Characterization of Recombinant Human Epidermal Growth Factor Produced in Yeast

Carlos George-Nascimento,* Alexander Gyenes, S. Mitchell Halloran, James Merryweather, Pablo Valenzuela, Kathelyn S. Steimer, Frank R. Masiarz, and Anne Randolph

Chiron Research Laboratories, Chiron Corporation, 4560 Horton Street, Emeryville, California 94608

Received June 25, 1987; Revised Manuscript Received September 24, 1987

ABSTRACT: Four different forms of human epidermal growth factor (h-EGF) are found in the culture medium of a recombinant strain of *Saccharomyces cerevisiae*. These forms were characterized after purification using reverse-phase high-performance liquid chromatography. The most abundant form of secreted recombinant h-EGF has leucine at the carboxyl terminus and is identical with γ-urogastrone. A second species is identical with the most abundant form except that it lacks the carboxyl-terminal leucine. This form appears to be the product of a carboxypeptidase found in the growth medium. The other two forms of recombinant h-EGF are the respective oxidation products of the above where the single methionine residue has been converted to methionine sulfoxide. These four forms of recombinant h-EGF are fully active; they bind to the EGF receptor of A431 cells as well as stimulate mitotic activity of human foreskin fibroblasts with equal specific activity. The location of the disulfide bonds in the predominant form of recombinant h-EGF was determined following digestion with thermolysin. The amino acid compositions of the resulting peptides showed that the placement of disulfide bonds in recombinant h-EGF is identical with that in murine EGF.

Epidermal growth factor (EGF)[1] is a single-chain polypeptide first isolated from the submaxillary glands of adult mice (Cohen, 1962). Epidermal growth factor stimulates the proliferation of many cell types, mainly those of epithelial and epidermal tissues. Several metabolic events, such as stimulation of protein and DNA synthesis (Carpenter & Cohen, 1979), stimulation of glucose intake (Barnes & Colowick, 1976), and ion fluxes as well as Ca^{2+} transport (Rozengurt & Heppel, 1975), have been described as specific effects of h-EGF. In 1975, Cohen and Carpenter (1975) and Gregory (1975) reported the purification and properties of human EGF (h-EGF) from concentrates of urine. Human EGF was identified later in most body fluids and is particularly high in milk, colostrum, urine, and seminal fluid (Oka & Orth, 1983; Carpenter, 1980, 1985). In most cases, more than one form of the EGF molecule has been described. A short version of the molecule (γ-urogastrone) has been found in urine. Milk appears to contain four molecular forms of EGF, all capable of competing with murine EGF for placental receptors. These four forms of h-EGF have not been characterized further (Petrides et al., 1985). Four forms of EGF have been purified from rat submaxillary glands (Simpson et al., 1985). All 4 rat-derived forms of EGF were found to differ in length, with the shortest having 44 amino acids and the longest having 48 amino acids. Three independent groups have reported the presence of two molecular forms of murine EGF (Burgess et al., 1984; Petrides et al., 1984; DiAugustine et al., 1985). In addition to the complete molecule (53 amino acids), an EGF lacking the amino-terminal residue (des-Asn) has been isolated; these 2 mouse-derived EGF forms are biologically active. Recently, full-length murine EGF as well as several truncated forms of the molecule has been chemically synthesized (Heath & Merrifield, 1986).

The h-EGF gene has been chemically synthesized and expressed in bacteria (Smith et al., 1982; Oka et al., 1985) as well as yeast (Urdea et al., 1983; Brake et al., 1984). The fusion product of the α-factor leader and the h-EGF gene is efficiently processed by yeast cells, resulting in the secretion of h-EGF into the medium. The production of large quantities of h-EGF by recombinant DNA techniques has allowed us to perform analyses on this polypeptide.

In this paper, we report the purification and characterization of h-EGF secreted by the yeast *Saccharomyces cerevisiae*. We demonstrate that this growth factor, synthesized by transformed yeast, has the correct placement of the three disulfide bridges with no free sulfhydryl groups. It is also shown that there are four molecular forms of h-EGF found in the growth medium of the transformed yeast and that these forms differ in their carboxyl-terminal residue as well as in the oxidation state of their single methionine residue.

Experimental Procedures

Growth of Cells. *Saccharomyces cerevisiae*, strain AB103.1, was used throughout this study (Brake et al., 1984). The cells were grown either in flasks containing 1 L of medium at 30 °C/300 rpm or in fermentors of 16 or 200 L.

Isolation of Recombinant h-EGF. Cell-free medium was obtained by centrifuging the culture in a continuous-flow centrifuge (CEPA LE or CEPA Z61). The culture supernatant was concentrated by using an Amicon ultrafiltration apparatus (Lexington, MA) equipped with a 2000 molecular weight cutoff membrane, and the proteins were chromatographed on a gel filtration column (Bio-Gel P-10/Bio-Rad Labs, Richmond, CA) as previously described (Savage & Harper, 1981). Separation of the four EGF species was carried out by reverse-phase HPLC using a Beckman Model 344 system equipped with a Vydac, 5-μm C4 column (4.6 × 250 mm). The column was operated at a flow rate of 0.8 mL/min and equilibrated in 75% mobile phase A/25% mobile phase B; mobile phase A is 5% acetonitrile (ACN) containing 0.05% trifluoroacetic acid (TFA), and mobile phase B is 80% ACN containing 0.05% TFA. Samples were eluted with a 55-min linear gradient from 25% mobile phase B to 40% mobile phase B, and the eluent was continuously monitored for protein at 280 nm.

[1] Abbreviations: h-EGF, human epidermal growth factor; HPLC, high-performance liquid chromatography; ACN, acetonitrile; TFA, trifluoroacetic acid; PITC, phenyl isothiocyanate; CNBr, cyanogen bromide; DMEM, Dulbecco's modified Eagle's medium.

0006-2960/88/0427-0797$01.50/0

Amino Acid Analysis. Amino acid compositions were obtained by using the Pico-Tag method of Waters Associates (Milford, MA) (Bidlingmeyer et al., 1984) which involves precolumn derivatization of the amino acids with phenyl isothiocyanate (PITC) followed by separation of the resulting phenylthiocarbamyl-amino acids by reverse-phase HPLC using a Waters 840 chromatography system. Analyses were performed on protein samples hydrolyzed in the vapors of constant-boiling HCl (Pierce Chemical Co., Rockford, IL) containing 1% (v/v) phenol for 22–24 h, in vacuo, at 110 °C, or on samples hydrolyzed in 4 M methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole (Pierce Chemical Co., Rockford, IL) for 20 h, in vacuo, at 110 °C. Samples hydrolyzed in 4 M methanesulfonic acid were neutralized with 4 M KOH and dried prior to derivatization with PITC as described by Bidlingmeyer et al. (1984).

Sequence Analysis. Automated Edman degradation was carried out by using an Applied Biosystems Model 470A gas-phase protein sequencer equipped with a 120A on-line phenylthiohydantoin–amino acid analyzer (Applied Biosystems, Foster City, CA).

Cyanogen Bromide Cleavage. Human EGF samples were dissolved in 70% formic acid (Eastman Kodak Co., Rochester, NY) (0.5 mg of h-EGF/mL), and crystalline CNBr (Pierce Chemical Co., Rockford, IL) was added (greater than 10-fold excess over h-EGF by weight). The reaction was allowed to proceed under argon in the dark at room temperature. An additional aliquot of CNBr was added to the reaction mixture after 24 h. After 48 h, the reaction mixture was dried in a Speed Vac (Savant Instruments Inc., Hicksville, NY) following dilution with water.

Carboxyl-Terminal Sequence Analysis. Carboxypeptidase A from bovine pancreas (Boehringer Mannheim Biochemicals, Indianapolis, IN) (20 μL) was solubilized by the addition of 1% sodium carbonate (40 μL) followed by the addition of 40 μL of freshly prepared 0.1 N NaOH. Forty microliters of this solution was added to 140 μg of h-EGF dissolved in 200 μL of 0.2 M *N*-ethylmorpholine acetate, pH 8.0, and the reaction mixture was incubated at 37 °C. Aliquots, removed at 0.5, 1, 2.5, 5, 10, and 20 min, were added to an equal volume of glacial acetic acid to stop the reaction. The aliquots were dried and derivatized with PITC as described by Bidlingmeyer et al. (1984), and the amino acids released were identified and quantitated (as for amino acid analysis). The chromatographic behavior of the recombinant h-EGF at each time point was also analyzed by reverse-phase HPLC (see Isolation of Recombinant h-EGF).

Localization of Disulfide Bonds. Purified recombinant h-EGF (2.5 mg) was dissolved in 250 μL of 0.1 M pyridine acetate, pH 6.5, containing 3% thermolysin (w/w) (Boehringer Mannheim), and the mixture was incubated at 42 °C for 40 h and then dried. The resulting peptides were separated by using a Techsphere Ultra, 5-μm C18 column (Phenomenex, Rancho Palos Verdes, CA) which was equilibrated in 0.1% TFA and operated at a flow rate of 0.8 mL/min. An aliquot of the digest (600 μg in 25 μL of 0.1% TFA) was injected on the column, and the peptides were eluted with a 2-h linear gradient to 80% ACN containing 0.1% TFA. Peptide-containing fractions were collected, and their amino acid compositions were determined as described above.

Mitogen Assay for EGF Biological Activity. Human foreskin fibroblasts (2×10^4 cells/well) were plated in 96-well microtiter plates in 0.1 mL of Dulbecco's modified Eagle's medium (DMEM) containing 5% fetal bovine serum. The plates were incubated for 5 days at 37 °C in a 7% CO_2 atmosphere to allow the cells to become confluent. Dilutions of samples containing EGF were made in DMEM (without serum) and added to the cultures. After incubation for 18 h, [^{3}H]thymidine (Amersham) was added to the wells (1 μCi/well). The plates were incubated for an additional 24 h; the cells were washed with phosphate-buffered saline and incubated for 15 min with 5% trichloroacetic acid. The precipitated cells were then treated with methanol for 15 min, air-dried, and solubilized in 0.3 N NaOH for scintillation counting. Background incorporation from wells that received diluent alone was subtracted from the data before calculating the average incorporation. Each data point is the average of triplicate assays. Background incorporation was usually in the range of 2000–4000 cpm/well, and the maximum incorporation was 40 000–60 000 cpm/well.

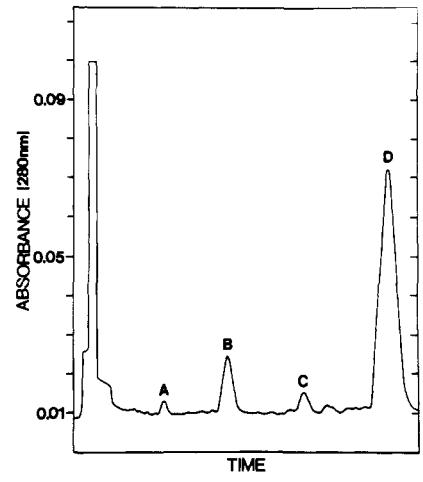

FIGURE 1: HPLC of recombinant h-EFG produced in yeast. Reverse-phase HPLC analysis of concentrated (14-fold) cell-free medium from a yeast culture producing recombinant h-EGF. The chromatography was performed on a 4.6 × 250 mm Vydac C4 column. Flow rate was 0.8 mL/min. Initial conditions are 75% A/25% B, where mobile phase A is 5% ACN/0.05% TFA and mobile phase B is 80% ACN/0.05% TFA. After the sample was loaded, a linear gradient was established from 25% B to 40% B in 55 min.

Receptor Binding Assays. Human EGF was assayed by competitive receptor binding using radiolabeled mouse EGF (Amersham) and human foreskin fibroblasts (Cohen & Carpenter, 1975).

RESULTS

Transformed yeast cells, strain AB103.1, actively secrete recombinant human EGF into the growth medium (Brake et al., 1984), and the secretion occurs at all times during cell growth. The recombinant h-EGF is purified by concentrating the medium by ultrafiltration followed by fractionation using gel filtration chromatography (Savage & Harper, 1981) and a final step using reverse-phase HPLC on a C4 column. Figure 1 shows a chromatogram of the yeast medium obtained 6 h after the cell culture reached stationary phase. Four distinct peaks (A, B, C, and D) of recombinant h-EGF were identified under these conditions. Each peak was isolated and analyzed for receptor binding activity (Cohen & Carpenter, 1975), mitogenic activity (Hirata & Orth, 1979), and amino acid composition.

As can be seen in Figure 2, all four h-EGF species are equivalent in stimulating [^{3}H]thymidine incorporation into foreskin fibroblasts. The concentrations of all four forms that stimulated 50% of the maximum [^{3}H]thymidine incorporation were similar to those found for murine EGF. The differences in incorporation observed among the h-EGF species are within the normal scatter of the mitogen assay. We have found that

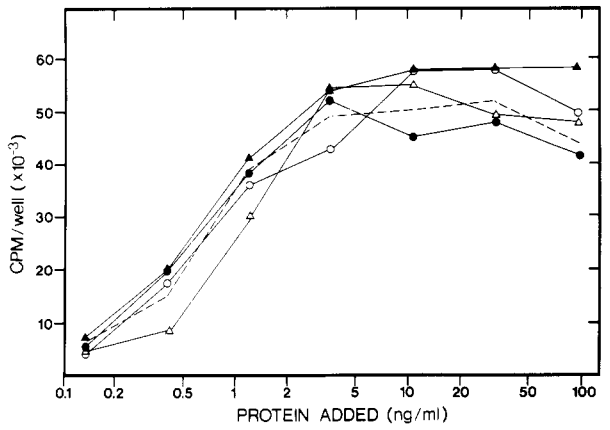

FIGURE 2: Mitogenic activity of purified recombinant human EGF fractions. Purified human EGF was tested for mitogenic activity using quiescent cultures of human foreskin fibroblasts (see Experimental Procedures). Each data point represents the average of triplicate assays. Samples assayed were recombinant human EGF-A (●), EGF-B (▲), EGF-C (△), and EGF-D (○) and natural EGF purified from mouse salivary glands (---) (Collaborative Research, Bedford, MA). Recombinant human EGF samples were quantitated by compositional analysis, and the amount of natural EGF was specified on the label.

Table I: Amino Acid Compositions of Four Forms of Recombinant h-EGF[a]

amino acid	expected	EGF-A	EGF-B	EGF-C	EGF-D
Asx	7	6.8	7.2	6.7	6.9
Glx	5	5.4	5.1	4.9	5.0
Ser	3	2.9	2.4	2.7	2.7
Gly	4	4.4	4.2	4.0	3.8
His	2	1.9	1.6	1.9	2.1
Arg	3	2.5	1.9	2.2	2.3
Thr	0	0.3	0.0	0.0	0.0
Ala	2	2.2	2.0	2.0	2.1
Pro	1	1.3	0.9	1.1	1.1
Tyr	5	4.5	4.8	5.0	5.1
Val	3	2.5	3.1	2.9	2.7
Met	1	0.7	1.1	1.3	1.0
Cys	6	ND[b]	ND	ND	ND
Ile	2	1.9	2.3	2.1	2.2
Leu	5	4.2	4.1	5.3	5.2
Phe	0	0.2	0.0	0.1	0.1
Trp	2	1.9	1.6	1.9	1.7
Lys	2	1.9	2.1	2.4	1.9

[a] The analyses were performed on samples hydrolyzed in vacuo in the vapors of constant-boiling HCl for 24 h at 110 °C. Tryptophan values were determined from samples hydrolyzed in methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole for 20 h, in vacuo, at 110 °C. Relative molar quantities are given for each amino acid. No corrections were made for time-dependent liberation or losses. The expected values were derived from the sequence of h-EGF (Gregory, 1975). [b] ND, not determined.

each one of these forms competes equally for binding to receptors present in A431 cells cultured in vitro (data not shown).

The amino acid compositions of the four h-EGF molecules (following hydrolysis with HCl) are shown in Table I. These data reveal that all four forms of recombinant h-EGF lack one arginine residue and that two forms (EGF-A and -B) contain only four leucine residues per molecule, while recombinant h-EGF species C and D contain the expected five leucine residues per molecule. No other differences in amino acid compositions were detected. Human EGF has a leucine residue at position 52 and an arginine residue at the carboxyl terminus (position 53) (Gregory & Preston, 1977). Therefore, the differences in amino acid composition among the four peaks of recombinant h-EGF can be explained by carboxyl-terminal processing, with all four forms missing the terminal

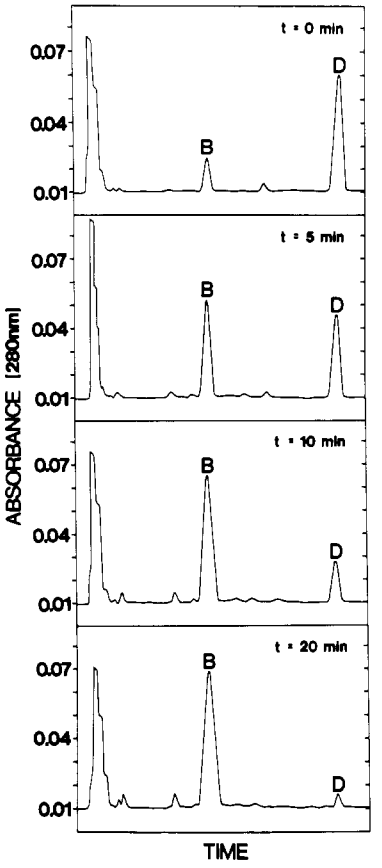

FIGURE 3: Kinetics of transformation of EGF-D into EGF-B by carboxypeptidase treatment. A mixture of EGF-D and EGF-B was treated with carboxypeptidase A as described under Experimental Procedures. Aliquots, taken at the indicated times, were analyzed by HPLC as described in the legend to Figure 1.

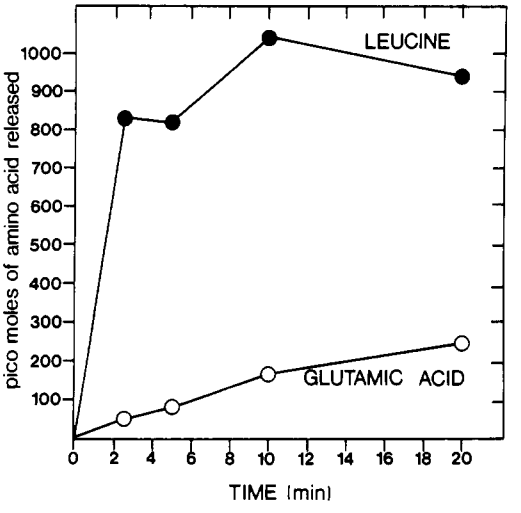

FIGURE 4: Release of amino acids from EGF-D by carboxypeptidase A. Purified EGF-D was treated with carboxypeptidase A. Aliquots (1 nmol) were removed at various times, and the amino acids released were identified and quantitated (see Experimental Procedures).

arginine residue and EGF-A and -B lacking the penultimate leucine residue.

We have observed that the ratio of EGF-D to EGF-B decreases with increasing time of fermentation. At the same time, the total amount of recombinant h-EGF produced remains approximately constant; therefore, it appears that EGF-D is converted to EGF-B by a carboxypeptidase-like activity in the medium. To substantiate further that EGF-D is the precursor of EGF-B, purified EGF-D was treated with

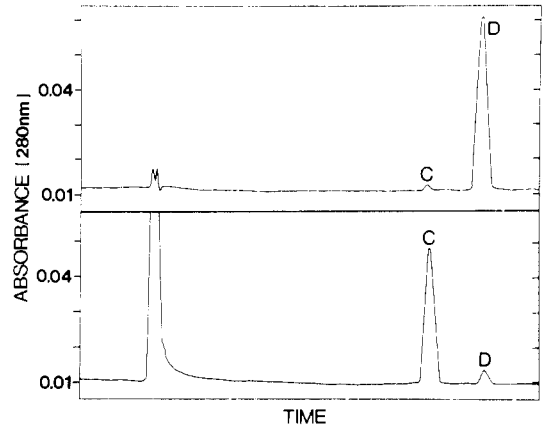

FIGURE 5: Formation of EGF-C by oxidation of EGF-D. EGF-D was treated with hydrogen peroxide (15% by volume) and immediately analyzed by HPLC as described in the legend to Figure 1.

carboxypeptidase A, and the kinetics of transformation of D into B is depicted in Figure 3. The data show that EGF-B is formed with time, and the concomitant reduction of EGF-D indicates a precursor–product relationship. Figure 4 shows the analysis of the amino acid residue being released by the carboxypeptidase treatment, revealing an increase of free leucine, derived from position 52 of the polypeptide, followed at a lower rate by the appearance of glutamic acid. The glutamic acid residue at position 51 is released upon further attack of the carboxypeptidase A on the recombinant h-EGF. It is clear that after 10 min of incubation, all of the leucine residues were hydrolyzed under these conditions. In addition, we have noticed an increase in EGF-A with time of fermentation and a corresponding decrease in EGF-C. Since amino acid analysis indicates that EGF-A (like EGF-B) lacks leucine-52, this effect is probably due to carboxyl-terminal processing in the medium. As expected, when purified EGF-C was treated with carboxypeptidase A, its transformation into EGF-A was observed (data not shown). In summary, the data indicate that EGF-D is a precursor of EGF-B and EGF-C is a precursor of EGF-A.

Highly purified EGF-D is stable when stored as a lyophilized powder; however, in solution, the appearance of EGF-C can be seen with time. In addition, more EGF-C is produced when the amount of aeration is increased during fermentation. These observations indicate that EGF-C might be the result of an oxidation of EGF-D. In order to test this hypothesis, purified EGF-D was subjected to oxidation. Figure 5 shows the kinetics of transformation of purified EGF-D into EGF-C upon oxidation with hydrogen peroxide. Identical results were obtained by using other oxidizing agents such as dimethyl sulfoxide. Clearly, EGF-D is transformed to EGF-C in the presence of the oxidizing reagent. When EGF-C (formed by the oxidation of EGF-D with dimethyl sulfoxide) is reduced with methyl sulfide, EGF-D is re-formed, demonstrating the reversibility of the oxidation (data not shown) (Shechter, 1986).

Human EGF contains a single methionine residue and two tryptophan residues that might be susceptible to oxidation. Amino acid analysis, following methansulfonic acid hydrolysis, shows the presence of two residues of tryptophan in all four forms of recombinant h-EGF, while the methionine content is considerably decreased in EGF-A and EGF-C (data not shown). These results provide a clear indication that only the methionine residue is oxidized. Since methionine sulfone is stable to HCl hydrolysis and amino acid analysis following HCl hydrolysis indicates the presence of a full residue of methionine in EGF-A, EGF-B, EGF-C, and EGF-D, the

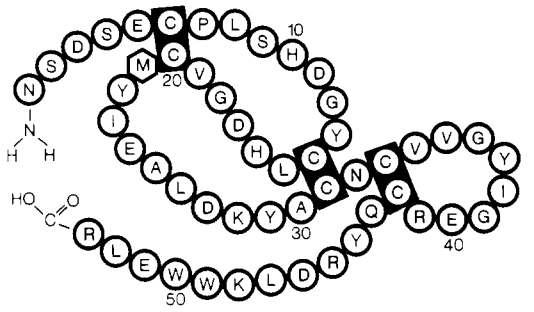

FIGURE 6: Representation of the structure of human EGF. Amino acid sequence and location of disulfide linkages of human EGF.

methionine residues in EGF-A and EGF-C must have been converted to methionine sulfoxide, not methionine sulfone.

To confirm the presence of oxidized methionine in EGF-A and EGF-C, we attempted cleavage of purified EGF-C and EGF-D with CNBr. After this treatment, both samples were subjected to automated Edman degradation. It is known that CNBr does not cleave after an oxidized methionine (Storring & Tiplady, 1984). Therefore, CNBr should not cleave EGF-C, resulting in only the normal amino-terminal sequence; cleavage following the methionine in EGF-D should give rise to a second amino terminus. This was found to be the case with the CNBr digest of EGF-C, showing only a single sequence, and that of EGF-D, having a double sequence. These sequences were the expected ones for the natural amino terminus as well as the new amino terminus formed after cleavage on the carboxyl-terminal side of the single methionine at position 21.

Location of Disulfide Bonds in Recombinant h-EGF. Figure 6 shows the amino acid sequence of h-EGF with disulfide bond placement by analogy with the pattern determined for murine EGF (Savage et al., 1973). In order to verify that the disulfide bonds of the recombinant h-EGF produced in yeast were correct, the predominant form (EGF-D) was digested with thermolysin, and the resulting peptides were separated by reverse-phase HPLC. The amino acid compositions of 10 peptide fractions are shown in Table II. The compositions of the peptides of peak 3 can be unambiguously assigned to a peptide containing cysteine-14 and cysteine-31 along with a non-cysteine-containing peptide (residues 44–46) in a ratio of 2:1. This confirms the disulfide bond between these two cysteine residues. Similarly, the compositions of peptides in peaks 5 and 6 confirm the presence of disulfide bonds connecting residues 33 to 44 and residues 6 to 20, respectively. Therefore, the data indicate that the location of disulfide bonds in recombinant h-EGF is identical with that found in murine EGF.

Discussion

Our study on the structure of the h-EGF molecules produced by transformed yeast cells revealed that four molecular forms of biologically active recombinant h-EGF are found in the culture supernatant. Previous work by Brake et al. (1984) showed that the EGF gene fused to the α-factor leader is efficiently expressed by yeast cells and the protein is secreted into the medium.

The reverse-phase HPLC procedure described here allows the purification of four peaks, each one containing a very specific form of EGF. The most abundant and most highly retained form is EGF-D (1–52) which has a leucine as a carboxyl-terminal residue. It is interesting to note that this species is identical with the natural product found in human urine (γ-urogastrone). Presumably, the carboxypeptidase that removes the basic residues of the spacer of the α-factor pre-

Table II: Amino Acid Compositions of EGF-D Peptides Generated by Thermolysin Digestion[a]

amino acid	peak 1	2	3	4	5	6	7	8	9	10
Asx	1.0 (1)	1.1 (1)	0.6 (0 + 1)[d]	0.1	1.3 (1)	2.0 (2)	0.3	0.1	0.1	
Glx	0.3		0.1	1.1 (1)	1.7 (2)	1.6 (1)	0.2	1.0 (1)	0.1	
Ser	0.1	0.3	0.1		0.2	1.5 (2)	0.2		0.1	
Gly	0.4	0.9 (1)	0.1	0.1	1.0 (1)					
His		1.0 (1)	0.1							
Arg			0.5 (0 + 1)	0.1	1.1 (1)	0.3				
Thr	0.1									
Ala			1.0 (1 + 0)	0.8 (1)	0.1	0.1	0.1			
Pro					0.1	0.8 (1)	0.2			
Tyr	1.0 (1)		1.5 (1 + 1)	0.1	0.1	0.1	1.3 (1)	0.2		
Val	0.4				0.2	1.1 (1)	0.2		0.1	
Met							0.7 (1)		0.1	
Cys[b]			1.6 (2 + 0)		1.8 (2)	1.6 (2)				
Ile	0.2			1.0 (1)	1.0 (1)	0.3	0.1			
Leu	1.2 (1)	1.0 (1)	0.1		0.4	0.3			1.1 (1)	
Phe										
Trp[b]								0.4 (1)		1.0 (1)
Lys	0.8 (1)	0.1							0.9 (1)	
identification by amino acid position in sequence	26–29	15–18[c]	13–14 and 30–31 + 44–46 (2:1 mixture)	23–25	32–33 and 38–43	1–7 and 19–20	21–22	50–51	47–48	49

[a] Relative molar quantities are given for each amino acid, and the figures in parentheses correspond to the expected composition for the peptide(s) identified in each fraction. [b] The yields of Cys and Trp are low due to their destruction during acid hydrolysis. [c] This fraction also contains a peptide (residues 8–12) which has the composition of the peptide identified (residues 15–18) plus one residue of Ser. [d] The numbers in parentheses correspond to the expected composition for the two peptides identified in peak 3 (residues 13–14 and 30–31 and residues 44–46), respectively.

cursor could be responsible for removal of the arginine residue at position 53 of the recombinant h-EGF (Julius et al., 1983). However, this hypothesis has not been tested, and the nature of this particular carboxypeptidase is unknown. The KEX2 gene product, responsible for processing the prepro-α-factor, would not remove the carboxyl-terminal residue of EGF since this enzyme is an endopeptidase that only cleaves on the carboxyl side of a pair of basic residues (Julius et al., 1984). Since we were never able to find EGF (1–53) in the medium, the carboxypeptidase responsible for the removal of arginine-53 must be very efficient and presumably well compartmentalized within the secretion pathway. This would allow secretion of those recombinant h-EGF polypeptides having leucine as the carboxyl-terminal residue and not full-length recombinant h-EGF.

Our observation that EGF-D (1–52) was further hydrolyzed at the carboxyl terminus, forming EGF-B (1–51) during fermentation, is explained by the action of a carboxypeptidase-like activity in the medium. Since the amount of EGF-B increases as the cell culture enters stationary phase, a plausible source of this peptidase activity is cell lysis, which may release vacuolar enzymes into the medium. The peptidase responsible for the removal of the leucine-52 residue of EGF-D and EGF-C has not yet been identified.

We have not found any modification at the amino terminus of these four species. Each molecular form has asparagine as the amino-terminal residue, like the natural product. Several groups (Petrides et al., 1984; DiAugustine et al., 1985) have isolated murine EGF which lacks the amino-terminal asparagine residue. Therefore, it seems that the aminopeptidase activity present in the murine system is not operative in the yeast system.

Of particular interest in this study was the demonstration that the four molecular species of EGF are equipotent in binding to the EGF receptor and promoting cell division in foreskin fibroblasts. It is not surprising that the EGF-D and EGF-B are fully active because it is known that shorter versions of mouse EGF are active. EGF-C and EGF-A, which are oxidized versions of the molecule, are also competent in biological function. The oxidized methionine at position 21 is in a region of the molecule believed to be involved in receptor binding (Hollenberg & Gregory, 1980). However, the extra oxygen present at this position does not appear to have a detrimental effect on the binding process. Chemically synthesized methionine sulfoxide-21 EGF has been shown to be active (Heath & Merrifield, 1986).

The oxidation of a methionine residue to its sulfoxide derivative has been found in several proteins, e.g., substance P, α- and γ-endorphins, gastrin, insulin-like growth factor, and nerve growth factor (Storring & Tiplady, 1984). It is interesting to note that the yeast α-factor peptide also suffers oxidation at its single methionine residue while retaining biological activity (Stötzler et al., 1976). The similarity between the posttranslational modification of the methionine residues of the secreted α-factor and the secreted recombinant h-EGF could be due to a similar oxidation reaction that takes place during culture of the yeast cells. In the case of the α-factor peptide, no tryptophan oxidation was found; likewise, we have not observed oxidation of the two tryptophan residues in the h-EGF molecule. Methionine-21 appears to be the only site susceptible to this oxidation within EGF, and its oxidation to the sulfoxide does not affect the biological activity of the molecule.

It is of paramount importance that a recombinant protein possess disulfide bonds that are identical with those of the natural product. The fact that recombinant h-EGF secreted from yeast is fully biologially active indicated that its three disulfide bonds were probably correct and this was indeed found to be the case in this study. This confirms that yeast can form the correct disulfide bonds during secretion of a recombinant human protein.

Registry No. EGF, 62229-50-9; γ-urogastrone, 96119-32-3.

REFERENCES

Barnes, D., & Colowick, S. P. (1976) *J. Cell. Physiol. 89*, 633–640.

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93–104.

Brake, A. J., Merryweather, J. P., Coit, D., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P., & Barr, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4642-4646.
Burgess, A. W., Knesel, J., Sparrow, L. G., Nicola, N. A., & Nice, E. C. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5753-5757.
Carpenter, G. (1980) *Science (Washington, D.C.) 210*, 198-199.
Carpenter, G. (1985) *J. Cell Sci., Suppl. 3*, 1-9.
Carpenter, G., & Cohen, S. (1979) *Annu. Rev. Biochem. 48*, 193-216.
Cohen, S. (1962) *J. Biol. Chem. 237*, 1555-1562.
Cohen, S., & Carpenter, G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1317-1321.
DiAugustine, R. P., Walker, M. P., Klapper, D. G., Grove, R. I., Willis, W. D., Harvano, D. J., & Hernandez, O. (1985) *J. Biol. Chem. 260*, 2807-2811.
Gregory, H. (1975) *Nature (London) 257*, 325-327.
Gregory, H., & Preston, B. M. (1977) *Int. J. Pept. Protein Res. 9*, 107-118.
Heath, W. F., & Merrifield, R. B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6367-6371.
Hirata, Y., & Orth, D. N. (1979) *J. Clin. Endocrinol. Metab. 49*, 481-483.
Hollenberg, M. P., & Gregory, H. (1980) *Mol. Pharmacol. 17*, 314-320.
Julius, D., Blair, L., Brake, A., Sprague, G., & Thorner, J. (1983) *Cell (Cambridge, Mass.) 32*, 839-852.
Julius, D., Brake, A., Blair, L., Kemisawa, R., & Thorner, J. (1984) *Cell (Cambridge, Mass.) 37*, 1075-1089.
Oka, Y., & Orth, D. N. (1983) *J. Clin. Invest. 72*, 249-259.
Oka, T., Sakamoto, S., Miyoshi, K. I., Fuwa, T., Yoda, K., Yamasaki, M., Tamura, G., & Miyake, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7212-7216.
Petrides, P. E., Böhlen, P., & Shively, J. E. (1984) *Biochem. Biophys. Res. Commun. 125*, 218-228.
Petrides, P. E., Hosang, M., Shooter, E., Esch, F. S., & Böhlen, P. (1985) *FEBS Lett. 187*, 89-95.
Rozengurt, E., & Heppel, L. A. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4492-4495.
Savage, C. R., & Harper, R. (1981) *Anal. Biochem. 111*, 195-202.
Savage, C. R., Hash, J. H., & Cohen, S. (1973) *J. Biol. Chem. 248*, 7669-7672.
Shechter, Y. (1986) *J. Biol. Chem. 261*, 66-70.
Simpson, R. J., Smith, J. A., Moritz, R. L., O'Hare, M. J., Rudland, P. S., Morrison, J. R., Lloyd, Ch. J., 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.
Storring, P. I., & Tiplady, R. J. (1984) *Anal. Biochem. 141*, 43-54.
Stötzler, D., Kiltz, H. H., & Duntze, W. (1976) *Eur. J. Biochem. 69*, 397-400.
Urdea, M. S., Merryweather, J. P., Mullenbach, G. T., Coit, T., Herberlein, U., Valenzuela, P., & Barr, P. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7461-7465.

Determination of Local Conformational Stability in Fragment 96-133 of Bovine Growth Hormone by High-Resolution ^{1}H NMR Spectroscopy

Paul R. Gooley,‡ Scott M. Plaisted, David N. Brems, and Neil E. MacKenzie*
Control Division, Building 259-2, The Upjohn Company, 7000 Portage Road, Kalamazoo, Michigan 49001
Received April 7, 1987; Revised Manuscript Received September 2, 1987

ABSTRACT: The specific assignment of resonances in the 400-MHz nuclear magnetic resonance (NMR) spectrum of fragment 96-133 (AII) of bovine growth hormone (bSt) is described. Assignments have been made with homonuclear two-dimensional techniques, in particular that of sequential resonance assignment. Complete assignments were possible for the spin systems of 16 residues out of a total of 38 and partial assignments for another 5. Assignment of resonances to either residue type or a class of residue was possible for a number of other spin systems. Analysis of the type of nuclear Overhauser effect (NOE) indicates that segments 96-110 and 130-133 are nonregular stable structures and that the segment 111-127, which putatively spans the α-helix, is not sufficiently stable to generate NOEs.

Pituitary bovine growth hormone (bSt) is a globular protein consisting of 191 amino acids. Partial tryptic digestion of bSt yields a fragment, AII (residues 96-133), which retains measurable biological activity (Sonenberg et al., 1968; Yamasaki et al., 1975). The primary sequence of the peptide AII is shown in Figure 1. On the basis of the primary sequence of intact bSt, Chou-Fasman predictions (Chou & Fasman, 1974a,b) find a helical region in the segment 111-127 and β-sheet structure in the pentapeptide 101-105. AII encompasses both these regions, and indeed, circular dichroism studies show that at pH 4 AII is composed of 32% helix and 10% β-sheet, which decreases to 7% helix and 6% β-sheet at pH 9 (Chen & Sonenberg, 1977). The helical content of AII has also been found to be concentration, pH, and ionic strength dependent (Brems et al., unpublished results).

Fragments derived from helical regions of intact protein are generally devoid of any ordered structure leading to the conclusion that tertiary interactions are necessary to stabilize the

* Address correspondence to this author.
‡ Present address: Department of Molecular Biophysics and Biochemistry, School of Medicine, Yale University, New Haven, CT 06510.

0006-2960/88/0427-0802$01.50/0