

TABLE XI
ECEPO SECTION 2

ECORI KpnI
A ATTCGGTACC ¹AGACCCCAAG GTAACTCT ³ACCGTGGAA ACGIATGGAA
GCCATGG TCGTGGTTC ²CAATTGAGA TGGGACCTT ⁴TGCATACCTT

⁵GTGGCAGC AAGCAGTGA AGTGGCCAG ⁷GGCTGGCC TCCTGACCC
⁶CAACCAATG TTCGCAACT TCAGCCCTC ⁸CCAGACCGTG ACGACTCGCT

CGCTGACIG ⁹CGTGGCAGG CACTGGTGGT AACTGGCTC CAGCCGTGG
CCGACATGAC ¹⁰GCACCGCTCC GTGACGACCA ¹¹TTGAGGAGA GTGGCCACC
¹²

¹³AACCCGTGA GCTGCATGT GACCAAGCAG ¹⁵TATCTGGCC GAGATCTG
¹⁴TTGGGACGT CGACGTACA CTGTTCTG ATAGCCCGA ¹⁶CTCTAGACCTAC
BamHI

A 6934

- 70 -

TABLE XII

ECEPO SECTION 3

- 1. GATCCAGATCTCTGACTACTCTGC
- 5 2. ACGCAGCAGAGTAGTCAGAGATCTG
- 3. TCCGTGCTCTGGGTGCACAGAAAGAGG
- 4. GATAGCCTCTTTCTGTGCACCCAGAGC
- 5. CTATCTCTCCGCCGGATGCTGCATCT
- 6. CAGCAGATGCAGCATCCGGCGGAGA
- 10 7. GCTGCACCGCTGCCATACCATCACTG
- 8. ATCAGCAGTGATGGTACGCAGCGGTG
- 9. CTGATACCTTCCGCAAACTGTTTCG
- 10. ATACACGAAACAGTTTGGGGAAGGT
- 11. TGTATACTCTAACTTCCTGGGTGGTA
- 15 12. CAGTTTACCCAGCAGGAAGTTAGAGT
- 13. AACTGAAACTGTATACTGGCGAAGC
- 14. GGCATGCTTCGCCAGTATACAGTTT
- 15. ATGCCGTACTGGTGACCCGCTAATAG
- 16. TCGACTATTAGCGGTCACCACTAC

20

25

A 6935

- 71 -

TABLE XIII
ECEPO SECTION 3

BamHI BglII
GA TCCAGATCTCTG
GTCTAGAGAC

5

ACTACTCTGC 1 TGGGTGCTCT 3 GGGTGCACAG AAAGAGGCTA 5 TCTCTCCGCC
TGATGAGACG 2 ACCCAGGAGA CCCACGTGTC TTCTCCGAT 4 AGAGAGGCCG

GGATGCTGCA TCTCTGAC 7 GGTGCGTAC CATCACTGCT 9 GATACCTCC
CCTACGACGT 6 AGACGAGCTG 8 GCGACCCATG GTAGTGACGA CTATGGAAGG

10

GCAAACCTGTT TCGTGTATAC 11 TCTAACTTCC TGGGTGGTAA 13 ACTCAAACCTG
CGTTTGACAA 10 AGCACATATG AGATTGAAGG 12 ACGCACCATT 11 TGACTTTGAC

15

TATACTGGCG 14 AAGCATGCCG 15 TACTGGTGAC CGCTAATAG SalI
ATATGACCGG TTCGTACGGC 16 ATGACCACTG GCGATTATC AGCT

20

25

A 6936

- 72 -

TABLE XIV
ECEPO GENE

	<u>XbaI</u>			-1	1
	<u>CTAG</u>	AAACCATGAG	GGTAATAAAA	TAATGGCTCC	MetAla
5		TTTGCTACTC	CCATTATTTT	ATTACCGAGG	GCCGCGTCTG
					CGGCGCAGAC
		ATCTGGCACT	CGAGAGTTCT	GGAACGTTAC	CTGCTGGAAG
		TAGACCGCTGA	GCTCTCAAGA	CCTTGCAATG	GACGACCTTC
					GATTTCTTCG
		TGAAAACATC	ACCACTGGTT	GTGCTGAACA	CTGTTCTTTG
		ACTTTTGTAG	TGGTGACCAA	CACGACTTGT	GACAAGAAAC
					TTGCTTTTGT
10		TTACGGTACC	AGACACCAAG	GTTAACTTCT	ACGCTTGGAA
		AATGCCATGG	TCTGTGGTTC	CAATTGAAGA	TGCGAACCTT
					TGCATACCTT
		GTTGGTCAAC	AAGCAGTTGA	AGTTTGGCAG	GGTCTGGCAC
		CAACCAGTTG	TTCGTCAACT	TCAAACCGTC	CCAGACCGTG
					ACGACTCGCT
		GGCTGTACTG	CGTGGCCAGG	CACTGCTGGT	AAACTCCTCT
15		CCGACATGAC	GCACCGGTCC	GTGACGACCA	TTTGAGGAGA
					CTCGGCACCC
		AACCGGTGCA	GCTGCATGTT	GACAAAGCAG	TATCTGGCCT
		TTGGCGACGT	CGACGTACAA	CTGTTTGGTC	ATAGACCGGA
					CTCTAGAGAC
		ACTACTCTGC	TGGTGTCTCT	GGGTGCACAG	AAAGAGGCTA
		TGATSAGACG	ACGCACGAGA	CCCACGTGTC	TTTCTCCGAT
					AGAGAGGCCG
20		GGATCCTGCA	TCTGCTGCAC	CGCTGCGTAC	CATCACTGCT
		CCTACGACGT	AGACGACGTG	GCGACGCATG	GTAGTGACGA
					CTATGGAAGG
		GCAAACTGTT	TGGTGTATAC	TCTAACTTCC	TGGTGGTAA
		CGTTTGACAA	AGCACATATG	AGATTGAAGG	ACGCACCATT
					TGACTTTGAC
		TATACTGGGG	AAGCATGCCG	TACTGGTGAC	CGCTAATAG
25		ATATGACCCG	TTGGTACGGC	ATGACCACTG	GCGATTATCA
					GCT

A 6937

10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10
10

More particularly, ^{Figure 10} ~~Table VIII~~ illustrates oligo-
 nucleotides employed to generate the Section 1 of the
 ECEPO gene encoding amino terminal residues of the human
 species polypeptide. Oligonucleotides were assembled
 5 into duplexes (1 and 2, 3 and 4, etc.) and the duplexes
 were then ligated to provide ECEPO Section 1 as ^{Figure 11} ~~in Table~~
~~IX~~. Note that the assembled section includes respective
 terminal EcoRI and BamHI sticky ends, that "downstream"
 of the EcoRI sticky end is a XbaI restriction enzyme
 10 recognition site; and that "upstream" of the BamHI sticky
 end is a KpnI recognition site. Section 1 could readily
 be amplified using the M13 phage vector employed for
 verification of sequence of the section. Some dif-
 ficulties were encountered in isolating the section as an
 15 XbaI/KpnI fragment from RF DNA generated in E.coli,
 likely due to methylation of the KpnI recognition site
 bases within the host. Single-stranded phage DNA was
 therefore isolated and rendered into double-stranded form
 in vitro by primer extension and the desired double-
 20 stranded fragment was thereafter readily isolated.
 ECEPO gene Sections 2 and 3 (^{Figures 13 and 15} ~~Tables XI and XIII~~)
 were constructed in a similar manner from the oligo-
 nucleotides of ^{Figures 12 and 14} ~~Tables X and XII~~, respectively. Each
 section was amplified in the M13 vector employed for
 25 sequence verification, and was isolated from phage DNA.
 As is apparent from ^{Figure 13} ~~Table XI~~, ECEPO Section 2 was con-
 structed with EcoRI and BamHI sticky ends and could be
 isolated as a KpnI/BglII fragment. Similarly, ECEPO
 Section 3 was prepared with BamHI and SalI sticky ends
 30 and could be isolated from phage RF DNA as a BglII/SalI
 fragment. The three sections thus prepared can readily
 be assembled into a continuous DNA sequence ^{as depicted in Figure 7} ~~(Table XIV)~~
 encoding the entire human species EPO polypeptide with an
 amino terminal methionine codon (ATG) for E.coli transla-
 35 tion initiation. Note also that "upstream" of the ini-
 tial ATG is a series of base pairs substantially

A 6938

- 74 -

duplicating the ribosome binding site sequence of the highly expressed OMP-f gene of E.coli.

Any suitable expression vector may be employed to carry the ECEPO. The particular vector chosen for expression of the ECEPO gene as the "temperature sensitive" plasmid pCFM536 -- a derivative of plasmid pCFM414 (A.T.C.C. 40076) -- as described in co-pending U.S. Patent Application Serial No. 636,727, filed August 6, 1984, ^(Published EPO Application No. 110,440) by Charles F. Morris. More specifically, pCFM536 was digested with XbaI and HindIII; the large fragment was isolated and employed in a two-part ligation with the ECEPO gene. Sections 1 (XbaI/KpnI), 2 (KpnI/BglII) and 3 (BglII/SalI) had previously been assembled in the correct order in M13 and the EPO gene was isolated therefrom as a single XbaI/HindIII fragment. This fragment included a portion of the polylinker from M13 mp9 phage spanning the SalI to HindIII sites therein. Control of expression in the resulting expression plasmid, p536, was by means of a lambda P_L promoter, which itself may be under control of the C₁₈₅₇ repressor gene (such as provided in E.coli strain K12ΔHtrp).

The manufactured ECEPO gene above may be variously modified to encode erythropoietin analogs such as [Asn², des-Pro² through Ile⁶]hEPO and [His⁷]hEPO, as described below.

A. [Asn², des-Pro² through Ile⁶]hEPO

Plasmid 536 carrying the ECEPO manufactured gene of ^{Figure 7} ~~Table IV~~ as a XbaI to HindIII insert was digested with HindIII and XhoI. The latter endonuclease cuts the ECEPO gene at a unique, 6 base pair recognition site spanning the last base of the codon encoding Asp⁸ through the second base of the Arg¹⁰ codon. A XbaI/XhoI "linker" sequence was manufactured having the following sequence:

35

A 6939

- 75 -

T0610X

	<u>XbaI</u>		+1	2	7	8	9		<u>XhoI</u>
		Met	Ala	Asn	Cys	Asp			
5'	-CTAG	ATG	GCT	AAT	TGC	GAC	-3'		
3'	-TAC	CGA	TTA	ACG	CTG	AGCT	-5'		

The XbaI/XhoI linker and the XhoI/HindIII ECEPO gene sequence fragment were inserted into the large fragment resulting from XbaI and HindIII digestion of plasmid pCFM526 -- a derivative of plasmid pCFM414 (A.T.C.C. 40076) -- as described in co-pending U.S. Patent Application Serial No. 636,727, filed August 6, 1984, by Charles F. Morris, to generate a plasmid-borne DNA sequence encoding E.coli expression of the Met⁻¹ form of the desired analog.

B. [His⁷]NEPO

Plasmid 536 was digested with HindIII and XhoI as in part A above. A XbaI/XhoI linker was manufactured having the following sequence:

T0611X

	<u>XbaI</u>		+1	2	3	4	5	6	7	8	9		<u>XhoI</u>
		Met	Ala	Pro	Pro	Arg	Leu	Ile	His	Asp			
20	5'-CTAG	ATG	GCT	CCG	CCA	CGT	CTG	ATC	CAT	GAC	-3'		
	3'-TAC	CGA	GGC	GGT	GCA	GAC	TAG	GTA	CTG	AGCT	-5'		

The linker and the XhoI/HindIII ECEPO sequence fragment were then inserted into pCFM526 to generate a plasmid-borne DNA sequence encoding E.coli expression of the Met⁻¹ form of the desired analog.

Construction of a manufactured gene ("SCEPO") incorporating yeast preference codons is as described in the following ~~Figures 16 through 21 and 2~~ ^{Figures 16 through 21}. As was the case with the ECEPO gene, the entire construction involved formation of three sets of oligonucleotides (~~Tables XV, XVI, and XIX~~ ^{Figures 16, 17, 18 and 19}) which were formed into duplexes and assembled into sections (~~Tables XVI, XVIII and XX~~ ^{Figures 17, 19 and 21}). Note that synthesis was facilitated in part by use of some sub-optimal codons in both the SCEPO and ECEPO construc-

35

A 6940

- 76 -

itions, i.e., oligonucleotides 7-12 of Section 1 of both
genes were identical, as were oligonucleotides 1-6 of
Section 2 in each gene.

5

10

15

20

25

A 6941

30

- 77 -

TABLE XV

SCEPD SECTION 1 OLIGONUCLEOTIDES

- 1. AATCAAGCTTGGATAAAAGAGCT
- 5 2. GTGGAGCTCTTTTATCCAAGCTTG
- 3. CCACCAAGATTGATCTGTGACTC
- 4. TCTCGAGTCACAGATCAATCTTG
- 5. GAGAGTTTTGGAAAGATACTTGTG
- 6. CTCCAACAAGTATCTTTCCAAAAC
- 10 7. GAAGCTAAAGAAGCTGAAAACATC
- 8. GTGGTGATGTTTTGAGCTTCTTTAG
- 9. ACCACTGGTTGTGCTGAACACTGTTT
- 10. CAAAGAACAGTGTTCAGCACATCCA
- 11. TTTGAACGAAAACATACGGTACCG
- 15 12. GATCCGGTACCGTAATGTTTTCGTT

TABLE XVI

SCEPD SECTION 1

EcoRI HindIII 1
 AATTCA AGCTTGGATA
 20 GT TCGAACCTAT
 2

AAAGAGCTCC ACCAAGATTG ATCTGTGACT CAGAGTTTT
 TTTCTCGAGG TGTCTTAAC TAGACACTGA GCTCTCAAAA
 4

5 7
 GGAAAGATAC TGTGGAAG CTAAGAAGC TGAAAACATC ACCACTGGTT
 25 CCTTCTATG AACAACTTC GATTTCTTCG ACTTTTGTAG TGGTACCAA
 6 8

9 11 KpnI BamHI
 GTGCTGAACA GTGTTCTTC AACCAAAAACA TTACGGTACC G
 CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG CCTAG
 12

A 6942

- 78 -

TABLE XVII

SCEPO SECTION 2 OLIGONUCLEOTIDES

- 1. AATTCCGTACCAGACACCAAGGT
- 5 2. GTTAACCTTGGTGTCTGGTACCG
- 3. TAACTTCTACGCTTGGAAACGTAT
- 4. TTCCATACGTTTCCAAGCGTAGAA
- 5. GGAAGTTGGTCAACAAGCAGTTGAAGT
- 6. CCAAACCTCAACTGCTTGTGACCAAC
- 10 7. TTGGCAAGGTTGGCCTTGTATCTG
- 8. GCTTCAGATAACAAGGCCAACCCTTG
- 9. AAGCTGTTTTGAGAGGTCAAGCCT
- 10. AACAAAGGCTTGACCTCTCAAAACA
- 11. TGTGGTTAACTCTTCTCAACCATGGG
- 15 12. TGGTCCCATGGTTGAGAAGAGTTAACC
- 13. AACCATTGCAATTGCACGTCGAT
- 14. CTTTATCGACGTGCAATTGCAA
- 15. AAAGCCGTCCTGGTTTGAGATCTG
- 16. GATCCAGATCTCAAACCAGAGACGG

20

25

A 6943