Structural Characterization of Natural Human Urinary and Recombinant DNA-derived Erythropoietin

IDENTIFICATION OF DES-ARGININE 166 ERYTHROPOIETIN*

(Received for publication, June 26, 1987)

Michael A. Recny, Hubert A. Scobell, and Yangkil Kim
From the Genetics Institute, Inc., Cambridge, Massachusetts 02140 and the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Recombinant human erythropoietin (rhEPO) has been purified to apparent homogeneity from a Chinese hamster ovary cell line expressing a cDNA clone of the human gene. NH2-terminal sequencing of the recombinant hormone indicates that the 27-residue leader peptide is correctly and consistently cleaved during secretion of the recombinant protein into conditioned medium, yielding the mature NH2 terminus (Ala-Pro-Pro-Arg...). Analysis of the COOH terminus of rhEPO by peptide mapping and fast atom bombardment mass spectrometry (FABMS) demonstrates that the arginyl residue predicted to be at the COOH terminus (based on confirmation of both genomic and cDNA sequences) is completely missing from the purified protein. The truncated form of the recombinant hormone, designated des-Arg166 rhEPO, displays an in vivo specific activity of greater than 200,000 units/mg protein. A structural characterization of natural human urinary EPO (uEPO) by peptide mapping and FABMS reveals that the urinary hormone is also missing the COOH-terminal Arg166 amino acid residue, a modification that remained undetected until now. There is no evidence of further proteolytic processing at the COOH terminus beyond specific removal of the Arg166 amino acid residue in either rhEPO or uEPO. On the basis of the FABMS data, we propose that the physiologically active form of the hormone circulating in plasma and interacting with target cells in vivo is des-Arg166 EPO.

The terminal differentiation of pre-erythroid colonies into mature red blood cells in the mammalian circulatory system is regulated by the hormone erythropoietin (Goldwasser, 1975; Graber and Krantz, 1978). The role of erythropoietin (EPO) as a physiological modulator of red cell production has been well established, although the precise mechanisms by which EPO interacts with erythroid target cells and influences the process of hematopoiesis are still unknown. The hormone is produced in the kidney (Sherwood and Goldwater, 1978) and liver (Fried, 1972; Naughton et al., 1977) of adults and in the liver of fetal mammals (Zanjani et al., 1977), and its production is stimulated by hypoxia (Erslev, 1955). Human EPO purified from the urine of patients with aplastic anemia reportedly consists of two forms (a and b), which have the same apparent specific activity in vivo (Miyake et al., 1977) but differ in overall carbohydrate content (Dordal et al., 1985). Recently, both cDNA clones (Jacobs et al., 1985) and genomic clones (Lin et al., 1985) of human EPO have been reported, as well as structural characterization of human urinary EPO (uEPO) by protein sequencing (Lai et al., 1986).

In this report we describe the initial characterization of recombinant human EPO (rhEPO) purified from a Chinese hamster ovary (CHO) cell line expressing a cDNA clone of the human gene. The recombinant glycoprotein displays an in vivo specific activity greater than 200,000 units/mg polypeptide when assayed in a murine model system. Structural characterization of rhEPO and uEPO by peptide mapping and fast atom bombardment mass spectrometry (FABMS) demonstrates that both the recombinant hormone and the natural urinary hormone are proteolytically processed at their COOH termini, resulting in truncated forms of the glycoprotein which are each missing the COOH-terminal Arg166 amino acid residue.

MATERIALS AND METHODS

Purification and Analysis of EPO Biological Activity

rhEPO was purified to apparent homogeneity from the conditioned medium of a CHO cell line expressing a cDNA clone of the human gene (Jacobs et al., 1985). Plasmid DNA expression vectors containing the EPO cDNA and a gene for dihydrofolate reductase were cotransfected into CHO dehydrofolate reductase-deficient cells, and resistant populations were selected for growth in the presence of methotrexate (Kaufman et al., 1985). Clone DN2-3 was chosen for further amplification, and transformants were selected for growth in increasing concentrations of methotrexate until a suitable level of EPO expression was observed. Stable transformants were maintained as confluent monolayers in roller bottles and as suspension cultures in deep tank bioreactors in both semi-defined and completely defined media. rhEPO was purified by sequential chromatography using a combination of procedures previously described for the purification of uEPO (Miyake et al., 1977; Krystal et al., 1986; Jacobs et al., 1985). Human uEPO, which had been purified to apparent homogeneity by sequential chromatography including reverse-phase high-performance liquid chromatography (RP-HPLC) as the final step, was a kind gift from Drs. N. Ochi and N. Imai (Chugai Pharmaceutical Co., Ltd., Tokyo, Japan). The in vitro biological activity of EPO was measured in the murine spleen cell [3H]thymidine uptake assay (Krystal, 1983), and the in vivo activity of EPO was measured using the polycythemic mouse assay (Erslev, 1983). The specific activities of the recombinant hormones were determined by comparing their relative bioactivity with reference standard preparations of partially purified uEPO obtained from Toyobo Biochemicals (Tokyo, Japan). Protein concen-

* This project was supported by Chugai Pharmaceuticals, Ltd. and Boehringer Mannheim GmbH. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Present address: Genetics Inst., Inc., Cambridge, MA 02140.

1 The abbreviations used are: EPO, erythropoietin; CHO, Chinese hamster ovary; rhEPO, recombinant human EPO; uEPO, human urinary EPO; FABMS, fast atom bombardment mass spectrometry; RP-HPLC, reverse-phase high-performance liquid chromatography; PTH, phenylthiohydantoine; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.
trations were determined by quantitative amino acid analysis using procedures described herein. Sodium doxycil sulfate-polycylylamide gel electrophoresis (SDS-PAGE) was performed as described (Laemmli, 1970).

Amino Acid Analysis/NH$_2$-terminal Sequencing

Standard Hydrolysis in 6 N HCl—Aliquots containing 0.1–3.0 nmol of protein or peptides were transferred to glass hydrolysis ampoules and dried by vacuum centrifugation. Samples were combined with 250 μl of 6 N constant boiling HCl (Pierce Chemical Co.), and the ampoules were repeatedly flushed with nitrogen and evacuated (to 10 μm) a total of three times. The ampoules were then sealed and hydrolysis was performed for 24 h at 110 °C. After cooling, the ampoules were opened and the contents were dried by vacuum centrifugation. The residue was reconstituted in 0.2 M sodium citrate sample buffer, pH 2.2 (Beckman), and aliquots were applied directly to a Beckman 6300 Amino Acid analyzer. Individual amino acids were detected as their ninhydrin derivatives by monitoring absorbance at 440 and 570 nm and quantitated versus known standards (Beckman) using a SICA Model 7000S computing integration system.

Modified Hydrolysis in Trifluoroacetic Acid/HCl Containing 2% Thiglycolic Acid—To address the accurate quantitation of methionine and cysteine, a method was developed that employs quick hydrolysis (3.5 h) at elevated temperature (140 °C) in trifluoroacetic acid/HCl containing 2% thiglycolic acid. Aliquots containing 0.1–3.0 nmol of peptide or protein were transferred to glass hydrolysis ampoules and lyophilized to dryness. The dried samples were reconstituted in 0.2 M sodium citrate sample buffer, pH 2.2 (Beckman), and aliquots were applied directly to a Beckman 6300 Amino Acid analyzer. Individual amino acids were detected as their ninhydrin derivatives by monitoring absorbance at 440 and 570 nm and quantitated versus known standards (Beckman) using a SICA Model 7000S computing integration system.

NH$_2$-terminal Sequence Analysis—Samples of protein or peptide fragments isolated by RP-HPLC were directly introduced into the sample compartment of an ABI Protein Sequenator and subjected to automated Edman degradation (Hewick et al., 1981). Phenylthiohydantoin (PTH)-derivatives were separated by narrow bore RP-HPLC in an ABI Model 120 A PTH Analyzer, using a gradient of acetonitrile in 0.3 M sodium citrate, pH 4.5, containing 5% tetrahydrofuran. Each derivative was identified and quantitated by comparison of the retention times and absorbance values to a mixture of standard PTH-derivatives (Pierce Chemical Co.).

Reduction and Pyridylethylation of EPO—The cysteine sulffhydril groups of either rhEPO or uEPO were pyridylethylated as follows. Approximately 5 nmol (100 μg) of rhEPO or 5 nmol (60 μg) of uEPO in RP-HPLC solvent were concentrated to near dryness by vacuum centrifugation. Samples were resuspended in 450 μl of 0.2 M N-ethylmorpholine acetate buffer, pH 8.0, containing 3 μl of neat triethylamine (Pierce Chemical Co.) and 1.5 μl of β-mercaptoethanol (Bio-Rad). The solution was flushed with nitrogen, capped, and incubated for 60 min at 36 °C. After reduction, the cysteine sulffhydril groups were pyridylethylated by adding 5 μl of 55% 4-vinylpyrdine (10 mM, Pierce Chemical Co.) and the reaction mixture was incubated for 90 min at 25 °C. The pH was then adjusted to 2.1 with 10% trifluoroacetic acid/H$_2$O and the reaction mixture was diluted to a final volume of 1.0 ml with 0.1% trifluoroacetic acid/H$_2$O. Desalting was accomplished via RP-HPLC by injecting the sample mixture directly onto a Beckman LC54 cartridge column (14.5 mm i.d. × 2 cm) equilibrated in 0.1% trifluoroacetic acid/H$_2$O. After flushing the column for 10 min at a flow rate of 1.0 ml/min, the column was eluted with a linear gradient of acetonitrile containing 0.1% trifluoroacetic acid, and fractions containing the pyridylethylated protein were collected and stored at 4 °C for subsequent analysis.

Endoproteinase Lys-C Digestion and Peptide Mapping—Desalted, pyridylethylated rhEPO (100 μg, 5 nmol) or uEPO (60 μg, 3 nmol) in RP-HPLC solvent was concentrated to near dryness by vacuum centrifugation and resuspended in 250 μl of 0.1 M ammonium bicarbonate, pH 8.5. After flushing the reaction mixture with nitrogen, an aliquot containing 0.3 mg of Endoproteinase Lys-C was then added (50 μl, 100 μg/ml in 0.1 M ammonium bicarbonate, pH 8.6; Boehringer Manneheim) and the reaction mixture was incubated for 4 h at 37 °C. A second 30-μg aliquot of enzyme was then added and digestion was allowed to proceed for an additional 16 h. Digestion was stopped by adjusting the pH of the solution to 2.1 with 10% trifluoroacetic acid/H$_2$O and diluting to a final volume of 1.0 ml with 0.1% trifluoroacetic acid/H$_2$O. Peptides from the resulting digest were separated by RP-HPLC using a Bio-Rad Hi-Pore RP 318 column (4.6 mm × 25 cm) equilibrated with a Bio-Rad Hi-Pore guard column. Chromatography was developed in a series of linear gradients from 0.1 trifluoroacetic acid/H$_2$O to 0.1% trifluoroacetic acid in 90% acetonitrile, 10% H$_2$O using a Beckman 421 gradient controller HPLC system as described in the text. Peptides were detected by their absorbance at 214 and 254 nm. The flow rate was 0.75 ml/min at 25 °C.

FAB Mass Spectrometry—FAB mass spectra were recorded using a JEOL HX10 high resolution mass spectrometer operated at 10-kV accelerating voltage (Bieman, 1986). Samples were introduced via a direct insertion probe through a vacuum lock into the ion source. The sample matrix was bombarded by xenon ion/atoms that had been accelerated to 8 kV, and the instrument was set at a resolution of 1:1400. Samples to be analyzed were dissolved in 1.0 μl of glycerol, 30% acetic acid (5:1, v/v) with 0.5 μl added to the probe tip. Limit mass-range single-scan spectra were recorded from 1250 to 1350 daltons using the JEOL DA5000 data system. The scan time over this mass range was approximately 8 s. In a similar manner full-range mass spectra (500 to 1500 daltons) were recorded with a scan time of 1.6 min.

RESULTS

Initial Characterization and Specific Activity Analysis—rhEPO was purified to apparent homogeneity from the conditioned medium of a CHO cell line transfected with a cDNA clone of the human gene (Jacobs et al., 1985). The recombinant protein is expressed from a single gene of 579 nucleotides encoding a protein of 193 amino acids in length. The first 27 amino acids consist of a hydrophobic leader sequence that is cleaved during secretion, yielding a mature protein with a predicted length of 166 amino acid residues and a molecular mass of 18,398 daltons. On the basis of multiple NH$_2$-terminal sequence analyses performed on various preparations of purified recombinant protein into conditioned medium and that no alternative NH$_2$-terminal processing occurs (data not shown).

Analysis of purified rhEPO by SDS-PAGE demonstrates that the recombinant hormone migrates as a broad, diffuse band displaying a molecular mass distribution between 32,000 and 38,000 daltons under both reducing and nonreducing conditions (Fig. 1). Since there are three potential N-linked glycosylation sites predicted by the cDNA sequence (Jacobs et al., 1985), the observed molecular weight of rhEPO is consistent with the presence of several highly branched oligosaccharide side chains attached to the polypeptide backbone. The "ladderlike" appearance of rhEPO is also characteristic of the behavior of heavily glycosylated proteins analyzed by SDS-PAGE (Westphal et al., 1975). Human uEPO has also been characterized as a heavily glycosylated protein having several complex-type, N-linked carbohydrate side chains and migrating with a molecular mass distribution between 34,000 and 38,500 daltons as measured by SDS-PAGE (Dordal et al., 1985; Krystal et al., 1988).

Analysis of the in vitro biological activity of rhEPO was performed by measuring the stimulation of $^{51}$H$^m$thymidine uptake into murine erythroid precursor cells by the addition of exogenous EPO (Krystal, 1983). The in vitro biological activity was determined in a murine model system by measuring the induction of $^{51}$Fe incorporation into mature erythrocytes in polycythemic mice (Erslev, 1983). Both assays were calibrated by establishing dose-response curves, using an EPO reference standard of known biological activity, which have been elaborated against the internationally recognized human EPO standard from the World Health Organization (Annabelle, 1972). On the basis of the data obtained from five individual assays (each performed in triplicate) on five separate prepara-
rations of rhEPO, the average in vitro specific activity of the recombinant hormone was measured at 234,000 ± 57,000 units/mg polypeptide and the average in vivo specific activity was measured at 216,000 ± 38,000 units/mg polypeptide.

Table I shows the results of quantitative amino acid composition analysis of rhEPO utilizing two separate hydrolysis methodologies. The number of residues per mole observed for most amino acid residues is in excellent agreement with the predicted values based on the cDNA sequence. To obtain the accurate quantification of methionine and tryptophan, we developed a methodology employing quick hydrolysis (3.5 h) at elevated temperature (140 °C) in trifluoroacetic acid/HCl containing 2% thioglycolic acid. This procedure maintains a reducing environment during hydrolysis and protects against cation (such as ω-N-methylation) of the COOH-terminal arginine were recovered in the hydrolysates, rather than 13 residues of arginine as predicted by the cDNA sequence. One possible explanation for these results is that the COOH-terminal arginine predicted at residue 166 was missing from the purified protein. Reexamination of the DNA coding sequence in the cell line expressing the recombinant protein demonstrated that a mutation did not occur in the plasmid DNA sequence that would lead to expression of a truncated form of the recombinant protein. Therefore, processing of the COOH-terminal arginyl residue by an endogenous carboxypeptidase present within the CHO cell culture system remained a likely possibility. Alternatively, an unusual modification (such as ω-N-methylation) of the COOH-terminal arginyl residue or another arginyl residue in the protein could have occurred, leading to a ninhydrin derivative that might not be readily identified in our standard Beckman 6300 Amino Acid Analyzer program.

An attempt was made to directly examine the COOH terminus of rhEPO by utilizing the broad specificity of carboxypeptidase P (from Penicillium janthinellum), which hydrolyzes nearly all COOH-terminal amino acids (Yokoyama et al., 1974, 1977, 1981). However, when pyridylethylated rhEPO was incubated with relatively high concentrations of carboxypeptidase P (enzyme to substrate ratio of 1:1 (w:w)), no arginine was detected in the digest above background (data not shown). In control experiments, carboxypeptidase P was shown to be very effective against synthetic peptides containing arginine at the COOH terminus (90% release of arginine within 10 min). Both pancreatic carboxypeptidase B and yeast-derived carboxypeptidase Y were also tried, and both gave negative results. Since none of the carboxypeptidase experiments provided interpretable data on the nature of the COOH terminus of rhEPO, we turned to direct analysis of the COOH-terminal peptide obtained from an Endoproteinase Lys-C digest of the recombinant hormone.

Complete digestion of rhEPO with the lysine-specific enzyme Endoproteinase Lys-C should produce nine peptides since the cDNA sequence predicts a total of 8 lysine residues.

TABLE I

| Amino acid | 6 N HCl hydrolysis | Trifluoroacetic acid/HCl hydrolysis | Predicted from cDNA |
|------------|--------------------|-----------------------------------|---------------------|
|            | Average ±         | Average ±                         |                     |
| Cys        | 3.5 ± 0.2         |                                   | 4                   |
| Asn        | 11.9 ± 0.1        | 12.1 ± 0.3                        | 12                  |
| Thr        | 10.5 ± 0.2        | 10.7 ± 0.8                        | 11                  |
| Ser        | 9.2 ± 0.1         | 8.3 ± 0.3                         | 10                  |
| Glx        | 19.0 ± 0.0        | 19.0 ± 0.0                        | 19                  |
| Pro        | 8.0 ± 0.1         | 8.4 ± 0.3                         | 8                   |
| Gly        | 9.0 ± 0.1         | 9.6 ± 0.1                         | 9                   |
| Ala        | 18.9 ± 0.2        | 19.5 ± 0.2                        | 19                  |
| Val        | 10.6 ± 0.1        | 11.0 ± 0.1                        | 11                  |
| Met        | 0.6 ± 0.1         | 1.0 ± 0.1                         | 1                   |
| Ile        | 4.6 ± 0.1         | 4.2 ± 0.1                         | 5                   |
| Leu        | 23.3 ± 0.1        | 23.1 ± 0.1                        | 23                  |
| Tyr        | 3.9 ± 0.1         | 4.1 ± 0.1                         | 4                   |
| Phe        | 4.0 ± 0.1         | 4.2 ± 0.1                         | 4                   |
| His        | 2.0 ± 0.1         | 2.1 ± 0.1                         | 2                   |
| Lys        | 7.9 ± 0.1         | 8.1 ± 0.1                         | 8                   |
| Trp        | 1.9 ± 0.1         |                                   | 3                   |
| Arg        | 12.0 ± 0.1        | 12.1 ± 0.1                        | 13                  |

* Composition data has been converted to express the number of individual amino acids normalized to Glx = 19 (predicted number of Glu + Gln based on the cDNA sequence).
* Standard 6 N HCl hydrolysis; average and standard deviation based on triplicate hydrolysates performed on five separate preparations.
* Modified trifluoroacetic acid/HCl hydrolysis; average and standard deviation based on triplicate hydrolysates performed on five separate preparations.
* Predicted number of residues based on cDNA sequence (Jacobs et al., 1985).
* Quantitated as the pyridylethylcysteine derivative.
* Asp + Asn.
* Glu + Gln.
FIG. 2. Endoproteinase Lys-C peptide maps of rhEPO and uEPO. RP-HPLC analysis of peptides resulting from Endoproteinase Lys-C digest of pyridylethylated rhEPO (A) and pyridylethylated uEPO (B) is depicted. Conditions for chemical modification, proteolytic digestion, and peptide mapping are described under "Materials and Methods." The upper trace shows absorbance at 214 nm (0.2 absorbance units at full scale), and the lower trace shows absorbance at 254 nm (0.05 absorbance units at full scale). Peptide fractions are numbered according to elution position. The COOH-terminal peptide is the most hydrophilic peptide in the digest and elutes at the position marked as peak 1.

Identification of Des-Arg\textsuperscript{166} Erythropoietin

A. Recombinant EPO

B. Human urinary EPO

in the mature protein (Jacobs et al., 1985). The peptide designated K9, corresponding to amino acid residues 155-166, is the COOH-terminal peptide having the predicted sequence Leu-Tyr-Thr-Gly-Glu-Ala-Cys-Arg-Thr-Gly-Asp-Arg-COOH. Fig. 2A illustrates a peptide map obtained from the RP-HPLC analysis of reduced, pyridylethylated rhEPO digested with Endoproteinase Lys-C. Every peak in the peptide map was identified as an EPO-related peptide fragment by a combination of NH\textsubscript{2}-terminal sequencing and amino acid composition analysis, and the relevant peak in the map corresponding to the COOH-terminal peptide (K9) is peak 1. An aliquot of this peptide was subjected to NH\textsubscript{2}-terminal sequence analysis and the results, shown in Table II, indicate the amino acid sequence corresponding to residues 155-165 predicted by the cDNA clone. No arginyl residue was detected in cycle 12 of the NH\textsubscript{2}-terminal sequence analysis, as predicted for amino acid residue 166 based on the cDNA sequence. An aliquot of this peptide also was subjected to quantitative amino acid analysis and the results are shown in Table III. The recovery of 1.0 mol of arginine/mol of COOH-terminal peptide is consistent with the NH\textsubscript{2}-terminal sequence data and represents the arginyl residue corresponding to amino acid residue 162 predicted by the cDNA sequence. The recovery of 1.0 mol of aspartate/asparagine/mol of peptide also verifies that the signals observed for PTH-Asp in cycles 12 and 13 of the NH\textsubscript{2}-terminal sequence analysis were due to carryover from PTH-Asp in cycle 11.

These two pieces of data, combined with data obtained from carboxypeptidase P digestion of pyridylethylated rhEPO strongly suggested that the COOH-terminal arginyl residue predicted at position 166 of the polypeptide chain was missing from the purified form of the recombinant protein. An alternative explanation would be that the COOH-terminal Arg\textsuperscript{166} residue in the recombinant molecule was modified in some unusual fashion (perhaps by \(\omega\)-N-methylation) and remained undetectable by both NH\textsubscript{2}-terminal sequence and amino acid analysis.

To investigate this possibility, FABMS was used to directly measure the protonated molecular weight of the COOH-terminal K9 peptide obtained from the Endoproteinase Lys-C digest of reduced, pyridylethylated rhEPO. These results,
Identification of Des-Arg<sup>166</sup> Erythropoietin

300 picomoles of K9 peptide isolated from an Endoproteinase Lys-C digest of rhEPO (Fig. "2A," peak 1) were subjected to NH<sub>2</sub>-terminal sequence analysis on an Applied Biosystems 470A Sequenator. PTH-derivatives were quantitated by comparison to known standards using an on-line Applied Biosystems 120A PTH Analyzer.

**Table II**

| Cycle | PTH-derivative | pmol  |
|-------|----------------|-------|
| 1     | Leu            | 210   |
| 2     | Tyr            | 170   |
| 3     | Thr            | 68    |
| 4     | Gly            | 130   |
| 5     | Glu            | 155   |
| 6     | Ala            | 120   |
| 7     | Cys            | 46    |
| 8     | Arg            | 10    |
| 9     | Thr            | 45    |
| 10    | Gly            | 64    |
| 11    | Asp            | 8     |
| 12    | Asp            | 6     |

*The signal for PTH-pyridylethylcysteine was observed but not quantitated.

**Table III**

Amino acid analysis of the COOH-terminal Endoproteinase Lys-C peptide of rhEPO

| Amino acid | pmol | Molar ratio | Predicted molar ratio |
|------------|------|-------------|-----------------------|
| Cys<sup>b</sup> | 268  | 0.73        | 1.00                  |
| Asx<sup>c</sup> | 369  | 1.00        | 1.00                  |
| Thr        | 684  | 1.86        | 2.00                  |
| Glx<sup>d</sup> | 368  | 1.00        | 1.00<sup>e</sup>      |
| Gly        | 658  | 1.79        | 2.00                  |
| Ala        | 425  | 1.15        | 1.00                  |
| Leu        | 366  | 0.95        | 1.00                  |
| Tyr        | 216  | 0.59        | 1.06                  |
| Arg        | 363  | 0.99        | 2.00                  |

<sup>*</sup>Normalized to Glx = 1.00 for the COOH-terminal Endoproteinase Lys-C peptide fragment as predicted by the cDNA sequence (residues 155-166).

<sup>**</sup>Quantitated as the pyridylethylcysteine derivative.

<sup>+</sup>Asn + Asp.

<sup>+</sup>Glu + Gln.

showed in Fig. 3A, demonstrate that the (M + H)<sup>+</sup> ion at m/z = 1290.6 is identical with the value predicted for the des-Arg<sup>166</sup> form of the pyridylethylated K9 peptide. There is no mass spectrometric evidence for the presence of any full-length, pyridylethylated Arg<sup>166</sup> K9 peptide at m/z = 1446.7. Control experiments with synthetic peptides corresponding to pyridylethylated Arg<sup>166</sup> K9 and des-Arg<sup>166</sup> K9 demonstrated that the two peptides nearly coeluted (within a change of 0.5% acetonitrile) using our RP-HPLC column and gradient system. Thus, any fraction of full-length Arg<sup>166</sup> K9 peptide in the rhEPO digest would have been pooled with the whole K9 peptide fraction taken from the HPLC chromatogram. These data prove that the arginyl residue predicted to be at the COOH terminus of rhEPO (based on the cDNA sequence) is not modified but absent from the purified protein. There is no evidence of partial removal of the COOH-terminal arginyl residue based on the amino acid composition and FABMS data. Moreover, only the COOH-terminal arginyl residue is missing; processing beyond Arg<sup>166</sup> further into the COOH terminus is not observed. Quantitative amino acid analysis performed on five different preparations of rhEPO (Table I) demonstrated the recovery of only 12 arginine residues per rhEPO molecule rather than 13 arginine residues predicted by the cDNA sequence. On the basis of these data and of analysis of the COOH-terminal K9 peptide from these preparations, processing of the recombinant hormone to des-Arg<sup>166</sup> EPO is consistently observed. Specific removal of the COOH-terminal amino acid is presumably due to the enzymatic activity of an endogenous carboxypeptidase present in the CHO cell culture system. Processing is not due to an enzymatic activity present in the fetal bovine serum supplement in media used to maintain CHO cells in suspension or roller bottle cultures, since we have observed that the recombinant hormone purified from CHO cell lines grown and maintained
under completely serum-free conditions is also des-Arg\textsuperscript{166} EPO (data not shown).

**Human uEPO Analysis**—The observation that CHO cells produce des-Arg\textsuperscript{166} rEPO led to investigation of the COOH terminus of the natural form of human EPO purified from the urine of patients with aplastic anemia. A 60-µg sample of uEPO was reduced, pyridylethylated, and digested with Endoproteinase Lys-C. The peptide digest was fractionated by RP-HPLC and the results, shown in Fig. 2B, illustrate a peptide map that is nearly identical with the peptide map obtained from an Endoproteinase Lys-C digest of rEPO. The only difference between the two profiles is in the relative peak shape of the two peptides containing N-linked carbohydrate.

One of these peptides elutes in peak 2 (corresponding to residues 21-45) and contains two consensus N-linked carbohydrate binding sites at Asn\textsuperscript{44} and Asn\textsuperscript{36}, and the other peptide elutes in peaks 10-12 (corresponding to residues 53-97) and contains one consensus N-linked carbohydrate binding site at Asn\textsuperscript{89}. The COOH-terminal K9 peptides from both rEPO and human uEPO migrate in exactly the same position in each of the RP-HPLC peptide maps. The urinary K9 peptide (Fig. 2B, peak 1) was subjected both to NH\textsubscript{2}-terminal sequencing (to verify that it was indeed the COOH-terminal K9 peptide) and to FABMS. These results, summarized in Table IV and Fig. 3B, reveal that the human urinary hormone is also des-Arg\textsuperscript{166} EPO. The (M + H\textsuperscript{+}) ions detected for both the uEPO K9 peptide and rEPO K9 peptide are experimentally identical. Furthermore, a scan over the predicted mass range for the full-length Arg\textsuperscript{166} K9 peptide demonstrated that no fraction of pyridylethylated Arg\textsuperscript{166} K9 peptide was detected at m/z = 1446.7.

**DISCUSSION**

We report here the initial characterization of recombinant human EPO that has been purified from the conditioned medium of a mammalian cell line expressing a cDNA clone of the human gene. The recombinant protein displays an in vivo specific activity of greater than 200,000 units/mg polypeptide when assayed in a murine model system. This value is nearly 3-fold higher than all values previously reported for human uEPO, which range from 70,400 units/mg polypeptide (Miyake et al., 1977) to 81,600 units/mg polypeptide (Yanagawa et al., 1984; Krystal et al., 1986). It should be noted that both our measurements of in vivo specific activity (and those previously reported) reflect only the mass of polypeptide backbone in the samples and neglect the contribution of the carbohydrate side chains to the overall mass of the glycoprotein hormone.

Our discovery that the natural hormone purified from urine and the recombinant hormone purified from CHO cell-conditioned media are both des-Arg\textsuperscript{166} EPO indicates that each is apparently processed by an enzyme that specifically removes COOH-terminal basic residues. Since natural EPO exerts its biological effect as a circulating plasma hormone prior to excretion into urine, COOH-terminal processing of the natural hormone to des-Arg\textsuperscript{166} EPO can occur at one of three stages.

1) **Intracellularly, Prior to, or Associated with Secretion of the Hormone into Plasma**—COOH-terminal processing of EPO by an intracellular enzyme at this stage might be facilitated by greater accessibility of the COOH terminus in the partially folded polypeptide chain prior to attainment of its native, fully folded conformation. This would mean that the physiologically active form of the hormone circulating in plasma is des-Arg\textsuperscript{166} EPO.

2) **Extracellularly, Due to the Activity of a Serum Carboxypeptidase That Specifically Removes COOH-terminal Basic Residues**—Processing at this stage could be an event that mediates the biological activity of the hormone perhaps by increasing (or decreasing) its affinity for the EPO receptor or else decreasing the effective half-life signaling for clearance of the truncated form of the hormone from the circulatory system.

3) **Extracellularly, as a Result of Exposure of the Excreted Hormone to a Urinary Carboxypeptidase**—The active form of EPO circulating in plasma would thus be the full-length hormone, and the generation of des-Arg\textsuperscript{166} EPO at this stage would simply be an unusual artifact lacking physiological relevance.

Given these various possibilities, we propose that the physiologically active form of the natural hormone circulating in plasma is des-Arg\textsuperscript{166} EPO and that COOH-terminal processing of the primary translation product occurs either intracellularly, prior to secretion of the hormone from its target cell, or during circulation of the hormone in plasma. This hypothesis is supported by evidence in the literature of intracellular, membrane-associated and serum-derived, arginine/lysine-specific carboxypeptidases that are present in mammalian systems (Erdös and Sloane, 1962; Bokisch and Müller-Eberhard, 1970; Skidgel et al., 1984a). Human carboxypeptidase N (arginine carboxypeptidase EC 3.4.17.3) hydrolyzes synthetic substrates containing arginine or lysine at the COOH terminus, exhibiting substrate specificity similar to pancreatic carboxypeptidase B but differing in peptidase and esterase activities (Erdös et al., 1967; Oshima et al., 1975). As a circulating plasma enzyme, carboxypeptidase N controls the activity of both complement-derived anaphylatoxins and kinins by specifically removing functional COOH-terminal arginyl residues. At physiological concentrations, C3a and C5a are inactivated within seconds by conversion to de-Arg\textsuperscript{171} C3a and des-Arg\textsuperscript{3} C5a, respectively (Bokisch and Müller-Eberhard, 1970; Gerard and Hugli, 1981). Other functions have been proposed for the enzyme, such as inactivation of vasoactive peptides released by plasmid degradation of fibrin (Bellev et al., 1980). Carboxypeptidase N activity has also been identified in membrane fractions of various human and animal tissues such as kidney and lung (Skidgel et al., 1984a). An intracellular carboxypeptidase isolated from porcine liver having the same specificity toward synthetic substrates as the serum-derived enzyme has also been reported (Oshima et al., 1984).

**Identification of Des-Arg\textsuperscript{166} Erythropoietin**

**DISCUSSION**

We report here the initial characterization of recombinant human EPO that has been purified from the conditioned medium of a mammalian cell line expressing a cDNA clone of the human gene. The recombinant protein displays an in vivo specific activity of greater than 200,000 units/mg polypeptide when assayed in a murine model system. This value is nearly 3-fold higher than all values previously reported for human uEPO, which range from 70,400 units/mg polypeptide (Miyake et al., 1977) to 81,600 units/mg polypeptide (Yanagawa et al., 1984; Krystal et al., 1986). It should be noted that both our measurements of in vivo specific activity (and those previously reported) reflect only the mass of polypeptide backbone in the samples and neglect the contribution of the carbohydrate side chains to the overall mass of the glycoprotein hormone.

Our discovery that the natural hormone purified from urine and the recombinant hormone purified from CHO cell-conditioned media are both des-Arg\textsuperscript{166} EPO indicates that each is apparently processed by an enzyme that specifically removes COOH-terminal basic residues. Since natural EPO exerts its biological effect as a circulating plasma hormone prior to excretion into urine, COOH-terminal processing of the natural hormone to des-Arg\textsuperscript{166} EPO can occur at one of three stages.

1) **Intracellularly, Prior to, or Associated with Secretion of the Hormone into Plasma**—COOH-terminal processing of EPO by an intracellular enzyme at this stage might be facilitated by greater accessibility of the COOH terminus in the partially folded polypeptide chain prior to attainment of its native, fully folded conformation. This would mean that the physiologically active form of the hormone circulating in plasma is des-Arg\textsuperscript{166} EPO.

2) **Extracellularly, Due to the Activity of a Serum Carboxypeptidase That Specifically Removes COOH-terminal Basic Residues**—Processing at this stage could be an event that mediates the biological activity of the hormone perhaps by increasing (or decreasing) its affinity for the EPO receptor or else decreasing the effective half-life signaling for clearance of the truncated form of the hormone from the circulatory system.

3) **Extracellularly, as a Result of Exposure of the Excreted Hormone to a Urinary Carboxypeptidase**—The active form of EPO circulating in plasma would thus be the full-length hormone, and the generation of des-Arg\textsuperscript{166} EPO at this stage would simply be an unusual artifact lacking physiological relevance.

Given these various possibilities, we propose that the physiologically active form of the natural hormone circulating in plasma is des-Arg\textsuperscript{166} EPO and that COOH-terminal processing of the primary translation product occurs either intracellularly, prior to secretion of the hormone from its target cell, or during circulation of the hormone in plasma. This hypothesis is supported by evidence in the literature of intracellular, membrane-associated and serum-derived, arginine/lysine-specific carboxypeptidases that are present in mammalian systems (Erdös and Sloane, 1962; Bokisch and Müller-Eberhard, 1970; Skidgel et al., 1984a). Human carboxypeptidase N (arginine carboxypeptidase EC 3.4.17.3) hydrolyzes synthetic substrates containing arginine or lysine at the COOH terminus, exhibiting substrate specificity similar to pancreatic carboxypeptidase B but differing in peptidase and esterase activities (Erdös et al., 1967; Oshima et al., 1975). As a circulating plasma enzyme, carboxypeptidase N controls the activity of both complement-derived anaphylatoxins and kinins by specifically removing functional COOH-terminal arginyl residues. At physiological concentrations, C3a and C5a are inactivated within seconds by conversion to de-Arg\textsuperscript{171} C3a and des-Arg\textsuperscript{3} C5a, respectively (Bokisch and Müller-Eberhard, 1970; Gerard and Hugli, 1981). Other functions have been proposed for the enzyme, such as inactivation of vasoactive peptides released by plasmid degradation of fibrin (Bellev et al., 1980). Carboxypeptidase N activity has also been identified in membrane fractions of various human and animal tissues such as kidney and lung (Skidgel et al., 1984a). An intracellular carboxypeptidase isolated from porcine liver having the same specificity toward synthetic substrates as the serum-derived enzyme has also been reported (Oshima et al., 1984).
Identification of Des-Arg166 Erythropoietin

It is intriguing to discover that des-Arg166 rhEPO purified from CHO-cell-conditioned medium is also processed at the COOH terminus of the hormone in vivo. We propose that the physiologically active form of the hormone circulating in plasma and interacting with target cells in vivo is des-Arg166 EPO.

Acknowledgments—We wish to thank Dr. Jaime Car0 for in vivo assay support; Ron Kaufman for peptide sequencing of the human EPO cDNA; Dr. Randy Kaufman, Dr. Bob Adamson, Pat Murtha, Ed Louie, and Barbara O’Connell for development of the CHO cell lines; Dr. Godfrey Amphlett, Jayashree Vomganti, and Susan Spielberg for providing the samples of purified rhEPO used in these studies; Dr. Randy Steinbrink for providing the synthetic peptides used as controls in the RP-HPLC peptide maps; Lisa Sperry and the Quality Control group for SDS-PAGE analysis and in vitro bioassay support; Dr. Edward Fritsch, Dr. Godfrey Amphlett, and Prof. Klaus Biemann for many helpful discussions and critical review of the manuscript; Dr. Charles Shoemaker and Katherine Smith for organizing the project; Douglas Owen for editorial support; and Gail Hockman for typing the manuscript.

REFERENCES

Belyaev, M., Gerdin, B., Lindberg, A., Porath, J., and Wallin, R. (1980) Biochim. Biophys. Acta 621, 169–178
Biemann, K. (1986) Anal. Chem. 58, 1288a–1295a
Bokisch, V. A., and Muller-Eberhard, H. J. (1970) J. Clin. Invest. 49, 2427–2436
Dijkman, R., Heremans, H., and Billiau, A. (1987) J. Biol. Chem. 262, 2528–2535
Dordal, M. S., Wang, F. F., and Goldwasser, E. (1985) Endocrinology 116, 2293–2299
Erdos, E. G., and Sloane, E. M. (1982) Biochem. Pharmacol. 11, 585–592
Fried, 1972; Naughton et al., 1977)

The identification of human urinary des-Arg166 EPO suggests that the natural form of the hormone is also a substrate for either an intracellular or a serum-derived carboxypeptidase N. No processing beyond the COOH-terminal arginyl residue is observed, presumably due to the steric restraints imposed by the remaining COOH-terminal sequence (Thr-Gly-Asp-COOH). This is consistent with reported observations that when glycine is in the penultimate position of the COOH-terminal sequence of a polypeptide, the rate of release of the COOH-terminal amino acid is significantly reduced when the substrate is digested with any one of a variety of carboxypeptidases (Smith, 1951; Neurath, 1960). This inhibition is accentuated by the presence of a charged residue (such as aspartic acid) at the COOH terminus. Our data obtained from carboxypeptidase P digestions performed on both intact rhEPO and the COOH-terminal K9 peptide confirm these observations. Unlike the anaphylatoxins, though, we do not believe that generation of des-Arg166 EPO signals either inactivation or cleavage of the truncated form of the hormone from plasma. If this were the case, then we would expect to see very low in vivo biological activity for both the recombinant and natural hormones. It is difficult to say whether the physiological activity of EPO is modulated in some other fashion by removal of the COOH-terminal arginyl residue since we have no full-length hormone available at this time for comparative studies.

The possibility also exists that the generation of des-Arg166 EPO could be due to exposure of the full-length protein to a urinary carboxypeptidase following clearance of the hormone through the kidney. A urinary carboxypeptidase N activity has also been purified and characterized (Skidgel et al., 1984b) and shown to be structurally and kinetically distinct from the serum enzyme. However, it seems likely that, if the COOH-terminal arginyl residue of EPO were susceptible to proteolysis, this modification would have already occurred either just prior to secretion of the hormone from target cells or during circulation of the hormone in plasma.

Our data obtained on this sample of uEPO contradicts a report in the literature that suggests that the COOH-terminal Arg166 residue is present in uEPO (Lai et al., 1986). However, in Lai et al., identification of the COOH-terminal Arg166 residue was based on a single experiment via NH2-terminal sequencing of a tryptic fragment isolated from a RP-HPLC peptide map that gave the reported sequence (Thr-Gly-Asp-Arg-COOH), which aligns with residues 163–166 predicted by the amino acid sequence, in concert with reported observa-

The generation of des-Arg166 rhEPO from the fully-length primary translation product is presumably due to post-translational proteolytic processing by either an intracellular carboxypeptidase, which

1975). It is interesting that both intracellular and membrane-associated carboxypeptidases have been identified in target cells thought to produce the hormone in vivo (Fried, 1972; Naughton et al., 1977).

The possibility also exists that the generation of des-Arg166 EPO could be due to exposure of the full-length protein to a urinary carboxypeptidase following clearance of the hormone through the kidney. A urinary carboxypeptidase N activity has also been purified and characterized (Skidgel et al., 1984b) and shown to be structurally and kinetically distinct from the serum enzyme. However, it seems likely that, if the COOH-terminal arginyl residue of EPO were susceptible to proteolysis, this modification would have already occurred either just prior to secretion of the hormone from target cells or during circulation of the hormone in plasma.

Our data obtained on this sample of uEPO contradicts a report in the literature that suggests that the COOH-terminal Arg166 residue is present in uEPO (Lai et al., 1986). However, in Lai et al., identification of the COOH-terminal Arg166 residue was based on a single experiment via NH2-terminal sequencing of a tryptic fragment isolated from a RP-HPLC peptide map that gave the reported sequence (Thr-Gly-Asp-Arg-COOH), which aligns with residues 163–166 predicted by the cDNA sequence. This assignment was not confirmed in their subsequent tryptic or Staphylococcus aureus V8 peptide maps, and their attempts to identify the COOH-terminal arginyl residue by carboxypeptidase digestion were also unsuccessful. Nevertheless, the peptide mapping and FABMS data presented in this report conclusively demonstrate that the sample of uEPO we characterized is entirely des-Arg166 uEPO.

It is intriguing to discover that des-Arg166 rhEPO purified from CHO-cell-conditioned medium is also processed at the COOH terminus in a manner similar to COOH-terminal processing of the natural hormone. The truncated form of the recombinant hormone is fully active in vivo, displaying a biological potency of greater than 200,000 units/mg polypeptide when assayed in a murine model system. The generation of des-Arg166 rhEPO from the fully-length primary translation product is presumably due to post-translational proteolytic processing by either an intracellular carboxypeptidase, which
Identification of Des-Arg$^{166}$ Erythropoietin

Erdős, E. G., Yang, H. Y. T., Tague, L. L., and Manning, N. (1967) Biochem. Pharmacol. 16, 1287–1297
Erslev, A. J. (1955) Blood 10, 954–961
Erslev, A. J. (1983) in Hematology (Williams, W. J., Beutler, E., Erslev, A. J., and Lichtman, M. A., eds) 3rd Ed., pp. 1634–1635, McGraw-Hill Publications, Minneapolis, MN
Fried, W. (1972) Blood 40, 671–677
Gerard, C., and Hugli, T. E. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 1833–1837
Goldwasser, E. (1975) Fed. Proc. 34, 2285–2292
Graber, S. E., and Krantz, S. B. (1978) Annu. Rev. Med. 29, 51–66
Gribaudo, G., Cofano, F., Prat, M., Baiocchi, C., Cavallo, G., and Landolfo, S. (1985) J. Biol. Chem. 260, 9936–9940
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., and Dryer, W. J. (1981) J. Biol. Chem. 256, 7990–7997
Huang, R., Bloor, C. M., Kawahara, M. S., and Hugli, T. E. (1983) Am. J. Pathol. 112, 48–60
Jacobs, K. J., Shoemaker, C., Rudersdorf, R., Neil, S., Kaufman, R. J., Mufson, A., Seehra, J., Jones, S., Hewick, R. M., Fritsch, E. F., Kawakita, M., Shimazu, T., and Miyake, T. (1985) Nature 313, 1833–1837
Krystal, G. (1983) Exp. Hematol. (N. Y.) 11, 649–660
Krystal, G., Pankratz, H. R. C., Farber, N. M., and Smart, J. E. (1986) Blood 67, 71–79
Laemmli, U. K. (1970) Nature 227, 680–685
Lai, P.-H., Everett, R., Wang, F. F., Arakawa, T., and Goldwasser, E. (1986) J. Biol. Chem. 261, 3112–3121
Lin, F. K., Suggs, S., Lin, C. H., Browne, J. K., Smalling, R., Egie, J. C., Chen, K. K., Fox, G. M., Martin, F., Stabinsky, Z., Badrawi, S. M., Lai, P. H., and Goldwasser, E. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 7580–7584
Miyake, T., Kung, C. K.-H., and Goldwasser, E. (1977) J. Biol. Chem. 252, 5558–5564
Naughton, T. A., Kaplan, S. M., Roy, M., Burdowski, A. J., Gordon, A. S., and Piliero, S. J. (1977) Science 196, 301–302
Neurath, H. (1960) in The Enzymes (Boyer, P. D., Lardy, H., and Myraback, K., eds) 2nd Ed., Vol. IV, pp. 793, Academic Press, New York
Ogle, J. D., and Ogle, C. K. (1983) Clin. Physiol. Biochem. 1, 194–213
Oshima, G., Kato, J., and Erdős, E. G. (1975) Arch. Biochem. Biophys. 170, 152–158
Sherwood, J. B., and Goldwasser, E. (1978) Endocrinology 103, 866–870
Skidgel, R. A., Johnson, A. R., and Erdős, E. G. (1984a) Biochem. Pharmacol. 33, 3471–3478
Skidgel, R. A., Davis, R. M., and Erdős, E. G. (1984b) Anal. Biochem. 140, 520–531
Smith, E. L. (1951) The Enzymes (Summer, J. B., and Mirbach, K., eds) Vol. I, Part 2, p. 793, Academic Press, New York
Westphal, U., Burton, R. M., and Harding, G. B. (1975) Methods Enzymol. 36, 91–104
Yanagawa, S., Hirade, K., Ohnota, H., Sasaki, R., Chiba H., Ueda, M., and Goto, M. (1984) J. Biol. Chem. 259, 2707–2710
Yokoyama, S., Oobayashi, A., Tanabe, O., Sugamara, S., Araki, E., and Ichishima, E. (1974) Appl. Microbiol. 27, 583–589
Yokoyama, S., Miyabe, T., Oobayashi, A., Tanabe, O., and Ichishima, E. (1977) Agric. Biol. Chem. 41, 1379–1383
Yokoyama, S., Miyabe, T., Tanabe, O., and Ichishima, E. (1981) Agric. Biol. Chem. 45, 311–312
Zanjani, E. D., Poster, J., Burlington, H., Mann, L. I., and Wasserman, L. R. (1977) J. Lab. Clin. Med. 88, 640–644