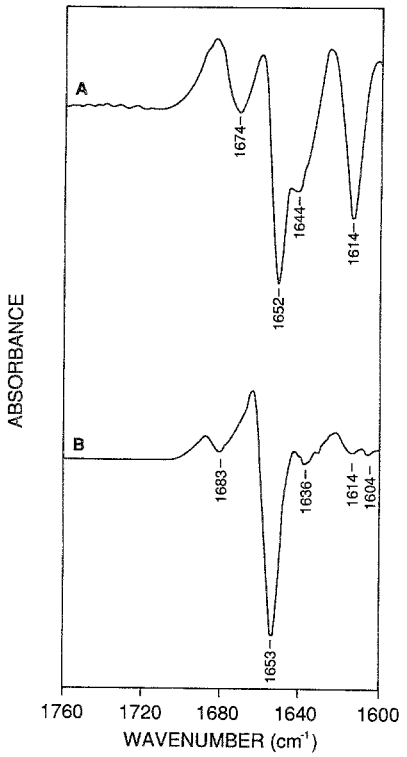

Fig. 3. Second derivative spectra of CHO-derived G-CSF in the amide I region after (A) 6 days and (B) 1 hr of incubation at room temperature and low *p*H.

at neutral *p*H. This is consistent with previous reports based on CD (Lu *et al.*, 1989; Narhi *et al.*, 1991*a*). Much of the decrease in amide I′ intensity arising from α-helical structure appears as increased intensity in the amide I′ component near 1644 cm^{-1}. This indicates an increase in irregular or disordered structure concomitant with the observed decrease in α-helical structure. Little or no change is observed among the other amide I′ components. The CD spectra of these proteins at both *p*H 4.5 and *p*H 7.1 were also taken, and were again identical between the CHO- and *E. coli*-derived proteins (data not shown). All of these results indicate that the conformations of *E. coli* and CHO-derived G-CSF are comparable.

During the course of infrared experiments originally designed to determine the deuteration rates of the two molecules, we observed spectral changes upon prolonged incubation, indicating the presence of a denatured form of the protein under the experimental conditions of FTIR (i.e., at higher protein concentrations and in D_2O solutions). We looked more carefully at the spectral changes at low *p*H, where sulfhydryl reactions are unlikely to occur. Figure 3 shows the derivative spectra of CHO-derived G-CSF after 1 hr of incubation and after 6 days of incubation at room temperature and low *p*H. The 6 day spectrum shows a strong peak near 1614 cm^{-1}, not present in the initial 1 hr spectrum, indicative of an unfolded species (Purcell and Susi, 1984; Surewicz *et al.*, 1990). Spectra of the *E. coli*-derived G-CSF taken after 1 hr and 6 days of incubation (not shown) show essentially no differences. Thus, it appears that at lower *p*H the CHO-derived protein undergoes a slow conformatinal change and that the *E. coli*-derived G-CSF is more stable.

3.3. Conformational Stability

We have previously reported that the conformation of G-CSF reversibly changes in a *p*H-dependent fashion, as demonstrated by CD and fluorescence (Narhi *et al.*, 1991*a*). Therefore, the conformation of these molecules were compared at different *p*H. The *p*H titration of the two molecules was carried out using fluorescence emission spectra as a conformational probe as described under Methods, and the ratio of fluorescence at 308 to that at 344 nm is plotted vs. *p*H in Fig. 4. This ratio was found to be a very sensitive indicator of the *p*H-dependent conformational change, as reported previously (Narhi *et al.*, 1991*a*). Little spectral change is observed in this parameter between *p*H 6.0 and 8.0 for both forms of G-CSF, suggesting that conformational change is not

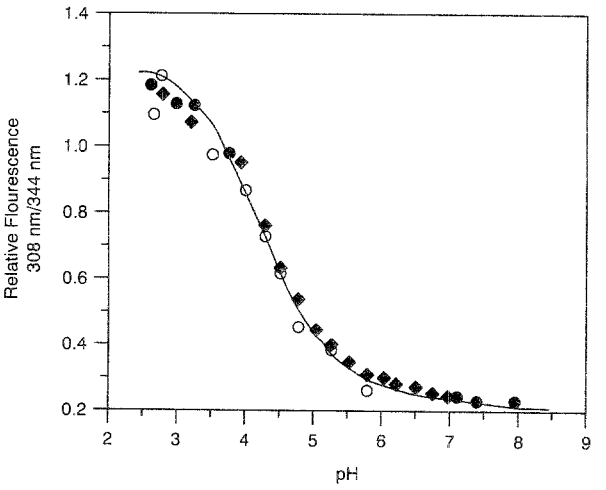

Fig. 4. Effect of *p*H on fluorescence of G-CSF species. Stock solutions of G-CSF in 0.3 mM HCl were diluted into 10 mM citrate, 100 mM NaCl at the indicated *p*H, and the fluorescence spectra recorded. The ratio of the fluorescence intensity at 308 nm to 344 nm is plotted for the *E. coli*-derived G-CSF (closed and open circles, two experiments) and CHO-derived G-CSF (closed triangles)

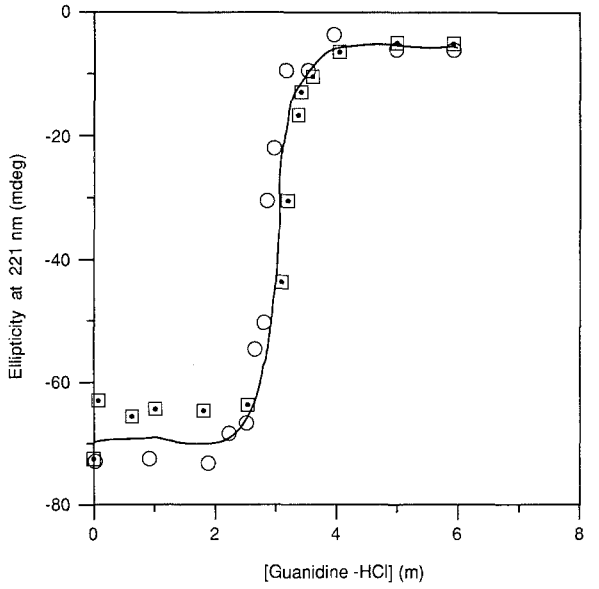

Fig. 5. Conformational stability of G-CSF species at neutral pH. Stock solutions of protein were diluted into 20 mM citrate, 100 mM NaCl, *p*H 7.1, containing the desired amount of GdnHCl, and the far UV CD spectra recorded. The ellipticity at 221 nm was determined and is plotted vs. GdnHCl concentration for *E. coli*-derived G-CSF (○) and CHO-derived G-CSF (□).

responsible for the observed instability of these proteins above *p*H 7.0. Moreover, over a broader *p*H range, these two proteins assume an identical course of conformational changes, indicating that they have comparable conformational stability.

The conformational stability of the *E. coli*- and CHO-derived G-CSF species was determined at neutral *p*H as described. When the near UV CD spectra of the CHO-derived G-CSF in different concentrations of GdnHCl are analyzed, they showed that the protein goes through several conformational changes prior to unfolding (data not shown), as was found previously for the *E. coli*-derived protein (Narhi *et al.*, 1991*a*). This makes it very difficult to determine a transition point from the near UV results, but is further evidence that the conformations of these molecules are comparable. The far UV CD spectra, however, reveal that the secondary structure of the molecule undergoes an apparent single, cooperative loss of structure. Figure 5 shows the effect of GdnHCl on the ellipticity at 221 nm for the *E. coli*-derived and CHO-derived G-CSF; identical results are obtained from following the loss of signal at 208 nm. As can be seen in the figure, these molecules have the same conformational stability, within the experimental error, to GdnHCl with a loss of half of the structure occurring at 3.1 M GdnHCl at 21°C. These results were further analyzed by determining the equilibrium constant, K = (unfolded)/(native), of unfolding along the curve, and then plotting the ln K vs. GdnHCl concentration, over the transition region of the curve, where errors in baseline extrapolation are minimized. From this curve, the ln K at 0 M GdnHCl is extrapolated to be −19.5, which corresponds to a free energy of unfolding of 11.3 kcal/mol. Thus, the sugar residues on the CHO-derived molecules do not appear to contribute to the conformational stability of the protein at neutral *p*H.

4. DISCUSSION

Development of suitable formulation is a difficult problem for protein pharmaceuticals, since most proteins are marginally stable in solution (Pace, 1970; Privalov and Khechinashvili, 1974). Recombinant G-CSF, examined at different *p*H and temperatures, is unstable at high temperature and above neutral *p*H, whether the protein is glycosylated or not (Oh-eda *et al.*, 1990), whereas the protein is much more stable at low *p*H. We find essentially no loss of activity upon prolonged storage at *p*H 4.0 and 4°C. It was revealed that the cysteine at position 17 is partially accessible to the sulfhydryl agent, suggesting that oxidation and disulfide exchange reactions due to this solvent-exposed thiol is involved in the observed instability of G-CSF. An analog of the protein in which this free cysteine has been replaced with serine confirmed the above hypothesis.

The CHO cell-derived G-CSF shows a lower degree of exposure of the free cysteine than the *E. coli*-derived molecule. Since the CHO cell- and *E. coli*-derived proteins have an apparently identical conformation at *p*H 7.0 to 8.0, the observed difference in the degree of solvent-exposure of thiol cannot be ascribed to a conformational difference. It might be possible that the carbohydrate sterically hinders the cysteine 17 from reacting with the sulfhydryl reagent or affects local conformation around cysteine 17 which cannot be detected by the techniques used in this study. The recently determined crystallographic structure of G-CSF shows that the carbohydrate would be far from cysteine 17, making its steric effect on the cysteine unlikely (T. Osslund, personal communication).

Oh-eda *et al.* (1990) used temperature changes to perturb the protein conformation. However, temperature-induced conformational changes of G-CSF is difficult to analyze, since the protein may aggregate

and even precipitate at higher temperatures. In this study, we used GdnHCl or *p*H as a protein perturbant and observed that the CHO cell- and *E. coli*-derived G-CSF have identical conformational stability.

REFERENCES

Byler, D. M., and Susi, H. (1986). *Biopolymers* **25**, 469–487.

Habeeb, A. F. S. A. (1972). *Methods Enzymol.* **25**, 457–465.

Kauppinen, J. R., Moffat, D. J., Mantsch, H. H., and Cameron, D. C. (1981). *Appl. Spectrosc.* **25**, 271–277.

Kitagawa, S., Yuo, A., Souza, L. M., Saito, M., Miura, Y., and Takaku, F. (1987). *Biochem. Biophys. Res. Commun.* **144**, 1143–1146.

Kolvenbach, C. G., Langley, K. E., Strickland, T. W., Kenney W. C., and Arakawa, T. (1991). *J. Biochem. Biophys. Meth.* **23**, 295–300.

Kubota, N., Orita, T., Hattori, K., Oh-eda, M., Ochi, N., and Yamazaki, T. (1990). *J. Biochem.* **107**, 486–492.

Lu, H.-S., Boone, T. C., Souza, L. M., and Lai, P.-H. (1989). *Arch. Biochem. Biophys.* **268**, 81–92.

Metcalf, D. (1984). *The Haemopoietic Colony Stimulating Factors.* Elsevier, Amsterdam.

Metcalf, D. (1985). *Science* **229**, 16–22.

Narhi, L. O., Kenney, W. C., and Arakawa, T. (1991*a*) *J. Protein Chem.* **10**, 359–367.

Narhi, L. O., Arakawa, T., Aoki, K. H., Elmore, R., Rohde, M. F., Boone, T., and Strickland, T. W. (1991*b*). *J. Biol. Chem.* **266**, 23,022–23,026.

Oh-eda, M., Hasegawa, M., Hattari, K., Kuboniwa, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T., and Ochi, N. (1990). *J. Biol. Chem.* **265**, 11,432–11,435.

Pace, N. (1970). *Trends Biochem. Sci.* **15**, 14–17.

Prestrelski, S. J., Byler, D. M., and Liebman, M. N. (1991). *Biochemistry* **30**, 133–143.

Prestrelski, S. J., Byler, D. M., and Thompson, M. P. (1991). *Int. J. Peptide Prot. Res.* **37**, 508–512.

Privalov, P. L., and Khechinashvili, N. N. (1974). *J. Mol. Biol.* **86**, 665–684.

Purcell, J. M., and Susi, H. (1984). *J. Biochem. Biophys. Meth.* **9**, 193–199.

Souza, L. M., Boone, T. C., Gabrilove, J., Lai, P.-H., Zsebo, K. M., Murdock, D. C., Chazin, V. R., Bruszewski, J., Lu, H., Chen, K. K., Barendt, J., Platzer, E., Moore, M. A. S., Mertelsmann, R., and Welte, K. (1986). *Science* **232**, 61–65.

Surewicz, W. R., Leddy, J. J., and Mantsch, H. H. (1990). *Biochemistry* **29**, 8106–8111.

Surewicz, W. K., and Mantsch, H. H. (1988). *Biochim. Biophys. Acta* **952**, 115–130.

Susi, H., and Byler, D. M. (1983). *Biochem. Biophys, Res. Commun.* **115**, 391–397.

You, A., Kitagawa, S., Okabe, T., Urabe, A., Komatsu, Y., Itoh, S., and Takaku, F. (1987). *Blood* **70**, 404–411.

Zsebo, K. M., Cohen, A. M., Murdock, D. C., Boone, T. C., Inoue, H., Chazin, V. R., Hines, D., and Souza, L. M. (1986). *Immunobiology* **172**, 175–184.