Erythropoietin Structure-Function Relationships

MUTANT PROTEINS THAT TEST A MODEL OF TERTIARY STRUCTURE*

(Received for publication, February 12, 1993, and in revised form, April 2, 1993)

Jean-Paul Boiselle, Woan-Ruoh Lee, Scott R. Presnell, Fred E. Cohen, and H. Franklin Bunn

From the Hematology/Oncology Division, Brigham and Women's Hospital, Harvard Medical School, Boston Massachusetts 02115 and the Departments of Medicine, Pharmaceutical Chemistry, Biochemistry, and Biophysics, University of California, San Francisco, California 94143-0446

On the basis of its primary sequence and the location of its disulfide bonds, we propose a structural model of the erythropoietic hormone erythropoietin (Epo) which predicts a four α-helical bundle motif, in common with other cytokines. To test this model, site-directed mutants were prepared by high level transient expression in Cos7 cells and analyzed by a radioimmunoassay and by biosays using mouse and human Epo-dependent cell lines. Deletions of 5 to 8 residues within predicted α-helices resulted in the failure of export of the mutant protein from the cell. In contrast, deletions at the NH₂ terminus (Δ2-5), the COOH terminus (Δ163-166), or in predicted interhelical loops (AB: Δ23-36, Δ53-57; BC: Δ78-82; CD: Δ111-119) resulted in the export of immunologically detectable Epo muteins that were biologically active. However, replacement of each of these deleted residues by serine resulted in Epo muteins with full biological activity. Replacement of Cys²⁰ and Cys²² by tyrosine residues also resulted in the export of fully active Epo. Therefore, this small disulfide loop is not critical to Epo's stability or function. The properties of the muteins that we tested are consistent with our proposed model of tertiary structure.

Humoral regulation of red blood cell production was first proposed at the beginning of this century (1). Convincing physiologic experiments documenting the existence of erythropoietin (Epo) (2-5) were followed by its purification (6) and partial structural characterization (7). The molecular cloning of this biologically and clinically important cytokine (8, 9) has led to further understanding of its properties (10, 11). The binding of Epo to its cognate receptor (12) on erythroid progenitors in the bone marrow results in salvaging these cells from apoptosis (13), allowing them to proliferate and differentiate into circulating erythrocytes. The Epo receptor is a member of an ever enlarging family of cytokine receptors (14). In like manner, Epo shares weak sequence homology with other members of a family of cytokines which also include growth hormone, prolactin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, G-CSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor, and ciliary neurotrophic factor (15-17). The genes encoding these proteins have similar numbers of exons as well as a clear relationship between intron-exon boundaries and predicted α-helical structure. These similarities have led to the prediction that this family of cytokines share a common pattern of folding into a compact globular structure consisting of four amphipathic α-helical bundles. Such theoretical models of the structures of human growth hormone (18) and IL-4 (19) have been in remarkably good agreement with subsequent structures established by x-ray diffraction (human growth hormone) (20, 21) or by multidimensional NMR (IL-4) (22, 23). Moreover, the crystal structures of GM-CSF (24) and monomeric M-CSF (25) are also in reasonable agreement with their predicted structures.

Thus far, the structure of Epo has not been analyzed by either x-ray diffraction or by NMR. In order to begin to gain an understanding of structure-function relationships, we have taken a three-pronged approach.

(a) Sequence determination of Epo from mammals of different orders in order to establish regions of homology (26).

(b) Construction of a model of the three-dimensional structure of Epo, followed by the design and preparation of muteins that test this model. These experiments are presented in this paper.

(c) Design and testing of muteins that provide information on receptor binding domain(s). This work will be presented in a subsequent paper.

MATERIALS AND METHODS

Computer-based Modeling of Structure

Prediction of Secondary Structure—Epo sequences from human, monkey, mouse, rat, sheep, pig, and cat were aligned (26) and examined using a hierarchical approach to secondary structure prediction that assumes that these proteins are members of the α/α folding class (27). First, the pattern-based method of Cohen et al. (28) for turn prediction was used to delimit sequence blocks likely to contain secondary structure. Predictions using the methods of Garnier et al. (29) and Chou and Fasman (30) suggested α-helical regions within these blocks. Finally, helical wheel projections were used to examine and then limit helix length based on preserving amphipathic character as codified in the work of Presnell et al. (31). The locations of glycosylation sites were also used to suggest helix boundaries.

Tertiary Structure Prediction—Earlier investigations have revealed the general principles of helix-to-helix packing in globular proteins (32). Exploring these principles, Cohen et al. (33) developed a method for the generation of three-dimensional protein structures from the secondary structure assignment. These methods have been applied to myoglobin, tobacco mosaic virus coat protein, growth hormone, prolactin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, G-CSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor, and ciliary neurotrophic factor (15-17). The genes encoding these proteins have similar numbers of exons as well as a clear relationship between intron-exon boundaries and predicted α-helical structure. These similarities have led to the prediction that this family of cytokines share a common pattern of folding into a compact globular structure consisting of four amphipathic α-helical bundles. Such theoretical models of the structures of human growth hormone (18) and IL-4 (19) have been in remarkably good agreement with subsequent structures established by x-ray diffraction (human growth hormone) (20, 21) or by multidimensional NMR (IL-4) (22, 23). Moreover, the crystal structures of GM-CSF (24) and monomeric M-CSF (25) are also in reasonable agreement with their predicted structures.

Thus far, the structure of Epo has not been analyzed by either x-ray diffraction or by NMR. In order to begin to gain an understanding of structure-function relationships, we have taken a three-pronged approach.

(a) Sequence determination of Epo from mammals of different orders in order to establish regions of homology (26).

(b) Construction of a model of the three-dimensional structure of Epo, followed by the design and preparation of muteins that test this model. These experiments are presented in this paper.

(c) Design and testing of muteins that provide information on receptor binding domain(s). This work will be presented in a subsequent paper.
Model of Erythropoietin Structure

The algorithm for tertiary structure generation is divided into four steps. The first step, aafold, generates all possible helix pairings according to the location and geometric preferences of the interaction sites. The second step, aafold generates three-dimensional models of all possible helix structures from the list of helix pairings (from aafold) and subject to steric restrictions and geometric constraints on chain folding. In the final step, aafold applies the user-defined distance constraints (e.g. disulfide bridges) to the structures generated. At this stage, the entire coding sequences of the Epo mutants were examined for the presence of unwanted mutation by sequencing or restriction analysis. Oligonucleotides (24-46-mer) were synthesized with their 5' and 3' ends complementary to the target wild type Epo sequence. A wide variety of mutations (base substitutions, deletions and insertions) have been specified only for residues in the core α-helices. For residues in sequentially distinct loops, lower bounds on the inter-residue distances can be inferred from the relevant helix terminus.

Preparation of Epo Mutants

Construction of the Mutagenic/Mammalian Expression Plasmid—A M13 plasmid, containing a 1.4-kilobase EcoRI-EcoRI human Epo cDNA insert (HEPO FL-12) was a gift from Genetics Institute (Cambridge, MA) (8). A 943-bp EcoRI-BglII fragment, corresponding to the complete coding sequence of the wild type human erthropoietin, including untranslated regions 216 bp upstream and 183 bp downstream, was inserted into the mammalian expression plasmid pSV2 (Stratagene) (37) and named pSV2/EPO/WT.

Site-directed—was carried out according to the protocol described by Kuneki et al. (38). Single-stranded DNA was rescued from the pSG5-EPO/WT phagemid grown overnight in Escherichia coli C3236, in 2XYT media containing M13K07 helper phage (In Vitrogen) and 70 μg/ml kanamycin (Sigma). The resulting uracil-containing single-stranded DNA was used as a template for mutagenesis. Oligonucleotides (24-46-mer) were synthesized with their 5’ and 3’ ends complementary to the target wild type Epo sequence. A wide variety of mutations (base substitutions, deletions and insertions) were created at the centers of the mutagenic primer sequences. Annealing of the phosphorylated primers (10:1 oligonucleotide/DNA template ratio) was performed in 10 μl of a 20 mM Tris- HCl, pH 7.4, 2 mM MgCl2, 50 mM NaCl solution. The reactions were incubated at 80 °C for 5 min and then allowed to cool slowly to room temperature over a 1-h period. The DNA polymerization was initiated by the addition of a mix of 1 μl of 10 × synthesis buffer (100 mM Tris-HCl, pH 7.4, 50 mM MgCl2, 10 mM ATP, 5 μm each dNTPs, 20 mM dithiothreitol), 0.5 μl (8 units) of T4 DNA ligase and 1 μl (1 unit) of T4 DNA polymerase (Boehringer Mannheim). After 2 h at 37 °C, 80 μl of 1 × Tris-EDTA was added. 5 μl of the diluted reaction mix was used to transform competent E. coli strain NM522 (unp, dut+)

Since a 40-95% mutation yield is normally obtained, four to five double-stranded plasmid DNA samples were sequenced by direct sequencing with primers flanking the mutation. The mutation at each site was checked by sequencing the clones from each experiment. Home computer programs of T. A. J. Schmid were used to analyze the DNA sequences. Oligonucleotide-directed mutagenesis was performed with the help of an automated DNA synthesizer (Applied Biosystems Inc.). The quality of the synthesized oligonucleotides was determined by combination of gel electrophoresis and automated fluorescence-based detection.

Production of Wild Type and Epo Mutants in Mammalian Cells—Cos-7 cells grown to ~70% confluence were transfected with 10 μg of recombinant plasmid DNA/10-cm dish using the calcium phosphate precipitation protocol (40). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia LKB Biotechnology Inc.) was cotransfected and β-galactosidase activity measured using the pBlue-β-gal (Pharmacia) for 21. Immunoprecipitations were performed overnight with our polyclonal antibody specific for human recombinant wild type Epo and immunoadsorbed with protein A-Sepharose CL-4B. Immunoprecipitates were run on 15% SDS-polyacrylamide gels (44) and analyzed by autoradiography after transfer to nitrocellulose by the procedure of Towbin et al. (45, 46).

Bioassays—The dose-dependent proliferation activities of WT and Epo mutants were assayed in vitro using three different target cells: murine spleen cells, following a modification of the method of Krystal et al. (47); and human Epo-dependent UT-7/Epo cell line, derived from the bone marrow of a patient with acute megakaryoblastic leukemia (48). After 22-72 h of incubation with increasing amounts of recombinant proteins, cellular growth was determined by [3H]thymidine incorporation (49). The [3H]thymidine incorporation was carried out by incubating cells for 24 h with [3H]thymidine (Du Pont-New England Nuclear) uptake and using the colorimetric MTT assay (Sigma) (49).

Production of Wild Type and Epo Mutants in Mammalian Cells—Cos-7 cells grown to ~70% confluence were transfected with 10 μg of recombinant plasmid DNA/10-cm dish using the calcium phosphate precipitation protocol (40). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia LKB Biotechnology Inc.) was cotransfected and β-galactosidase activity measured using the pBlue-β-gal (Pharmacia) for 21. Immunoprecipitations were performed overnight with our polyclonal antibody specific for human recombinant wild type Epo and immunoadsorbed with protein A-Sepharose CL-4B. Immunoprecipitates were run on 15% SDS-polyacrylamide gels (44) and analyzed by autoradiography after transfer to nitrocellulose by the procedure of Towbin et al. (45, 46).

Bioassays—The dose-dependent proliferation activities of WT and Epo mutants were assayed in vitro using three different target cells: murine spleen cells, following a modification of the method of Krystal et al. (47); and human Epo-dependent UT-7/Epo cell line, derived from the bone marrow of a patient with acute megakaryoblastic leukemia (48). After 22-72 h of incubation with increasing amounts of recombinant proteins, cellular growth was determined by [3H]thymidine incorporation (49). The [3H]thymidine incorporation was carried out by incubating cells for 24 h with [3H]thymidine (Du Pont-New England Nuclear) uptake and using the colorimetric MTT assay (Sigma) (49).

RESULTS

Construction of a Model of the Three-dimensional Structure of Erythropoietin

From an analysis of the putative Epo helix sequences, aapatch identified eight possible helix-helix interaction sites.
Table I

<table>
<thead>
<tr>
<th>Helix</th>
<th>NH₂ terminus</th>
<th>COOH terminus</th>
<th>Potential helix-helix interaction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>B</td>
<td>59</td>
<td>76</td>
<td>63, 67, 70, 71</td>
</tr>
<tr>
<td>C</td>
<td>90</td>
<td>107</td>
<td>95, 102</td>
</tr>
<tr>
<td>D</td>
<td>132</td>
<td>152</td>
<td>141</td>
</tr>
</tbody>
</table>

In principle, these sites could be used to generate $1.6 \times 10^4$ structures. Of these, only 706 maintained the connectivity of the chain and were sterically sensible. These structures resembled four helix bundles, an increasingly common motif in protein structure (53). The structures that were not compatible with the native disulfide bridge between Cys²⁷ and Cys³⁶ were eliminated. This reduced the total number of structures from 706 to 184 (total computer time approximately 1 h on a Silicon Graphics IRIS 4D/35G). The remaining structures were then ranked ordered by solvent-accessible surface contact area, a measure of the validity of model structures. The most compact structures were right-handed, all anti-parallel four-helix bundles with no overhand connections, but this may be an artifact of a failure to add the polypeptide chain that forms the loops to the helical core constructed by abuild. The other less compact structures were left-handed four-helix bundles with two overhand loops, a topology previously seen in the structures of IL-4 and growth hormone. We suspect that this is the likely structure for Epo. The consensus for assignments of putative α-helices in human Epo are summarized in Table I. First, analysis of the topological distribution of known four-helix bundle structures indicates that nearly all examples have an antiparallel orientation (53). Second, the left-handed four-helix bundles with two overhand connections arrange the four amphipathic helices to form a compact hydrophobic core. Finally, the AB and CD loop regions of Epo are predicted to have β-sheet segments analogous to IL-4 and growth hormone that preserves the compact globular nature of the Epo model structure. Fig. 1 shows schematic representations of predicted topological interactions between the four anti-parallel α-helical bundles.

Several authors have suggested that the helical cytokines form a structural superfamily (34, 54-57). On the basis of both the mature protein and the individual α-helices, Epo seems to be more closely related to growth hormone, prolactin, IL-6, and GM-CSF rather than the other members of the helical cytokine superfamily. Nevertheless, recent improvements in algorithms for the identification of distant evolutionary relationships between proteins from structural fingerprints suggested that it might be possible to align the IL-4 structure to the Epo sequences. The Eisenberg et al. (58) structural environment and 3D-1D profile methods are a powerful tool for recognizing that a sequence is compatible with a known structure, e.g. a four-helix bundle. The NMR structure of IL-4 from Smith et al. (59) was used to construct a 3D-1D profile. A mixture of sequences including four helix bundles, globins, and non-helical structures were aligned against the IL-4 profile. Not surprisingly, the IL-4 structures from human and mouse gave the highest scores ($Z^2 = 22.8$ and 8.1). However, the other known four-helix bundle cytokines known to share a similar fold with IL-4, e.g. human growth hormone (60) ($Z = 2.3$) and GM-CSF (61) ($Z = 2.3$) fared no better than some globin sequences (Kuroda’s and slug sea hare globin, $Z = 5.0$ and 4.8) that adopt a distinct tertiary structure. The results for the human and sheep Epo sequences were also ambiguous ($Z = 1.6$ and 0.8). These results suggest that while profile methods are a powerful tool for recognizing structural similarity, their failure to identify homology does not exclude the possibility that two proteins share a common fold. For distantly related or unrelated structures, current profile methods cannot replace de novo methods for tertiary structure prediction.

Design and Expression of Epo Muteins That Test the Proposed Structure

To test the proposed four α-helical bundle structure of erythropoietin and at the same time to attempt to locate functional domains, we created by site-directed mutagenesis a series of deletions, insertion, and replacement mutants. These muteins were designed to analyze the principal predicted structural features of the molecule: α-helices, interconnecting loops, as well as the NH₂ and COOH termini. Structural and functional implications of the disulfide bridges and the glycosylation sites were also investigated.

- **α-Helices**—Short amino acid deletions were prepared in, or close to, the predicted A, B, C, and D α-helices. Human wild type and muteins were transiently expressed in Cos7 cells. Northern blot analyses demonstrated that all the mutant plasmids produced about the same amount of mRNA as that of the wild type (data not shown). Yet, no detectable amount of Epo protein could be found in the Cos7 supernatants, either by radioimmunoprecipitation or by bioassay using various Epo-dependent cell lines. Table II summarizes these findings.

An example of SDS-polyacrylamide gel electrophoresis of immunoprecipitation bands from vivo [³⁵S]labeled is presented in Fig. 2. As expected, when Cos7 cells were transfected with pSG5-EPO/WT, a 35-37 kDa band was detected in the supernatant. In contrast, the deletion mutants (Table II) could be detected in cellular extracts but were not exported from the cells. Fig. 2 shows the cytoplasmic retention of the mutein Δ140-144, lacking 4 residues in the middle of the predicted D-helix. The apparent molecular mass (~28 kDa) is less than expected for a 5-amino acid deletion. Therefore, not only the secretion, but also the glycosylation, seem to be impaired. None of these muteins had deletion of glycosylation sites. It is likely that full glycosylation of Epo requires conservation of its molecular architecture. Similar results (reported in Table II) were obtained for all the muteins having partial deletion of an α-helical peptide segment.

Because contaminants in crude Cos7 cellular extracts severely interfere with the radioimmunoprecipitation, no direct Epo quantitation was possible. However, aliquots of hypotonic extracts of Cos7 transfected with wild type Epo were able to sustain HDC57 proliferation. No similar biological activity was found for muteins with limited deletion of α-helices.

- **Interconnecting Loops**—The peptide segment joining A- and B-helices presents several interesting features (Fig. 3A). The AB loop consists of 36 amino acids. Two N-glycosylation sites and a small disulfide bridge are located in the first half and their biological implications will be discussed later. The COOH end of the AB loop contains a stretch of amino acids that is strongly conserved among mammals (26). Alignments of human, monkeys, cat, mouse, rat, pig, and sheep Epos showed a consensus sequence: DTKYNFAYWKR(M/I)(E/D)VYG (residues 43-57). Three deletions were constructed:
FIG. 1. Model of the three-dimensional structure of erythropoietin. A, ribbon diagram of the predicted Epo tertiary structure. The four α-helices are labeled A–D (magenta); Loops between helices are named for the helices they interconnect. Two regions of extended structure which could form hydrogen bonds between Loop AB and Loop CD are also present (cyan). N- and O-glycosylation sites are indicated in green and blue, respectively. Disulfide bonds bridge residues 29–33 in Loop AB, and 7–161 on the NH2-terminal side of Helix A and the COOH-terminal side of Helix D are not shown. N.B.: The loop tracing shown does not represent predicted coordinates. B, schematic representation of Epo's primary structure depicting predicted up-up-down-down orientation of the four antiparallel α-helices (boxes with arrowhead). This folding pattern is strongly suggested by the large size of the two interconnecting loops AB and CD. The limits of each helix were drawn accordingly to Table I. A predicted short region of β-sheet is delineated by the dashed rectangle. The N-glycosylation sites are represented by the dotted diamonds, and the O-glycosylation site by the dashed oval. The locations of the two disulfide bridges are shown. C, cross-section of the Epo molecule at the level of the four α-helices. The helical wheel projections are viewed from the NH2 end of each helix. The hydrophobic residues, localized inside the globular structure, are indicated by filled circles. The charged and neutral residues (open and gray circles, respectively) are exposed at the surface of the molecule.

Δ43–47, Δ48–52, and Δ53–57, and transiently expressed in Cos7. The amount of muteins detected by RIA in the supernatants of transfected cells was 10–40% lower than observed with wild type Epo (Fig. 3B). Nevertheless, the three secreted muteins were biologically active. However, because Δ48–52 exhibited a marked decrease of the specific bioactivity, this site was studied in more detail by means of serine replace-ments. Krystal ex vivo bioassay as well as HCD57 and UT7-Epo in vitro bioassays showed that these Ser mutants had biological activities similar to that of wild type (Fig. 3C). Therefore, the observed decreases in both RIA and bioassay for the three deletion mutants are likely to be the result of changes of structural conformation. The long length of loop AB may be critical for the up-up-down-down topography. A
shorter AB segment may impose a strain on the interhelical connection. Chou and Fasman (30) algorithms predicted a short $\beta$-sheet structure from residues 44 to 51 (<Pa>;<Pb>, 1.005 < 1.196). The presence of a short region of $\beta$-sheet in the connection between helices 1 and 2 (A and B) have been documented in the analyses of the three-dimensional structures of IL-4 (22, 23), GM-CSF (24), and monomeric M-CSF (25). In contrast, in human GH a short segment of $\alpha$-helix is found at the same location (20). The structure/function implications of these short features are not yet understood.

Helix B is linked to helix C by a much shorter segment (residues 77–89) and contains in its center the third N-glycosylation site (Asn$^{68}$). When the $\Delta 78$–82 mutein was expressed, a secreted protein was detected in the conditioned medium and conferred proliferative bioactivity on Epo-dependent cell lines (see Fig. 8).

A similar long crossover connection (23 amino acids) is found between helix C and helix D. In contrast to what we previously observed for loop AB, a large deletion of 9 residues at position 111–119 or a 7-amino-acid insertion of a myc epitope after residue 116 did not affect the secretion of these muteins (Fig. 4). Furthermore, these two proteins had normal specific activity, as seen by the ratio of bioassay to RIA. Our rabbit polyclonal antibody raised against the native form of the human wild type fully recognized the two mutants, demonstrating that the overall spatial conformation of Epo was well preserved. According to the algorithm of Emini et al. (62), the residues 111–119 are predicted to be at the surface of the molecule. Since the $\Delta 111$–119 mutein is readily secreted and has full biological activity, it seems unlikely that the putative $\beta$-sheet segment in the CD loop is an important determinant of molecular stability. Primary amino acid alignments of mammalian Epo showed a large variation in the sequence of residues 116–130, including amino acid deletion, insertion, and substitution (26). Surprisingly, when the deletion $\Delta 122$–126 mutein, which removed the O-glycosylation site (Ser$^{126}$), was transiently expressed in monkey cells, protein secretion was inhibited. Both rodents, rat and mouse, lack the O-glycosylation site because of a Ser$^{126}$ to Pro replacement. Furthermore, when a Ser$^{126}$ replacement mutein was expressed in normal Chinese hamster ovary cells, (63) or when wild type Epo was expressed in cells having a defect in O-linked glycosylation (64), neither secretion nor biological activity were impaired. Therefore, failure of secretion of the $\Delta 122$–126 mutein may be the result of some other structural alteration. In particular, the proline residue at position 122 is invariant among mammals.

$NH_2$ and COOH Termini—Deletion of residues 2 to 5 only slightly affected the processing of a biologically active protein (see Fig. 8). This deletion may impair cleavage of the propeptide, therefore explaining the lower yield of secreted Epo mutein in comparison to that of wild type. The fact that the mature monkey protein has an elongated (Val-Pro-Gly) NH$_2$ terminus strongly suggests that the NH$_2$-terminal part is not involved in the bioactivity of the molecule. Further evidence comes from the results, reported below, on the N-poly-His-Epo fusion protein expressed in E. coli, and also from the identical binding of in vitro translated $^{35}$S-labeled wild type Epo onto EREx–glutathione agarose beads (Fig. 5), with or without addition of canine pancreatic microsomal membranes which permit cleavage of the propeptide.

The COOH-terminal sequence following helix D can clearly be divided into two distinct domains, separated by Cys$^{165}$. The residues 151–161 were of special interest because they are highly conserved among mammals (26). There are only two substitutions: Lys$^{164}$ is replaced by a Thr in artiodactyls and cat, and Ala$^{165}$ is replaced by a Val in mouse Epo. Both the $\Delta 152$–155 and the $\Delta 156$–160 muteins remained in the cytosol of the transfected Cos7 (Table II). One possible explanation is that the residues 152–160 may, in fact, participate in the D helix. We predict that Glu$^{153}$ is the break point of the struct-

---

**TABLE II**

**Short deletions in, or close to, $\alpha$-helices**

<table>
<thead>
<tr>
<th>Mutants</th>
<th>NH$_2$ Terminus of each $\alpha$-helix are indicated in the vertical boxes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta 12$–16</td>
<td>9 A helix</td>
</tr>
<tr>
<td>$\Delta 65$–69</td>
<td>59 B helix, RNA levels comparable to WT.</td>
</tr>
<tr>
<td>$\Delta 96$–100</td>
<td>90 C helix, No detectable Epo in the Cos7 supernatant, both by RIA and bioassay.</td>
</tr>
<tr>
<td>$\Delta 105$–109</td>
<td>107</td>
</tr>
<tr>
<td>$\Delta 122$–126</td>
<td></td>
</tr>
<tr>
<td>$\Delta 131$–135</td>
<td>132 D helix</td>
</tr>
<tr>
<td>$\Delta 140$–144</td>
<td></td>
</tr>
<tr>
<td>$\Delta 152$–155</td>
<td>152</td>
</tr>
<tr>
<td>$\Delta 156$–160</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Immunoprecipitations of wild type Epo and the $\Delta 140$–144 mutein. Cos7 cells were transfected with pSG5, pSG5-EPO/wt, or pSG5-EPO/$\Delta 140$–144. After 3 days, the cells were metabolically labeled with $[^{35}]$S-methionine and $[^{35}]$S-cysteine. Immunoprecipitates of cellular extracts and supernatants were performed with our polyclonal antibody, raised in rabbit against the native form of human Epo. The immunoprecipitates were analyzed by SDS-polyacrylamide gel electrophoresis. Lanes 2–4 correspond to the cellular extracts; lanes 5–7 in the culture supernatants from transformed Cos7 with: plasmid without insert (lanes 2 and 5), wild type Epo (lanes 3 and 6), and $\Delta 140$–144 (lanes 4 and 7). Lane 1 represents the protein molecular weight standard. The two arrows show the normal secretion of the wild type Epo (35–37 kDa) and the cytoplasmic retention of the mutein $\Delta 140$–144 (~28 kDa).

---

1 All the mutants described in this paper were subcloned into pSPG4T plasmid. Studies of the binding of their translation products to EREx are in process.
Fig. 3. Interconnecting loop AB. A, schematic representation of the loop AB showing the localization of muteins with various deletions and amino acid replacements. The dashed arrows point to the positions of the serine substitutions (in Δ48–52). The two N-glycosylation sites are represented by the gray diamonds. The small Cys²⁸–Cys³⁰ disulfide bridge is indicated. B, amount and biological activities of secreted
Model of Erythropoietin Structure

A LOOP CD

B RIA mU/ml

Specific activity %

WT Δ105-109 Δ111-119 116/myc Δ122-126

C

D

FIG. 4. Interconnecting loop CD. A, schematic representation of the loop CD showing the location of three deletion muteins: Δ105-109, Δ111-119, Δ122-126, and the insertion of 7 residues after Lys166 (myc epitope). The O-glycosylation site is indicated by the dashed oval. B, secretion and biological activities of the muteins located in loop CD. The two bar graphs were created as described in Fig. 3B. The two mutants Δ111-119 and 116/myc were normally secreted and had full biological activities. mU, milliunit.

However, it is possible that this residue causes only a bend in the α-helical structure and helix D may extend to Gly186. The COOH-terminal part of the protein (residues 162-166) is clearly not involved in any structural or functional feature. Thus, the deletion of the 4 last amino acids or the replacement of residues 162-166 by either a KDEL sequence or a poly-histidine sequence did not modify the specific activity of the erythropoietin (Fig. 6). Radioimmunoassay revealed that the secretion of the KDEL-tail mutein in the media of transfected cells was 45% less than normally obtained with the wild type Epo. However, when compared to the wild type, this mutein had more biological activity in the hypotonic Cos7 cell extracts. The KDEL COOH-terminal sequence has been shown to be essential for the retention of several proteins in the lumen of the endoplasmic reticulum. Nevertheless, because of overproduction in transiently expressed cells, a large percentage of recombinant protein escaped into the media.

Disulfide Bridges—Wang et al. (66) demonstrated that the biological activity of Epo was lost irreversibly if the sulfhydryl

muteins. The upper bar graphs show the relative secretion of wild type and loop AB muteins as determined by radioimmunoassay. The lowest bar graphs display the calculated specific activity (ratio bioassay/RIA) for each mutein, in comparison with the value obtained for the wild type Epo (ratio = 100%). C, HCD57 cell proliferation as a function of increasing concentration of wild type and serine-substituted Epo muteins. HCD57 cells (10^5/ml) were cultured for 3 days in a 96-well microtiter plate with media containing increasing concentrations of secreted proteins. The line graphs show the cellular growth as measured by [3H]thymidine uptake. The number of viable cells was also measured with the MTT colorimetric assay and gave similar curves. In vitro proliferation experiments using the human UT-7 cell line (48) and the Krystal assay (45) produced identical results. cpm, counts/minute; mU, milliunit.
Model of Erythropoietin Structure

Fig. 5. *In vitro* translation of the Epo wild type. *A,* analysis of the 35S-labeled translation products by SDS-polyacrylamide gel electrophoresis. One-step transcription/translation reactions were performed in the SP6-TnT rabbit reticulocyte lysate system. 1/30 of each reaction was resolved on a 15% polyacrylamide gel. *Lane 1,* low $M_\text{r}$ standard from Amersham Corp.; *lanes 2, 3, and 4,* translation products obtained after incubation of 1 $\mu$g of circular p64T-Epo, respectively, in the presence or absence of canine pancreatic microsomal membranes. *B,* binding of the *in vitro* translated Epo wild type onto Epo receptor-GTS-agarose beads. 6 $\times$ 106 cpm/min of purified 35S-labeled erythropoietin products, processed with microsomes (+) or not (−) were incubated in the presence of EREx, following the protocol described by Harris et al. (52). Identical binding demonstrated that the conservation of the propeptide did not impair the hormone-receptor interaction.

Fig. 6. COOH end of Epo. *A,* schematic representation of the analyzed muteins, corresponding to the deletion of the four last amino acids Δ163-166 and the replacements of the residues 162-166 by a KDEL or poly (His) sequences. *B,* relative secretion of these muteins. The bioactivities in the supernatants of the propeptide did not impair the hormone-receptor interaction.

The bioactivities in the supernatants of the propeptide did not impair the hormone-receptor interaction.
**Model of Erythropoietin Structure**

DISCUSSION

Currently, the accrual rate of new protein sequences through gene cloning far outstrips the rate of determination of three-dimensional structure. Epo is among a large number of biologically important proteins which have not yet been analyzed by x-ray diffraction or NMR. The problem is simplified by cumulative evidence that the structures of most proteins are likely to be variations on existing themes (27). Indeed, as mentioned above, Epo appears to share common structural features with a large group of cytokines (15-17).

Computer-based prediction of structure can be reduced to a three-stage process: secondary structure is predicted from the primary amino acid sequence and, when available, optical measurements. Analysis of Epo by circular dichroism reveals about 50% α-helix and no detectable β-sheet (7, 11). With the knowledge of disulfide bonds, secondary structural elements are then packed into a set of alternative tertiary structures. The number of plausible arrangements can be reduced by empirical knowledge of preferred helix-helix packing geometries and the need for globular structure to form a hydrophobic core. The putative tertiary structure is then refined by standard force field calculations. Since there are a large number of alternate tertiary structures, the availability of experimentally determined structure of a homologous protein is critically important. Thus, the predicted model of Epo structure gains considerable validity by knowledge of the structures of growth hormone (20, 21) and IL-4 (22, 23).

We have tested the predicted four anti-parallel α-helical bundle structure by means of site-directed mutagenesis. Deletions within predicted α-helices would be expected to destabilize tertiary structure, whereas deletions or insertions in non-helical segments should be permitted unless they impose undue strain on the structure. For example, a deletion in an overhand inter-helical loop may result in insufficient length to connect the two helices. Results that we have obtained on muteins produced in mammalian (Cos7) cells are summarized in Fig. 8. Our measurements of the quantities of processed mutein by RIA may underestimate the true amount of secreted Epo. Even a small deletion or insertion can result in a conformational change that may lead to impaired binding by our polyclonal antibody, raised against native human Epo. Thus the values for specific activity (biologic activity/RIA) that we report must be regarded as approximations. This caveat notwithstanding, our mutagenesis results are in good agreement with our proposed four α-helical model of erythropoietin. The proper folding of Epo into its native tertiary structure is necessary for stability and biological function. Muteins with short deletions inside predicted α-helices were not processed and exhibited no biological activity. In contrast, when deletions were created in predicted interconnecting loops, secreted proteins were detected, to varying degrees, both by radioimmunoassay and bioassay. Furthermore, additions or deletions at the NH₂ or COOH termini did not markedly impair the secretion and the biological activity of the Epo protein. Moreover, mutations at Cys²⁹ and Cys³⁵ showed that the small disulfide loop is not critical for biological activity. In order to delineate Epo’s functionally important residues involved in the direct binding onto the Epo receptor, we have prepared and tested a series of amino acid replacements on the surfaces of the predicted α-helices. These experiments will be described in a subsequent paper.

**Acknowledgments**—We thank Elizabeth Eldridge for technical assistance and Hanna Bandes for secretarial and editorial assistance. We also thank Dr. David Hankins for providing the HCD57 cell line,
Fig. 8. Relationship between production of mutants and proposed secondary structure. This bar graph shows the amount of secreted proteins in the supernatants of transiently expressed Epo mutants, as detected by radioimmunoassay. The mutants were aligned over a schematic representation of the native Epo molecule. Each deletion is shown as a stippled bar, the width of which is proportional to the number of residues deleted. The four α-helices are represented by the black rectangles. The two disulfide bridges are indicated. These mutagenesis results are in good agreement with our proposed four α-helical model of Epo.

REFERENCES

Model of Erythropoietin Structure