Conformation and Unfolding Thermodynamics of Epidermal Growth Factor and Derivatives†

Leslie A. Holladay,*,‡ C. Richard Savage, Jr.,§ Stanley Cohen, and David Puett¶

ABSTRACT: Mouse submaxillary epidermal growth factor (EGF) is a 53-residue single chain peptide hormone of known amino acid sequence which contains three disulfides, five tyrosines, and two tryptophans. Circular dichroic (CD) spectra have been obtained and resolved for EGF, several well-characterized chemical and enzymic derivatives, and related low molecular weight model compounds. Assignments have been made to most of the resolved bands; these include the peptide, aromatic, and disulfide chromophores. From a comparison of the rotational strength of the 213-nm resolved CD band in native EGF with that of standard proteins, EGF is estimated to contain about 22% β structure and no α helicity. A derivative of EGF lacking the five carboxyl-terminal residues (prepared by limited trypsin digestion) and the cyanogen bromide derivative, in which there is a single main-chain cleavage at residue 21, have spectral properties indicative of approximately 10 and 12% β structure, respectively. The near-ultraviolet CD spectra of the derivatives are similar to, albeit not identical with, that of EGF. The rotational strengths characteristic of the side-chain chromophores in EGF and these derivatives are several-fold higher than the corresponding values in low molecular weight model compounds. Thus, it appears that EGF and these modified forms contain a stable (and similar) tertiary structure. In contrast, the S-aminoethylated derivative of EGF exhibits a drastically altered CD spectrum relative to EGF indicating a different conformation(s). Equilibrium studies on the guanidinium hydrochloride (GdmCl) mediated reversible unfolding of EGF showed that the transition midpoint is quite high (i.e., 6.89 M GdmCl at 25.0 °C), thus indicating considerable stability. From these data a rough estimate of 16 kcal/mol can be made for the unfolding free energy ($\Delta G°$) of EGF in the absence of denaturant. Interestingly, EGF exhibits greater stability characteristics than several proteins two to four times its size. The cyanogen bromide derivative of EGF exhibited greatly reduced stability characteristics, e.g., the transition midpoint occurred at 4.19 M GdmCl (25.0 °C) and $\Delta G°$ was estimated to be approximately 4 kcal/mol. Thus, a single main-chain cleavage reduced the stability of EGF by about 70%. Thermal transitions of EGF and the cyanogen bromide derivative in the presence of concentrated GdmCl are characterized by a relatively high enthalpy of about 25 kcal/mol at 40 °C and a low (probably zero) heat capacity. From these thermodynamic parameters one can calculate that the large reduction in $\Delta G°$ due to scission of the single peptide bond between residues 21 and 22 can be attributed almost completely to a change in entropy; e.g., at 40 °C the apparent entropy of unfolding of EGF is 20.4 cal mol^{-1} deg^{-1} while that of the cyanogen bromide derivative is 66.4 cal mol^{-1} deg^{-1}.

Epidermal growth factor is a single-chain 53-residue polypeptide that has been isolated from the submaxillary glands of adult male mice (Cohen, 1962), and a polypeptide with similar biological activities, but with a different amino acid composition, has been isolated from human urine (Cohen and Carpenter, 1975). The amino acid sequence and disulfide positions of mouse EGF[1] are known (Savage et al., 1972, 1973) and both the hormone and a binding protein have been characterized in a number of physicochemical studies (Taylor et al., 1970, 1972, 1974a,b).

EGF stimulates the growth and differentiation of various epidermal and epithelial tissues, both in vivo and in vitro (Cohen and Elliott, 1963; Cohen, 1965; Turkington, 1969; Savage and Cohen, 1973). The hormone also binds specifically and with high affinity to fibroblasts maintained in tissue culture and stimulates the growth of these cells (Armelin, 1973; Hollenberg and Cuatrecasas, 1973; Cohen et al., 1975). A new aspect of the biological activity of EGF has recently been reported (Gregory, 1975). Mouse EGF shows remarkable structural and functional similarities to urogastrone, a gastric antisecretory polypeptide isolated from human urine (Gregory, 1975). Urogastrone may, in fact, be identical with human EGF (Cohen and Carpenter, 1975). Recent reviews on the chemistry and biological actions of EGF are available (Cohen and Taylor, 1974; Cohen and Savage, 1974; Cohen et al., 1975).

The far-ultraviolet CD spectrum of EGF has been reported by Cohen and co-workers and the results indicate the presence of β structure (Taylor et al., 1972). However, no information is available on the near-ultraviolet CD spectrum of EGF or on the conformational aspects of well-defined derivatives. In view of the fairly high cystine content of this relatively small peptide hormone, i.e., three disulfides per 53 residues, and the presence of five tyrosines, two tryptophans, and zero phenylalanines (Savage et al., 1972), a detailed analysis and comparison of

† From the Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232. *Received October 24, 1975.* A portion of this work was presented at the 59th Annual Meeting of the Federation of American Societies of Experimental Biology, Atlantic City, April 13-18 ((1975) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 659). This research was supported in large part by the National Institutes of Health (Research Grants AM-15838 and HD-00700) and the Vanderbilt University Diabetes-Endocrinology Center (AM-17026). L.A.H. received partial support from the Vanderbilt University Research Council and D.P. was a Dreyfus Teacher-Scholar Awardee during the course of these studies.

‡ Recipient of an Andrew W. Mellon Teacher-Scientist Award.

§ Recipient of a Damon Runyon Postdoctoral Fellowship (DRF-610). Present address: Department of Biochemistry, Temple University, Philadelphia, Pennsylvania 19140.

¶ Recipient of a Research Career Development Award (AM-00055).

[1] Abbreviations used are: CD, circular dichroic; CNBr-EGF, the cyanogen bromide derivative of EGF; des-(49-53)-EGF, an EGF derivative lacking the five carboxyl-terminal residues prepared by limited trypsin digestion of EGF; EGF, epidermal growth factor; GdmCl, guanidinium chloride; SAE-EGF, the S-aminoethylated derivative of EGF; uv, ultraviolet.

TABLE I: Biological Activity of EGF and Derivatives.[a]

Compd	In Vivo Bioassay[b]	Binding Affinity[d]
EGF	1.0	1.0
des-(49–53)-EGF	1.0[c]	0.05–0.10
CNBr-EGF	<0.1	[e]
SAE-EGF	0	0

[a] Much of these data have been reported elsewhere (cf. Savage and Cohen, 1972; Carpenter et al., 1975). The values for the derivatives are relative to that of EGF which is taken as unity. [b] Based on the eyelid opening bioassay described by Cohen (1962). [c] This derivative exhibits very low biological activity in an in vitro bioassay involving the stimulation of human fibroblast growth (Cohen et al., 1975). [d] Estimated by determining the amount of derivative required to inhibit binding of [^{125}I]EGF to human fibroblasts in cell culture by 50% (Carpenter et al., 1975). [e] No data are available in this system.

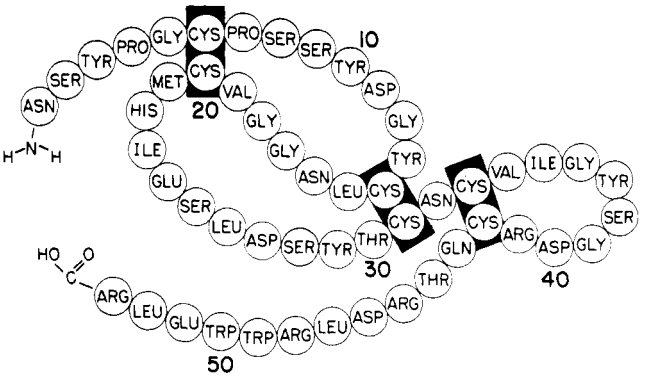

FIGURE 1: Amino acid sequence of mouse EGF (Savage et al., 1972, 1973). Reproduced by permission of the American Association of Biological Chemists, Inc.

the CD spectrum of EGF with those of derivatives and model compounds is of interest both in extending our understanding of protein CD spectroscopy and in clarifying hormone structure–function relationships.

It has been amply demonstrated that CD spectroscopy is sufficiently sensitive to study conformational differences and local environmental differences in related proteins and protein derivatives (Puett, 1972a,b; Holladay et al., 1974; Holladay and Puett, 1976a). Also, such studies are essential in order to complement advances being made in various aspects and applications of protein CD spectroscopy (Chen et al., 1972, 1974; Adler et al., 1973; Sears and Beychok, 1973; Strickland, 1974; Holladay et al., 1974; Robinson et al., 1974; Moore et al., 1974; Puett et al., 1974; Holladay and Puett, 1975a,b, 1976a,b).

Herein, the resolved CD spectrum of EGF and of three well-defined derivatives is reported and interpreted. Also, quantitative estimates are made of the (conformational) unfolding free energies, transition enthalpies, and heat capacities of EGF and the CNBr derivative. This work represents one of the few studies that have been conducted on the unfolding thermodynamics of proteins of molecular weight less than 10 000 (cf. Pace, 1975); Vincent et al. (1971) have investigated the conformational stability of the pancreatic trypsin-inhibitor protein which contains 58 residues.

Materials and Methods

Mouse EGF was isolated and purified as described earlier (Savage and Cohen, 1972). Des-(49–53)-EGF, the carboxyl-terminal pentapeptide of EGF, and SAE-EGF were prepared using the methods given elsewhere (Savage et al., 1972). EGF was treated with CNBr using conditions identical with those reported for the CNBr cleavage of [^{14}C]SAE-EGF (Savage et al., 1972). No detectable methionine remained in this derivative as shown by the amino acid analysis of a 22-h, 110 °C hydrolysate of CNBr-EGF using 4 N methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole (Liu and Chang, 1970; Moore, 1973).

The compounds L-cystine and L-cystinylbis(L-tyrosine) were from Schwarz/Mann and Fox Chemical Co., respectively. Spectrophotometric grade GdmCl was purchased from Heico, Inc.

Absorption spectra were obtained with an Hitachi Perkin-Elmer EPS-3T spectrophotometer using matched pairs of either 1.00- or 0.437-cm cells.

CD spectra were obtained using a Cary 60 spectropolarimeter with a Model 6002 CD attachment; calibration was with camphor-d_{10}-sulfonic acid. Path lengths of 0.5, 1, 5, and 10 mm and concentrations ranging from 0.1 to 1 mg/ml were used in order to obtain the CD spectrum of EGF and the various derivatives over the wavelength range of 197–310 nm. The detailed methodology for obtaining and resolving the CD spectra have been given elsewhere (Zahler et al., 1972; Puett, 1972a; Holladay et al., 1974). The mean residue ellipticity of EGF is based on an average residue molecular weight of 114.

Concentrations of EGF and derivatives were based either on the extinction coefficient of 3.09 (1 mg/ml, 1 cm) at 280 nm reported for EGF (Taylor et al., 1972), dry weight, or additive absorptivity of the side-chain chromophores (Beaven and Holiday, 1952). The agreement, using two or more methods, was good in all cases. The absorption and CD spectra of EGF, des-(49–53)-EGF, CNBr-EGF, and the carboxyl-terminal pentapeptide were determined using a 10 mM potassium phosphate buffer, pH 7.5, with 50 mM KCl. The spectrum of SAE-EGF was obtained in 10 mM acetic acid. In all cases, optically clear solutions were obtained.

Equilibrium GdmCl-mediated unfolding studies were conducted at 25.0 °C by monitoring $[\theta]_{232}$ at various denaturant concentrations. EGF stock solutions were prepared in 0.3 M potassium phosphate, pH 7.5, in 6 M GdmCl, and in 8.5 M GdmCl, 60 mM potassium phosphate, pH 7.5; CNBr-EGF stock solutions were prepared in 60 mM potassium phosphate, pH 7.5, in 6 and 8 M GdmCl. All data were obtained on dilutions from these solutions. Essentially the same protocol given earlier was followed in data reduction (Puett 1972a, 1973a,b; Holladay et al., 1974). Thermal studies on EGF and CNBr-EGF were performed using a thermostated cell holder.

Results

Characterization of the Derivatives. In order to assist in relating the nature of the derivatives to EGF, the amino acid sequence of mouse EGF (Savage et al., 1972, 1973) is given in Figure 1. Limited trypsin action yields a specific cleavage between Arg-48 and Trp-49 (Savage et al., 1972), thus producing des-(49–53)-EGF and the carboxyl-terminal pentapeptide. CNBr produces a cleavage of the peptide bond between Met-21 and His-22 and converts the methionyl residue to homoserine lactone. SAE-EGF represents the product obtained upon reduction of the three cystines followed by the conversion of the resulting cysteines to *S*-aminoethylcysteine using ethylenimine.

The in vivo biological activity and the in vitro binding characteristics to fibroblasts of the three derivatives relative to EGF are summarized in Table I. Interestingly, des-(49–

TABLE II: Mean Residue Rotational Strengths (cgs Units) and Tentative Assignments of the Resolved Gaussian Bands of EGF and Derivatives.[a]

Tentative Assignment	(EGF)[c]		(des-(49-53)-EGF)[d]		(CNBr-EGF)[f]		(SAE-EGF)[g]	
	λ (nm)	$R \times 10^{42}$	λ (nm)	$R \times 10^{42}$	λ (nm)	$R \times 10^{42}$	λ (nm)	$R \times 10^{42}$
Trp								
0–0 ^1L_b	289.5	+1.01		[e]	289	+0.38	292.5	+0.95
0+850 ^1L_b	283	+0.22		[e]	282	+0.18	285	+0.27
Tyr								
0–0	277.8	−0.67	280.5	−2.42	280	−3.30	280.6	−1.94
0+800	271	−1.64	273.5	−1.56	272	−1.36	274.4	−0.50
0+2(800)	264.7	−1.98	267.5	−1.61	265	−1.40	268.5	−1.74
Cys	280	−10.3	280	−7.05	280	−3.08		[h]
	253	−27.7	251	−18.6	252	−25.9		[h]
[b]	230	+64.4	228	+125	230	+74.0		
Peptide								
β structure	213	−162	214.5	−73.6	215	−89.5		[i]
Aperiodic	198.7	−643	199.5	−637	199.5	−841	198.5	−807

[a] The rotational strength, R, of a Gaussian band is estimated from the relationship, $R \simeq (1.234 \times 10^{-42}[\Theta°])(\Delta/\lambda_0)$, cgs units, where $[\Theta°]$ is the extremum ellipticity of the resolved band at wavelength λ_0 and Δ is the bandwidth (Sears and Beychok, 1973). Assignments are based on the CD spectra of proteins and model compounds (Chen et al., 1972, 1974; Strickland, 1974; Puett et al., 1974; Holladay and Puett, 1975a,b, 1976a,b). R is based on the mean residue ellipticity. [b] This band may have contributions from both the aromatics and the disulfides. [c] A band at 223 nm ($R = +11.1 \times 10^{-42}$) was also resolved. [d] Bands, probably from Tyr, were also observed at 270.5 and 262 nm with respective R's of -0.45 and -2.67×10^{-42}. [e] No Trp is present in this derivative. [f] A weak band ($R = -0.22 \times 10^{-42}$) was noted at 257.5 nm. [g] Bands at 270.5, 261, and 238.5 nm were resolved with respective R's of -0.08, -0.32, and -21.7×10^{-42}. [h] There are no disulfides in this derivative. [i] Two bands at 223 and 211.5 nm with R's of -268 and -12.2×10^{-42}, respectively, were found. These may reflect β structure (see text).

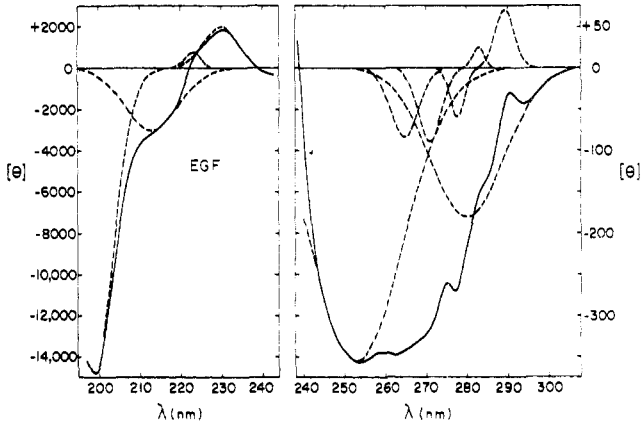

FIGURE 2: The resolved CD spectrum (mean residue ellipticity) of EGF in 50 mM KCl, 10 mM potassium phosphate, pH 7.5. (—) Experimental spectrum; (- - -) resolved Gaussian bands. Rotational strengths and tentative assignments are given in Table II.

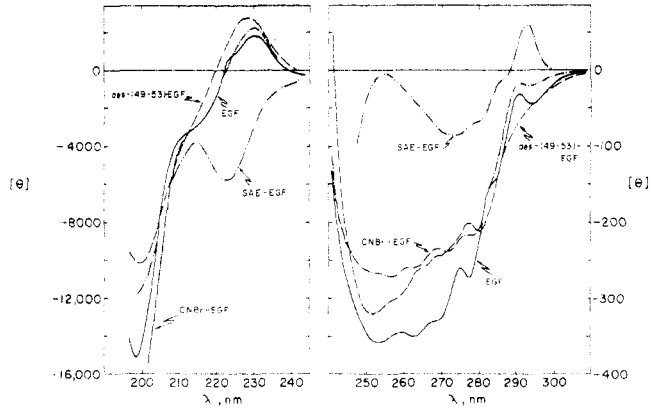

FIGURE 3: CD spectra (mean residue ellipticity) of des-(49–53)-EGF, CNBr-EGF and, for comparison, EGF, in 50 mM KCl, 10 mM potassium phosphate, pH 7.5. The spectrum of SAE-EGF in 10 mM acetic acid is also shown. Rotational strengths and tentative assignments of the resolved Gaussian bands (not shown) are given in Table II.

53)-EGF is fully active in the in vivo bioassay but demonstrates much less activity in the in vitro fibroblast growth and binding assay. The other two derivatives appear to be biologically inactive; the carboxy-terminal pentapeptide also had no demonstrable activity.

CD of EGF, Derivatives, and Model Compounds. Figure 2 shows the resolved CD spectrum of EGF and the rotational strengths and tentative assignments of the resolved bands are given in Table II. The spectrum above 225 nm arises mainly from side-chain chromophores (e.g., tyrosines, tryptophans, and disulfides in EGF), while below 225 nm the peptide groups make the major contribution.

It is important to point out that both the position and the rotational strength of the longest wavelength disulfide CD band are very dependent on the dihedral angle of the C–S–S–C chromophore (Kahn, 1972). Thus, in order to calculate meaningful ratios of R/R_0 it is important to use a reference model compound which has a value for λ_0 close to that for the resolved band in the protein CD spectrum. Therefore we have used L-cystine (λ_0 257 nm) to give an R_0 for the 251–253 nm –S–S– CD band of EGF, and L-cystinylbis(L-tyrosyl) (λ_0 275 nm) to give an R_0 for the 280-nm –S–S– CD band of EGF.

The CD spectra of the three derivatives are shown in Figure 3. These curves have also been resolved into Gaussian constituents (for brevity the Gaussian bands are not shown) and pertinent data are given in Table II. Some differences exist in

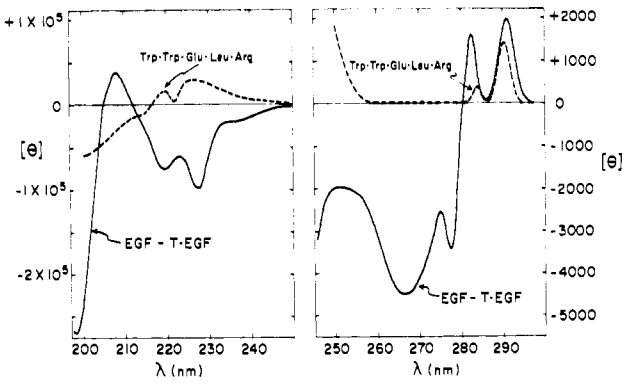

FIGURE 4: CD spectra (molar ellipticity) of the EGF carboxyl-terminal pentapeptide in phosphate buffered KCl (- - -) and of the CD difference spectrum (generated from the spectra in Figure 3) of EGF minus that of des-(49-53)-EGF (for brevity, the latter is abbreviated T-EGF).

TABLE III: Molar Rotational Strengths (cgs Units) and Gaussian Band Parameters for Cystine and Cystinylbis(tyrosine).

Tentative Assignment	λ_0 (nm)	[θ°] (deg cm²/dmol)	R
	Cystine		
-S-S-	257.3	−1 490	-1.74×10^{-40}
-S-S-	235	+720	$+3.77 \times 10^{-41}$
-S-S-[a]	200.5	+15 250	$+1.74 \times 10^{-39}$
-S-S-[a]	197	−40 000	-1.68×10^{-39}
	Cystinylbis(tyrosine)		
0-0 Tyr[b]	280.5	+1 080	$+3.09 \times 10^{-41}$
0+800 Tyr[b]	274	+660	$+1.19 \times 10^{-41}$
0+2(800)Tyr[b]	269	+460	$+7.39 \times 10^{-42}$
-S-S-	275	−1 300	-7.58×10^{-41}
-S-S-	254.5	−640	-1.86×10^{-41}
π-π* Tyr[b]	225	+39 000	$+2.03 \times 10^{-39}$
π-π* Tyr[b]	201.5	+90 000	$+3.03 \times 10^{-39}$
-S-S-	215	+31 000	$+1.51 \times 10^{-39}$

[a] The rotational strengths for these two bands represent only a minimal estimate due to the considerable overlap of the two far ultraviolet bands. [b] To obtain the molar rotational strength per tyrosyl residue, these values of R should be halved.

all the derivatives relative to EGF but, not surprisingly, the major effects are noted with SAE-EGF. It is noteworthy, however, that even SAE-EGF in 10 mM acetic acid appears to have some stable structure since in 4.8 M GdmCl the CD spectrum is reduced by about 50% over the wavelength range 212-230 nm (data now shown).

Based on the CD spectra of EGF, des-(49-53)-EGF, and CNBr-EGF below 225 nm, there is no evidence of α helicity. On the other hand, the resolved 213-215-nm band is indicative of β structure (Adler et al., 1973; Chen et al., 1972, 1974). A comparison of the rotational strength of this band in EGF and these two derivatives with that of standard proteins (Chen et al., 1972; Puett et al., 1974; Moore et al., 1974; Robinson et al., 1974) indicates the presence of about 22% EGF, 10% des-(49-53)-EGF), and 12% CNBr-EGF β structure, respectively.

Figure 4 shows the CD spectrum of the carboxy-terminal pentapeptide of EGF and the CD difference spectrum obtained by subtracting the spectrum of des-(49-53)-EGF from that of EGF. The two spectra differ both qualitatively and quantitatively and there is reasonable agreement only above 285 nm. This suggests a somewhat different environment of the

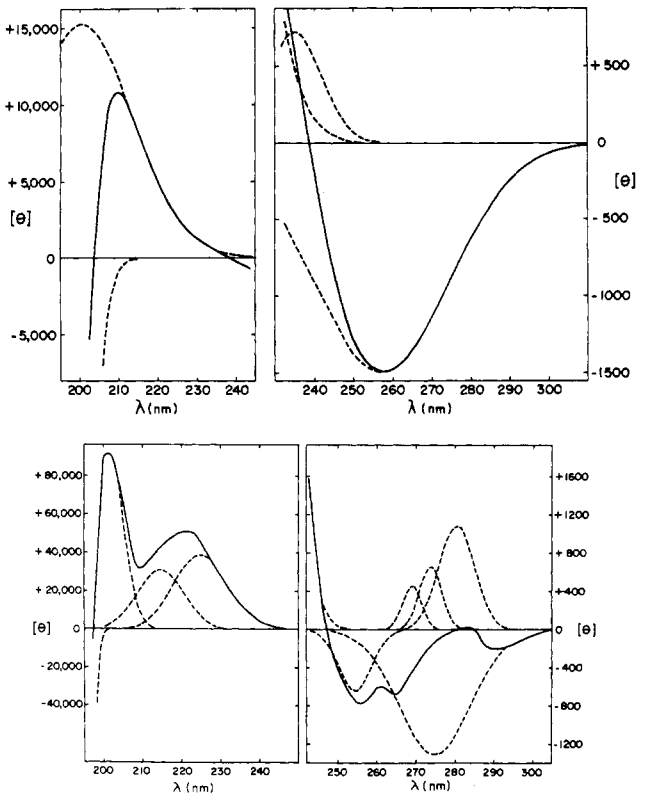

FIGURE 5: Resolved CD spectra (molar ellipticity) of L-cystine in 1 M KF, 20 mM Tris-HCl, pH 7.5, and L-cystinylbis(L-tyrosine) 0.1 M KCl, 10 mM Tris-HCl, pH 7.0. (—) Experimental spectrum; (- - -) resolved Gaussian bands.

chromophores in the core containing residues 1-48 in the absence of the five carboxyl terminal residues.

The CD contributions of side-chain chromophores are very dependent on the local chemical environment which is, of course, defined by the amino acid sequence and the resulting higher orders of structure. With EGF one needs to consider only β structure and tertiary structure since earlier studies have shown that in aqueous solution EGF exists primarily in the monomeric form (Taylor et al., 1972). In order to compare in at least a semi-quantitative fashion the contributions of side-chain chromophores in EGF and derivatives, one needs comparable data on model compounds which lack higher order structures (cf. Strickland, 1974; Holladay et al., 1974; Holladay and Puett, 1976a).

The resolved CD spectra of L-cystine and L-cystinylbis(L-tyrosine) are presented in Figure 5. As also found by others, the CD spectrum of the disulfide group varies considerably depending on the solvent, structure, and, of course, the screw sense (Sears and Beychok, 1973). Table III gives the resolved Gaussian band parameters for the disulfide bond in these two model compounds.

Using the major rotational strengths of Trp from the pentapeptide (Figure 4), the spectrum reported elsewhere for Gly-Tyr-Gly (Holladay and Puett, 1976a), and the spectra in Figure 5 for Cys chromophores, the contributions (relative to these model compounds) of Trp, Tyr, and Cys in EGF and derivatives have been calculated and the results are given in Table IV. There is clearly an enhancement of the aromatic contributions to the near-uv CD spectra of EGF and the appropriate derivatives. The comparison of disulfides is much less certain due to the variable nature of the resolved bands but, in general, it appears that an enhancement also occurs in Cys of EGF, des-(49-53)-EGF, and CNBr-EGF.

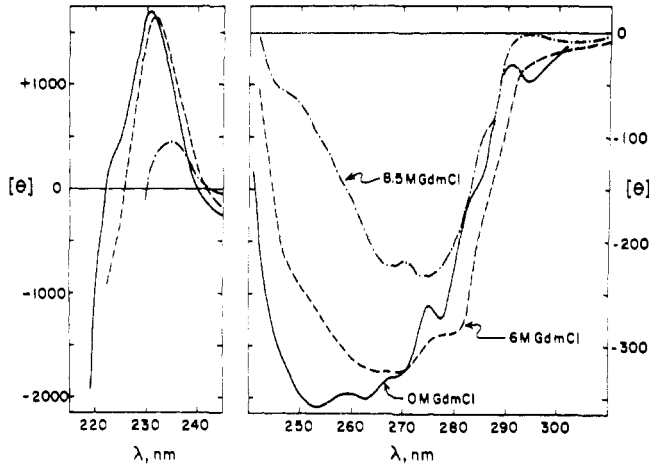

FIGURE 6: CD spectra (mean residue ellipticity) of native EGF and EGF in 6.0 and 8.5 M GdmCl.

While Gly-Tyr-Gly has negative near-ultraviolet CD bands and positive far-ultraviolet CD bands, L-cystinylbis(L-tyrosyl) appears to have positive tyrosyl CD bands in both the near and far ultraviolet. This is an interesting result since the absolute magnitude of the rotational strength for the tyrosyl CD bands are nearly the same (Holladay and Puett, 1975a). Thus, it appears that a knowledge of the near-ultraviolet CD spectrum of an aromatic chromophore will not allow unequivocal predictions to be made about the far-ultraviolet CD spectrum and vice versa.

Conformational Stability of EGF. Of special interest is the question of conformational stability for such a small protein. Much of the interpretation of the CD spectra is consistent with the existence of a stable tertiary structure for EGF. Equilibrium unfolding profiles obtained with GdmCl offer a useful method for examining the stability, or lack of stability, of proteins. Following optical or hydrodynamic properties of proteins as a function of GdmCl concentration yields information on their conformational stability and several methods are available for estimating the free energy of unfolding in the absence of denaturant (Tanford, 1970; Pace, 1975).

Figure 6 shows the CD spectrum of EGF between 230 and 310 nm in 0, 6, and 8.5 M GdmCl. There are no major changes in the conformation as monitored by circular dichroism between 0 and 6 M GdmCl; the observed differences may be solvent perturbation effects or local conformational shifts. In 8.5 M GdmCl the CD spectrum is significantly altered, suggesting that dramatic conformational changes are occurring as the GdmCl concentration is increased from 6 to 8.5 M.

The 232-nm extremum is particularly sensitive to GdmCl. A plot of $[\Theta]_{232}$ vs. GdmCl concentration for EGF at 25.0 °C is shown in Figure 7. The four arrows represent the midpoints for GdmCl denaturation of sperm whale ferrimyoglobin (Puett, 1973b), bovine pancreatic ribonuclease (Puett, 1972a), bovine β-lactoglobulin A (Pace, 1975), and bovine growth hormone (Holladay et al., 1974). These four proteins are representative of those previously characterized by GdmCl unfolding. The most obvious characteristics of the EGF (reversible) unfolding profile is that the midpoint occurs at 6.89 M GdmCl at 25.0 °C. The solubility of GdmCl is about 8.6 M at 25.0 °C (Nozaki, 1972), and thus we are unable to reach the end point of the unfolding profile. It does appear, however, that at 8.5 M GdmCl nearly complete unfolding is achieved. The value of $[\Theta]_{232}$ for denatured EGF at 25 °C is estimated to be about +350 deg cm²/dmol while the corresponding value for the native protein is +1658 deg cm²/dmol.

TABLE IV: Comparison of Molar Rotational Strengths (cgs Units) per Group of Some Major Tryptophanyl, Tyrosyl, and Cystinyl Bands of EGF and Derivatives with Corresponding Data from Low Molecular Weight Model Compounds.[a]

	R/R_0			
Transition	EGF	des-(49–53)-EGF	CNBr-EGF	SAE-EGF
Trp[b]				
0-0 ^1L_b	4.3	0	1.6	4.0
0+850 ^1L_b	4.4	0	3.6	5.5
Av:	(4.4)	(0)	(2.6)	(4.8)
Tyr[c]				
0-0	0.4	1.4	2.1	1.2
0+800	3.3	2.9	2.8	1.0
0+2(800)	3.2	2.3	2.3	2.8
Av:	(2.3)	(2.2)	(2.4)	(1.7)
Cys				
280 nm[d]	2.4	1.5	0.7	0
251–253 nm[e]	2.8	1.7	2.6	0
Av:	(2.6)	(1.6)	(1.7)	(0)

[a] The data are presented as the ratio of the molar R per group to that found in model compounds, R_0. [b] The EGF carboxyl-terminal pentapeptide (cf. Figure 4) was used as reference ($R_0 = +6.2 \times 10^{-42}$ and $+1.3 \times 10^{-42}$ for the 0-0 ^1L_b and 0+850 ^1L_b transition, respectively). [c] R_0 was from studies with Gly-Tyr-Gly in 20 mM Tris-HCl, pH 7 (Holladay and Puett, 1976a). [d] R_0 was from the 275-nm band in the resolved spectrum of Cys-bis(Tyr) in 0.1 M KCl, 10 mM Tris-HCl, pH 7 (cf. Figure 5). [e] R_0 was from the 257-nm band in the resolved spectrum of Cys in 1 M KF, 20 mM Tris-HCl, pH 7.5 (cf. Figure 5).

The extrapolation procedures which may be used for estimating the free energy of unfolding in the absence of denaturant, $\Delta G°$, have recently been reviewed (Pace, 1975), and only a brief summary will be given here. By assuming a two-state transition, the observable parameter $y = [\Theta]_{232}$ is given by eq 1

$$y = f_N y_N + f_D y_D \quad (1)$$

where f_N and f_D denote the fraction of native and denatured molecules, respectively, and y_N and y_D represent the value of $[\Theta]_{232}$ for fully native and fully denatured EGF extrapolated, if necessary, to the proper GdmCl concentration. Thus, an apparent equilibrium denaturation constant may be calculated by eq 2:

$$K_{app} = e^{-\Delta G_{app}/RT} = (y_N - y)/(y - y_D) \quad (2)$$

The simplest method for estimating $\Delta G°$ is the linear extrapolation based on eq 3:

$$\Delta G_{app} = \Delta G° + m[\text{GdmCl}] \quad (3)$$

This method appears to yield a minimum estimate for $\Delta G°$ (Pace, 1975) and does not involve any type of model for the denaturation process.

The second procedure involves a model of unfolding in which the denaturant preferentially binds to the unfolded state (Aune and Tanford, 1969). The relationship is given by eq 4, where Δn is the difference in the number of binding sites between states D and N, k is the average equilibrium constant for

TABLE V: Estimates for the Free Energy of Unfolding (kcal/mol) of EGF and CNBr-EGF in the Absence of Denaturant Using Various Extrapolation Methods.

EGF: Equation 3, Simple Extrapolation

$\Delta G° = 10.0 \pm 1.0$; slope $= -1.44$ kcal mol^{-1} M^{-1}

EGF: Equation 4, Binding Model

$\Delta G°$	Δn	k (M^{-1})
16.4	19	1.2
10.9	17	2.0
19.1	40	0.2

EGF: Equation 5, Side-Chain Transfer Model

$\Delta G° = 23.4 \pm 1.8$, $\bar{\alpha} = 0.63$

CNBr-EGF: Equation 3, Simple Extrapolation

$\Delta G° = 3.06$; slope $= -0.73$ kcal mol^{-1} M^{-1}

CNBr-EGF: Equation 4, Binding Model

$\Delta G°$	Δn	k (M^{-1})
4.27	13	0.45
5.27	9	1.20
5.86	7	1.80

GdmCl binding at each site, and a represents the mean ion activity of GdmCl.

$$\Delta G_{app} = \Delta G° - \Delta nRT \ln (1 + ka) \quad (4)$$

In general, the precise value of k cannot be readily estimated and has generally been assumed to be about 1.2 M^{-1} (Pace, 1975).

The third procedure involves a model for denaturation based on the solubilities of amino acid side chains and peptide groups in water and in GdmCl solutions (Tanford, 1970). The pertinent relationship is eq 5, where $\bar{\alpha}$ represents the average fractional change in exposure of side chains and peptide groups in going from state N to state D, and $\delta g_{tr,i}$ is the free energy of transfer of a particular class of side chains or the peptide group from water to denaturant.

$$\Delta G_{app} = \Delta G° + \bar{\alpha} \sum_i n_i \delta g_{tr,i} \quad (5)$$

Since the unfolding profile lies beyond 6 M GdmCl, values of $\delta g_{tr,i}$ were not available and thus were extrapolated by a parabolic fit to data points at 2, 4, and 6 M GdmCl (Robinson and Jencks, 1965; Nozaki and Tanford, 1970) with values for ½ Cys estimated from calculations on other side chains (Pace, 1975).

Results obtained for EGF with these three equations are given in Table V with standard errors wherever applicable. It does not appear possible to fix a precise value for k, so a range of k values was explored and reasonable upper and lower limits were taken as 0.2 and 2.0. From these data it appears that $\Delta G° = 16 \pm 7$ kcal/mol. The dangers of extrapolation using these three methods have been previously discussed (Pace, 1975). It seems clear that extrapolation will be even more hazardous in this case, where the midpoint is at 6.89 M GdmCl and transfer free energy data must first be extrapolated beyond 6 M GdmCl. Nevertheless, it is obvious that EGF possesses a remarkable degree of conformational stability regardless of the extrapolation procedure used.

The GdmCl unfolding profile of CNBr-EGF at 25.0 °C is also shown in Figure 7. Cleavage at the methionyl residue reduces the transition midpoint to 4.19 M GdmCl and the transition, like that of EGF, is reversible. The estimated values of $\Delta G°$ are given in Table V; from these data an average $\Delta G°$ of 4.2 kcal/mol was obtained for CNBr-EGF.

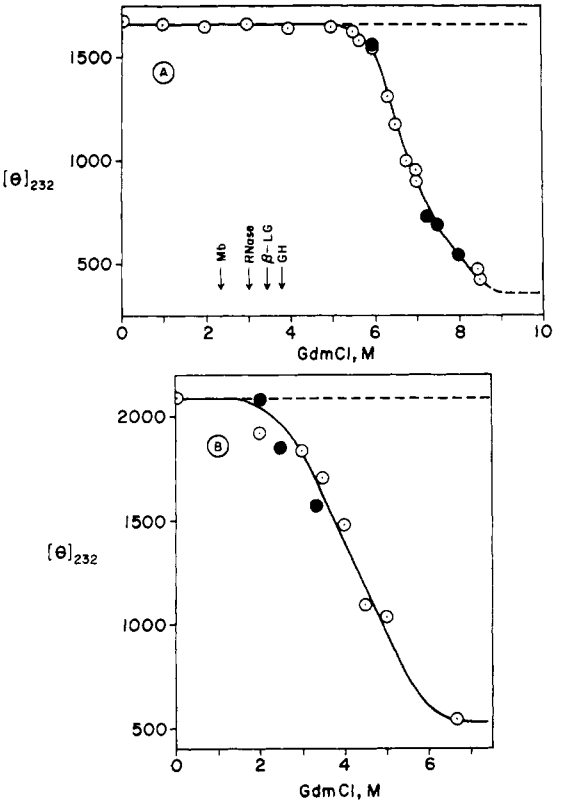

FIGURE 7: (A) Dependence of $[\theta]_{232}$ on GdmCl concentration at 25.0 °C. Open circles denote dilutions of EGF stock solutions (in 60 mM potassium phosphate, pH 7.5, or in 5.5 M GdmCl, 60 mM potassium phosphate, pH 7.5) with concentrated GdmCl; filled circles refer to dilutions made from an EGF stock in 8.5 M GdmCl, 60 mM potassium phosphate, pH 7.5. In all cases the final EGF concentration was 0.15 mg/ml and the solutions contained 10 mM potassium phosphate, pH 7.5. Spectra were recorded between 228 and 250 nm in a 5-mm cell. The arrows indicate the transition midpoint of sperm whale ferrimyoglobin (Mb) (Puett, 1973b), bovine pancreatic ribonuclease (RNase) (Puett, 1972a), bovine β-lactoglobulin A (β-LG) (Pace, 1975), and bovine growth hormone (GH) (Holladay et al., 1974). (B) The same as in A, except it is for CNBr-EGF instead of EGF. Protein concentrations, buffers, and other conditions were as in part A.

Thermal Stability of EGF. Careful analyses of the thermal transition of proteins can yield valuable information on the apparent enthalpy of unfolding, ΔH_{app}, and the apparent heat capacity, ΔC_p. Figure 8 shows a plot of $[\theta]_{232}$ vs. temperature for EGF in 5.5 M GdmCl and for CNBr-EGF in 2.0 M GdmCl. (Data, not shown, were also obtained at other GdmCl concentrations.) These thermal transitions are reversible. It is noteworthy that thermally denatured EGF and CNBr-EGF exhibit small negative values of $[\theta]_{232}$ above 60 °C, while at 25.0 °C the unfolded proteins appear to have small positive $[\theta]_{232}$ values. In fact, in 8.46 M GdmCl (data not shown), $[\theta]_{232}$ of EGF was positive (e.g., +225 deg cm^2/dmol) at 45.0 °C. Thus, it appears that an additional thermal state can be achieved at high temperatures as observed for other proteins (Salahuddin and Tanford, 1970).

Figure 9 shows the dependence of ln K_{app} on $1/T$ for EGF at two concentrations of GdmCl (5.50 and 5.75 M) and for CNBr-EGF in 2.0 M GdmCl. $[\theta]_{232}$ for thermally denatured EGF and CNBr-EGF was estimated to be −300 and −95 deg

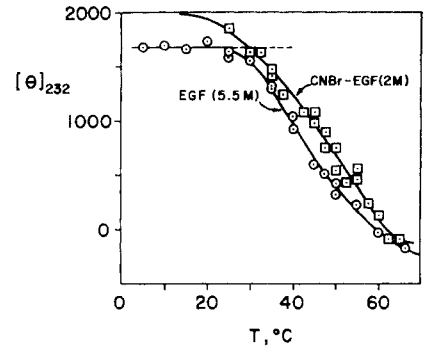

FIGURE 8: Thermal dependence of $[\theta]_{232}$ for EGF in 5.5 M GdmCl and for CNBr-EGF in 2.0 M GdmCl. Protein concentrations were 0.15 mg/ml and both solutions contained 10 mM potassium phosphate, pH 7.5. Complete reversibility was noted.

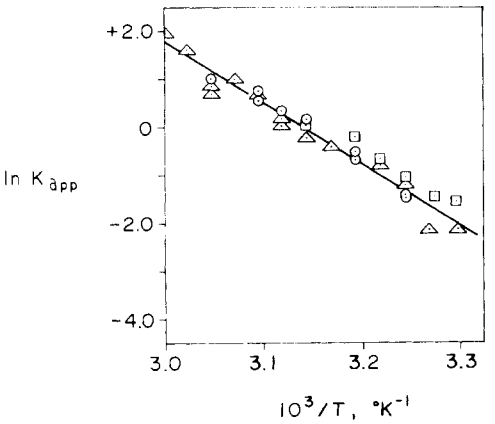

FIGURE 9: van't Hoff plots for EGF in GdmCl concentrations of 5.5 M (○) and 5.75 M (□) and for CNBr-EGF in 2.0 M GdmCl (△). The line was obtained by least-squares analysis using all the points and corresponds to $\Delta H_{app} = 25.37 \pm 1.27$ kcal/mol, with $\Delta C_p = 0$.

cm^2/dmol, respectively. The van't Hoff plots were examined for curvature by fitting the data to a parabola in $1/T$. For EGF in 5.75 M GdmCl and for CNBr-EGF in 2.0 M GdmCl, it was found that the coefficient of the $(1/T)^2$ term was not statistically different from zero. For EGF in 5.5 M GdmCl, there appeared to be a statistically significant downward curvature ($p < 0.02$). However, since only one of the three plots appeared nonlinear, it was assumed that the magnitude of ΔC_p is sufficiently small over this temperature interval to be ignored in all three cases, e.g., the absolute magnitude of ΔC_p is estimated to be less than 0.5 kcal/(mol deg^{-1}) for all three sets of data. Table VI summarizes the results obtained from the data in Figure 9. There appears to be no significant difference in the values of ΔH_{app} obtained from the three plots.

Discussion

The far-uv CD spectrum of EGF suggests the presence of some β structure, but not α helicity. This conclusion is in agreement with an earlier study using a different method of analysis (Taylor et al., 1972). However, the amount of ordered secondary structure is small and most of the polypeptide backbone (e.g., about 80%) appears to be in an aperiodic conformation. Relatively small changes in the primary structure of EGF, e.g., cleavage of the peptide bond at Met-21 or removal of the five carboxyl-terminal residues, seem to reduce the degree of β structure. It must be pointed out, however, that an accurate determination of the extent of ordered secondary structure is especially difficult in proteins that are mainly aperiodic.

TABLE VI: Apparent Enthalpies of Unfolding of EGF and CNBr-EGF in GdmCl at 40 °C.

	EGF	EGF	CNBr-EGF
GdmCl (M)	5.5	5.75	2.0
ΔH_{app} (kcal/mol)	25.8 ± 1.5	23.0 ± 2.0	26.3 ± 1.4

We have applied one set of predictive rules for secondary structure and bends (Chou and Fasman, 1974) to the sequence of EGF. The following regions of ordered secondary structure are predicted: two sections of β structure, residues 19–23 and residues 29–37, one section of α helix, residues 46–53, and chain reversals (β turns) beginning at residues 3, 6, 10, 15, 27, 37, and 40. The predicted α-helical segment also has a high $\langle P_\beta \rangle$ value and, thus, may exist as β structure in EGF. The modification at residue 21 (CNBr-EGF) and at residue 49 (des-(49–53)-EGF), appears to result in a loss of β structure; this is in accord with the predicted conformation if the C-terminal sequence forms β structure. We have constructed a space-filling model of EGF with β reversals at the aforementioned positions which contains three segments of β sheet very nearly in the positions indicated by the Chou–Fasman method. The ambivalent region consisting of residues 45–48 is tentatively assigned to the β form since this permitted the completion of a three-tiered β-sheet region on one face of the molecule. The resulting model had a very compact tertiary structure.

The enhancement noted in the near-uv CD spectrum of EGF, relative to low molecular weight chromophores and to denatured EGF, suggests that the protein contains a stable tertiary structure. Curve resolution of the near-uv CD spectra of EGF and the three derivatives yield interesting results on the conformational constraints of the aromatic and disulfide chromophores. The CD spectra of the derivatives lacking either the tryptophanyl chromophores (i.e., des-(49–53)-EGF) or the disulfide chromophores (i.e., SAE-EGF) support the assignments of the resolved near-uv CD bands of EGF to the particular classes.

The CNBr derivative appears to possess essentially the same near-uv CD spectrum as does EGF with the notable exception of the 280-nm -S-S- CD band. Given the inherent difficulties in curve resolution, part of the apparent change may also result from an altered conformation of one or more tyrosyl residues. The peptide bond cleaved in CNBr-EGF is only one residue removed from the disulfide bond connecting residues 6 and 20. It seems plausible to suggest that the 280-nm CD band may result in part from this disulfide because the rotational strength of the 280-nm band in CNBr-EGF is less than or equal to the rotational strength for the corresponding band in the model compound. Thus, the threefold decrease in the rotational strength of the 280-nm -S-S- CD band of EGF, accompanying cleavage of the peptide bond between residues 21 and 22, may arise from the fact that the disulfide bond between residues 6 and 20 possesses greater conformational flexibility in CNBr-EGF and the net rotational strength of the CD band drops. A threefold decrease in rotational strength would be consistent with the change of the -C-S-S-C- dihedral angle from a fixed position to an equilibrium mixture of left- and right-hand screw sense conformers (Sears and Beychok, 1973; Kahn, 1972).

This analysis suggests that the 253-nm CD band of EGF

results from disulfide bridges 14–31 and 33–42. The average molecular contribution of each of these two disulfides would be approximately −9000 deg cm^2/dmol. This value is in good agreement for a disulfide chromophore which exists as a single conformer with regard to the dihedral angle of the -C-S-S-C bond (Kahn, 1972). The fact that these two disulfide bonds are separated only by a single residue (Asn-32) may introduce conformational constraints not present in less hindered disulfide bonds and may account for their existence as a single conformer. As pointed out by Kahn (1972), the separation between the first two disulfide long wavelength transitions is more indicative of the dihedral angle than is the position of the longest wavelength disulfide CD band. Unfortunately, the curve resolution of the CD spectrum of EGF does not reveal the position of the lower wavelength disulfide CD bands and the determination of the dihedral angle by this method is not possible.

The data given in Table IV suggest that, as a class, the five tyrosyl residues of EGF, des-(49–53)-EGF, and CNBr-EGF are in a relatively constrained and asymmetric environment, as the average rotational strength per tyrosyl residue is roughly twofold higher than that of the model compound Gly-Tyr-Gly. For SAE-EGF, the tyrosyl residues have rotational strengths more nearly comparable to those of the model compound, suggesting that the conformational constraints on the tyrosyl residues in this derivative are much less than for EGF. This analysis is only approximate, however, since the tyrosyl CD bands of EGF, des-(49–53)-EGF, and CNBr-EGF overlap considerably with bands from the other chromophores in the near uv. The resolution of a CD spectrum containing many overlapping bands is somewhat uncertain and the rotational strength of any one tyrosyl band is subject to error. It should be pointed out in defense of curve resolution that the λ_0 values for the resolved tyrosyl bands are separated by nearly the expected amount for the 0–0, 0+800, and 0+2(800) transitions (Strickland, 1974).

The conformational stability of EGF is quite high as judged by the transition midpoint in GdmCl. As discussed in the previous section, our estimates of $\Delta G°$ are subject to considerable error. Nevertheless, even with the wide error limits, EGF appears to be one of the most stable proteins yet described (cf. Pace, 1975).

Our assumption of a two-state transition is only a working hypothesis and additional studies are required to elucidate the exact mechanism(s) of unfolding and refolding. Monitoring the ellipticity at wavelengths other than 232 nm was not feasible for the isothermal GdmCl-mediated unfolding study. For example, at longer wavelengths, higher protein concentrations are necessary and the amount of material required for a complete study was prohibitive. At shorter wavelengths, the low signal-to-noise ratio, coupled with the fact that signal changes are small in tranforming a protein with 20% β structure to a completely disordered chain, is not conducive for accurate measurements. As pointed out by both Tanford (1970) and Pace (1975), the simplifying assumption of a two-state transition for multi-state transitions will lead to low values for $\Delta G°$. Hence, we conclude that regardless of the conformational transition mechanism(s), EGF exhibits considerable conformational stability.

Presently, it is not possible to rationalize the $\Delta G°$ of proteins even knowing the sequence and three-dimensional structure (Tanford, 1970). For EGF, the three disulfides should contribute significantly to the stability by decreasing the entropy of the unfolded state (relative to the unfolded, reduced state). In this context, we have recently suggested that cyclic somatostatin (a tetradecapeptide with a single disulfide) contained β structure and formed a stable conformation (Holladay and Puett, 1975a, 1976b).

Previous quantitative unfolding studies on globular proteins have dealt with myoglobin, ribonuclease, lysozyme, cytochrome *c*, β-lactoglobulin, α-chymotrypsin, chymotrypsinogen, and the immunoglobulin light chains (cf. the review by Pace, 1975), i.e., proteins which have molecular weights between about 10 000 and 30 000. The present results on EGF, those on pancreatic trypsin inhibitor (Vincent et al., 1971), the study on somatostatin (Holladay and Puett, 1975a, 1976b), and of course the structure of insulin (Blundell et al., 1972) demonstrate that stable tertiary structures can form in small proteins. Additional studies with such materials should prove quite useful in elucidating on a molecular basis sequence-conformation-stability relationships.

Transition enthalpies of other proteins, e.g., β-lactoglobulin, lysozyme, and ribonuclease (Pace and Tanford, 1968; Tanford and Aune, 1970; Salahuddin and Tanford, 1970), vary between about −20 and +40 kcal/mol. Interestingly, we find the apparent enthalpy of unfolding to be about the same for both EGF and CNBr-EGF. The apparent ΔC_p has been determined for several proteins (all larger than EGF) in the absence of denaturant or in the presence of limited concentrations of denaturant. The range of ΔC_p values is generally between 1 and 3 kcal/(mol deg^{-1}) (cf. Tanford, 1970). Such relatively large heat capacities can be explained in terms of the exposure of masked hydrophobic side chains to an aqueous environment upon unfolding (Tanford, 1970). The small ΔC_p found for EGF and CNBr-EGF (i.e., 0 ± 0.5 kcal/mol) may arise in part from a relatively high degree of exposure of the hydrophobic side chains in native EGF to the solvent. Also, since our data were obtained in fairly concentrated GdmCl solutions, it is possible that a lower ΔC_p is expected upon exposure of hydrophobic groups to a GdmCl solvent as opposed to water. Pertinent model compound data are not available.

It is not unlikely that in a 53-residue protein complete masking of hydrophobic side chains is quite difficult to achieve for more than a few residues. The value of $\bar{\alpha}$ = 0.63 (cf. eq 5 and Table V) for GdmCl-mediated EGF unfolding indicates a greater degree of exposure of the side chains to solvent in native EGF than has normally been observed for larger proteins (Pace, 1975). Also germane to this issue is our finding on the λ_0 values for the aromatics in EGF and CNBr-EGF. For example, the resolved CD band positions for the 0–0 ^1L_b tryptophanyl and the tyrosyl bands strongly suggest that, as a class, these chromophores are not totally buried in a hydrophobic environment, but are at least partially exposed to the solvent (Strickland, 1974; Holladay and Puett, 1975a).

Thus, it becomes of considerable interest to understand the nature of the forces which stabilize the native conformation of EGF in the absence of complete burial of aromatic residues, e.g., complete burial of a single Trp would contribute about 3 kcal/mol (Tanford, 1970). At present we are not able to offer any substantial insight into this question except to again mention the possible role of the disulfides in reducing the entropy of the denatured protein (relative to linear unfolded proteins). The polarity of EGF as determined from the amino acid sequence (Savage et al., 1972) using the method of Capaldi and Vanderkooi (1972) is 43%. Thus, based on this parameter, EGF is a "typical" soluble protein and does not contain an unusually high content of hydrophobic side chains.

From the estimated thermodynamic parameters in Tables V and VI, the changes in the apparent free energy and entropy of unfolding at 40 °C have been determined. For these calcu-

TABLE VII: Apparent Thermodynamic Parameters of Unfolding[a] of EGF and CNBr-EGF at 40 °C.

	$\Delta G_{app}°$ (kcal/mol)	ΔH_{app} (kcal/mol)	ΔS_{app} (cal mol^{-1} deg^{-1})
EGF	18.0	24.4	20.4
CNBr-EGF	5.5	26.3	66.4

[a] It is assumed that ΔC_p is zero for both EGF and CNBr-EGF (cf. text). The $\Delta G_{app}°$ values were obtained by integrating the van't Hoff equation using $\Delta G°$ values of 16 and 4 kcal/mol, respectively, for EGF and CNBr-EGF at 25 °C. The ΔH_{app} values are from Table VI.

lations we assume that ΔH is independent of GdmCl concentration; this appears to be true for the limited data presented in Table VI and has been shown to hold for other proteins in urea and GdmCl solutions (Pace and Tanford, 1968; Tanford and Aune, 1970). These results, obtained by integration of the van't Hoff equation, are presented in Table VII. From these data one can conclude that the large reduction in the unfolding free energy of EGF upon cleaving the peptide bond between residues 21 and 22 results almost entirely from entropy changes. This is not surprising since the formation of two additional chain ends is expected to lead to a state of greater conformational flexibility, particularly when comparing denatured EGF and denatured CNBr-EGF. Pertinent to this issue is the finding reported earlier that both EGF and des-(49-53)-EGF will spontaneously refold and reoxidize to give full biological activity after reduction in 8 M urea (Savage et al., 1973). Thus, assuming complete unfolding under these conditions, it appears that the information necessary for refolding and correct positioning of the disulfide bonds is present in the 1-49 fragment of EGF.

Lastly, a comment on estimating ΔH_{app} is in order. Previously, ln K_{app} vs. T data have been reduced with the equation, $\ln K_{app} = a_0 + a_1 \ln T + a_2(1/T)$ (Tanford and Aune, 1970). While attempting to fit our data to this equation, it was found that upon reducing the least-squares matrix a divide by zero operation occurred in the Hewlett-Packard 9100B computer. This was found to result from the high degree of correlation of ln T with $1/T$. On examining the series expansion of ln T, it becomes evident why ln T and $1/T$ should correlate well when T is in the range of about 300-330 K. In fact, it is very likely that the use of the above equation will lead to serious round-off errors in estimating values of ΔH_{app} and ΔC_p as it is well known that reduction of a least-squares matrix for which two of the independent variables are highly correlated is an ill-conditioned numerical process. Thus, an expansion of ln K_{app} in powers of $1/T$ as reported herein, or an expansion in powers of T (Brandts and Hunt, 1967), appears to be preferable.

Acknowledgments

It is a pleasure to thank Ms. Brenda Crews, Ms. Marty Reich, and Ms. Betty Kay Wasserman for skillful assistance. Also, we thank Dr. C. N. Pace for providing a pertinent manuscript prior to publication. We would also like to acknowledge Mr. R. Glenn Hammonds, Jr., for allowing us to use his computer program for the Chou-Fasman prediction method. In addition, we appreciate the excellent services of the Vanderbilt University Computer Center.

References

Adler, A. J., Greenfield, N. J., and Fasman, G: D. (1973), *Methods Enzymol. 27D,* 675-735.

Armelin, H. A. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2702-2706.

Aune, K. C., and Tanford, C. (1969), *Biochemistry 8,* 4586-4590.

Beaven, G. H., and Holiday, E. R. (1952), *Adv. Protein Chem. 7,* 319-386.

Blundell, T., Dodson, G., Hodgkin, D., and Mercola, D. (1972), *Adv. Protein Chem. 26,* 279-402.

Brandts, J. F., and Hunt, L. (1967), *J. Am. Chem. Soc. 89,* 4826-4838.

Capaldi, R. A., and Vanderkooi, G. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 930-932.

Carpenter, G., Lembach, K. J., Morrison, M. M., and Cohen, S. (1975), *J. Biol. Chem. 250,* 4297-4304.

Chen, Y. H., Yang, J. T., and Chau, K. H. (1974), *Biochemistry 13,* 3350-3359.

Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972), *Biochemistry 11,* 4120-4131.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 222-245.

Cohen, S. (1962), *J. Biol. Chem. 237,* 1555-1562.

Cohen, S. (1965), *Dev. Biol. 12,* 394-407.

Cohen, S., and Carpenter, G. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 1317-1321.

Cohen, S., Carpenter, G., and Lembach, K. J. (1975), *Adv. Metab. Disord. 8,* 265-284.

Cohen, S., and Elliott, G. A. (1963), *J. Invest. Dermatol. 40,* 1-5.

Cohen, S., and Savage, C. R., Jr. (1974), *Recent Prog. Horm. Res. 30,* 551-572.

Cohen, S., and Taylor, J. M. (1974), in Thirtieth Symposium of the Society of Developmental Biology, Hay, E. D., King, T. J., and Papaconstantinou, J., Ed., New York, N.Y., Academic Press, pp 25-42.

Gregory, H. (1975), *Nature (London) 257,* 325-327.

Holladay, L. A., Hammonds, R. G., Jr., and Puett, D. (1974), *Biochemistry 13,* 1653-1661.

Holladay, L. A., and Puett, D. (1975a), in Peptides: Chemistry, Structure, and Biology, Walter, R., and Meienhofer, J., Ed., Ann Arbor, Mich., Ann Arbor Science Publishers, Inc., pp 175-179.

Holladay, L. A., and Puett, D. (1975b), *Arch. Biochem. Biophys. 171,* 708-720.

Holladay, L. A., and Puett, D. (1976a), *Biopolymers 15,* 43-59.

Holladay, L. A., and Puett, D. (1976b), *Proc. Natl. Acad. Sci. U.S.A. 73,* 1199-1202.

Hollenberg, M. D., and Cuatrecasas, P. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2964-2968.

Kahn, P. C. (1972), Ph.D. Dissertation, Columbia University.

Liu, T. Y., and Chang, Y. H. (1970), *J. Biol. Chem. 246,* 2842-2848.

Moore, S. (1973), in Chemistry and Biology of Peptides, Meienhofer, J., Ed., Ann Arbor, Mich., Ann Arbor Science Publishers, Inc., pp 642-643.

Moore, W. M., Holladay, L. A., Puett, D., and Brady, R. N. (1974), *FEBS Lett. 45,* 145-149.

Nozaki, Y. (1972), *Methods Enzymol. 26C,* 43-50.

Nozaki, Y., and Tanford, C. (1970), *J. Biol. Chem. 245,* 1648-1652.

Pace, N. (1975), *CRC Crit. Rev. Biochem. 3,* 1-43.

Pace, N., and Tanford, C. (1968), *Biochemistry 7,* 198–208.
Puett, D. (1972a), *Biochemistry 11,* 1980–1990.
Puett, D. (1972b), *Biochemistry 11,* 4304–4307.
Puett, D. (1973a), *J. Biol. Chem. 248,* 3566–3572.
Puett, D. (1973b), *J. Biol. Chem. 248,* 4623–4634.
Puett, D., Ascoli, M., and Holladay, L. A. (1974), in Hormone Binding and Target Cell Activation in the Testis, Dufau, M. L., and Means, A. R., Ed., New York, N.Y., Plenum Publishing Corp., pp 109–124.
Robinson, D. R., and Jencks, W. D. (1965), *J. Am. Chem. Soc. 87,* 2462–2470.
Robinson, J. P., Holladay, L. A., Picklesimer, J. B., and Puett, D. (1974), *Mol. Cell. Biochem. 5,* 147–151.
Savage, C. R., Jr., and Cohen, S. (1972), *J. Biol. Chem. 247,* 7609–7611.
Savage, C. R., Jr., and Cohen, S. (1973), *Exp. Eye Res. 15,* 361–366.
Savage, C. R., Jr., Hash, J. H., and Cohen, S. (1973), *J. Biol. Chem. 248,* 7669–7672.
Savage, C. R., Jr., Inagami, T., and Cohen, S. (1972), *J. Biol. Chem. 247,* 7612–7621.
Salahuddin, A., and Tanford, C. (1970), *Biochemistry 9,* 1342–1347.
Sears, D. W., and Beychok, S. (1973), in Physical Principles and Techniques of Protein Chemistry, Part C, Leach, S. J., Ed., New York, N.Y., Academic Press, pp 445–593.
Strickland, E. H. (1974), *CRC Crit. Rev. Biochem. 2,* 113–175.
Tanford, C. (1970), *Adv. Protein Chem. 24,* 1–95.
Tanford, C., and Aune, K. C. (1970), *Biochemistry 9,* 206–211.
Taylor, J. M., Cohen, S., and Mitchell, W. M. (1970), *Proc. Natl. Acad. Sci. U.S.A. 67,* 164–171.
Taylor, J. M., Mitchell, W. M., and Cohen, S. (1972), *J. Biol. Chem. 247,* 5928–5934.
Taylor, J. M., Mitchell, W. M., and Cohen, S. (1974a), *J. Biol. Chem. 249,* 2188–2194.
Taylor, J. M., Mitchell, W. M., and Cohen, S. (1974b), *J. Biol. Chem. 249,* 3198–3203.
Turkington, R. W. (1969), *Exp. Cell Res. 57,* 79–85.
Vincent, J. P., Chicheportiche, R., and Lazdunski, M. (1971), *Eur. J. Biochem. 23,* 401–411.
Zahler, W. L., Puett, D., and Fleischer, S. (1972), *Biochim. Biophys. Acta 255,* 365–379.

Oxidation–Reduction Properties of Several Low Potential Iron–Sulfur Proteins and of Methylviologen†

N. A. Stombaugh,‡ J. E. Sundquist,§ R. H. Burris, and W. H. Orme-Johnson*

ABSTRACT: Apparent oxidation–reduction potentials at pH 7.0 and 25 °C were determined using the H_2–hydrogenase system with ferredoxins from the following sources: *Clostridium pasteurianum,* −403 mV; *C. tartarovorum,* −424 mV; *C. acidi-urici,* −434 mV; *Peptococcus aerogenes,* −427 mV; *Chromatium* D, −482 mV (pH 8.0); *B. polymyxa,* Fd I, −377 mV, and Fd II, −422 mV; and spinach, −428 mV. The pH dependence of these values was variable, ranging from −2 to −24 mV/pH unit increase for different ferredoxins. Over the range of buffer concentrations between 0.05 and 0.2 M, the potentials did not vary significantly. The number of electrons transferred during reduction (as determined by integrations of EPR spectra and by dithionite titration) is 2 for the first five proteins, while potentiometric data for all the cases fit a Nernst equation for which $n = 1$. The $E°'$ value for the redox indicator methylviologen at pH 7.4 was found to be −460 mV, according to both the H_2–hydrogenase system and cyclic voltammetry, significantly different from the value previously reported at higher pH's. Additionally, the presence of *C. pasteurianum* ferredoxin appears to shift the $E°$ value of methylviologen to even more negative values. An analysis of sources of error inherent with potential determinations with H_2 and hydrogenase is presented. The electronic and EPR spectra of *P. aerogenes* ferredoxin, for which the x-ray structure has been published, are given here. It appears that the determination of potentials of ferredoxin and other low-potential proteins with the H_2–hydrogenase system affords certain experimental advantages over alternative methods currently employed with these and similar substances.

Among the iron–sulfur proteins, two distinct classes of these proteins transfer electrons at oxidation–reduction potentials near the potential of the H_2/H^+ couple. These classes are (a) the plant-type ferredoxins, which contain a cluster of two nonheme iron and two labile sulfur atoms per molecule, and (b) the bacterial-type ferredoxins, which appear to contain one or two clusters of four iron and four labile sulfur atoms per molecule. The distinguishing characteristics of these proteins have been extensively reviewed (see Orme-Johnson, 1973, and references therein). In particular, the amino acid sequences of many of these proteins are known, several crystallographic structure studies on the iron–sulfur proteins are near the atomic resolution stage, and a number of synthetic analogues of the metal clusters

† From the Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706. *Received July 28, 1975.* Supported by National Institutes of Health Grants GM 17170 and AI 00848.

‡ Present address: Central Research Department, Experimental Station, E. I. du Pont de Nemours and Company, Wilmington, Delaware 19898.

§ Deceased January 19, 1972.