Laboratory, Cold Spring Harbor, NY.
Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499-560.
McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171-204.
Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Misra, T. K., Brown, N. L., Fritzinger, D. C., Pridmore, R. D., Barnes, W. M., Haberstroh, L., & Silver, S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5975-5979.
Ni'Bhriain, N. N., Silver, S., & Foster, T. J. (1983) *J. Bacteriol. 155*, 690-713.
Ni'Bhriain, N., Dorman, C. J., & Higgins, C. F. (1989) *Mol. Microbiol. 3*, 933-942.
O'Halloran, T. V., Frantz, B., Shin, M. K., Ralston, D. M., & Wright, J. G. (1989) *Cell 56*, 119-129.
Ralston, D. M., & O'Halloran, T. V. (1990) *Proc. Natl. Acad. Sci. U.S.A. 56*, 3846-3850.
Reitzer, L. J., Bueno, R., Cheng, W. D., Abrams, S. A., Rothstein, D. M., Hunt, T. P., Tyler, B., & Magasanik, B. (1987) *J. Bacteriol. 169*, 4279-4284.
Ross, W., Park, S.-J., & Summers, A. O. (1989) *J. Bacteriol. 171*, 4009-4018.
Sasse-Dwight, S., & Gralla, J. D. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8934-8938.
Sasse-Dwight, S., & Gralla, J. D. (1989) *J. Biol. Chem. 264*, 8074-8081.
Shanblatt, S. H., & Revzin, A. (1986) *J. Biol. Chem. 261*, 10885-10890.
Shewchuk, L. M., Helmann, J. D., Ross, W., Park, S.-J., Summers, A. O., & Walsh, C. T. (1989a) *Biochemistry 28*, 2340-2344.
Shewchuk, L. M., Verdine, G. L., & Walsh, C. T. (1989b) *Biochemistry 28*, 2331-2339.
Shewchuk, L. M., Verdine, G. L., Nash, H., & Walsh, C. T. (1989c) *Biochemistry 28*, 6140-6145.
Siebenlist, U. (1979) *Nature 279*, 651-652.
Siebenlist, U., Simpson, R. B., & Gilbert, W. (1980) *Cell 20*, 269-281.
Sippel, A., & Hartmann, G. (1968) *Biochim. Biophys. Acta 157*, 218-219.
Stefano, J. E., & Gralla, J. D. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1069-1072.
Straney, S. B., & Crothers, D. M. (1987) *Cell 51*, 699-707.
Summers, A. O. (1986) *Annu. Rev. Microbiol. 40*, 607-634.
Walsh, C. T., Distefano, M. D., Moore, M. J., Shewchuk, L. M., & Verdine, G. L. (1988) *FASEB J. 2*, 124-130.
Wang, C. W. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 200-203.

Replacement of a Labile Aspartyl Residue Increases the Stability of Human Epidermal Growth Factor†

Carlos George-Nascimento,*,‡ Jonathan Lowenson,§ Michael Borissenko,‡ María Calderón,‡ Angélica Medina-Selby,‡ Jane Kuo,‡ Steven Clarke,§ and Anne Randolph‡

Chiron Research Laboratories, Chiron Corporation, Emeryville, California 94608, and Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90024

Received May 2, 1990; Revised Manuscript Received July 20, 1990

ABSTRACT: Long-term storage of recombinant human epidermal growth factor (EGF), an important promoter of cell division, results in its conversion to a new species that elutes later than native EGF on a reverse-phase column. This new species, called EGF-X, has only 20% of the biological activity of native EGF. Peptide mapping indicated that the primary structure of EGF-X differs from that of native EGF solely within the first 13 residues. N-Terminal sequencing of EGF-X revealed that about 30% of the polypeptides have been cleaved at the Asp-3/Ser-4 bond. In addition, the yields after the His residue at position 10 were extremely low, indicating that a chemical modification occurs at residue 11 that is incompatible with Edman degradation. We hypothesized that aspartic acid 11 had been converted to an isoaspartyl residue, and this was confirmed with L-isoaspartyl/D-aspartyl methyltransferase, an enzyme that methylates the side-chain carboxyl group of L-isoaspartyl residues but does not recognize normal L-aspartyl residues. EGF-X, but not EGF, was found to be a substrate of this enzyme, and proteolytic digestion of EGF-X with thermolysin localized the site of methylation to a nine-residue peptide containing position 11. We did not observe formation of the isoaspartyl derivative in EGF that had been denatured by reduction of its disulfide bonds. In addition, replacement of the aspartyl residue at position 11 with glutamic acid resulted in a fully active EGF derivative that does not form detectable amounts of EGF-X. We propose that conversion of this aspartyl residue to isoaspartate is a significant nonenzymatic degradation reaction affecting this growth factor. Replacing this residue with a glutamyl residue, however, prevents this degradation reaction, producing a biologically active EGF molecule with greater stability.

Epidermal growth factor (EGF)[1] is a potent mitogen and an inhibitor of gastric acid secretion. Due to these biological activities, EGF has been studied extensively since its discovery (Cohen, 1962). Human EGF has 53 amino acid residues and three disulfide bridges, the positions of which are conserved

† This work was partially supported by a grant from the National Science Foundation (DMB 89-04170) to S.C. and a U.S. Public Health Training Grant (GM 07185) to J.L.

‡ Chiron Corp.
§ University of California.

[1] Abbreviations: EGF, epidermal growth factor; HPLC, high-performance liquid chromatography; IAA, iodoacetic acid; D11E-EGF, variant of epidermal growth factor having glutamic acid at position 11.

0006-2960/90/0429-9584$02.50/0

with respect to those of murine EGF (Gregory, 1975). Several posttranslational modifications have been observed in EGF from a variety of organisms. EGF derivatives truncated at the N-terminus have been detected in rats (Simpson et al., 1985), while EGF extracted from mouse submaxillary glands can be modified at both the C- and N-termini (O'Keefe et al., 1984; Petrides et al., 1984; Smith et al., 1984). DiAugustine et al. (1987) proposed that the N-terminal asparaginyl residue in murine EGF is deamidated via a mechanism that results in the formation of an isoaspartyl residue, in which the peptide backbone is redirected through the side-chain β-carboxyl group. Aspartyl and asparaginyl residues are particularly susceptible to nonenzymatic degradation because intramolecular succinimide formation can lead to isomerization and/or racemization. The factors that contribute to this reaction are not totally understood; it has been suggested, however, that the sequence (Geiger & Clarke, 1987; Stephenson & Clarke, 1989) and the conformation (Clarke, 1987; Kossiakoff, 1988) of a polypeptide can markedly affect the rate of succinimide formation.

Human EGF found in urine contains a significant amount of a derivative lacking the C-terminal Arg residue (Oka & Orth, 1983), but no N-terminal modifications have as yet been described in human EGF. Bacterial-produced recombinant human EGF contains all 53 of the expected amino acid residues (Smith et al., 1982). In contrast, we have previously found that yeast-produced recombinant human EGF lacks the C-terminal arginyl residue, presumably due to enzymatic processing in the yeast secretory pathway, and is thus identical with the γ-urogastrone found in human urine (George-Nascimento et al., 1988). In the same study, another modification was reported: the lone methionine at position 21 of this EGF is quite susceptible to oxidation to the sulfoxide form.

Since recombinant human EGF is currently in clinical trials as a potential drug to be used for wound healing, it is important to study in detail the possible modifications affecting EGF upon storage in either a lyophilized form or in solution. These studies are important for understanding the stability of biologically active peptides, as well as the effect of an amino acid residue's environment on its behavior within a polypeptide. We wish to report here a posttranslational modification affecting human EGF that results in the replacement of the aspartyl residue at position 11 with an isoaspartyl residue and a partial loss of biological activity. To determine if an active EGF could be synthesized that would not undergo such a reaction, a variant was constructed in which a glutamic acid replaced the aspartyl residue at position 11. We have found that this altered EGF is fully active and is much more stable than native EGF during aging.

Experimental Procedures

Recombinant human epidermal growth factor was prepared and characterized as previously described (George-Nascimento et al., 1988). Mouse EGF was a gift from Dr. S. Cohen, Vanderbilt University, Nashville, TN.

Purification of the EGF Degradation Product (EGF-X). Lyophilized human recombinant EGF was dissolved at 1 mg/mL final protein concentration in 0.1 M citrate, pH 3.0. This solution was incubated at 45 °C, and the formation of EGF-X was monitored daily by reverse-phase HPLC (George-Nascimento et al., 1988). For purification purposes, the protein peak containing EGF-X was collected. For long-term storage of either EGF or EGF-derived peptides, a lyophilized powder was prepared and kept at 4 °C.

Peptide Mapping Using Chymotrypsin. A total of 100 μg of protein was dissolved in 0.2 M *N*-ethylmorpholineacetate, 6 M guanidine hydrochloride, and 3 mM EDTA, pH 8.6. One microliter of β-mercaptoethanol was added, and the reaction mixture was sealed under argon and incubated at 50 °C. After 1 h, 14 μL of a solution containing 2.7 mg of iodoacetic acid (IAA) in 1.0 N NaOH (140 μL of NaOH for 27 mg of IAA) was added. The reaction mixture was sealed under argon and placed in the dark for 30 min. The reduced and carboxymethylated protein was dialyzed on a microdialyzer (BRL) against 50 mM *N*-ethylmorpholine, pH 8.6, with Spectrapor 3 membrane (3500 molecular weight cutoff). The sample was lyophilized and dissolved in 100 μL of 0.2 M Tris-HCl, 2 M urea, and 0.1 mM $CaCl_2$, pH 8.1; 4 μg of chymotrypsin-TLCK was added, and digestion was allowed to proceed overnight at 37 °C. The digest was loaded directly onto a reversed-phase HPLC C18 (4.6 × 250 mm) Vydac column, and the column effluent was monitored at 230 nm.

Peptide Mapping for Methylation Analysis. Intact EGF and EGF-X were digested with thermolysin according to a procedure similar to that of Savage et al. (1973). Thermolysin (Calbiochem, 50 units/mg) was reconstituted to a concentration of 0.4 mg/mL in 10 mM sodium borate and 2 mM $CaCl_2$, pH 10, and 5 μL of this solution (2 μg of thermolysin) was added to 60 μg of EGF or EGF-X in 100 μL of 0.4 M Bis-Tris–acetate, pH 6.5. This digestion mixture was incubated at 45 °C for 24 h with occasional mixing. Half of this solution was then acidified with 10% trifluoroacetic acid, and the resulting peptides were separated by reverse-phase HPLC (Alltech Econosphere C18 column, 4.6 × 250 mm) as described elsewhere (Gilbert et al., 1988).

Performic Acid Oxidation and Amino Acid Analysis. An aliquot of each of the collected peptides, as well as intact EGF and EGF-X, was oxidized with performic acid to convert cystines to cysteic acid residues. Performic acid was made by addition of 1 volume of 30% hydrogen peroxide to 19 volumes of 90% formic acid, followed by 2 h at room temperature (Savage et al., 1972). After being cooled on ice, 20 μL of performic acid was added to 20 pmol of EGF or 1–5 nmol of each thermolytic peptide that had been lyophilized to dryness in a 1.5-mL microcentrifuge tube. These samples were then incubated on ice for 2 h. The reaction was stopped by diluting each sample with 1 mL of water. These solutions were lyophilized to dryness, reconstituted in water, and lyophilized again. The compositions of both oxidized and unoxidized thermolytic peptides were determined by precolumn derivatization of acid hydrolysates with *o*-phthalaldehyde and HPLC analysis (Gilbert et al., 1988).

Purification of L-*Isoaspartyl*/D-*Aspartyl Protein Carboxyl Methyltransferase.* Isozyme I of this enzyme was purified from fresh human erythrocytes essentially as described by Gilbert et al. (1988). The specific activity was 673 units/mg of protein (1 unit = 1 pmol of methyl groups transferred to the substrate ovalbumin per minute); this represents a 358-fold purification of the enzyme over crude erythrocyte cytosol.

Preparation of Synthetic Peptides. The octapeptide L-Leu-L-Ser-L-His-L-isoAsp-Gly-L-Tyr-L-Cys-L-Leu was synthesized by Dr. Janis Young at the UCLA Peptide Synthesis Facility on an Applied Biosystems Model 340A instrument and was cleaved from the resin with hydrogen fluoride. The isoaspartyl-containing peptide was purified from normal aspartyl- and succinimidyl-containing contaminants according to reverse-phase HPLC procedures similar to those previously reported (Murray & Clarke, 1984). The structure of the octapeptide was confirmed by amino acid analysis and positive-ion fast atom bombardment mass spectroscopy (parent ion m/z 907; performed by Dr. Dilip Sensharma, UCLA

Department of Chemistry and Biochemistry).

Methylation Assay of Peptides and EGF. The number of methyl groups transferred from *S*-adenosylmethionine to the isoaspartyl-containing substrates was determined by measuring base-labile radioactivity in a vapor diffusion assay (Murray & Clarke, 1984). Peptide or EGF (final concentration 0.25–32 μM) was incubated at 37 °C for 15 min to 4 h in 50 μL of 0.2 M sodium citrate, pH 6.0, containing 1.67 units (for K_m assays) or 8.35 units (for time course) of methyltransferase and 10 μM *S*-adenosyl-L-[*methyl*-^{14}C]methionine (ICN 314007, 47 mCi/mmol). The reaction was quenched by the addition of 50 μL of 0.2 M NaOH and 1% (w/v) sodium dodecyl sulfate, and immediately, a 75-μL aliquot was spotted on a 1 × 8 cm piece of thick filter paper (Bio-Rad No. 165-090) that had been prefolded in an accordion pleat (Gilbert et al., 1988). This was wedged into the neck of a 20-mL scintillation vial containing 10 mL of ACS II counting fluor (Amersham), which was then capped and allowed to equilibrate at room temperature for 2 h. During this time, the base-labile aspartyl methyl esters were hydrolyzed to [^{14}C]-methanol, which diffused into the fluor. The filter paper was then removed, and the vials were counted.

Genesis and Purification of D11E-EGF. A mutant EGF clone was generated by starting with the clone pGAP/EGF, where the expression of the wild-type human EGF is driven by the GAP promoter and includes the α-factor leader sequence (Brake et al., 1984; Travis et al., 1985). Correct recombinants in the proper orientation were identified on the basis of their restriction pattern. Positive recombinants were transformed into *Saccharomyces cerevisiae*, strain JSC-20 competent yeast cells. Transformed yeast cells were grown in a minimal medium, and after 72–96 h the medium was collected by centrifugation. The medium was concentrated by membrane filtration on an Amicon concentrator (membrane YM2) and applied to a P-10 column (gel filtration chromatography) equilibrated with 0.1 M acetic acid. The EGF-containing fractions were pooled and concentrated by membrane filtration as described above. The final purification was accomplished by reverse-phase HPLC on a semipreparative C4 column (10 mm × 25 cm) (George-Nascimento et al., 1988).

Results

Formation of a New Species of EGF upon Aging. The reverse-phase HPLC profile of recombinant human EGF aged in solution showed the appearance of a new molecular form of EGF (EGF-X). Figure 1 shows the kinetics of transformation of the growth factor in solution to this new form of EGF, depicting a precursor–product relationship between EGF and EGF-X. EGF-X is formed preferentially in the pH range from 2 to 6.5, and this transformation is temperature dependent (data not shown). We have observed EGF-X formation at −20 and 2–8 °C only after 2–16 months (data not shown). The maximum transformation of EGF into EGF-X occurred when the EGF was in citrate buffer, pH 3.0, at 45 °C. Longer incubation times, under the same conditions, induced transformation of EGF and EGF-X into other species that were not characterized in this study.

To determine whether mouse EGF underwent a reaction similar to that seen with human EGF, the murine growth factor was incubated under the conditions optimal for human EGF-X formation described above. We did not observe any change in the elution pattern of mouse EGF by reverse-phase HPLC. We conclude from this result that mouse EGF is not capable of forming, under these conditions, a modified form analogous to the one observed with human EGF.

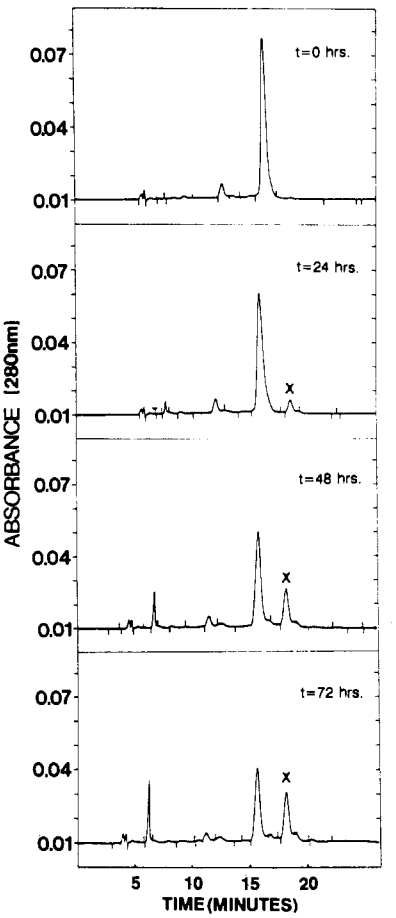

FIGURE 1: Analytical HPLC showing EGF-X formation as a function of time. Native EGF incubated in citrate buffer at pH 3.0 at 45 °C elutes at 16 min when analyzed by reverse-phase HPLC on a C4 Vydac column. Each chromatogram was performed under conditions identical with those previously reported (George-Nascimento et al., 1988). The letter X denotes the elution of EGF-X.

Since reverse-phase HPLC resolved EGF from EGF-X, we were able to purify sufficient quantities of EGF-X for its characterization. The biological activity of the EGF-X was determined with both a receptor binding assay and a mitogenic assay (George-Nascimento et al., 1988). The results of the mitogenic assay, measured as [^{3}H]thymidine incorporation into fibroblast cells upon the addition of human EGF and EGF-X, are depicted in Figure 2. Within the normal scatter of the mitogen assay, EGF-X was found to be partially active, having approximately 20% of the specific activity of the normal molecule. The receptor binding activity of EGF-X, compared with that of the native molecule, was found to be decreased to a similar extent (data not shown).

To localize the site(s) of modification in EGF-X, we generated a chymotryptic peptide map of this new EGF species and compared it with the peptide profile of the native growth factor. Figure 3 shows the reverse-phase HPLC chromatograms of peptides generated by cleavage with chymotrypsin of reduced and carboxymethylated EGF and EGF-X. The elution time of a single peptide, labeled C(1-2), was the only major difference between the chymotryptic maps of EGF and EGF-X. The amino acid composition of this peptide showed that it corresponds to the first 13 residues of EGF. These results indicate that within the first 13 residues there is a difference in peptide structure between EGF and EGF-X. In addition, we have also observed reduced levels of Ser and Asp residues in those peptides [C(1) comprising residues 1–8 and C(1-3) comprising residues 1–22] containing the N-terminal portion of EGF, indicated by an asterisk in Figure 3. Since

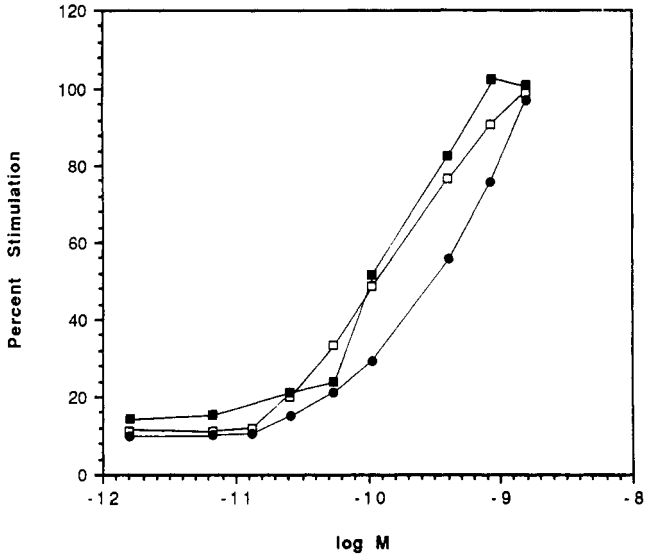

FIGURE 2: Mitogenic activity of purified EGF, EGF-X, and D11E-EGF. Stimulation of [^{3}H]thymidine uptake by confluent fibroblast cells. Points represent the percentage of counts, at each protein concentration, of the maximum stimulation. Recombinant human epidermal growth factor (□), EGF-X (●), and D11E-EGF (■), a mutant of EGF having a glutamic acid residue at position 11.

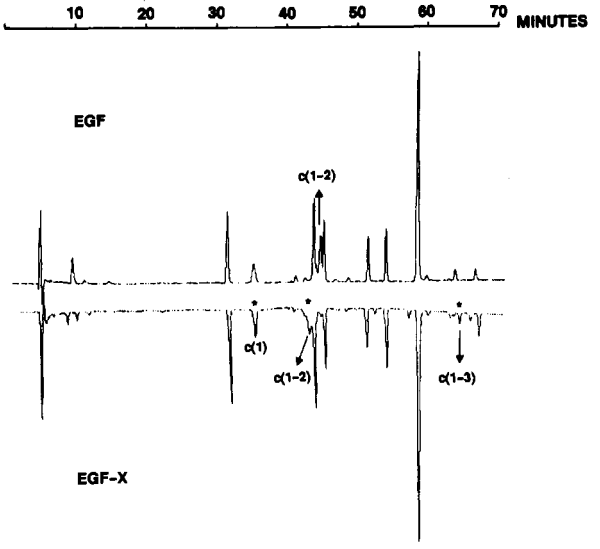

FIGURE 3: Chymotryptic peptide mapping of native EGF and EGF-X. Samples prepared as described under Experimental Procedures were loaded onto a Vydac C18 column (0.46 × 25 cm) equilibrated in 90% A buffer at 0.8 mL/min, and the peptides were eluted with a linear gradient to 100% B in 100 min. Buffer A is 0.1% (w/v) trifluoroacetic acid (TFA), and B buffer is 0.1% (w/v) TFA in 80% acetonitrile. Sequence of peptides: C(1), Asn-Ser-Asp-Ser-Glu-Cys-Pro-Leu; C(1-2), Asn-Ser-Asp-Ser-Glu-Cys-Pro-Leu-Ser-His-Asp-Gly-Tyr; C(1-3), Asn-Ser-Asp-Ser-Glu-Cys-Pro-Leu-Ser-His-Asp-Gly-Tyr-Cys-Leu-His-Asp-Gly-Val-Cys-Met-Tyr. The column effluent was monitored at 230 nm.

the first three residues of EGF are Asn-Ser-Asp, the loss of these residues would be reflected in low levels of Asp and Ser by amino acid analysis. On the basis of the hydrophilicity of these terminal residues, it is not unexpected that the complete and the truncated peptides coelute on reverse-phase HPLC (Guo et al., 1986). Because the other peptides in the proteolytic maps are identical for EGF and EGF-X, it appears that the difference between EGF and EGF-X is within the first 13 residues of these molecules.

To characterize further the differences between EGF and EGF-X, isolated EGF-X was subjected to amino-terminal sequencing, and the data obtained are presented in Table I.

Table I: Edman Sequence Data for Native EGF and for the Major and Minor Sequences of EGF-X[a]

	native EGF		EGF-X major		EGF-X minor	
cycle	residue	pmol	residue	pmol	residue	pmol
1	Asn	713	Asn	576	Ser	98
2	Ser	344	Ser	281*	Glu	72
3	Asp	561	Asp	388	(Cys-Cys)	
4	Ser	193	Ser	187	Pro	69
5	Glu	376	Glu	251	Leu	63
6	(Cys-Cys)		(Cys-Cys)		Ser	31
7	Pro	277	Pro	196		
8	Leu	250	Leu	183		
9	Ser	112	Ser	86		
10	His	84	His	50		
11	Asp	161	(Asp)	−1		
12	Gly	92	(Gly)	−2		
13	Tyr	123	(Tyr)	−1		

[a] In cycles 2–12, the number of picomoles of amino acid residue recovered is adjusted by subtracting the amount of that amino acid observed in the previous cycle. This has not been done for the serine in the second cycle of the major sequence of EGF-X [denoted by an asterisk (*)] because the minor sequence contains serine in the previous cycle. Reliable data were not available for the minor EGF-X sequence beyond cycle 6, and thus the configuration of the aspartyl residue at position 11 in this N-terminal processed form could not be determined.

Two significant observations can be made from these data. First, there was a secondary sequence observed, albeit in lower yield, whose first residue corresponded to the fourth residue in native EGF. This result indicated that N-terminal degradation had occurred in about 30% of the EGF-X polypeptides. Second, the normal Edman degradation reaction terminated after the histidine residue at position 10, possibly due to a modification at residue 11.

To study the effect of the conformation of EGF on the formation of EGF-X, we unfolded the polypeptide by reducing and alkylating the disulfide bonds prior to incubation at pH 3. No formation of a new EGF-like species was observed following aging of this unfolded EGF under conditions described previously for the generation of EGF-X (data not shown). These experiments indicate that the changes leading to the formation of this new EGF species are influenced by the nature of the three-dimensional conformation of the growth factor.

The data obtained thus far indicate that EGF-X is formed by a chemical change in the aspartyl residue at position 11 in EGF. It is known that some age-damaged polypeptides can be substrates for L-isoaspartyl/D-aspartyl methyltransferase [EC 2.1.1.77] (Aswad, 1984; Clarke, 1985; DiDonato et al., 1986; Ota & Clarke, 1989a). This enzyme is a type II protein carboxyl methyltransferase, catalyzing the transfer of a methyl group from *S*-adenosyl-L-methionine to the α-carboxyl group of L-isoaspartyl residues and the β-carboxyl group of D-aspartyl residues in polypeptides and proteins. Because of its ability to methylate L-isoaspartyl residues in a wide variety of different peptide sequences (McFadden & Clarke, 1986; Aswad & Johnson, 1987; Lowenson & Clarke, 1990), this methyltransferase can be used in vitro as an analytical probe for the presence of these residues in polypeptides. We tested both native EGF and the EGF-X material for the ability to act as methylation substrates. While native EGF was not recognized as a substrate, we found that EGF-X was methylated, albeit poorly. For example, when 20 pmol of EGF-X (40 μM) was incubated for 240 min with 10 μM *S*-adenosyl[*methyl*-^{14}C]-methionine and 8.35 units of methyltransferase, conditions that would result in the stoichiometric methylation of most L-isoaspartyl-containing peptides (McFadden & Clarke, 1986;

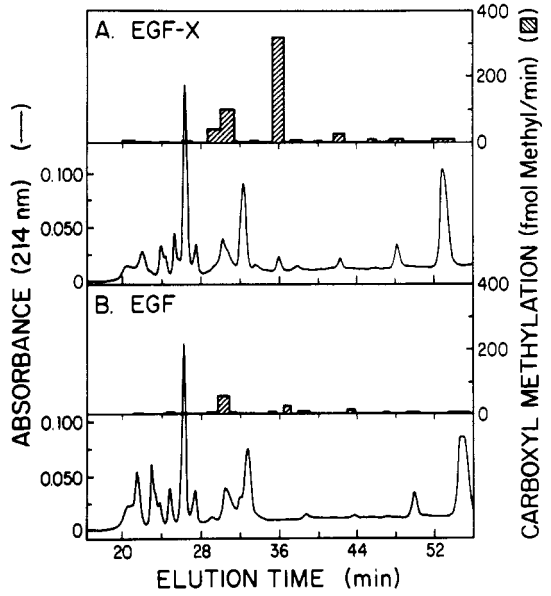

FIGURE 4: Thermolytic digests of native EGF and EGF-X and methylation of these peptides. Thermolytic digests prepared as described under Experimental Procedures were loaded onto an Alltech Econosphere C18 column equilibrated in 0.1% (w/v) trifluoroacetic acid (buffer A). The peptides were eluted with a linear increase in buffer B [0.1% (w/v) trifluoroacetic acid and 90% acetonitrile] of 1%/min and monitored at 214 nm. The peptide peaks were collected individually, and 6% of each peptide was assayed for its activity as a methyl-accepting substrate for the methyltransferase. Base-labile methyl groups were measured by a vapor diffusion assay as described under Experimental Procedures. The major methyl-acceptor peptide eluting at 36 min in the EGF-X digest is L-S-H-isoD-G-Y-C A-C. The minor methyl acceptor eluting at 31 min is the N-terminal peptide, N-S-D-S-E-C-(P) V-C. Given the flexibility of this portion of the polypeptide (Makino et al., 1987), the susceptibility of Asn-1 and Asp-3 to succinimide formation (Geiger & Clarke, 1987; Stephenson & Clarke, 1989), and the presence of this site of methylation in both digests, it is possible that a small number of isoaspartyl residues arose in this peptide during the 24-h proteolytic incubation.

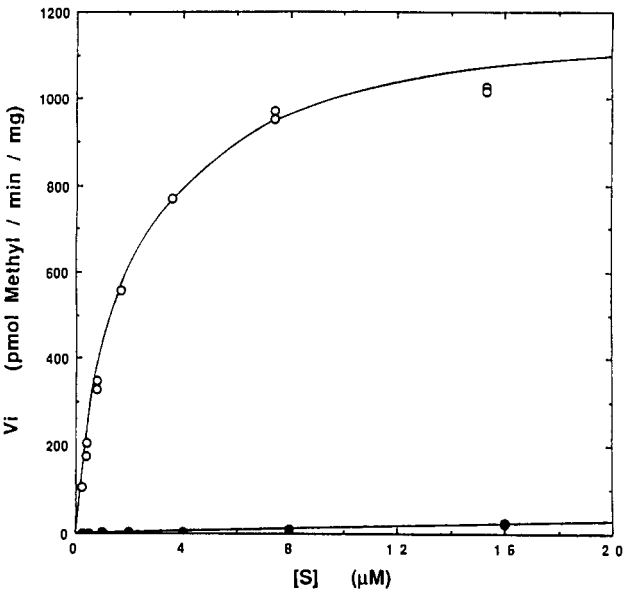

FIGURE 5: An octapeptide derived from EGF-X is a much better methylation substrate than EGF-X itself. The initial velocity of methyltransferase activity is measured at varying concentrations of either octapeptide (Leu-Ser-His-isoAsp-Gly-Tyr-Cys-Leu) (○) or EGF-X (●). The background from incubations lacking substrate has been subtracted from the experimental values. The curve drawn through the octapeptide points is calculated by assuming Michaelis–Menten kinetics, a K_m of 2.5 μM, and a V_{max} of 1208 pmol of methyl min^{-1} (mg of protein)$^{-1}$.

Aswad & Johnson, 1987; Lowenson & Clarke, 1990), only 4% of the EGF-X molecules were methylated (data not shown).

To examine whether secondary structure was responsible for the poor methylation of EGF-X, performic acid was used to oxidize the disulfide bonds and denature the molecule. Oxidized EGF-X was found to be an approximately 2-fold better methyl acceptor than unoxidized EGF-X (data not shown), suggesting that the folding of the polypeptide does have an inhibitory effect on methylation. To clarify this result, two different approaches were taken. The first method involved digestion with thermolysin, a protease that cleaves EGF molecules containing intact disulfide bonds (Savage et al., 1973). This eliminates the need to oxidize or carboxymethylate the polypeptide prior to proteolysis, procedures which may themselves generate methylatable residues. The reverse-phase HPLC maps of thermolytic peptides of EGF and EGF-X are quite similar with the exception of a peak at 36 min unique to EGF-X (Figure 4). When a small aliquot (approximately 6%) of each of the collected peptides was tested as a substrate for the methyltransferase, the 36-min peptide peak was found to be the major methyl acceptor (Figure 4). Amino acid analyses of this peptide indicated that it had the structure L-S-H-D-G-Y-C A-C, consisting of Leu-8 through Cys-14 and Ala-30 through Cys-31, joined by the disulfide bond. The K_m for methylation of this peptide was 5.3 μM (data not shown), a value which is comparable to the K_m values of other small isoaspartyl-containing peptides previously reported (McFadden & Clarke, 1986; Aswad & Johnson, 1987; Lowenson & Clarke, 1990). The absence of a corresponding peak in the native EGF digest indicates that the peptide with a normal aspartyl residue at position 11 was probably further cleaved by thermolysin (possibly at Tyr-13) into smaller fragments.

The second approach to the problem of assessing the effect of secondary structure on the methylation of isoaspartyl 11 involved the synthesis of an octapeptide, Leu-Ser-His-isoAsp-Gly-Tyr-Cys-Leu, equivalent to residues 8–15 in reduced EGF-X. As seen in the plot of initial velocity of methylation versus substrate concentration (Figure 5), this octapeptide is a much better methylation substrate than intact EGF-X and is comparable to the proteolytic peptide containing isoaspartyl 11. The K_m of the octapeptide for the methyltransferase is 2.5 μM, demonstrating that the additional residues in the proteolytic peptide attached through the disulfide bond have little, if any, effect on the methylation of the isoaspartyl residue in this sequence. Although the K_m of EGF-X methylation cannot be calculated from the data presented in Figure 5, other experiments indicate that this K_m is greater than 800 μM (data not shown). These results further support the hypothesis that an isoaspartyl residue exists at position 11 in EGF-X and that the secondary structure of the polypeptide around this site inhibits its recognition by the methyltransferase. A similar pattern has been seen by Galleti et al. (1988) with a form of bovine seminal ribonuclease containing an isoaspartyl residue at position 67.

Substitution of Aspartyl 11 in Human EGF. To determine if EGF can be stabilized by the removal of the aspartyl residue at position 11, a variant of EGF having a glutamic acid residue at this position was constructed. This material, called D11E-EGF, was produced by site-directed mutagenesis in transformed yeast and is secreted into the growth medium.

Purified D11E-EGF was resolved into two peaks by reverse-phase HPLC; the major peak, eluting 15 s later than that of wild-type EGF, was unmodified D11E-EGF, while the

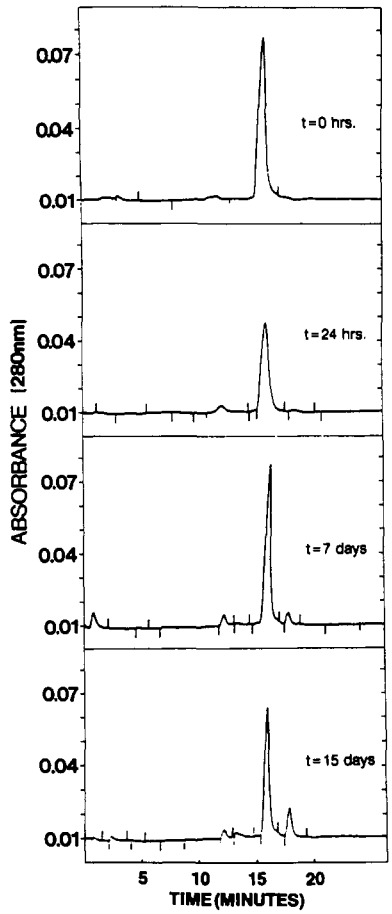

FIGURE 6: Reverse-phase HPLC monitoring of the stability of D11E-EGF upon incubation at pH 3.0 at 45 °C. The EGF variant (main peak) containing a Glu residue at position 11 instead of Asp was analyzed at the times indicated in the figure. The conditions used for each chromatogram were identical with those described in the legend of Figure 1. The new peak appearing 2 min after the main peak is a cleavage product lacking the first three residues.

minor peak contained D11E-EGF in which the methionine at position 21 is oxidized to the sulfoxide, as was seen in the wild-type recombinant EGF (George-Nascimento et al., 1988). Amino-terminal sequence analysis of D11E-EGF showed a sequence identical with that of native human EGF with the exception of the presence of a glutamic acid at position 11 (data not shown). A chymotryptic peptide map of D11E-EGF yielded the same pattern as native EGF with the exception of a single new peptide peak corresponding to the first 15 residues of EGF. This peptide contained one extra Glu residue and one less Asp residue, as expected for the Asp → Glu mutation at position 11. Taken together, these results indicated that D11E-EGF has the expected mutation, namely, a Glu residue instead of Asp at position 11. Mitogenic assays of D11E-EGF showed equipotency with EGF in stimulating [^{3}H]thymidine incorporation into DNA (Figure 2).

In order to check for the formation of an EGF-X-like species from D11E-EGF, the polypeptide was dissolved in 0.1 M citrate buffer, pH 3.0, and incubated at 45 °C. As previously established, these are the optimal conditions for formation of EGF-X from the native growth factor. The reaction was monitored by reverse-phase HPLC for a period of 15 days, which is much longer than the 3-day study presented in Figure 1. The results of this experiment, depicted in Figure 6, reveal that D11E-EGF degrades at a much slower rate than native human EGF. The amount of degradation of D11E-EGF observed after 7 days is approximately the same as occurs with wild-type EGF in 1 day. Amino-terminal sequencing of the major D11E-EGF degradation product eluting at about 18 min (Figure 6) shows that, unlike with EGF-X, there is no modification present that blocks the Edman reaction (data not shown). Most of this degradation product, however, is lacking the three N-terminal residues as was seen in fraction of the EGF-X molecules. Furthermore, the mitogenic activity of the D11E-EGF did not change with time, in agreement with previous data indicating that changes in the first few residues of EGF do not affect its biological activity (Simpson et al., 1985). Thus, the presence of the glutamyl residue, by preventing imide formation and isomerization at position 11, stabilizes the biological activity of the growth factor but does not alter the N-terminal processing.

Discussion

We have found that recombinant human EGF is susceptible to a nonenzymatic degradation reaction when stored in solution, resulting in the conversion of the aspartyl residue at position 11 to an isoaspartyl residue. Although other mechanisms cannot be ruled out, this reaction probably involves the attack of the aspartyl side-chain carbonyl carbon by the peptide-bond nitrogen of the glycyl residue at position 12 to form a succinimidyl residue (Clarke, 1987; Geiger & Clarke, 1987; Ota et al., 1987; Galletti et al., 1988; Kossiakoff, 1988; Johnson et al., 1989; Ota & Clarke, 1989b; Stephenson & Clarke, 1989; Martin et al., 1990). This succinimide then undergoes spontaneous hydrolysis to either the normal aspartyl or isoaspartyl form. The isoaspartyl-containing EGF (EGF-X) elutes several minutes later than native EGF from a reverse-phase HPLC column, presumably because the isoaspartyl residue perturbs the folding of the polypeptide, increasing its hydrophobicity. The only other covalent modification detectable in EGF-X is the cleavage of three residues from the N-terminus in a fraction of the polypeptides, probably due to spontaneous acid-catalyzed aspartyl-specific peptide hydrolysis (Schultz, 1967). Because of the hydrophilic nature of these hydrolyzed residues, however, their loss does not affect the elution time of EGF-X (Guo et al., 1986).

To prove that the altered residue at position 11 is actually an isoaspartyl, we used the L-isoaspartyl/D-aspartyl methyltransferase from human erythrocytes as an analytical probe. We found that EGF-X, but not native EGF, is a substrate for this methyltransferase, confirming the presence of an isoaspartyl residue in EGF-X. The site of methylation was pinpointed to position 11 with a thermolytic peptide of EGF-X in which residue 11 contained the only carboxyl side chain (i.e., the only potential methylation site). Interestingly, whereas intact EGF-X is very poorly recognized by the methyltransferase, its affinity was increased by unfolding the polypeptide with performic acid oxidation. The thermolytic peptide, which contains the isoaspartyl residue but presumably no secondary structure, is an even better methylation substrate, having a K_m of 5.3 μM. Thus, it appears that the folded structure of EGF-X inhibits the recognition of the isoaspartyl residue by the methyltransferase. Similar conformational effects have been seen with isoaspartyl-containing forms of mouse EGF (Galletti et al., 1989), human growth hormone (Johnson et al., 1989), and bovine seminal ribonuclease (Galletti et al., 1988). It has been demonstrated that the methylation of an isoaspartyl residue can lead to either its "repair" to the normal configuration (Clarke, 1985; Johnson et al., 1987; McFadden & Clarke, 1987) or its degradation (Momand & Clarke, 1987) in vitro. Our findings with EGF-X, however, join several previous reports (Galletti et al., 1988, 1989; Lowenson & Clarke, 1990) in suggesting that a subset of the damaged residues will be poorly recognized by

the methyltransferase and thus may accumulate in vivo.

When one examines the crystal structures of a variety of proteins, it becomes clear that nearly all aspartyl and asparaginyl residues are constrained in conformations that minimize succinimide formation (Clarke, 1987). In fact, recently identified isoaspartyl residues in polypeptides have been shown to occur primarily in regions of suspected backbone flexibility (Johnson et al., 1989; Ota & Clarke, 1989b). In contrast, our results with EGF suggest that secondary and/or tertiary structure is critically involved with isoaspartyl formation at position 11. First, denaturation of EGF by reduction of its disulfide bonds prevents this isoaspartyl formation, meaning that the lability of the aspartyl residue is not an inherent property of the amino acid sequence around position 11. The second factor is the unexpected pH dependence of this reaction. The larger the fraction of peptide that contains both a protonated aspartyl side-chain carboxyl group and a deprotonated backbone amide nitrogen, the faster should be the rate of succinimide formation. This fraction should be smaller at pH 3 than at higher pH values. The measured rate with EGF, however, was found to be fastest at pH 3, indicating that some other factor, most likely the polypeptide's conformation, is affecting succinimide formation in a pH-dependent fashion. Finally, the residues in positions 17–19 are identical with those in positions 10–12, and yet, no isoaspartate has been detected at position 18. According to NMR spectroscopy studies (Carver et al., 1986; Cooke et al., 1987; Makino et al., 1987), aspartyl 11 is in close proximity to tyrosyl residues 22 and 29, which may enclose it in hydrophobic pocket. In contrast, aspartyl 18 is believed to be more exposed to the solvent. The residues surrounding aspartyl 11 might chemically assist succinimide formation (e.g., by protonating the aspartyl side chain) or physically bring the reactive groups into closer contact, accounting for the greater lability of residue 11. A better understanding of the factors promoting succinimide formation in this relatively simple polypeptide should be useful in identifying labile aspartyl residues in larger protein structures.

When recombinant proteins are to be used as pharmaceuticals, their stability during storage is of critical importance. We have observed that the biological activity of stored EGF slowly decays with time. To determine whether this loss in activity might be due to the formation of EGF-X, the activity of this degradation product was assayed. We found that EGF-X has only 20% of the mitogenic activity of native EGF, apparently because its ability to bind to the EGF receptor is diminished by the same amount. This is consistent with a recent report that described similar studies of the stability of recombinant EGF in solution (Araki et al., 1989). These authors found that the ability of their major EGF degradation product to bind to the EGF receptor was only 18% that of the native polypeptide. They suggested that an isoaspartyl residue might be present at position 11, but gave no experimental evidence to support this. Another recent report shows that the major degradation reactions in stored recombinant human growth hormone are the formation of two isoaspartyl residues (Johnson et al., 1989). While one of these residues arises from the deamidation of an asparaginyl residue, the other originates from an aspartylglycine sequence like that in EGF at positions 11 and 12. It is thus clear that isoaspartyl formation is one of the primary degradation reactions in stored polypeptides that must be contended with if the biological potency is to be maintained.

To investigate the possibility of stabilizing EGF, we used site-directed mutagenesis to replace the aspartyl residue at position 11 with a glutamyl residue. The biological activity of this altered EGF is the same as that of the wild-type polypeptide; the extra methylene group in the glutamyl side chain apparently does not affect interaction with the EGF receptor. Incubation of D11E-EGF in citrate buffer, followed by HPLC analysis, showed that the altered form is much more stable than wild-type EGF under these conditions. Thus, the D11E-EGF could be stored for a longer period of time before its eventual use as a pharmaceutical. Furthermore, the stability of other recombinant polypeptides might also be improved by the replacement of labile aspartyl and asparaginyl residues by glutamyl and glutaminyl residues, respectively.

Acknowledgments

We thank Ellen Freiberg and Diane Lee for performing the mitogenic assays and Lane Copley for peptide mapping D11E-EGF. We also thank Lawrence Cousens, Frank Marcus, and Pablo Valenzuela for help in preparing the manuscript and Frank Masiarz for suggesting the formation of isoaspartic acid in growth factors.

References

Araki, F., Nakamura, H., Nojima, N., Tsukumo, K., & Sakamoto, S. (1989) *Chem. Pharm. Bull. 37*, 404–406.
Aswad, D. W. (1984) *J. Biol. Chem. 259* (17), 10714–10721.
Aswad, D. W., & Johnson, B. A. (1987) *Trends Biochem. Sci. 12*, 155–158.
Brake, A. J., Merryweather, J. P., Coit, D. S., Heberlein, U. A., Masiarz, F. R., Mullenback, G. T., Urdea, M. S., Valenzuela, P., & Barr, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4642–4646.
Carver, J. A., Cooke, R. M., Esposito, G., Campbell, I. D., Gregory, H., & Sheard, B. (1986) *FEBS Lett. 205*, 77–81.
Clarke, S. (1985) *Annu. Rev. Biochem. 54*, 479–506.
Clarke, S. (1987) *Int. J. Pept. Protein Res. 30*, 808–821.
Cohen, S. (1962) *J. Biol. Chem. 237*, 1555–1562.
Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature 327*, 339–341.
DiAgustine, R. P., Gibson, B. W., Aberth, W., Kelly, M., Ferrua, A. M., Tomooka, Y., Brown, C. F., & Walker, M. (1987) *Anal. Biochem. 165*, 420–429.
DiDonato, A., Galletti, P., & D'Alessio, G. D. (1986) *Biochemistry 25*, 8361–8368.
Galletti, P., Ciardiello, A., Ingrosso, D., DiDonato, A., & D'Alessio, G. (1988) *Biochemistry 27*, 1752–1757.
Galletti, P., Iardino, P., Ingrosso, D., Manna, C., & Zappia, V. (1989) *Int. J. Pept. Protein Res. 33*, 397–402.
Geiger, T., & Clarke, S. (1987) *J. Biol. Chem. 262*, 785–794.
George-Nascimento, C., Gyenes, A., Halloran, S. M., Merryweather, J., Valenzuela, P., Steimer, K. S., Masiarz, F. R., & Randolph, A. (1988) *Biochemistry 27*, 797–802.
Gilbert, J. M., Fowler, A., Bleibaum, J., & Clarke, S. (1988) *Biochemistry 27*, 5227–5233.
Gregory, H. (1975) *Nature 257*, 325–327.
Guo, D., Mant, C. T., Taneja, A. K., Parker, J. M. R., & Hodges, R. S. (1986) *J. Chromatogr. 359*, 499–517.
Johnson, B. A., Langmack, E. L., & Aswad, D. W. (1987) *J. Biol. Chem. 262*, 12283–12287.
Johnson, B. A., Shirokawa, J. M., Hancock, W. S., Spellman, M. W., Basa, L. J., & Aswad, D. W. (1989) *J. Biol. Chem. 264*, 14262–14271.
Kossiakoff, A. A. (1988) *Science 240*, 191–194.
Lowenson, J., & Clarke, S. (1990) *J. Biol. Chem. 265*, 3106–3110.

Makino, K., Morimoto, M., Nishi, M., Sakamoto, S., Tamura, A., Inooka, H., & Akasaka, K. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7841–7845.
Martin, B. L., Wu, D., Tabatabi, L., & Graves, D. J. (1990) *Arch. Biochem. Biophys. 276*, 94–101.
McFadden, P. N., & Clarke, S. (1986) *J. Biol. Chem. 261*, 11503–11511.
McFadden, P. N., & Clarke, S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2595–2599.
Momand, J., & Clarke, S. (1987) *Biochemistry 26*, 7798–7805.
Murray, E. D., Jr., & Clarke, S. (1984) *J. Biol. Chem. 259*, 10722–10732.
Oka, Y., & Orth, D. N. (1983) *J. Clin. Invest. 72*, 249–259.
O'Keefe, J. H., Fifis, T., Bender, V. J., & Moss, B. A. (1984) *J. Chromatogr. 25*, 45–59.
Ota, I., & Clarke, S. (1989a) *Biochemistry 28*, 4020–4027.
Ota, I., & Clarke, S. (1989b) *J. Biol. Chem. 264*, 54–60.
Ota, I., Ding, L., & Clarke, S. (1987) *J. Biol. Chem. 262*, 8522–8531.
Petrides, P. E., Böhlen, P., & Shively, J. E. (1984) *Biochem. Biophys. Res. Commun. 125*, 218–228.
Savage, C. R., Jr., Inagami, T., & Cohen, S. (1972) *J. Biol. Chem. 247*, 7612–7621.
Savage, C. R., Jr., Hash, J. H., & Cohen, S. (1973) *J. Biol. Chem. 248*, 7669–7672.
Schultz, J. (1967) *Methods Enzymol. 11*, 255–263.
Simpson, R. J., Smith, J. A., Moritz, R. L., O'Hare, M. J., Rudland, P. S., Morrison, J. R., Lloyd, C. L., Grego, B., Burgess, A. W., & Nice, E. C. (1985) *Eur. J. Biochem. 153*, 629–637.
Smith, J., Cook, E., Fotheringham, I., Pheby, S., Derbyshire, R., Eaton, M. A. W., Doel, M., Lilly, D. M., Pardon, J. F., Patel, T., Lewis, M., & Bell, L. D. (1982) *Nucleic Acids Res. 10*, 4467–4482.
Smith, J. A., Ham, J., Winslow, D. P., O'Hare, M. J., & Rudland, P. S. (1984) *J. Chromatogr. 305*, 295–308.
Stephenson, R. C., & Clarke, S. (1989) *J. Biol. Chem. 264*, 6164–6170.
Travis, J., Owen, M., George, P., Carrel, R., Rosenberg, S., Hallewell, R. A., & Barr, P. J. (1985) *J. Biol. Chem. 260*, 4384–4389.

Folding of a Predominantly β-Structure Protein: Rat Intestinal Fatty Acid Binding Protein†

Ira J. Ropson,‡ Jeffrey I. Gordon,‡,§ and Carl Frieden*,‡
Department of Biochemistry and Molecular Biophysics and Department of Medicine, Washington University Medical School, St. Louis, Missouri 63110
Received June 20, 1990

ABSTRACT: The equilibrium and kinetic properties of the unfolding–refolding transitions of *Escherichia coli* derived rat intestinal fatty acid binding protein have been examined using several different denaturants. This protein, which contains 2 tryptophans but no prolines or cysteines, has a predominantly β-structure: its 10 antiparallel β-strands are organized into 2 orthogonal sheets surrounding a large solvent-filled internal cavity. For urea and guanidine hydrochloride, the completely reversible transition was monitored by circular dichroism, absorbance, and fluorescence spectroscopy. Each of these data sets was best fit by a simple, two-state model involving only native and unfolded forms. However, linear extrapolation to determine the free energy of folding in the absence of denaturant resulted in different values for the free energy of folding depending upon which denaturant was used. When fluorescence was used to monitor the transition, the extrapolated free energy estimates for the two denaturants were markedly different: 10.03 ± 0.24 kcal mol^{-1} for urea versus 5.22 ± 0.33 kcal mol^{-1} for guanidine hydrochloride. The midpoints of these transitions were 5.51 and 1.36 M, respectively. The transition caused by either denaturant as monitored by circular dichroism and absorbance spectroscopy was virtually coincident with that monitored by fluorescence, further supporting the assignment of a two-state model for the equilibrium results. The addition of a 2-fold molar excess of ligand (oleate) increased the extrapolated estimates approximately 2.5 kcal mol^{-1} for both denaturants. Stop-flow kinetic studies of the guanidine hydrochloride induced transitions showed that the unfolding and folding processes are rapid (complete in less than a minute under the slowest conditions) and complex, indicating the presence of intermediates on both pathways. The unfolding process was biphasic: the two phases accounted for the entire amplitude of the transition, suggesting that some type of sequential unfolding of the protein occurs. Two phases were also observed for the folding reaction, but at least one additional phase must occur in the dead time of the instrument (<10 ms) to account for the expected amplitude change. The changes in the amplitudes of the folding phases suggest the existence of multiple pathways for the folding of rat intestinal fatty acid binding protein and that the proportion of protein molecules following any particular pathway is dependent on the final denaturant concentration.

One of the fundamental questions of biochemistry is how the primary sequence of a protein encodes the information leading to the formation of an organized structure capable of function. Although considerable progress has been made, particularly for a few carefully studied proteins, the general rules of protein folding have not been determined. This is particularly true for proteins that consist primarily of β-sheet structure, since so few have been examined. This paper begins

† Supported by NIH Grants DK13332 to C.F. and DK30292 to J.I.G.
* Author to whom correspondence should be addressed.
‡ Department of Biochemistry and Molecular Biophysics.
§ Department of Medicine.

0006-2960/90/0429-9591$02.50/0