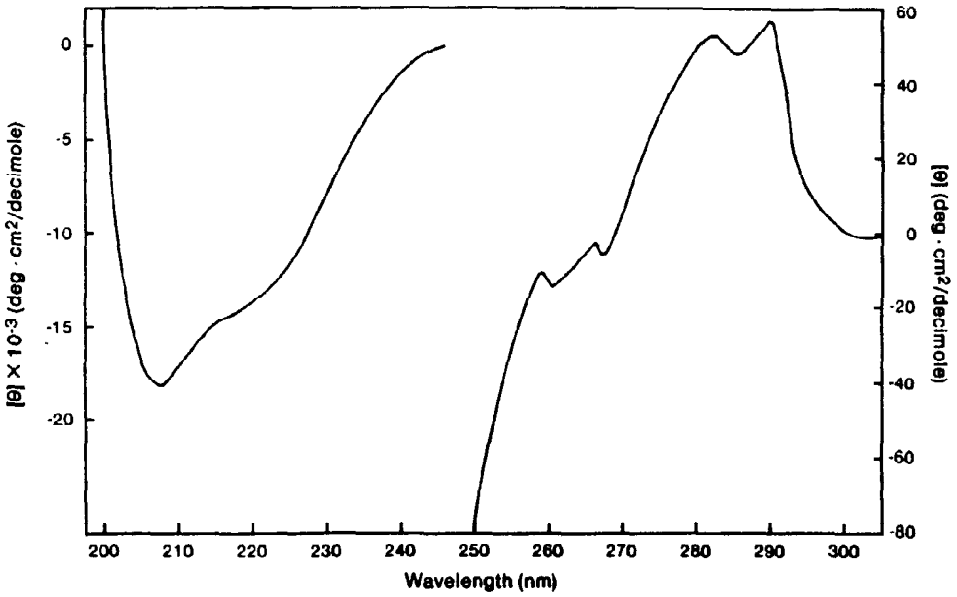


FIG. 4. The near- and far-UV spectra of human erythropoietin.

TABLE III
*Analysis of secondary structure of human EPO*

| Method | Helix | Antiparallel and parallel β-sheet | Turns and others |
|---|---|---|---|
| | % | % | % |
| Chou and Fasman (14) | 36 | 28 | 36 |
| Garnier *et al.* (15) | 42 | 21 | 37 |
| CD (this work) | 50 | 0 | 50 |

peptide T30 (Table 2) and at the twelfth step for peptide 1S32a (not shown in Table 2).

Based on these same observations, we also concluded that human EPO contains two disulfide bonds, one formed between Cys 7 and Cys 161, the other between Cys 29 and Cys 33. This conclusion also supports the previous report (6) that no free thiol is present in the EPO molecule. It is interesting to note that the second disulfide bond is sandwiched between two nearby glycosylation sites, i.e. Asn 24 and Asn 38.

*Secondary Structure*—The near- and far-UV CD spectra are shown in Fig. 4. The far-UV spectrum shows a minimum at 207.5 nm and a shoulder around 218 nm. The secondary structure of the protein was examined according to the method of Greenfield and Fasman (13). The α-helix content was calculated to be about 50% from the observed mean residue ellipticity at 208 nm. It seems that the remaining structure is mainly random and no obvious β-sheet structure could be observed.

Analysis of the sequence by a computer program based on the method of Chou and Fasman (14) suggests an α-helix content of about 36% and a β-sheet content of about 28%. Similar analysis by the method of Garnier *et al.* (15) predicts the α-helix content to be 42% with a β-sheet content of about 21%. Analysis of secondary structure of human EPO by prediction and CD measurements is summarized in Table III. The agreement with respect to α-helix is satisfactory but we do not yet know exactly whether there is any significant β-structure. However, it is noteworthy that the absence of obvious β-sheet structure may be expected from the distribution of proline, aspartic acid, and glutamic acid residues in the EPO molecule as shown in Fig. 1. These residues are highly unfavorable for β-sheet structure (16).

As reported in this study, EPO contains two disulfide bonds; however, the CD analysis of EPO showed no apparent CD signals between 300 and 350 nm where disulfide CD usually can be observed as a broad band (17). It may be possible that the EPO disulfide bonds have unfavorable configurations and give no CD extrema in this wavelength range or there is microenvironmental perturbation caused by interfering groups such as carbohydrate moieties.

The near-UV CD spectrum in Fig. 4 shows two strong positive bands at 282 and 290 nm and two weak negative bands at 260.5 and 267 nm. The observed positive CD bands can be assigned to the $^1L_b$ transition of tryptophan (18). The negative CD bands may be assigned to the transition of phenylalanine. Because of overlapping with the strong tryptophan transitions, the tyrosine CD bands, the maximum of which is usually located between 275 and 282 nm, are not apparent. These CD results clearly indicate that the protein has a distinct tertiary structure, providing asymmetric environments for the aromatic residues (17).

In an effort to understand possible structural relations between EPO and other known protein and nucleic acid sequences, we used a computer homology search which covers the Genbank and Dayhoff data bases. This analysis revealed no easily discernible homology with any proteins. Comparison with the recently published (19) sequence of another hemopoietic growth stimulator, colony-stimulating factor, also shows no significant homology.

*Acknowledgments*—We acknowledge Nowell Stebbing and Dan Vapnek for critical review of the manuscript and Joan Bennett for preparation of the manuscript.

REFERENCES

1. Carnot, P., and Deflandre (1906) *C. R. Acad. Sci.* **1434,** 384–387
2. Goldwasser, E., and Kung, C. K.-H. (1971) *Proc. Natl. Acad. Sci. U. S. A.* **68,** 697–698

3. Goldwasser, E., and Kung, C. K.-H. (1972) *J. Biol. Chem.* **247,** 5159–5160
4. Goldwasser, E. (1981) in *Control of Cellular Division and Development* (Cunningham, D., Goldwasser, E., Watson, J., and Fox, F., eds) pp. 487–494, A. R. Liss, New York
5. Miyake, T., Kung, C. K.-H., and Goldwasser, E. (1977) *J. Biol. Chem.* **252,** 5558–5564
6. Dordal, M. S., Wang, F. F., and Goldwasser, E. (1985) *Endocrinology* **116,** 2293–2299
7. Wang, F. F., Kung, C. K.-H., and Goldwasser, E. (1985) *Endocrinology* **116,** 2286–2292
8. Brauer, A. W., Oman, C. L., and Margolies, M. N. (1984) *Anal. Biochem.* **137,** 134–142
9. Yanagawa, S., Hirade, K., Ohnota, H., Sasaki, R., Chiba, H., Ueda, M., and Goto, M. (1984) *J. Biol. Chem.* **259,** 2707–2710
10. Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S. D., Kaufman, R. J., Mufson, A., Seehra, J., Jones, S. S., Hewick, R., Fritsch, E. F., Kawakita, M., Shimizu, T., and Miyake, T. (1985) *Nature* **313,** 806–810
11. Lin, F. K., Suggs, S., Lin, C.-H., Browne, J., Smalling, R., Egrie, J., Chen, K., Fox, M., Stabinsky, Z., Badrawi, S., Lai, P.-H., and Goldwasser, E. (1985) *Proc. Natl. Acad. Sci. U. S. A.*, **82,** 7580–7584
12. Neuberger, A., Gottschalk, A., Marshall, R. D., and Spiro, R. G. (1972) in *The Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., ed) pp. 450–490, Elsevier/North Holland, Amsterdam
13. Greenfield, N., and Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4116
14. Chou, P. Y., and Fasman, G. D. (1977) *J. Mol. Biol.* **115,** 135–175
15. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120
16. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276
17. Timasheff, S. N. (1970) *Enzymes* **2,** 371–443
18. Strickland, E. H. (1974) *CRC Crit. Rev. Biochem.* **2,** 113–175
19. Wong, G. G., Witek, J. S., Temple, P. A., Wilkens, K. M., Leary, A. C., Luxenberg, D. P., Jones, S. S., Brown, E. L., Kay, R. M., Orr, E. C., Shoemaker, C., Golde, D. W., Kaufman, R. J., Hewick, R. M., Wang, E. A., and Clark, S. C. (1985) *Science* **228,** 810–815

Additional references are found on p. 3120.

Supplemental Material to
Structural Characterization of Human Erythropoietin
by
Por-Hsiung Lai, Richard Everett, Fung-Fang Wang, Tsutomu Arakawa and Eugene Goldwasser

## MATERIALS AND METHODS

Human urinary erythropoietin was purified as previously described (1,2).

Protease digestion and peptide separation. α-EPO, 120 μg, was lyophilized in a reacti-vial (Pierce) and dissolved in 25 μl of 10 mM calcium chloride, 0.1 M Tris-Cl, pH 8.0. TPCK-treated trypsin (Worthington) (2.4 μg) was added and the digestion was carried out at 37°C for 25 minutes. The reaction was stopped by adding phenylmethylsulfonylfluoride (Sigma) to a final concentration of 0.2 mM. This digest is designated digest A. Another tryptic digest was prepared using 80 μg of α-EPO (digest B). This digestion was performed at 37°C for 6 hours in 100 μl of 0.1 M ammonium bicarbonate, pH 8.0, using 2 μg of TPCK-treated trypsin (Worthington). Digestion of 185 μg of α-EPO (digest C) with 6.5 μg of S. aureus V8 protease (Miles) was performed in 0.1 M ammonium bicarbonate, pH 8.0 at 37°C for 42 hr. A second batch of S. aureus V8 protease digest using 50 μg of α-EPO (digest D) was prepared similarly.

All protease digests were separated by reverse-phase HPLC immediately after digestion. Peptides were eluted by a gradient formed by an aqueous mobile phase and an organic mobile phase. The aqueous mobile phase was either 0.06% TFA in water (solvent A) or 0.1% TFA in water (solvent B). The organic mobile phase was one of the following solvents: 0.05% TFA in 70% acetonitrile (solvent C), 0.1% TFA in 90% acetonitrile (solvent D), or 0.1% TFA in 80% acetonitrile (solvent E). Fractions were manually collected, dried, and kept at -20°C.

Digest A was separated on a Varian 5000 liquid chromatograph system equipped with a Synchropak RP-P (0.41 x 5 cm) column (Syn Chrom). The peptides were eluted with a linear gradient of 100% solvent A: 0% solvent C to 40% solvent A: 60% solvent C over a period of 60 minutes. The flow rate was 0.5 ml/min. The column was monitored at 220 nm by a Jasco Unidec-100-III UV Spectrophotometer.

Digests B and D were separated on a Vydac 5 μm C4 HPLC column (0.46 x 25 cm) using a Waters gradient HPLC system. The peptides were eluted with a linear gradient of 97% solvent B: 3% solvent D to 45% solvent B: 55% solvent D over a period of 95 minutes. The flow rate was 0.8 ml/min.

Digest C was separated on a Rainin Microsorb Short-Ones 3um, C18 HPLC column (0.46 x 10 cm) using a Waters gradient HPLC system. The peptides were eluted with a linear gradient of 100% solvent B to 35% solvent B: 65% solvent E over 90 min then to 100% solvent E over 20 min. The flow rate was 1 ml/min.

## Amino Acid Sequence Determinations

Automated sequence analyses (3, 4) of intact protein and peptide fragments isolated by HPLC were performed with a gas-phase sequenator using either a standard protein program or a new program designated HHNVAC supplied by H. Hunkapiller of Applied Biosystems. All peptide solutions were made in 50% formic acid (Fluka) in water before being applied to the sequenator. The amount of peptide samples used for sequence analysis varied from 30- to 80% of the materials recovered in the peptide fractions obtained from HPLC of protein digests. In the later phase of this study, the polybrene treated glass-fiber disc described previously (2) was replaced with a TFA-activated glass-fiber disc for the analysis of intact protein. The procedure[3] for the activation of the glass-fiber disc is as follows: the glass fiber disc is immersed in TFA in a covered glass container and kept for one hour at 22°±2°C. This TFA is then decanted and the activated disc is first air dried and further dried under vacuum over KOH. The protein sample previously reduced with 2-mercaptoethanol at 35°C for 30 min is applied in 50% formic acid directly onto the activated filter without polybrene. The disc loaded with sample is further dried under argon before sequencing.

The PTH-amino acid obtained from each sequenator cycle was identified by reverse-phase HPLC (5).

In situ CNBr (Eastman Kodak) cleavage of the remainder of the protein molecule after initial extended sequence analysis was performed as follows: after extended sequence analysis, the sequenator was paused at the end of a cycle leaving the PITC coupled protein uncleaved and the delivery tubing was disconnected. The cartridge reaction cell with the sample disc was removed from the reaction chamber. 30 μl of 10% (w/v) CNBr solution in 70% formic acid was quickly applied to the sample disc which was kept in place in the top piece of the cartridge. After loading CNBr solution, the cartridge reaction cell containing the disc was sealed using Teflon tape, wrapped with aluminum foil and placed in the reaction chamber of sequenator for one hour at 44°C.

[3] S. Kent, unpublished procedure.

Cartridge was reassembled at the end of CNBr cleavage. Before resuming sequence analysis using the same program, the sample disc was dried with argon for 10 min, washed with S2 (ethylacetate) for 2 min and then dried with argon for another 10 min.

## Determination of Protein Disulfide Structure

Assignments of protein disulfide bonds were based on results of sequence analysis of peptide fragments and native and reduced intact protein, and HPLC mapping of tryptic and S. aureus V8 protease digests.

## Prediction of Secondary Structure from Sequence

The amino acid sequence for human EPO was taken from this study. The prediction methods used are those of Chou and Fasman (6) and Garnier et al. (7).

## Peptide Compositional Analysis by PTC-Amino Acids Methods

Compositional analysis of peptide hydrolysates derived from T28 and Z563 were performed according to the improved method (8) of a published procedure (9). In this method, PTC-galactosamine is eluted between PTC-serine and PTC-glycine.

## CD Measurements and Analysis

Circular dichroic spectra were determined at room temperature on a Jasco J-500C spectropolarimeter. Spectral band width was set at 1 nm. Cuvettes used were 0.1 and 1 cm in light path length for 190 to 260 nm and, 240 to 340 nm, respectively. The solvent spectrum was manually subtracted from the protein spectrum. CD measurements were made with the purified EPO in 2 mM $K_2HPO_4$-$KH_2PO_4$ (pH 7.0) at a protein concentration of 0.3 mg/ml. The results were expressed as mean residue ellipticity, [θ], calculated from the mean residue weight of 111. This value was obtained as the molecular weight of polypeptide/number of amino acid residues.

## RESULTS

### Primary Structure

Two separate runs of automated sequence analysis of the intact protein were performed. In one run, 30 μg of native α-EPO was carried out through 50 cycles, and 42 residues were positively assigned. Since a previous study (10) indicated that α-EPO contains only two methionine residues, it was possible to obtain more sequence information by performing in situ CNBr cleavage on the remaining part of the EPO molecule after initial extended N-terminal sequencing. Thus, 100 μg of 2-mercaptoethanol reduced α-EPO were sequenced on a TFA activated filter through 47 cycles. Forty-five residues were positively assigned after prompt HPLC analyses of products derived from Edman degradation cycles. After CNBr cleavage of the residual material, automated sequence analysis was continued through another 44 cycles, and degradation products were again promptly analyzed by HPLC. Only one major PTH-amino acid residue could be detected in each cycle analyzed except one which was later assigned to a glycosylation site (Asn 83). A total of 35 out of 44 cycles were positively identified such that a single sequence could be assigned. This sequence represents the region covering residues 55–89. The results of sequencing intact protein and sequence analysis after in situ CNBr cleavage are shown in Table 1. Although amino acid composition analysis (10) indicated the presence of two methionine residues, the in situ CNBr cleavage of the protein lacking the N-terminal 47 residues did not yield three fragments. Only one fragment, residues 55–166 was sequenced as judged from the sequencing results (Table 1). Fragment $Phe_{88}$-Tyr-Ala-Trp-Lys-Arg-$Met_{54}$ was not sequenced possibly due to loss of the peptide during washing or due to destruction, during CNBr cleavage, of the PTC moiety which was left on the residual fragment when the sequenator was paused for CNBr treatment. In the latter case, this peptide would not have a free α-amino group for coupling reaction.

Most of the tryptic peptides analyzed were obtained from digest A (Figure 2). From this digest, 17 tryptic fractions were isolated and identified. Two fractions, i.e., fractions 4 and 26, are mixtures of two peptides of equal recovery. Fraction 4 consists of two tetrapeptides, i.e., T4a and T4b; the sequence of the latter peptide was determined by subtracting the known residues from the previously sequenced N-terminal region. The sequences of the two peptides collected in fraction 26, i.e., T26a and T26b were similarly determined.

The primary structure of EPO was established by aligning sequences of tryptic peptides with those obtained from the intact protein and the S. aureus V8 protease peptides as shown in Figure 1. The peptide contains residues 54–76 (T-81), was not isolated by HPLC of digest A. Since it consists of many hydrophobic residues, it was probably lost on the reverse phase column. A separation procedure that involves weaker hydrophobic interaction between peptide and column matrix was designed for the specific isolation of this peptide. Digest B was chromatographed on a $C_4$ HPLC column and peptide T81 was found in the last fraction eluted from the $C_4$ column (chromatogram not shown). Another peptide shown in Figure 1 which was also isolated from digest B is T45. T45 which had an intact Lys-Arg bond at its carboxy terminal was not found in digest A. It is interesting to note that 142 residues out of 166 residues (86%) of the whole molecule were entirely sequenced with the peptides

isolated from digest A. The average repetitive yield of sequencing runs was about 94%. All small peptides could be completely sequenced if the glass fiber disc contained polybrene. Results of sequence analysis of all peptide fragments used in establishing the primary structure of the protein are summarized in Table 2.

Sequence data obtained from peptides isolated from HPLC of digest D (Figure 3) provided further necessary information for reconstruction of the complete sequence by overlapping the tryptic fragments. From digest D, 10 *S. aureus* V8 protease peptide fractions were isolated and identified. Fraction 33 in Figure 3 contains two peptides of equal recovery, i.e., 2533a and 2533b as shown in Figure 1. Peptide 2547 overlaps tryptic peptides T33, T9, and T2e. Peptide 2563 established the overlap between fragments T28 and T27. Fragments T27, T21 and T31 are overlapped by peptide 2545. The rest of the C-terminal tryptic peptides including T31, T4b, T26b and T2 were overlapped by two peptides, i.e., 2545 and 2533b obtained from digest D. Fragments T38, T16, and T33 were overlapped by the *S. aureus* V8 protease peptide, 1532b which together with peptide 1532a were isolated from digest C (chromatogram not shown).

All of the residues were assigned positions by sequencing and positive identification except the asparagines at positions 24, 38, and 83 and one serine at position 126 which was identified and assigned by composition analysis of peptide T28 (data not shown). Ninety-nine percent of the residues have been assigned after more than one determination. The only two residues which were assigned based on single determination are Arg53 and Arg166. Determination of the C-terminal residue was based on sequence analysis and alignment of peptide T2 and confirmed by DNA sequencing. We did not detect any peptide whose sequence is not shown in Figure 1. Attempts to confirm the C-terminal residue by carboxypeptidase digestion failed possibly due to a sterically hindered C-terminal end.

Two peptide bonds which do not involve any aspartic acid or glutamic acid residues but involve serine residues were unexpectedly hydrolyzed by *S. aureus* V8 protease as evidenced by isolation and identification of peptides 1532b and 2547 and 2544 and 2545 which are linked by Ser104 - Leu105 bond and Ser146 - Asn147 bond, respectively.

## REFERENCES

1. Miyake, T., Kung, C.K.-H. and Goldwasser, E. (1977) J. Biol. Chem. 252, 5558-5564.
2. Dordal, M.S., Wang, F.F. and Goldwasser, E. (1985) Endocrinol., 116, 2293-2299.
3. Hewick, R.M., Hunkapiller, M.W., Hood, L.E., and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990-7997.
4. Lai, P.-H. (1984) Anal. Chim. Acta 163, 243-248.
5. Hunkapiller, M.W. and Hood, L.E. (1983) Science 219, 650-659.
6. Chou, P.Y. and Fasman, E.D. (1977) J. Mol. Biol. 115, 135-175.
7. Garnier, J., Osguthorpe, D.J., and Robson, B. (1978) J. Mol. Biol. 120, 97-120.
8. Lu, H.S., Klein, M.L., Everett, R.R. and Lai, P.-H. in Modern Methods in Protein Chemistry (L'Italien, J., ed., Plenum, New York) in press.
9. Heinrikson, R.L. and Meredith, S.C. (1984) Anal. Biochem. 136, 65-74.
10. Goldwasser, E. (1981) in Control of Cellular Division and Development Cunningham, D., Goldwasser, E., Watson, J., and Fox, F., eds.) pp. 487-494. A.R. Liss, New York.

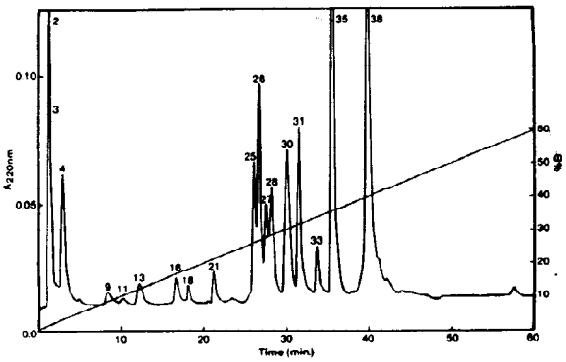


Figure 2. Tryptic map of Digest A.

Dash line indicates the gradient of solvent B. Number indicates source of the respective T-peptide.

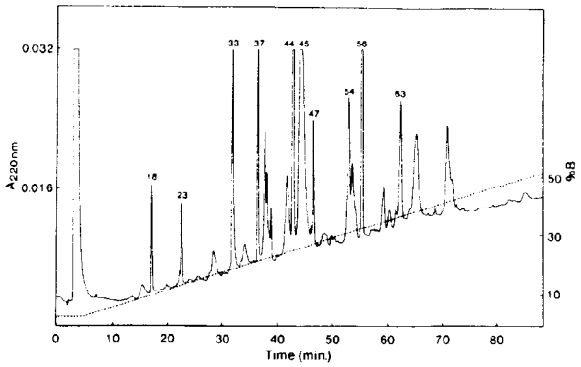


Figure 3. *S. aureus* V8 protease map of Digest D.

Dash line indicates the gradient of solvent B. Number indicates source of the respective S-peptide.

Table 1. Gas-phase sequence analysis of 130 ug of intact human EPO. No., residue number shown in Figure 1; I.D., PTH-amino acid identification by the one-letter code; yields are in pmols; NC, not calculated, identified qualitatively; (C), tentative assignments for cysteine based on absence of assignable PTH-amino acid; (N), tentative assignments for glycosylated Asn according to the rules of Asn-X-Ser/Thr;

(1) Sequenator was paused for *in situ* CNBr cleavage at methionines after sequencing through this residue.

(2) Known from cleavage chemistry and from T81.

(3) Automated sequence analysis resumed after CNBr cleavage.

Table I

| No. | I.D. | Yield | No. | I.D. | Yield | No. | I.D. | Yield |
|---|---|---|---|---|---|---|---|---|
| 1 | A | 520 | 31 | E | 93 | 61 | Y | 259 |
| 2 | P | 378 | 32 | H | NC | 62 | E | 148 |
| 3 | P | 297 | 33 | (C) | - | 63 | V | NC |
| 4 | R | NC | 34 | S | NC | 64 | W | NC |
| 5 | L | 333 | 35 | L | 66 | 65 | Q | 144 |
| 6 | I | 297 | 36 | N | 35 | 66 | G | 131 |
| 7 | (C) | - | 37 | E | 81 | 67 | L | 239 |
| 8 | D | 96 | 38 | (N) | - | 68 | A | 249 |
| 9 | S | 24 | 39 | I | 26 | 69 | L | 242 |
| 10 | R | NC | 40 | T | 16 | 70 | L | 249 |
| 11 | V | 228 | 41 | V | 60 | 71 | S | 32 |
| 12 | L | 222 | 42 | P | 124 | 72 | E | 104 |
| 13 | E | 130 | 43 | D | 36 | 73 | A | 179 |
| 14 | R | NC | 44 | T | NC | 74 | V | 96 |
| 15 | Y | 120 | 45 | K | 52 | 75 | L | 202 |
| 16 | L | 230 | 46 | V | 81 | 76 | R | NC |
| 17 | L | 219 | 47 | N(1) | 23 | 77 | G | 78 |
| 18 | E | 125 | 48 | | | 78 | Q | 80 |
| 19 | A | 136 | 49 | | | 79 | A | 193 |
| 20 | K | 132 | 50 | | | 80 | L | 160 |
| 21 | E | 125 | 51 | | | 81 | L | 158 |
| 22 | A | 177 | 52 | | | 82 | V | 83 |
| 23 | E | 122 | 53 | | | 83 | (N) | - |
| 24 | (N) | - | 54 | H(2) | | 84 | S | 24 |
| 25 | I | 101 | 55 | E(3) | 237 | 85 | S | 23 |
| 26 | T | 24 | 56 | V | 287 | 86 | Q | 40 |
| 27 | T | 58 | 57 | G | 217 | 87 | P | NC |
| 28 | G | 96 | 58 | Q | 238 | 88 | W | NC |
| 29 | (C) | - | 59 | Q | 220 | 89 | E | 44 |
| 30 | A | 102 | 60 | A | 284 | 90 | X | - |
| | | | | | | 91 | L | 93 |

Table 2. Automated Edman degradation of peptides derived from tryptic and *S. aureus* V8 protease digests: yield of PTH-amino acids in pmols. No., residue number shown in Figure 1; I.D., PTH-amino acid identification by the one-letter code; NC, not calculated, identified qualitatively; -, not identified as PTH-amino acid.

(1) Identified as PTH-cystine, yield not calculated.

(2) Known as N-glycosylation site as described in text.

(3) Tentatively assigned as O-glycosylation site as described in text.

Table 2

| No. | I.D. | 2S33a | T26a | T13 | T25 | T30 |
|---|---|---|---|---|---|---|
| 1 | A | 173 | | | | |
| 2 | P | 162 | | | | |
| 3 | P | 165 | | | | |
| 4 | R | 78 | | | | |
| 5 | L | 150 | 220 | | | |
| 6 | I | 142 | 210 | | | |
| 7 | C | NC(1) | - | | | |
| 8 | D | 41 | 52 | | | |
| 9 | S | | 49 | | | |
| 10 | N | | 21 | | | |
| 11 | V | | | 450 | | |
| 12 | L | | | 433 | | |
| 13 | E | | | 156 | | |
| 14 | R | | | 26 | | |
| 15 | Y | | | | 155 | |
| 16 | L | | | | 149 | |
| 17 | L | | | | 152 | |
| 18 | E | | | | 96 | |
| 19 | A | | | | 98 | |
| 20 | K | | | | 45 | |
| 21 | E | | | | | 78 |
| 22 | A | | | | | 75 |
| 23 | E | | | | | 42 |
| 24 | N | | | | | -(2) |
| 25 | I | | | | | 36 |
| 26 | T | | | | | 32 |
| 27 | T | | | | | 34 |
| 28 | G | | | | | 46 |
| 29 | C | | | | | - |
| 30 | A | | | | | 49 |
| 31 | E | | | | | 22 |
| 32 | H | | | | | NC |
| 33 | C | | | | | NC(1) |
| 34 | S | | | | | 6 |
| 35 | L | | | | | 15 |
| 36 | N | | | | | 8 |
| 37 | E | | | | | 16 |
| 38 | M | | | | | -(2) |
| 39 | I | | | | | 12 |
| 40 | T | | | | | 7 |

| No. | I.D. | T30 | T35 | T45 | T81 | 2S56 | T38 |
|---|---|---|---|---|---|---|---|
| 41 | V | 13 | | | | | |
| 42 | P | 8 | | | | | |
| 43 | D | 8 | | | | | |
| 44 | T | 4 | | | | | |
| 45 | K | 6 | | | | | |
| 46 | V | | 210 | 58 | | | |
| 47 | N | | 195 | 42 | | | |
| 48 | F | | 139 | 45 | | | |
| 49 | Y | | 105 | 22 | | | |
| 50 | A | | 155 | 30 | | | |
| 51 | W | | 58 | NC | | | |
| 52 | R | | 54 | 12 | | | |
| 53 | R | | | 6 | | | |
| 54 | M | | | | 55 | | |
| 55 | E | | | | 46 | | |
| 56 | V | | | | 43 | | |
| 57 | G | | | | 38 | | |
| 58 | Q | | | | 22 | | |
| 59 | Q | | | | 25 | | |
| 60 | A | | | | 19 | | |
| 61 | V | | | | 14 | | |
| 62 | E | | | | 8 | | |
| 63 | V | | | | 12 | 320 | |
| 64 | W | | | | NC | 185 | |
| 65 | Q | | | | 11 | 267 | |
| 66 | G | | | | 20 | 251 | |
| 67 | L | | | | 24 | 279 | |
| 68 | A | | | | 21 | 243 | |
| 69 | L | | | | 19 | 254 | |
| 70 | L | | | | 21 | 260 | |
| 71 | S | | | | 9 | 121 | |
| 72 | E | | | | 5 | 44 | |
| 73 | A | | | | 11 | | |
| 74 | V | | | | 9 | | |
| 75 | L | | | | 6 | | |
| 76 | R | | | | NC | | |
| 77 | G | | | | | | 235 |
| 78 | Q | | | | | | 262 |
| 79 | A | | | | | | 278 |
| 80 | L | | | | | | 155 |

| No. | I.D. | T38 | 1S32b | T15 | T33 | T9 | 2S47 |
|---|---|---|---|---|---|---|---|
| 81 | L | 181 | | | | | |
| 82 | V | 172 | | | | | |
| 83 | N | -(2) | | | | | |
| 84 | S | 60 | | | | | |
| 85 | S | 62 | | | | | |
| 86 | Q | 78 | | | | | |
| 87 | P | 35 | | | | | |
| 88 | W | 48 | | | | | |
| 89 | E | 38 | | | | | |
| 90 | P | 22 | | | | | |
| 91 | L | 39 | | | | | |
| 92 | Q | 27 | | | | | |
| 93 | L | 31 | | | | | |
| 94 | H | 9 | | | | | |
| 95 | V | 29 | | | | | |
| 96 | D | 12 | | | | | |
| 97 | K | 10 | 90 | | | | |
| 98 | A | | 121 | 210 | | | |
| 99 | Y | | 115 | 202 | | | |
| 100 | S | | 20 | 35 | | | |
| 101 | G | | 66 | 75 | | | |
| 102 | L | | 75 | 178 | | | |
| 103 | R | | 31 | 12 | | | |
| 104 | S | | 12 | | 40 | | |
| 105 | L | | | | 155 | | 226 |
| 106 | T | | | | 45 | | 78 |
| 107 | T | | | | 46 | | 75 |
| 108 | L | | | | 98 | | 94 |
| 109 | L | | | | 112 | | 98 |
| 110 | R | | | | 17 | | 22 |

| No. | I.D. | T9 | 2S47 | T28 | 2S63 | T27 | 2S44 | T21 |
|---|---|---|---|---|---|---|---|---|
| 111 | A | 317 | 42 | | | | | |
| 112 | L | 305 | 28 | | | | | |
| 113 | G | 292 | 33 | | | | | |
| 114 | A | 279 | 39 | | | | | |
| 115 | Q | 267 | 35 | | | | | |
| 116 | K | 75 | 38 | | | | | |
| 117 | E | | 10 | 92 | | | | |
| 118 | A | | | 105 | 42 | | | |
| 119 | I | | | 101 | 50 | | | |
| 120 | S | | | 55 | 24 | | | |
| 121 | P | | | 88 | 48 | | | |
| 122 | P | | | 91 | 38 | | | |
| 123 | D | | | 49 | 35 | | | |
| 124 | A | | | 69 | 38 | | | |
| 125 | A | | | 77 | 35 | | | |
| 126 | S | | | -(3) | -(3) | | | |
| 127 | A | | | 77 | 33 | | | |
| 128 | A | | | 81 | 29 | | | |
| 129 | P | | | NC | 21 | | | |
| 130 | L | | | 15 | 18 | | | |
| 131 | R | | | NC | 15 | | | |
| 132 | T | | | | 6 | 75 | | |
| 133 | I | | | | 18 | 155 | | |
| 134 | T | | | | 6 | 72 | | |
| 135 | A | | | | 15 | 136 | | |
| 136 | D | | | | 9 | 130 | | |
| 137 | T | | | | | 48 | 45 | |
| 138 | F | | | | | 75 | 75 | |
| 139 | R | | | | | 22 | 40 | |
| 140 | K | | | | | | 98 | 160 |

| No. | I.D. | 2S44 | T21 | T31 | 2S45 | T46 | T26b | 2S33b | T2 |
|---|---|---|---|---|---|---|---|---|---|
| 141 | L | 86 | 155 | | | | | | |
| 142 | F | 70 | 152 | | | | | | |
| 143 | R | 32 | 22 | | | | | | |
| 144 | V | 72 | | 155 | | | | | |
| 145 | Y | 68 | | 78 | | | | | |
| 146 | S | 12 | | 45 | | | | | |
| 147 | M | | | 135 | 28 | | | | |
| 148 | F | | | 128 | 45 | | | | |
| 149 | L | | | 130 | 48 | | | | |
| 150 | R | | | 42 | 14 | | | | |
| 151 | G | | | | 33 | 45 | | | |
| 152 | K | | | | 42 | 48 | | | |
| 153 | L | | | | 30 | 50 | | | |
| 154 | K | | | | 38 | 12 | | | |
| 155 | L | | | | 29 | | 220 | | |
| 156 | Y | | | | 16 | | 207 | | |
| 157 | T | | | | 5 | | 45 | | |
| 158 | G | | | | 18 | | 172 | | |
| 159 | E | | | | 5 | | 68 | | |
| 160 | A | | | | | | 80 | 173 | |
| 161 | C | | | | | | NC(1) | - | |
| 162 | R | | | | | | 25 | 70 | |
| 163 | T | | | | | | | 55 | 48 |
| 164 | G | | | | | | | 122 | 65 |
| 165 | D | | | | | | | 33 | 44 |
| 166 | R | | | | | | | | 21 |

# EXHIBIT 5

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1987 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 262, No. 35, Issue of December 15, pp. 17156–17163, 1987
Printed in U.S.A.

# 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. Scoble¶‖, 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. $NH_2$-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 $NH_2$ 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-Arg$^{166}$ rhEPO, displays an *in vivo* specific activity of greater than 200,000 units/mg protein. Structural characterization of natural human urinary EPO (uEPO) by peptide mapping and FABMS reveals that the urinary hormone is also missing the COOH-terminal Arg$^{166}$ 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 Arg$^{166}$ 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-Arg$^{166}$ 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)[1] 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 Goldwasser, 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 ($\alpha$ and $\beta$), 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 Arg$^{166}$ 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 [$^3$H]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.

§ To whom correspondence should be addressed.
‖ 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, phenylthiohydantoin; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

trations were determined by quantitative amino acid analysis using procedures described herein. Sodium dodecyl sulfate-polyacrylamide 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 peptide 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 7000 S computing integration system.

*Modified Hydrolysis in Trifluoroacetic Acid/HCl Containing 2% Thioglycolic Acid*—To address the accurate quantitation of methionine and tryptophan, a method was developed that employs quick hydrolysis (3.5 h) at elevated temperature (140 °C) in trifluoroacetic acid/HCl containing 2% thioglycolic 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 combined with 250 μl of a 1:5 mixture of neat trifluoroacetic acid/6 N constant boiling HCl (Pierce Chemical Co.) containing 2% thioglycolic acid (Sigma), 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 3.5 h at 140 °C. After cooling, the ampoules were opened and the contents were dried by vacuum centrifugation. The residue was reconstituted with 0.2 M sodium citrate sample buffer, injected on the Beckman Amino Acid Analyzer, and analyzed as described previously.

*$NH_2$-terminal Sequence Analysis*—Samples of protein or peptide fragments isolated by RP-HPLC were applied directly to the reaction cartridge of an ABI Protein Sequenator and subjected to automated Edman degradation (Hewick *et al.*, 1981). Phenylthioydantoin (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 acetate, 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 sulfhydryl groups of either rhEPO or uEPO were pyridylethylated as follows. Approximately 5 nmol (100 μg) of rhEPO or 3 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 sulfhydryl groups were pyridylethylated by adding 5 μl of 95% 4-vinylpyridine (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/$H_2O$ and the reaction mixture was diluted to a final volume of 1.0 ml with 0.1% trifluoroacetic acid/$H_2O$. Desalting was accomplished via RP-HPLC by injecting the sample mixture directly onto a Supelcosil LC304 cartridge column (4.6 mm × 2 cm) equilibrated in 0.1% trifluoroacetic acid/$H_2O$. 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 8.6. After flushing the reaction mixture with nitrogen, an aliquot containing 3.0 μg of Endoproteinase Lys-C was added (30 μl, 100 μg/ml in 0.1 M ammonium bicarbonate, pH 8.6; Boeringher Manneheim) and the reaction mixture was incubated for 4 h at 37 °C. A second 3.0-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/$H_2O$ and diluting to a final volume of 1.0 ml with 0.1% trifluoroacetic acid/$H_2O$. Peptides from the resulting digest were separated by RP-HPLC using a Bio-Rad Hi-Pore RP 318 column (4.6 mm × 25 cm) combined with a Bio-Rad Hi-Pore guard column. Chromatography was developed in a series of linear gradients from 0.1 trifluoroacetic acid/$H_2O$ to 0.1% trifluoroacetic acid in 90% acetonitrile, 10% $H_2O$ 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 HX110 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 applied to the probe tip. Limited 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 rhEPO, we determined that the signal peptide is correctly and consistently cleaved during secretion of the 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.*, 1986).

Analysis of the *in vitro* biological activity of rhEPO was performed by measuring the stimulation of [$^3$H]thymidine uptake into murine erythroid precursor cells by the addition of exogenous EPO (Krystal, 1983). The *in vivo* biological activity was determined in a murine model system by measuring the induction of $^{59}$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 calibrated against the internationally recognized human EPO standard from the World Health Organization (Annable, 1972). On the basis of the data obtained from five individual assays (each performed in triplicate) on five separate prepa-

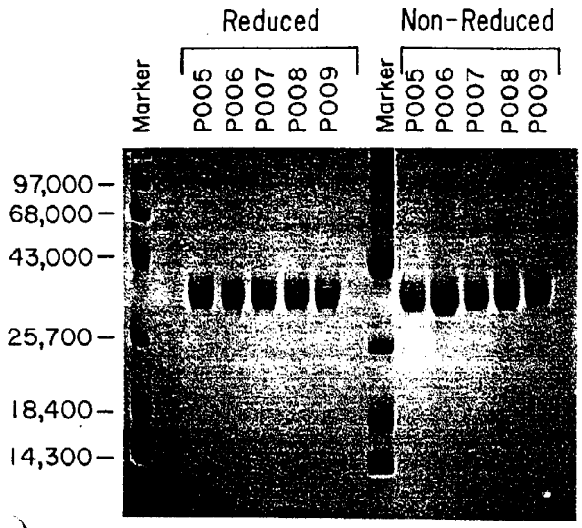


FIG. 1. **SDS-PAGE analysis of rhEPO.** Five separate preparations of purified rhEPO (P005–P009) were analyzed on 10% SDS-polyacrylamide gels as described (Laemmli, 1970). Samples were incubated in the presence (*reduced*) or absence (*non-reduced*) of 2-mercaptoethanol prior to electrophoresis, and the gel was stained with Coomassie Brilliant Blue. Molecular weight markers correspond to phosphorylase *b* (97,000), bovine serum albumin (68,000), ovalbumin (43,000), α-chymotrypsinogen (25,700), β-lactoglobulin (18,400), and lysozyme (14,300).

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 quantitation 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 the oxidative destruction of tryptophan and the oxidation of methionine to methionine sulfoxide and methionine sulfone. The recovery of methionine is nearly quantitative in the system, although tryptophan values are somewhat lower (2 residues) than the predicted value (3 residues). This modified hydrolysis system also gives very accurate quantitation of all other residues in the hydrolysate compared with the standard 6 N HCl hydrolysis protocol. The one notable observation in both sets of analyses, though, is the consistently low recovery of arginine from the protein hydrolysates.

*COOH-terminal Sequence Analysis*—The results obtained from quantitative amino acid composition analysis performed on five separate preparations of rhEPO purified from CHO cell-conditioned media indicated that only 12 residues of 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

TABLE I
*Quantitative amino acid analysis of rhEPO*

| Amino acid | Number of residues/molecule[a] | | | | Predicted from cDNA[d] |
|---|---|---|---|---|---|
| | 6 N HCl hydrolysis[b] | | Trifluoroacetic acid/HCl hydrolysis[c] | | |
| | Average | ± | Average | ± | |
| Cys[e] | 3.5 | 0.2 | | | 4 |
| Asx[f] | 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[g] | 19.0 | | 19.0 | | 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 |

[a] 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).
[b] Standard 6 N HCl hydrolysis; average and standard deviation based on triplicate hydrolysates performed on five separate preparations.
[c] Modified trifluoroacetic acid/HCl hydrolysis; average and standard deviation based on triplicate hydrolysates performed on five separate preparations.
[d] Predicted number of residues based on cDNA sequence (Jacobs *et al.*, 1985).
[e] Quantitated as the pyridylethylcysteine derivative.
[f] Asp + Asn.
[g] Glu + Gln.

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

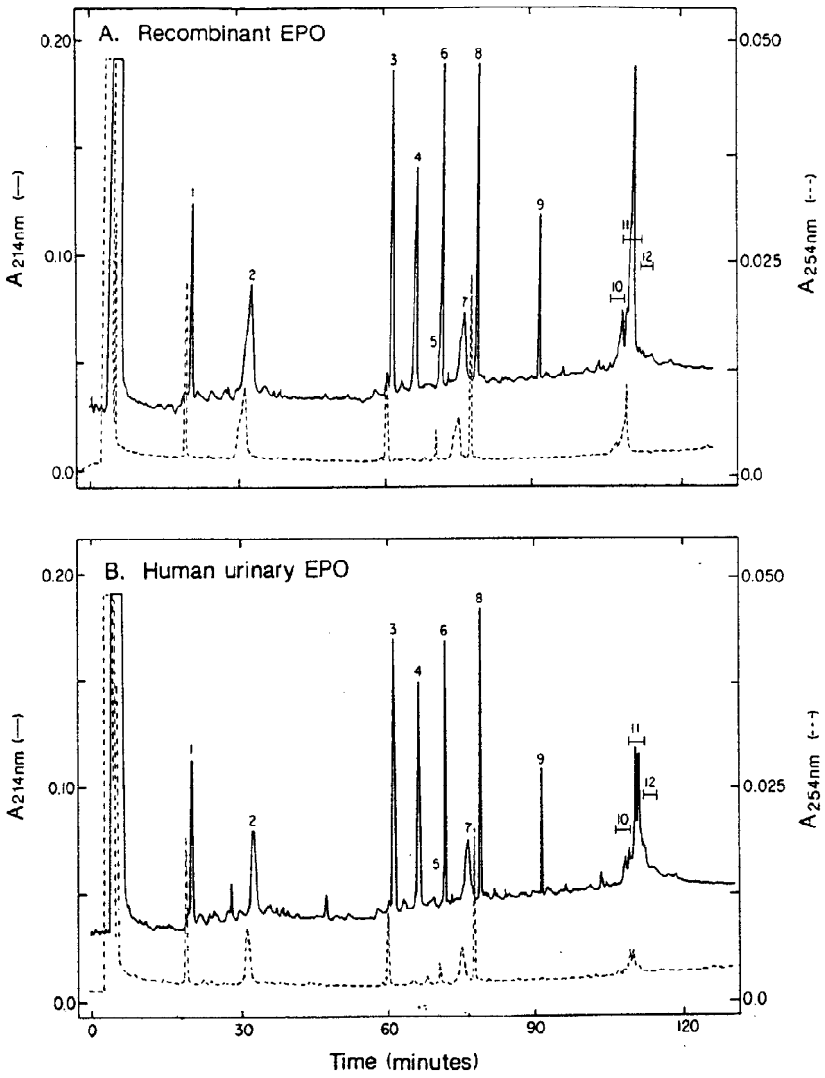


FIG. 2. **Endoproteinase Lys-C peptide maps of rhEPO and uEPO.** RP-HPLC analysis of peptides resulting from Endoproteinase Lys-C digest of pyridylethylated rhEPO (*A*) and pyridylethylated uEPO (*B*) is depicted. Conditions for chemical modification, proteolytic digestion, and peptide mapping are described 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*.

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. 2*A* 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 $NH_2$-terminal sequencing and amino acid composition analysis,[2] 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 $NH_2$-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 $NH_2$-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 $NH_2$-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 $NH_2$-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$^{166}$ residue in the recombinant molecule was modified in some unusual fashion (perhaps by $\omega$-*N*-methylation) and remained undetectable by both $NH_2$-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,

[2] Y. Kim and M. Recny, manuscript in preparation.

TABLE II

*$NH_2$-terminal sequence analysis of the COOH-terminal Endoproteinase Lys-C peptide of rhEPO*

300 picomoles of K9 peptide isolated from an Endoproteinase Lys-C digest of rhEPO (Fig. "2A," *peak 1*) were subjected to $NH_2$-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.

| Cycle | PTH-derivative | pmol |
|---|---|---|
| 1 | Leu | 210 |
| 2 | Tyr | 170 |
| 3 | Thr | 68 |
| 4 | Gly | 130 |
| 5 | Glu | 105 |
| 6 | Ala | 120 |
| 7 | Cys | —[a] |
| 8 | Arg | 46 |
| 9 | Thr | 45 |
| 10 | Gly | 64 |
| 11 | Asp | 10 |
| 12 | Asp | 8 |
| 13 | Asp | 6 |

[a] 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*

By quantitative amino acid analysis, 1.74 nmol of rhEPO Endoproteinase Lys-C digest were injected onto the RP-HPLC column (Fig. 2A). The recovery of 1.47 nmol (84.5%) of the COOH-terminal peptide was obtained in peak 1. No other peptides related to the COOH terminus could be detected in the peptide map.

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

[a] Normalized to Glx = 1.00 for the COOH-terminal Endoproteinase Lys-C peptide fragment as predicted by the cDNA sequence (residues 155–166).
[b] Quantitated as the pyridylethylcysteine derivative.
[c] Asn + Asp.
[d] Glu + Gln.

shown in Fig. 3A, demonstrate that the $(M + H)^+$ ion at $m/z$ = 1290.6 is identical with the value predicted for the des-Arg[166] form of the pyridylethylated K9 peptide. There is no mass spectrometric evidence for the presence of any full-length, pyridylethylated Arg[166] K9 peptide at $m/z$ = 1446.7. Control experiments with synthetic peptides corresponding to pyridylethylated Arg[166] K9 and des-Arg[166] 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[166] 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[166] further into the COOH terminus is not observed. Quantitative amino acid analysis

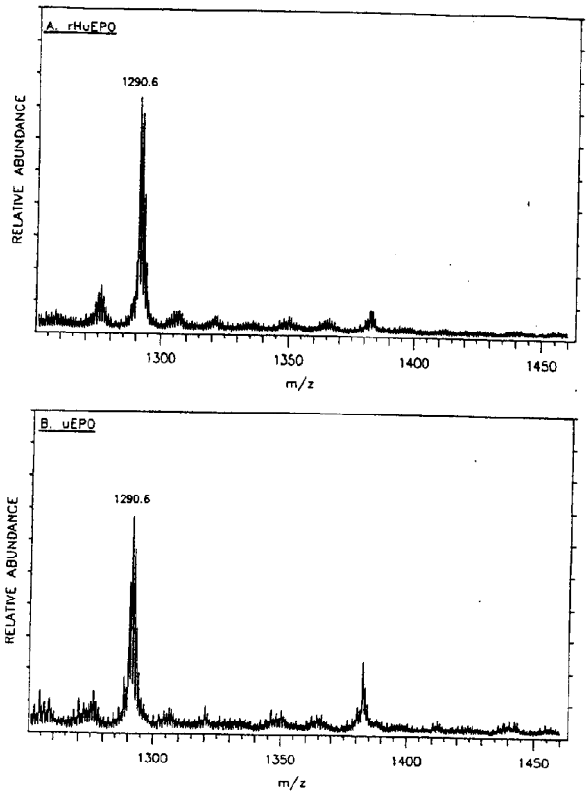


FIG. 3. **Fast atom bombardment mass spectrometry: analysis of the Endoproteinase Lys-C K9 peptides from rhEPO and uEPO.** FABMS was performed in the two-sector mode as described under "Materials and Methods." *A*, an aliquot containing 500 pmol of the pyridylethylated K9 peptide isolated from the Endoproteinase Lys-C peptide map of rhEPO (Fig. 2*A*, *peak 1*) was scanned from $m/z$ = 500 to $m/z$ = 1500. The peak corresponding to the $(M + H)^+$ ion of the K9 peptide was detected at $m/z$ = 1290.6. This value is in excellent agreement with the $(M + H)^+$ ion predicted for the pyridylethylated des-Arg[166] K9 peptide (1290.58). The small peaks observed at $m/z$ = 1289.7 and $m/z$ = 1381.8 correspond to cluster ions $[(C_3H_8O_3)_{14} + H]^+$ and $[(C_3H_8O_3)_{15} + H]^+$, respectively, formed from the glycerol matrix. The small cluster ion spanning the mass range from $m/z$ = 1274.6 to 1275.6 presumably represents a fragmentation product that corresponds to the loss of a $CH_3$ or $NH_2$ group from the parent ion at $m/z$ = 1290.6. *B*, an aliquot containing 500 pmol of the pyridylethylated K9 peptide isolated from the Endoproteinase Lys-C peptide map of uEPO (Fig. 2*B*, *peak 1*) was scanned in an identical manner as in *A*. The $(M + H)^+$ of the uEPO K9 peptide (1290.6) is in excellent agreement with that predicted for the pyridylethylated des-Arg[166] K9 peptide (1290.58). The glycerol cluster ions at $m/z$ = 1289.7 and $m/z$ = 1381.8 were also detected in this case.

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[166] 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$^{166}$ EPO (data not shown).

*Human uEPO Analysis*—The observation that CHO cells produce des-Arg$^{166}$ rhEPO 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-$\mu$g 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. 2*B*, illustrate a peptide map that is nearly identical with the peptide map obtained from an Endoproteinase Lys-C digest of rhEPO. 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$^{24}$ and Asn$^{36}$, 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$^{83}$. The COOH-terminal K9 peptides from both rhEPO and human uEPO migrate in exactly the same position in each of the RP-HPLC peptide maps. The urinary K9 peptide (Fig. 2*B*, *peak 1*) was subjected both to $NH_2$-terminal sequencing (to verify that it was indeed the COOH-terminal K9 peptide) and to FABMS. These results, summarized in Table IV and Fig. 3*B*, reveal that the human urinary hormone is also des-Arg$^{166}$ EPO. The $(M + H)^+$ ions detected for both the uEPO K9 peptide and rhEPO K9 peptide are experimentally identical. Furthermore, a scan over the predicted mass range for the full-length Arg$^{166}$ K9 peptide demonstrated that no fraction of pyridylethylated Arg$^{166}$ 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.

TABLE IV

*$NH_2$-terminal sequence analysis of the COOH-terminal Endoproteinase Lys-C peptide of human uEPO*

An aliquot containing 10% of the peptide eluting in Fraction 1 of the Endoproteinase Lys-C peptide map of uEPO (Fig. 2*B*) was subjected directly to $NH_2$-terminal sequence analysis as described under "Materials and Methods."

| Cycle | PTH-derivative | pmol |
|---|---|---|
| 1 | Leu | 85 |
| 2 | Tyr | 84 |
| 3 | Thr | 43 |
| 4 | Gly | 45 |
| 5 | Glu | 42 |
| 6 | Ala | 43 |
| 7 | Cys[a] | |
| 8 | Arg | 27 |
| 9 | Thr | 16 |
| 10 | Gly | 17 |
| 11 | Asp | 8 |
| 12 | Asp | 6 |
| 13 | Asp | 4 |

[a] The signal for PTH-pyridylethylcysteine was observed but not guaranteed.

Our discovery that the natural hormone purified from urine and the recombinant hormone purified from CHO cell-conditioned media are both des-Arg$^{166}$ 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$^{166}$ 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$^{166}$ 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$^{166}$ 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$^{166}$ 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 des-Arg$^{77}$ C3a and des-Arg$^{74}$ 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 plasmin degradation of fibrin (Belew *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.*,

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 identification of human urinary des-Arg[166] 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 confirms these observations. Unlike the anaphylatoxins, though, we do not believe that generation of des-Arg[166] EPO signals either inactivation or clearance 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-Arg[166] 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 Arg[166] residue is present in uEPO (Lai *et al.*, 1986). However, in Lai *et al.*, identification of the COOH-terminal Arg[166] residue was based on a single experiment via $NH_2$-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 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-Arg[166] uEPO.

It is intriguing to discover that des-Arg[166] 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-Arg[166] rhEPO from the fully-length primary translation product is presumably due to post-translational proteolytic processing by either an intracellular carboxypeptidase, which modifies the recombinant protein prior to its secretion from CHO cells, or a secreted CHO-cell-derived carboxypeptidase that hydrolyzes the COOH-terminal arginyl residue from rhEPO as the recombinant hormone accumulates in conditioned medium. The existence of a rodent enzyme similar to human carboxypeptidase N has also been identified in the sera of guinea pig and rat and is responsible for regulating the spasmogenic activity of complement-derived anaphylatoxins in these species (Huey *et al.*, 1983; Ogle and Ogle, 1983). We have consistently observed COOH-terminal processing of the recombinant hormone to des-Arg[166] EPO in CHO cell cultures maintained in either semi-defined or completely defined media. Therefore, processing is clearly not due to a residual carboxypeptidase N activity in the fetal bovine serum supplement used to maintain CHO cells in semi-defined cultures. Molecular heterogeneity of recombinant murine $\gamma$-interferon expressed in CHO cells was also recently reported (Dijkmans *et al.*, 1987) and was ascribed, in part, to post-translational proteolytic processing of the COOH terminus either before, during, or directly after secretion of recombinant murine $\gamma$-interferon from CHO cells. This result is also consistent with the observation of multiple COOH termini in natural murine $\gamma$-interferon (Gribaudo *et al.*, 1985). Therefore, the extent to which recombinant proteins expressed in CHO cells are proteolytically processed at their COOH termini may be governed by the nature of their primary amino acid sequence, in concert with either secondary or tertiary structure within the polypeptide backbone in a manner analogous to COOH-terminal processing of the natural protein produced in its own target cell.

In summary, we have demonstrated by peptide mapping and FABMS analyses that both human urinary and recombinant CHO-cell-derived EPO are truncated by a single arginyl residue at their COOH termini. The purified recombinant hormone, herein designated des-Arg[166] rhEPO, displays an *in vivo* potency of greater than 200,000 units/mg protein when assayed in a murine model system. From the FABMS data presented in this paper, we propose that the physiologically active form of the hormone circulating in plasma and interacting with target cells *in vivo* is des-Arg[166] EPO.

*Acknowledgments*—We wish to thank Dr. Jaime Caro for *in vivo* assay support; Ron Kriz for DNA sequence analysis 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 Voruganti, 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 Bieman 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

Annable, L., Cotes, P. M., and Mussett, M. V. (1972) *Bull. W. H. O.* **47**, 99–112

Belew, M., Gerdin, B., Lindberg, A., Porath, J., Saldeen, T., and Wallin, R. (1980) *Biochim. Biophys. Acta* **621**, 169–178

Bieman, K. (1986) *Anal. Chem.* **58**, 1288a–1295a

Bokisch, V. A., and Müller-Eberhard, H. J. (1970) *J. Clin. Invest.* **49**, 2427–2436

Dijkmans, 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

Erdös, E. G., and Sloane, E. M. (1962) *Biochem. Pharmacol.* **11**, 585–592

Erdös, 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
Huey, 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**, 806–810
Kaufman, R. J., Wasley, L. C., Spiliotes, A. J., Gossels, S. D., Latt, S. A., Larsen, G. R., and Kay, R. M. (1985) *Mol. Cell. Biol.* **5**, 1750–1759
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**, 3116–3121
Lin, F. K., Suggs, S., Lin, C. H., Browne, J. K., Smalling, R., Egrie, 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, B. 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, p. 11, Academic Press, New York
Ogle, J. D., and Ogle, C. K. (1983) *Clin. Physiol. Biochem.* **1**, 194–213
Oshima, G., Kato, J., and Erdös, E. G. (1975) *Arch. Biochem. Biophys.* **170**, 132–138
Sherwood, J. B., and Goldwasser, E. (1978) *Endocrinology* **103**, 866–870
Skidgel, R. A., Johnson, A. R., and Erdös, E. G. (1984a) *Biochem. Pharmacol.* **33**, 3471–3478
Skidgel, R. A., Davis, R. M., and Erdös, E. G. (1984b) *Anal. Biochem.* **140**, 520–531
Smith, E. L. (1951) *The Enzymes* (Summer, J. B., and Mirback, 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**, 953–959
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.* **89**, 640–644