EXHIBIT 4
(PART 1 OF 2)

Erythropoietin: Gene Cloning, Protein Structure, and Biological Properties

J.K. BROWNE, A.M. COHEN, J.C. EGRIE, P.H. LAI, F.-K. LIN, T. STRICKLAND, E. WATSON, AND N. STEBBING
Amgen, Thousand Oaks, California 91320

The successful cloning of the gene for human erythropoietin (EPO) (Jacobs et al. 1985; Lin et al. 1985) has yielded information on the genetic organization and protein structure of this hormone and has allowed assessment of its biological properties. Biological studies have clearly indicated the clinical potential for this hormone in treatment of various anemias, and initial clinical studies of recombinant-DNA-produced human EPO (r-hEPO)[1] are now under way.

EPO, a sialylglycoprotein hormone, is responsible for regulating the rate of red blood cell formation and for maintaining the red blood cell mass (Krantz and Jacobson 1970; Graber and Krantz 1978; Spivak and Graber 1980). EPO is produced primarily by the kidney in adults and by the liver during fetal life and is secreted into the circulation (Jacobsen et al. 1957; Fried 1972; Zanjani et al. 1981). Circulating levels of EPO are approximately 20 mU/ml (Koeffler and Goldwasser 1981; Cotes 1982; Garcia et al. 1982). Serum levels of EPO increase under conditions of tissue hypoxia and decrease under conditions of hyperoxia. The kidney responds to anemia by increasing the rate of EPO production, resulting in an increase of as much as 100-fold in serum EPO levels (Eschbach and Adamson 1985). Damage to the kidney, as found in chronic renal failure, results in anemia primarily due to a deficiency in EPO production (Brown 1965; Naets 1975; Erslev et al. 1980). Although postulated at the turn of the century (Carnat and Defandre 1906), EPO was first partially purified in 1971 from anemic sheep plasma (Goldwasser and Kung 1971). Human EPO was first purified to homogeneity in 1977 from the urine of aplastic anemia patients (Miyake et al. 1977). Purified human urinary EPO has an apparent molecular weight of about 34,000 and can be separated into two forms, termed α and β, which differ in their carbohydrate content (Dordal et al. 1985). A specific activity of 70,000 U/mg has been reported for purified human urinary EPO (Miyake et al. 1977).

Isolation of a gene for human EPO proved to be particularly problematical because there was no known source of mRNA. No cell lines had been characterized that produced significant amounts of EPO that could provide enriched sources of mRNA. Moreover, although the kidney had been identified as the probable site of synthesis, there remained considerable uncertainty as to whether induction of EPO represented de novo synthesis, release, or activation of an inactive precursor (Fyhrquist et al. 1984). In addition, there was no reliable, convenient assay that could be used to screen cell lines or clones rapidly in an expression system. Thus, the overall cloning strategy was based on obtaining some amino acid sequence information from the very limited amount of purified human urinary EPO that was available. Mixed DNA probes, based on all possible coding sequences, could then be used to screen a human genomic library, and candidate DNA would be expressed in mammalian cells. A cDNA clone, isolated separately, would also be required to unambiguously assign intron/exon boundaries.

Due to the lack of a suitable source of human tissue mRNA, a heterologous approach was employed in which mixed oligonucleotide probes, based on human EPO amino acid sequence information, were used to identify EPO mRNA isolated from various tissues from normal animals or animals made anemic experimentally. This approach required that the amino acid sequence be essentially identical over the region for which the oligonucleotide probes were constructed. For this reason, primates, namely, cynomolgus monkeys, were used because of their relatively close relationship to man.

Gas-phase microsequencing of a small amount of purified human urinary EPO yielded some definitive amino acid assignments but also some uncertainties in the first amino-terminal 23 positions. Due to the redundancy in the genetic code and uncertainties in the amino acid sequence, oligonucleotide probes corresponding to this region proved not to be useful for isolating the EPO gene. Amino acid sequences from internal peptide regions with lower genetic code redundancy were obtained by trypsin digestion of EPO. The amino-terminal sequence did allow, however, preparation of a monoclonal antibody directed against a synthetic peptide from this region (Egrie 1983), and this material, together with antibodies against intact EPO, proved useful in the overall cloning strategy. Because of the heterogeneous nature of urinary EPO preparations, it seemed that either there was considerable secondary

[1] r-hEPO is being jointly developed by Amgen (Thousand Oaks, California), Cilag (Schaffhausen, Switzerland), Kirin Brewery Co., Ltd. (Tokyo, Japan), and Ortho Pharmaceutical Corp. (Raritan, New Jersey).

Cold Spring Harbor Symposia on Quantitative Biology, Volume LI. ©1986 Cold Spring Harbor Laboratory 0-87969-052-6/86 $1.00

modification of the protein or the material was heterogeneous in amino acid sequence and was perhaps coded by a gene family. Thus, considerable effort was directed toward characterization of urinary EPO with the realization that, at the very least, it could confirm primary structure features predictable from the coding gene sequence(s). As it transpired, almost complete characterization of the primary structure of human urinary EPO was achieved by direct analysis within the time taken to characterize the genes (Lai et al. 1986). The direct structural analyses proved important in resolving various structural features of this hormone. EPO proved to be a posttranslationally modified product of a single-copy gene that is highly conserved within the mammals.

EXPERIMENTAL PROCEDURES

Isolation of EPO. Human EPO was purified from the urine of aplastic anemia patients as described previously (Miyake et al. 1977). r-hEPO was produced by Chinese hamster ovary (CHO) cells stably transformed with the human gene (Lin et al. 1985) and purified to homogeneity from conditioned culture media.

Isolation and analyses of gene clones. Cloning of the human and cynomolgus monkey EPO genes, RNA isolation, cDNA cloning, and Northern and Southern blot analyses have been described recently (Lin et al. 1985, 1986). Nucleic acid sequence analysis was carried out primarily by the dideoxy method (Sanger et al. 1977), with a few regions sequenced by the chemical cleavage method (Maxam and Gilbert 1980).

DNA-mediated gene transfer. DNA-mediated gene transfer was carried out using the calcium phosphate microprecipitation method (Graham and van der Eb 1973) as modified by Wigler et al. (1978). For gene transfer into 293 cells (Graham et al. 1977), cells were treated for 4 hours with the calcium-DNA precipitate. COS-1 cells (Gluzman 1981) (ATCC no. CRL 1650) were incubated for 16 hours with the precipitate. Media were sampled for EPO 3–7 days posttransfection.

Amino acid sequence analysis. The sequencing of human EPO was described by Lai et al. (1986).

EPO assays. A radioimmunoassay (RIA) for EPO using purified urinary ^{125}I-labeled EPO and rabbit sera raised against a partially purified preparation of urinary EPO was described recently by Egrie et al. (1986). RIAs using recombinant reagents were performed as described by J.C. Egrie et al. (in prep.). The in vivo exhypoxic polycythemic mouse bioassay for EPO (Cotes and Bangham 1961) and the in vitro rat bone-marrow assay (Goldwasser et al. 1975) were used to assess the biological activity of various EPO preparations. Western blot analysis of EPO, which used a mouse monoclonal antibody raised to a synthetic peptide corresponding to the amino terminus of EPO (Egrie 1983), was described previously (Egrie et al. 1985, 1986).

Physical analyses of EPO. The carbohydrate composition of r-hEPO was determined by methanolysis, trifluoroacetylation, and separation by gas chromatography (Zanetta et al. 1972). Endoglycosidase F (New England Nuclear), neuraminidase, and *O*-glycanase (Genzyme) digestions were performed according to the manufacturers procedures. Analytical ultracentrifuge analysis, dry-weight determination, and measurement of the partial specific volume of EPO were performed as described by J. Davis et al. (in prep.).

Treatment of dogs with r-hEPO. Young adult male and female beagles (six of each sex per group) were dosed intravenously via the cephalic vein for 3 weeks, three times per week, with purified r-hEPO. The animals received either excipient control or a dose of 280 or 2800 U/kg EPO. Blood samples were taken 1 week prior to the study and weekly thereafter.

Pharmacokinetic studies. Pharmacokinetic analysis of r-hEPO in rats will be described in detail separately (A.C. Cohen et al., in prep.). CD rats (Charles Rivers Breeding Laboratories) had a polyethylene cannula implanted into the left carotid artery and a silastic cannula implanted into the right jugular vein 3 days prior to the study. To determine the effect of renal failure on the pharmacokinetics of EPO, both renal pedicles were ligated in one group of animals on the day of the study. r-hEPO was metabolically labeled with [^{35}S]methionine and cysteine and purified. ^{35}S-labeled r-hEPO was administered via the jugular cannula at 1 μCi (370 units r-hEPO) per kilogram of body weight. Blood samples were collected periodically via the carotid cannula, and ethanol precipitable counts were determined. The best-fit lines were determined by a Guass-Newton curve-fitting algorithm, and pharmacokinetic parameters were determined by standard methods.

Subtotal nephrectomized animal study. The system developed by Anagnostou et al. (1977) was used to study r-hEPO in an animal model of renal failure. One group of rats were subjected to a two-step surgical removal of all of one kidney and one half of the second kidney. Sham-operated controls were subjected to concurrent laparotomy. Starting 1 week after surgery, the two groups of animals were injected intramuscularly with 0.1 ml of saline control solution or 10 units of r-hEPO five times a week for 2 weeks. Body weight, hematocrit, and plasma urea nitrogen were determined at the beginning and end of the study (A.C. Cohen, in prep.).

Inflammatory disease animal model. The anemia of adjuvant inflammation was induced according to the procedure of Lukens et al. (1967). Rats were given a single injection of Freund's complete adjuvant in the left hind footpad. Ten days later, treatments were begun wih 0.1-ml intraperitoneal injections of either saline (control solution) or 40 units of r-hEPO. After 10 days of treatment, blood samples were obtained for determination of hematocrit, red cell mass, hemoglobin

concentration, plasma iron, and iron-binding capacity. Similar determinations were made on nonadjuvant treated rats concurrently given injections of either saline or r-hEPO.

RESULTS

Cloning of the Human EPO Gene

The strategy employed to isolate a human EPO genomic gene clone and a cynomolgus monkey cDNA is outlined in Figure 1. A human EPO gene clone was isolated from a λ bacteriophage-borne human genomic library using oligonucleotide probes (Lin et al. 1985). Two sets of mixed oligonucleotide probes were used, each containing a pool of 128 sequences: One pool was a mixture of 20-nucleotide-long oligonucleotides containing all possible coding sequences for an internal hexapeptide, and the other pool was a mixture of 17-nucleotide-long oligonucleotides directed against the coding sequence for a heptapeptide. Using probes directed to two nonoverlapping regions of the EPO gene allowed rapid confirmation of putative clones, eliminating a great number of the false positives obtained with either probe alone.

DNA sequence analysis demonstrated that three of four independent genomic clones had sequences corresponding to the known EPO tryptic peptide amino acid sequences. Proof that a clone contained a complete functional gene encoding human EPO was obtained by its expression in mammalian cells, to produce a gene product with the immunological and biological properties of EPO. A 5.4-kb *Bam*HI-*Hin*dIII restriction fragment from clone λHE1 was identified as potentially carrying the entire EPO gene by Southern blot analysis (Southern 1975) using mixed nucleotide probes. This fragment was subcloned into the plasmid pUC8 (Vieira and Messing 1982) and transiently transfected into 293 cells. These cells are a human embryonic kidney line stably transformed with adenovirus type 5 (Ad5). 293 cells constitutively express the Ad5 EIA gene products that act as *trans*-acting enhancers of expression of DNA introduced into these cells by transfection.

RIA analysis demonstrated the presence of EPO in culture medium samples from transfected cells; the expressed EPO produced a dose-response curve identical to that of urinary EPO. Medium samples from cultures transfected with the plasmid lacking the EPO gene insert were uniformly negative in the RIA. The *Bam*HI-*Hin*dIII fragment was inserted into a shuttle vector, containing the SV40 origin of replication, and transiently transfected into COS-1 cells. RIA analysis again demonstrated that EPO was secreted into the culture media; this material was also shown to be biologically active as determined by the in vitro rat bone-marrow assay and the in vivo exhypoxic polycythemic mouse bioassay. Control cultures were uniformly negative in

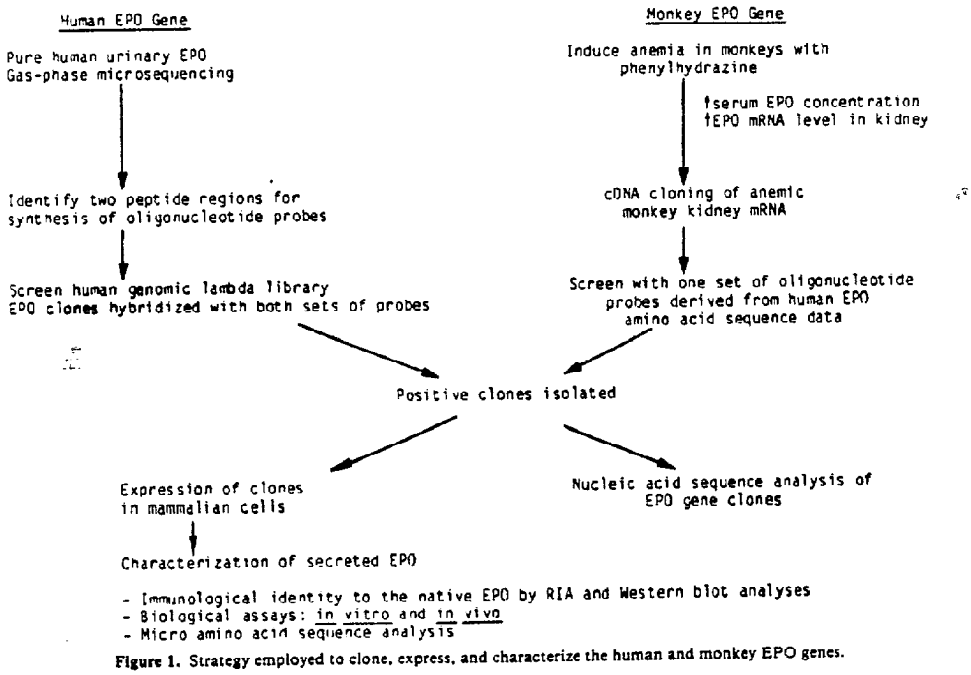

Figure 1. Strategy employed to clone, express, and characterize the human and monkey EPO genes.

these assays. The r-hEPO produced in COS-1 cells is also indistinguishable from urinary EPO by Western blot analysis (Egrie et al. 1985). Polyadenylated RNA isolated from COS-1 cells 48 hours after transfection with this construction was used to prepare a cDNA library from which human EPO cDNA clones were isolated.

The amino acid sequence of the 193-amino-acid primary translation product was deduced from the nucleic acid sequence as shown in Figure 2. The mature hormone is 166 amino acids in length (calculated M_r of 18,399). There are three potential N-linked glycosylation sites, as indicated. The first 27 amino acids predicted by the DNA coding sequence are consistent with this being a hydrophobic leader peptide. The coding portion of the gene is divided by four intervening sequences. The amino terminus of the mature protein was assigned directly from amino-terminal amino acid sequencing of urinary and r-hEPO. Intron/exon junction assignments, which all conform to consensus splice rules (Mount 1982), were made by comparison of a monkey cDNA clone, a human EPO cDNA clone prepared from mRNA isolated from COS-1 cells transfected with the genomic gene clone, and, ultimately, complete amino acid sequence analysis of human urinary EPO (Lai et al. 1986) and recombinant EPO produced in cell culture (P.H. Lai, unpubl.).

Southern blot analysis (Southern 1975) was used to analyze the human EPO gene. The restriction fragment pattern of human DNA probes with a human EPO cDNA clone and the results of low-stringency hybridizations with this probe demonstrate that there is a single copy of the human EPO gene and that there are no apparent closely related genes or pseudogenes (Lin et al. 1985). Computer searches of protein and nucleic acid databases failed to reveal a significant homology with any published sequence. In particular, there is no homology with angiotensinogen, which has been suggested as a possible precursor of EPO (Fyhrquist et al. 1984). In addition, the structure of the human EPO gene and cDNA clones and the amino acid sequence results are consistent with EPO being secreted as an active hormone rather than an inactive precursor.

Cloning of a Monkey EPO cDNA

A cynomolgus monkey EPO cDNA was isolated using the strategy outlined in Figure 1 (Lin et al. 1986). Anemia was induced in monkeys by treatment with phenylhydrazine. As a result of the anemia, serum EPO levels were found to increase in anemic monkeys as measured by RIA and Western blot analyses (Egrie et al. 1985).

A Northern blot analysis of polyadenylated mRNA isolated from normal and anemic monkey kidneys is shown in Figure 3. One of the pools of mixed oligonu-

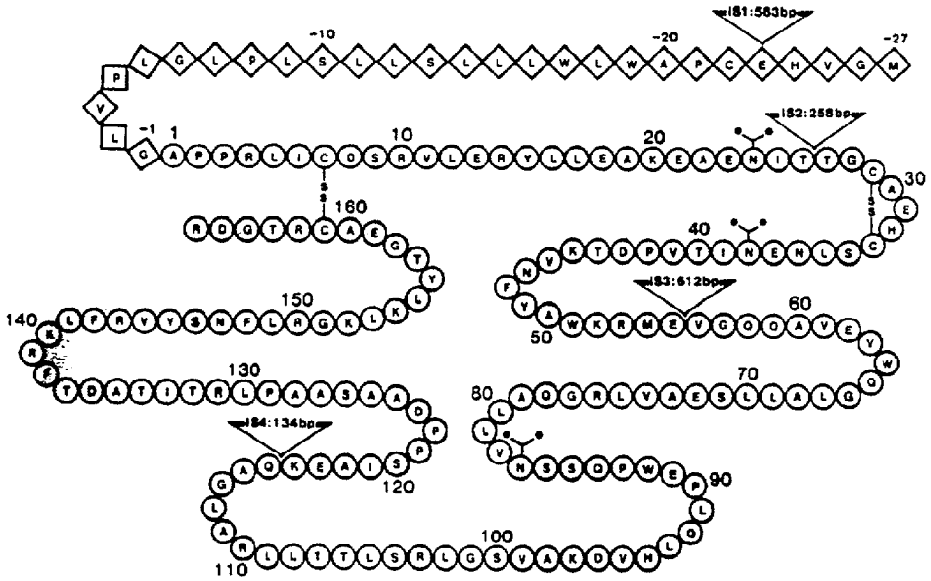

Figure 2. Amino acid sequence of the 193-amino-acid primary translation product of the human EPO gene is presented in one-letter code. The 27 amino-terminal amino acids of the putative signal peptide are boxed, and the residues in the 166-amino-acid mature hormone are circled. Positions of the two disulfide bands (S-S) and the three N-linked glycosidation sites (Y) are noted. Positions at which the four intervening sequences (IS) interrupt the coding position of the gene are given, along with the length of each IS sequence.

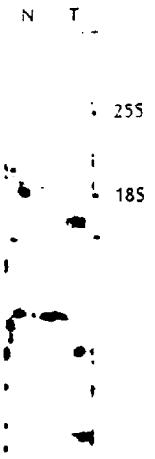

Figure 3. Northern blot analysis of cynomolgus monkey kidney poly(A) containing RNA. ^{32}P-end-labeled mixed oligonucleotide probes to EPO were used to develop the blot. (N) RNA isolated from normal monkey kidneys; (T) RNA isolated from the kidneys of monkeys made anemic by phenylhydrazine treatment. Positions of the ribosomal RNA markers are indicated.

cleotides corresponding to the human EPO amino acid sequence was used as the probe. Positive results were obtained only with mRNA from phenylhydrazine-treated animals (Fig 3, lane T). There was no detectable EPO-specific mRNA in normal monkey kidneys (Fig. 3, lane N). The size of the EPO-specific mRNA is approximately 1600 nucleotides in length. These Northern blot analysis results demonstrate that the steady-state level of EPO mRNA in the kidney is dramatically increased upon induction of anemia. The rise in serum EPO levels is therefore apparently due to increased EPO synthesis mediated by an increased transcription rate of the EPO gene and/or an increase in EPO mRNA stability as opposed to release of previously synthesized and sequestered EPO.

A monkey EPO cDNA clone was isolated from a cDNA library prepared from mRNA from anemic monkey kidneys. Only one of the two pools of mixed oligonucleotide probes proved useful for Northern blot analysis or cDNA cloning. It was subsequently determined that the second probe mixture failed to hybridize with monkey EPO due to a single-amino-acid difference between human and monkey EPO in the region corresponding to the probe. The cDNA clone sequence codes for a 192-amino-acid protein, differing at 15 residues from the human EPO gene sequence. Monkey EPO lacks lysine at position 116, when aligned with the human EPO sequence. Lysine at position 116 in the human gene is the first amino acid in the fourth exon of the human gene. The monkey gene presumably uses an alternative splice junction for the fourth exon, and the DNA sequence is consistent with this. The amino terminus of mature-cell-culture-produced recombinant monkey EPO is at position −3 relative to human EPO. Thus, monkey EPO has a 24-amino-acid signal peptide, and the mature hormone is 168 amino acids in length. A substitution of a proline for a leucine at position −2 (relative to the human gene) in monkey EPO may be the cause of the difference in the signal peptide cleavage site.

Structural Features of EPO

To verify the primary amino acid sequence determined by gene cloning and to assess the nature of secondary modifications, purified human urinary EPO was characterized beyond the initial sequencing involved in designing gene probes. A total of 565 μg of urinary EPO was used to determine the primary structure of the molecule (Lai et al. 1986). About 30 μg was used to determine the amino-terminal amino acid sequence and this allowed assignments at 42 of 50 cycles of degradation. Cyanogen bromide cleavage fragments (100 μg) allowed assignment of 45 additional positions from a total of 77 cycles. About 200 μg was then utilized for sequencing tryptic digests, and about the same amount was used for V8-protease digests. The only residues not assigned directly were the asparagines at positions 24, 38, and 83. The absence of signals for these three residues during direct sequencing is consistent with the presence of carbohydrate moieties at these positions.

The positions of the disulfides were determined by the copurification of peptides predicted from the primary amino acid sequence to be separate after protease treatments (see Fig. 2). A PTH-cysteine was detected in the seventh step of Edman degradation, and sequencing of a reduced preparation of the peptide indicated the presence of the 7–161 disulfide. Assignment of the 29–33 disulfide was indirect and based on the following observations: (1) by Ellman's reaction and attempted labeling with [^{3}H]iodoacetic acid, EPO was found to contain no free thiol residues; (2) sequencing of EPO showed no residues at positions 29 and 33, but performic-acid-oxidized material showed cysteic acid at these positions; and (3) the Glu-31, His-32 bond in oxidized r-hEPO was not hydrolyzed by V8 protease, indicating some conformational abnormality in the structure around these residues.

The molecular weight of purified r-hEPO, produced in CHO cells stably transfected with the human EPO gene inserted in an expression vector, was determined by analytical ultracentrifuge analysis. To carry out this analysis, the extinction coefficient of r-hEPO was measured by dry-weight determination, and the partical-specific volume was determined by the mechanical oscillation technique. This analysis yielded a result of 29,900 ± 400 daltons for the mass of r-hEPO (J. Davis et al., in prep.). The carbohydrate portion of the molecule, assuming a molecular weight of 18,399 for the