Amgen Inc. v. F. Hoffmann-LaRoche LTD et al
Doc. 441 Att. 3
Case 1:05-cv-12237-WGY Document 441-4 Filed 05/21/2007 Page 1 of 20

FIG.6C

```
                                                                  27                  30
TGGTGGCCCCAAACCATACCTGAAACTAGGCAAGGAGCAAAGCCAGCAGATCCTACGCCTGTGGGCCAGGG
                                                                  Thr Gly Cys Ala Glu
CCAGAGCCTTCAGGGACCCTTGACTCCCCGGGCTGTGTGCATTTCAG    ACG GGC TGT GCT GAA
                                *       40
His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
CAC TGC AGC TTG AAT GAG AAT ATC ACT GTC CCA GAC ACC AAA GTT AAT TTC TAT

50              55
Ala Trp Lys Arg Met Glu
GCC TGG AAG AGG ATG GAG GTGAGTTCCTTTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCTCATT

TGCGAGCCTGATTTTGGATGAAAGGGAGAATGATCGGGGGAAAGGTAAAATGGAGCAGCAGAGATGAGGCT

GCCTGGGCGCAGAGGCTCACGTCTATAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCT

GGAGTTTCAGACCAACCTAGGCAGCATAGTGAGATCCCCCATCTCTACAAACATTTAAAAAAATTAGTCAG

GTGAAGTGGTGCATGGTGGTAGTCCCAGATATTTGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAA

TTTGAGGCTGCAGTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGCCCTGTCTCA
```

Dockets.Justia.com

FIG. 6D

```
AAAAAGAAAAGAAAAAAGAAAAATAATGAGGGCTGTATGGAATACATTCATTATTCATTCACTCACTCACT

CACTCATTCATTCATTCATTCATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTGG

GGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCAGCTCGACTCCCAGAGTCCACTCCCTGTAG

 56                 60                                               70
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
GTC GGG CAG CAG GCC GTA GAA GTC TGG CAG GGC CTG GCC CTG CTG TCG GAA GCT

                         80          *                           90
Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
GTC CTG CGG GGC CAG GCC CTG TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG

                                100
Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
CAG CTG CAT GTG GAT AAA GCC GTC AGT GGC CTT CGC AGC CTC ACC ACT CTG CTT

110                 115
Arg Ala Leu Gly Ala Gln
CGG GCT CTG GGA GCC CAG GTGAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGA

GAAGGGTCTTGCTAAGGAGTACAGGAACTGTCCGTATTCCTTCCCTTTCTGTGGCACTGCAGCGACCTCCT

                       116             120
                       Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
GTTTTCTCCTTGGCAG       AAG GAA GCC ATC TCC CCT CCA GAT GCG GCC TCA GCT GCT
```

FIG. 6E

```
    130                                              140
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
CCA CTC CGA ACA ATC ACT GCT GAC ACT TTC CGC AAA CTC TTC CGA GTC TAC TCC

            150                                         160
Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
AAT TTC CTC CGG GGA AAG CTG AAG CTG TAC ACA GGG GAG GCC TGC AGG ACA GGG

    166
Asp Arg OP
GAC AGA TGA CCAGGTGTGTCCACCTGGGCATATCCACCACCTCCCTCACCAACATTGCTTGTGCCACA

CCCTCCCCCGCCACTCCTGAACCCCGTCGAGGGGCTCTCAGCTCAGCGCCAGCCTGTCCCATGGACACTCC

AGTGCCAGCAATGACATCTCAGGGGCCAGAGGAACTGTCCAGAGAGCAACTCTGAGATCTAAGGATGTCAC

AGGGCCAACTTGAAGGGCCCAGAGCAGGAAGCATTCAGAGAGCAGCTTTAAACTCAGGGACAGAGCCATGC

TGGGAAGACGCCTGAGCTCACTCGGCACCCTGCAAAATTTGATGCCAGGACACGCTTTGGAGGCGATTTAC

CTGTTTTCGCACCTACCATCAGGGACAGGATGACCTGGAGAACTTAGGTGGCAAGCTGTGACTTCTCCAGG

TCTCACGGGCATGGGCACTCCCTTGGTGGCAAGAGCCCCCTTGACACCGGGGTGGTGGGAACCATGAAGAC

AXGATXGGGGCTGGCCTCTGGCTCTCATGGGGTCCAAGTTTTGTGTATTCTCAACCTATTGACAGACTGAA

ACACAATATGAC
```

FIG. 7

```
                                        -1  1
          XbaI                          MetAla
        CTAG AAACCATGAG GGTAATAAAA TAATGGCTCC GCCGCGTCTG
             TTTGGTACTC CCATTATTTT ATTACCGAGG CGGCGCAGAC

ATCTGCGACT CGAGAGTTCT GGAACGTTAC CTGCTGGAAG CTAAAGAAGC
TAGACGCTGA GCTCTCAAGA CCTTGCAATG GACGACCTTC GATTTCTTCG

TGAAAACATC ACCACTGGTT GTGCTGAACA CTGTTCTTTG AACGAAAACA
ACTTTTGTAG TGGTGACCAA CACGACTTGT GACAAGAAAC TTGCTTTTGT

TTACGGTACC AGACACCAAG GTTAACTTCT ACGCTTGGAA ACGTATGGAA
AATGCCATGG TCTGTGGTTC CAATTGAAGA TGCGAACCTT TGCATACCTT

GTTGGTCAAC AAGCAGTTGA AGTTTGGCAG GGTCTGGCAC TGCTGAGCGA
CAACCAGTTG TTCGTCAACT TCAAACCGTC CCAGACCGTG ACGACTCGCT

GGCTGTACTG CGTGGCCAGG CACTGCTGGT AAACTCCTCT CAGCCGTGGG
CCGACATGAC GCACCGGTCC GTGACGACCA TTTGAGGAGA GTCGGCACCC

AACCGCTGCA GCTGCATGTT GACAAAGCAG TATCTGGCCT GAGATCTCTG
TTGGCGACGT CGACGTACAA CTGTTTCGTC ATAGACCGGA CTCTAGAGAC

ACTACTCTGC TGCGTGCTCT GGGTGCACAG AAAGAGGCTA TCTCTCCGCC
TGATGAGACG ACGCACGAGA CCCACGTGTC TTTCTCCGAT AGAGAGGCGG

GGATGCTGCA TCTGCTGCAC CGCTGCGTAC CATCACTGCT GATACCTTCC
CCTACGACGT AGACGACGTG GCGACGCATG GTAGTGACGA CTATGGAAGG

GCAAACTGTT TCGTGTATAC TCTAACTTCC TGCGTGGTAA ACTGAAACTG
CGTTTGACAA AGCACATATG AGATTGAAGG ACGCACCATT TGACTTTGAC

                                                SalI
TATACTGGCG AAGCATGCCG TACTGGTGAC CGCTAATAG
ATATGACCGC TTCGTACGGC ATGACCACTG GCGATTATCA GCT
```

FIG. 8

```
                -1 +1
HindIII         ArgAla
AGCTTGGATA AAAGAGCTCC ACCAAGATTG ATCTGTGACT CGAGAGTTTT
    ACCTAT TTTCTCGAGG TGGTTCTAAC TAGACACTGA GCTCTCAAAA

GGAAAGATAC TTGTTGGAAG CTAAAGAAGC TGAAAACATC ACCACTGGTT
CCTTTCTATG AACAACCTTC GATTTCTTCG ACTTTTGTAG TGGTGACCAA

GTGCTGAACA CTGTTCTTTG AACGAAAACA TTACGGTACC AGACACCAAG
CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG TCTGTGGTTC

GTTAACTTCT ACGCTTGGAA ACGTATGGAA GTTGGTCAAC AAGCTGTTGA
CAATTGAAGA TGCGAACCTT TGCATACCTT CAACCAGTTG TTCGACAACT

AGTTTGGCAA GGTTTGGCCT TGTTATCTGA AGCTGTTTTG AGAGGTCAAG
TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC TCTCCAGTTC

CCTTGTTGGT TAACTCTTCT CAACCATGGG AACCATTGCA ATTGCACGTC
GGAACAACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG

GATAAAGCCG TCTCTGGTTT GAGATCTTTG ACTACTTTGT TGAGAGCTTT
CTATTTCGGC AGAGACCAAA CTCTAGAAAC TGATGAAACA ACTCTCGAAA

GGGTGCTCAA AAGGAAGCCA TTTCCCCACC AGACGCTGCT TCTGCCGCTC
CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG

CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
GTAACTCTTG GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG

TCCAACTTCT TGAGAGGTAA ATTGAAGTTG TACACCGGTG AAGCCTGTAG
AGGTTGAAGA ACTCTCCATT TAACTTCAAC ATGTGGCCAC TTCGGACATC

AACTGGTGAC AGATAAGCCC GACTGATAAC AACAGTGTAG
TTGACCACTG TCTATTCGGG CTGACTATTG TTGTCACATC

            SalI
ATGTAACAAA G
TACATTGTTT CAGCT
```

FIG. 9

```
                -20       -10       +1        10        20        30        40
Human   MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTK
        ********************* ********** ********** ********** * ** * **********
Monkey  MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLICDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTK

                50        60        70        80        90        100       110
Human   VNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKE
        ********** ********** ********** * ********** *** ********** ********** *
Monkey  VNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-E

                120       130       140       150       160
Human   AISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGDR
        *** ********** ********** ********** ********** ***
Monkey  AISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGDR
```

FIG. 10

1. AATTCTAGAAACCATGAGGGTAATAAAATA

2. CCATTATTTTATTACCCTCATGGTTTCTAG

3. ATGGCTCCGCCGCGTCTGATCTGCGAC

4. CTCGAGTCGCAGATCAGACGCGGCGGAG

5. TCGAGAGTTCTGGAACGTTACCTGCTG

6. CTTCCAGCAGGTAACGTTCCAGAACT

7. GAAGCTAAAGAAGCTGAAAACATC

8. GTGGTGATGTTTTCAGCTTCTTTAG

9. ACCACTGGTTGTGCTGAACACTGTTC

10. CAAAGAACAGTGTTCAGCACAACCA

11. TTTGAACGAAAACATTACGGTACCG

12. GATCCGGTACCGTAATGTTTTCGTT

FIG. 11

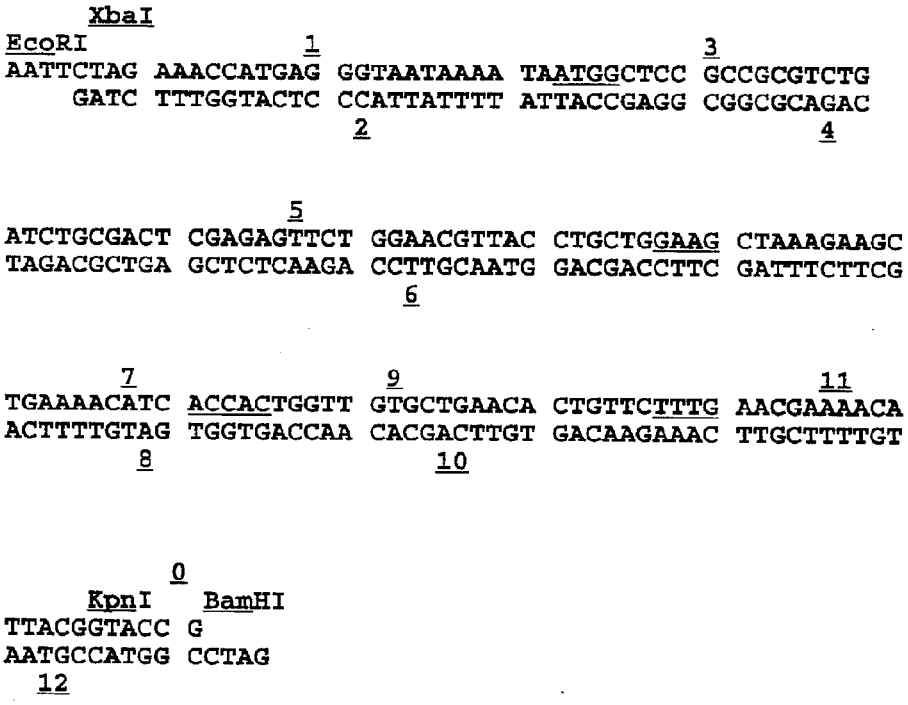

XbaI
EcoRI
1
3
AATTCTAG AAACCATGAG GGTAATAAAA TAATGGCTCC GCCGCGTCTG
GATC TTTGGTACTC CCATTATTTT ATTACCGAGG CGGCGCAGAC
2
4
5
ATCTGCGACT CGAGAGTTCT GGAACGTTAC CTGCTGGAAG CTAAAGAAGC
TAGACGCTGA GCTCTCAAGA CCTTGCAATG GACGACCTTC GATTTCTTCG
6
7
9
11
TGAAAACATC ACCACTGGTT GTGCTGAACA CTGTTCTTTG AACGAAAACA
ACTTTTGTAG TGGTGACCAA CACGACTTGT GACAAGAAAC TTGCTTTTGT
8
10
0
KpnI
BamHI
TTACGGTACC G
AATGCCATGG CCTAG
12

FIG. 12

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAACTTCTACGCTTGGAAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACTTCAACTGCTTGTTGACCAAC
7. TTGGCAGGGTCTGGCACTGCTGAGCG
8. GCCTCGCTCAGCAGTGCCAGACCCTG
9. AGGCTGTACTGCGTGGCCAGGCA
10. GCAGTGCCTGGCCACGCAGTACA
11. CTGCTGGTAAACTCCTCTCAGCCGT
12. TTCCCACGGCTGAGAGGAGTTTACCA
13. GGGAACCGCTGCAGCTGCATGTTGAC
14. GCTTTGTCAACATGCAGCTGCAGCGG
15. AAAGCAGTATCTGGCCTGAGATCTG
16. GATCCAGATCTCAGGCCAGATACT

FIG. 13

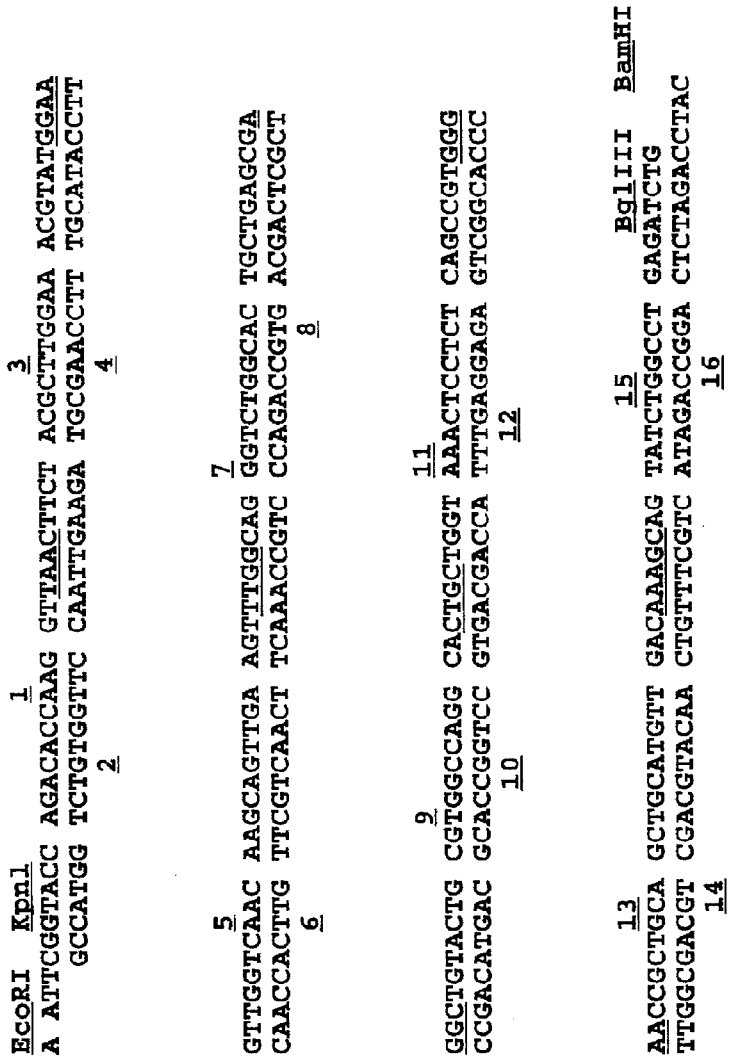

EcoRI KpnI
1
A ATTCGGTACC AGACACCAAG GTTAACTTCT ACGCTTGGAA ACGTATGGAA
GCCATGG TCTGTGGTTC CAATTGAAGA TGCGAACCTT TGCATACCTT
2
3
4
5
GTTGGTCAAC AAGCAGTTGA AGTTTGGCAG GGTCTGGCAC TGCTGAGCGA
CAACCACTTG TTCGTCAACT TCAAACCGTC CCAGACCGTG ACGACTCGCT
6
7
8
9
GGCTGTACTG CGTGGCCAGG CACTGCTGGT AAACTCCTCT CAGCCGTGGG
CCGACATGAC GCACCGGTCC GTGACGACCA TTTGAGGAGA GTCGGCACCC
10
11
12
13
AACCGCTGCA GCTGCATGTT GACAAAGCAG TATCTGGCCT GAGATCTG
TTGGCGACGT CGACGTACAA CTGTTTCGTC ATAGACCGGA CTCTAGACCTAC
14
15
16
BglIII BamHI

FIG. 14

1. GATCCAGATCTCTGACTACTCTGC
2. ACGCAGCAGAGTAGTCAGAGATCTG
3. TGCGTGCTCTGGGTGCACAGAAAGAGG
4. GATAGCCTCTTTCTGTGCACCCAGAGC
5. CTATCTCTCCGCCGGATGCTGCATCT
6. CAGCAGATGCAGCATCCGGCGGAGA
7. GCTGCACCGCTGCGTACCATCACTG
8. ATCAGCAGTGATGGTACGCAGCGGTG
9. CTGATACCTTCCGCAAACTGTTTCG
10. ATACACGAAACAGTTTGCGGAAGGT
11. TGTATACTCTAACTTCCTGCGTGGTA
12. CAGTTTACCACGCAGGAAGTTAGAGT
13. AACTGAAACTGTATACTGGCGAAGC
14. GGCATGCTTCGCCAGTATACAGTTT
15. ATGCCGTACTGGTGACCGCTAATAG
16. TCGACTATTAGCGGTCACCAGTAC

FIG. 15

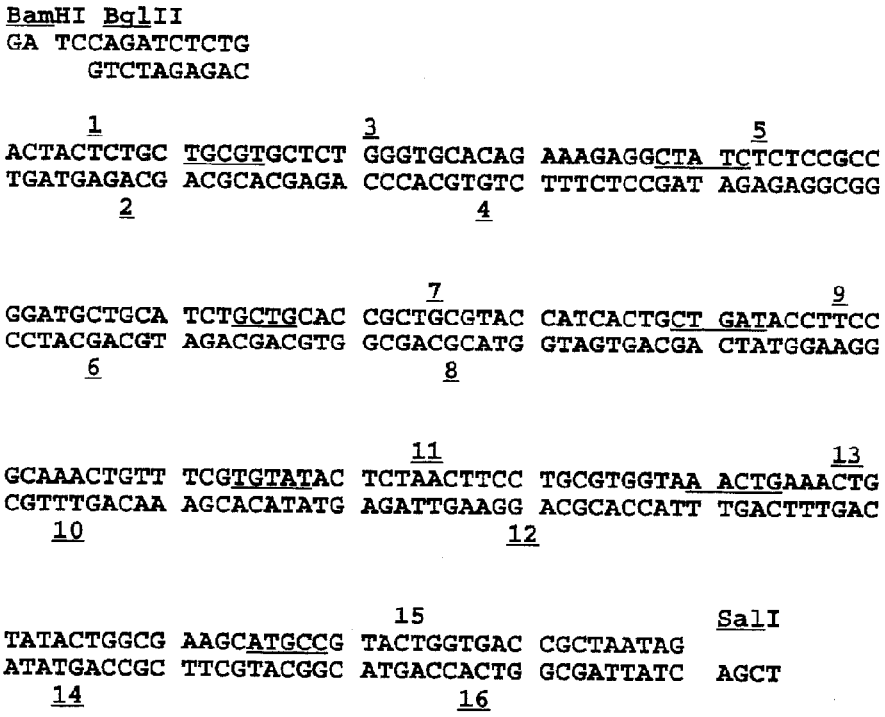

BamHI BglII
GA TCCAGATCTCTG
GTCTAGAGAC
1
3
5
ACTACTCTGC TGCGTGCTCT GGGTGCACAG AAAGAGGCTA TCTCTCCGCC
TGATGAGACG ACGCACGAGA CCCACGTGTC TTTCTCCGAT AGAGAGGCGG
2
4
7
9
GGATGCTGCA TCTGCTGCAC CGCTGCGTAC CATCACTGCT GATACCTTCC
CCTACGACGT AGACGACGTG GCGACGCATG GTAGTGACGA CTATGGAAGG
6
8
11
13
GCAAACTGTT TCGTGTATAC TCTAACTTCC TGCGTGGTAA ACTGAAACTG
CGTTTGACAA AGCACATATG AGATTGAAGG ACGCACCATT TGACTTTGAC
10
12
15
SalI
TATACTGGCG AAGCATGCCG TACTGGTGAC CGCTAATAG
ATATGACCGC TTCGTACGGC ATGACCACTG GCGATTATC AGCT
14
16

FIG. 16

1. AATTCAAGCTTGGATAAAAGAGCT
2. GTGGAGCTCTTTTATCCAAGCTTG
3. CCACCAAGATTGATCTGTGACTC
4. TCTCGAGTCACAGATCAATCTTG
5. GAGAGTTTTGGAAAGATACTTGTTG
6. CTTCCAACAAGTATCTTTCCAAAAC
7. GAAGCTAAAGAAGCTGAAAACATC
8. GTGGTGATGTTTTCAGCTTCTTTAG
9. ACCACTGGTTGTGCTGAACACTGTTC
10. CAAAGAACAGTGTTCAGCACAACCA
11. TTTGAACGAAAACATTACGGTACCG
12. GATCCGGTACCGTAATGTTTTCGTT

FIG. 17

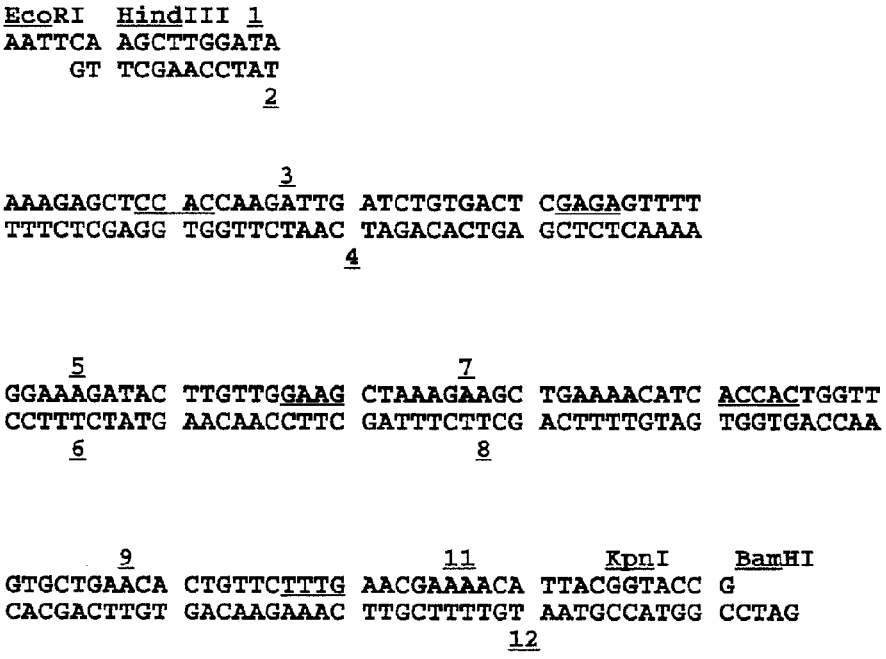

EcoRI HindIII 1
AATTCA AGCTTGGATA
GT TCGAACCTAT
2
3
AAAGAGCTCC ACCAAGATTG ATCTGTGACT CGAGAGTTTT
TTTCTCGAGG TGGTTCTAAC TAGACACTGA GCTCTCAAAA
4
5
7
GGAAAGATAC TTGTTGGAAG CTAAAGAAGC TGAAAACATC ACCACTGGTT
CCTTTCTATG AACAACCTTC GATTTCTTCG ACTTTTGTAG TGGTGACCAA
6
8
9
11
KpnI
BamHI
GTGCTGAACA CTGTTCTTTG AACGAAAACA TTACGGTACC G
CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG CCTAG
12

FIG. 18

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAACTTCTACGCTTGGAAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACTTCAACTGCTTGTTGACCAAC
7. TTGGCAAGGTTTGGCCTTGTTATCTG
8. GCTTCAGATAACAAGGCCAAACCTTG
9. AAGCTGTTTTGAGAGGTGAAGCCT
10. AACAAGGCTTGACCTCTCAAAACA
11. TGTTGGTTAACTCTTCTCAACCATGGG
12. TGGTTCCCATGGTTGAGAAGAGTTAACC
13. AACCATTGCAATTGCACGTCGAT
14. CTTTATCGACGTGCAATTGCAA
15. AAAGCCGTCTCTGGTTTGAGATCTG
16. GATCCAGATCTCAAACCAGAGACGG

FIG. 19

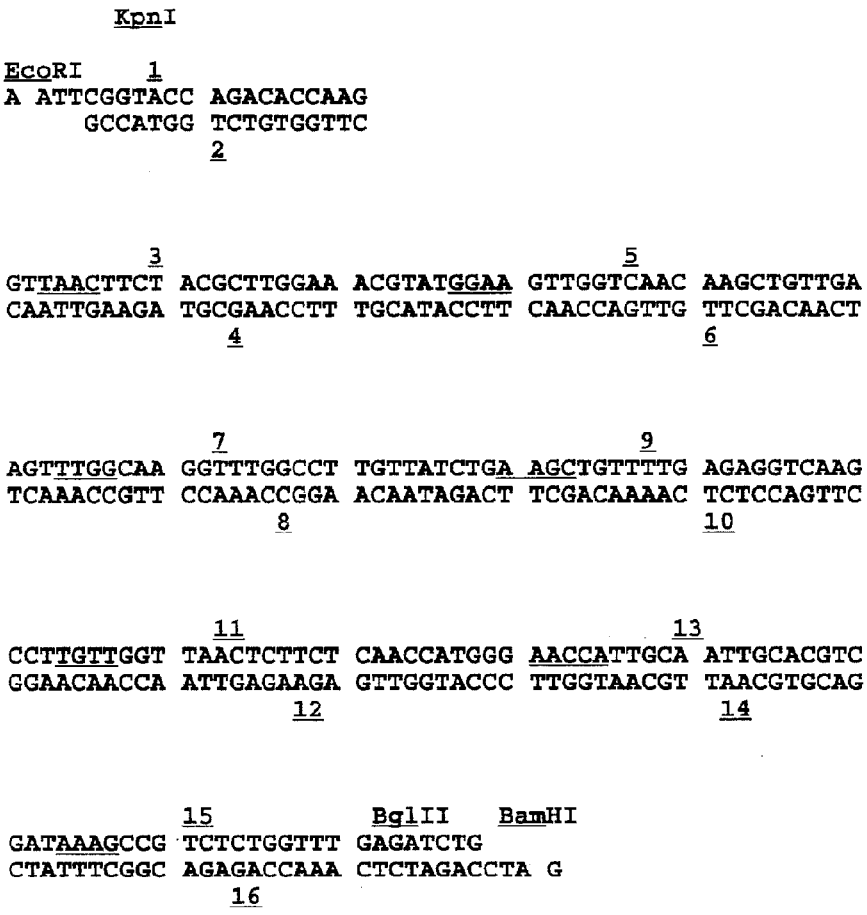
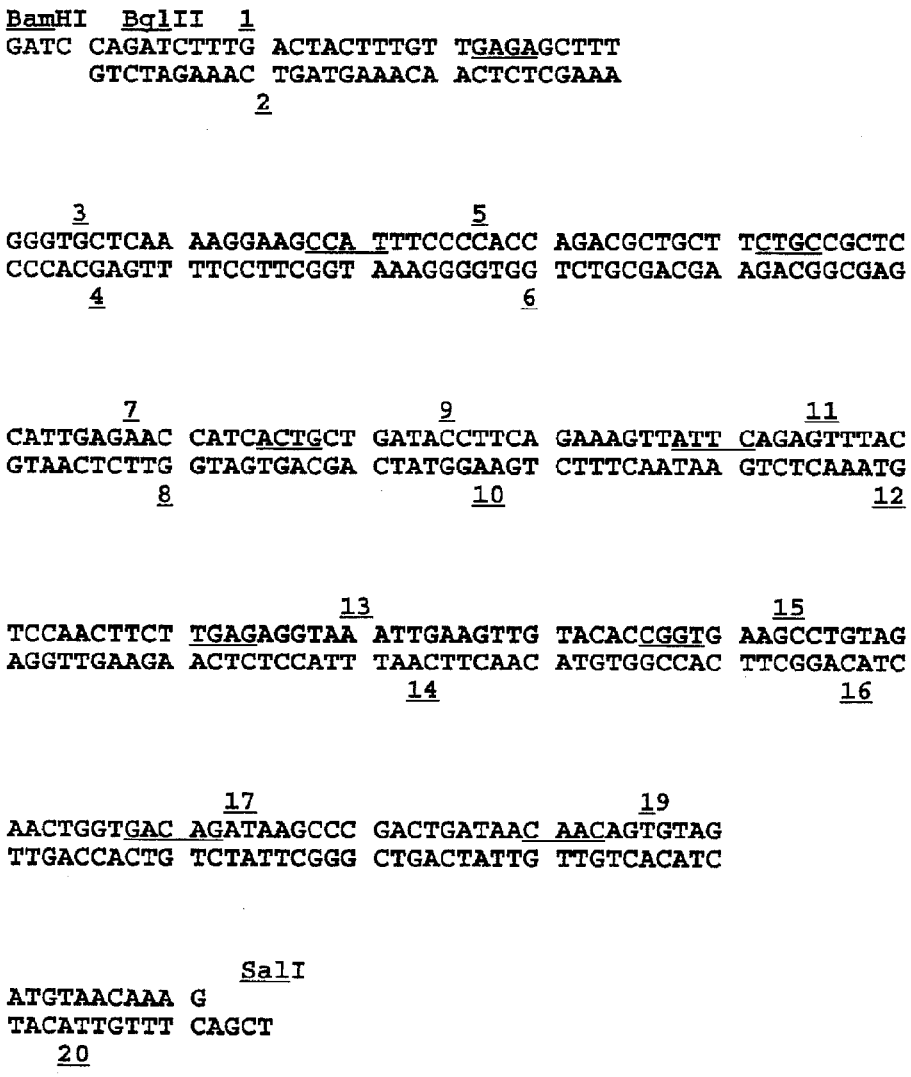

KpnI
EcoRI
1
A ATTCGGTACC AGACACCAAG
GCCATGG TCTGTGGTTC
2
3
5
GTTAACTTCT ACGCTTGGAA ACGTATGGAA GTTGGTCAAC AAGCTGTTGA
CAATTGAAGA TGCGAACCTT TGCATACCTT CAACCAGTTG TTCGACAACT
4
6
7
9
AGTTTGGCAA GGTTTGGCCT TGTTATCTGA AGCTGTTTTG AGAGGTCAAG
TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC TCTCCAGTTC
8
10
11
13
CCTTGTTGGT TAACTCTTCT CAACCATGGG AACCATTGCA ATTGCACGTC
GGAACAACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG
12
14
15
BglII
BamHI
GATAAAGCCG TCTCTGGTTT GAGATCTG
CTATTTCGGC AGAGACCAAA CTCTAGACCTA G
16

FIG. 20

1. GATCCAGATCTTTGACTACTTTGTT
2. TCTCAACAAAGTAGTCAAAGATCTG
3. GAGAGCTTTGGGTGCTCAAAAGGAAG
4. ATGGCTTCCTTTTGAGCACCCAAAGC
5. CCATTTCCCCACCAGACGCTGCTT
6. GCAGAAGCAGCGTCTGGTGGGGAA
7. CTGCCGCTCCATTGAGAACCATC
8. CAGTGATGGTTCTCAATGGAGCG
9. ACTGCTGATACCTTCAGAAAGTT
10. GAATAACTTTCTGAAGGTATCAG
11. ATTCAGAGTTTACTCCAACTTCT
12. CTCAAGAAGTTGGAGTAAACTCT
13. TGAGAGGTAAATTGAAGTTGTACAC
14. ACCGGTGTACAACTTCAATTTACCT
15. CGGTGAAGCCTGTAGAACTGGT
16. CTGTCACCAGTTCTACAGGCTTC
17. GACAGATAAGCCCGACTGATAA
18. GTTGTTATCAGTCGGGCTTAT
19. CAACAGTGTAGATGTAACAAAG
20. TCGACTTTGTTACATCTACACT

FIG. 21


BamHI BglII 1
GATC CAGATCTTTG ACTACTTTGT TGAGAGCTTT
GTCTAGAAAC TGATGAAACA ACTCTCGAAA
2
3
5
GGGTGCTCAA AAGGAAGCCA TTTCCCCACC AGACGCTGCT TCTGCCGCTC
CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG
4
6
7
9
11
CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
GTAACTCTTG GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG
8
10
12
13
15
TCCAACTTCT TGAGAGGTAA ATTGAAGTTG TACACCGGTG AAGCCTGTAG
AGGTTGAAGA ACTCTCCATT TAACTTCAAC ATGTGGCCAC TTCGGACATC
14
16
17
19
AACTGGTGAC AGATAAGCCC GACTGATAAC AACAGTGTAG
TTGACCACTG TCTATTCGGG CTGACTATTG TTGTCACATC
SalI
ATGTAACAAA G
TACATTGTTT CAGCT
20

PRODUCTION OF ERYTHROPOIETIN

This is a continuation of my U.S. patent application Ser. No. 08/202,874 filed Feb. 28, 1994, now abandoned, which was a continuation of U.S. patent application Ser. No. 07/113,178, filed Oct. 23, 1987, now abandoned, which was a continuation of U.S. patent application Ser. No. 06/675,298 filed Nov. 30, 1984 and issued as U.S. Pat. No. 4,703,008 on Oct. 27, 1987, which was a continuation-in-part of U.S. patent application Ser. No. 06/655,841, filed Sep. 28, 1984, now abandoned, which was a continuation-in-part of U.S. patent application Ser. No. 06/582,185, filed Feb. 21, 1984, now abandoned, and which was a continuation-in-part of U.S. patent application Ser. No. 06/561,024, filed Dec. 13, 1983, now abandoned.

BACKGROUND

The present invention relates generally to the manipulation of genetic materials and, more particularly, to recombinant procedures making possible the production of polypeptides possessing part or all of the primary structural conformation and/or one or more of the biological properties of naturally-occurring erythropoietin.

A. Manipulation of Genetic Materials

Genetic materials may be broadly defined as those chemical substances which program for and guide the manufacture of constituents of cells and viruses and direct the responses of cells and viruses. A long chain polymeric substance known as deoxyribonucleic acid (DNA) comprises the genetic material of all living cells and viruses except for certain viruses which are programmed by ribonucleic acids (RNA). The repeating units in DNA polymers are four different nucleotides, each of which consists of either a purine (adenine or guanine) or a pyrimidine (thymine or cytosine) bound to a deoxyribose sugar to which a phosphate group is attached. Attachment of nucleotides in linear polymeric form is by means of fusion of the 5' phosphate of one nucleotide to the 3' hydroxyl group of another. Functional DNA occurs in the form of stable double stranded associations of single strands of nucleotides (known as deoxyoligonucleotides), which associations occur by means of hydrogen bonding between purine and pyrimidine bases [i.e., "complementary" associations existing either between adenine (A) and thymine (T) or guanine (G) and cytosine (C)]. By convention, nucleotides are referred to by the names of their constituent purine or pyrimidine bases, and the complementary associations of nucleotides in double stranded DNA (i.e., A-T and G-C) are referred to as "base pairs". Ribonucleic acid is a polynucleotide comprising adenine, guanine, cytosine and uracil (U), rather than thymine, bound to ribose and a phosphate group.

Most briefly put, the programming function of DNA is generally effected through a process wherein specific DNA nucleotide sequences (genes) are "transcribed" into relatively unstable messenger RNA (mRNA) polymers. The mRNA, in turn, serves as a template for the formation of structural, regulatory and catalytic proteins from amino acids. This mRNA "translation" process involves the operations of small RNA strands (tRNA) which transport and align individual amino acids along the mRNA strand to allow for formation of polypeptides in proper amino acid sequences The mRNA "message" derived from DNA and providing the basis for the tRNA supply and orientation of any given one of the twenty amino acids for polypeptide

"expression" is in the form of triplet "codons"—sequential groupings of three nucleotide bases. In one sense, the formation of a protein is the ultimate form of "expression" of the programmed genetic message provided by the nucleotide sequence of a gene.

"Promoter" DNA sequences usually "precede" a gene in a DNA polymer and provide a site for initiation of the transcription into mRNA. "Regulator" DNA sequences, also usually "upstream" of (i.e., preceding) a gene in a given DNA polymer, bind proteins that determine the frequency (or rate) of transcriptional initiation. Collectively referred to as "promoter/regulator" or "control" DNA sequence, these sequences which precede a selected gene (or series of genes) in a functional DNA polymer cooperate to determine whether the transcription (and eventual expression) of a gene will occur. DNA sequences which "follow" a gene in a DNA polymer and provide a signal for termination of the transcription into mRNA are referred to as transcription "terminator" sequences.

A focus of microbiological processing for the last decade has been the attempt to manufacture industrially and pharmaceutically significant substances using organisms which either do not initially have genetically coded information concerning the desired product included in their DNA, or (in the case of mammalian cells in culture) do not ordinarily express a chromosomal gene at appreciable levels. Simply put, a gene that specifies the structure of a desired polypeptide product is either isolated from a "donor" organism or chemically synthesized and then stably introduced into another organism which is preferably a self-replicating unicellular organism such as bacteria, yeast or mammalian cells in culture. Once this is done, the existing machinery for gene expression in the "transformed" or "transfected" microbial host cells operates to construct the desired product, using the exogenous DNA as a template for transcription of mRNA which is then translated into a continuous sequence of amino acid residues.

The art is rich in patent and literature publications relating to "recombinant DNA" methodologies for the isolation, synthesis, purification and amplification of genetic materials for use in the transformation of selected host organisms. U.S. Pat. No. 4,237,224 to Cohen, et al., for example, relates to transformation of unicellular host organisms with "hybrid" viral or circular plasmid DNA which includes selected exogenous DNA sequences. The procedures of the Cohen, et al. patent first involve manufacture of a transformation vector by enzymatically cleaving viral or circular plasmid DNA to form linear DNA strands. Selected foreign ("exogenous" or "heterologous") DNA strands usually including sequences coding for desired product are prepared in linear form through use of similar enzymes. The linear viral or plasmid DNA is incubated with the foreign DNA in the presence of ligating enzymes capable of effecting a restoration process and "hybrid" vectors are formed which include the selected exogenous DNA segment "spliced" into the viral or circular DNA plasmid.

Transformation of compatible unicellular host organisms with the hybrid vector results in the formation of multiple copies of the exogenous DNA in the host cell population. In some instances, the desired result is simply the amplification of the foreign DNA and the "product" harvested is DNA. More frequently, the goal of transformation is the expression by the host cells of the exogenous DNA in the form of large scale synthesis of isolatable quantities of commercially significant protein or polypeptide fragments coded for by the foreign DNA. See also, e.g., U.S. Pat. Nos. 4,264,731 to Shine), 4,273,875 (to Manis), 4,293,652 (to Cohen), and

European Patent Application 093,619, published Nov. 9, 1983.

The development of specific DNA sequences for splicing into DNA vectors is accomplished by a variety of techniques, depending to a Great deal on the degree of "foreignness" of the "donor" to the projected host and the size of the polypeptide to be expressed in the host. At the risk of over-simplification, it can be stated that three alternative principal methods can be employed: (1) the "isolation" of double-stranded DNA sequence from the genomic DNA of the donor; (2) the chemical manufacture of a DNA sequence providing a code for a polypeptide of interest; and (3) the in vitro synthesis of a double-stranded DNA sequence by enzymatic "reverse transcription" of mRNA isolated from donor cells. The last-mentioned methods which involve formation of a DNA "complement" of mRNA are generally referred to as "cDNA" methods.

Manufacture of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. DNA manufacturing procedures of co-owned, co-pending U.S. patent application Ser. No. 483,451, by Alton, et al., (filed Apr. 15, 1983and corresponding to PCT US83/00605, published Nov. 24, 1983 as W083/04053), for example, provide a superior means for accomplishing such highly desirable results as: providing for the presence of alternate codons Commonly found in genes which are highly expressed in the host organism selected for expression (e.g., providing yeast or *E.coli* "preference" codons); avoiding the presence of untranslated "intron" sequences (commonly present in mammalian genomic DNA sequences and mRNA transcripts thereof) which are not readily processed by procaryotic host cells; avoiding expression of undesired "leader" polypeptide sequences commonly coded for by genomic DNA and cDNA sequences but frequently not readily cleaved from the polypeptide of interest by bacterial or yeast host cells; providing for ready insertion of the DNA in convenient expression vectors in association with desired promoter/regulator and terminator sequences; and providing for ready construction of genes coding for polypeptide fragments and analogs of the desired polypeptides.

When the entire sequence of amino acid residues of the desired polypeptide is not known, direct manufacture of DNA sequences is not possible and isolation of DNA sequences coding for the polypeptide by a cDNA method becomes the method of choice despite the potential drawbacks in ease of assembly of expression vectors capable of providing high levels of microbial expression referred to above. Among the standard procedures for isolating cDNA sequences of interest is the preparation of plasmid-borne cDNA "libraries" derived from reverse transcription of mRNA abundant in donor cells selected as responsible for high level expression of genes (e.g., libraries of cDNA derived from pituitary cells which express relatively large quantities of growth hormone products). Where substantial portions of the polypeptide's amino acid sequence are known, labelled, single-stranded DNA probe sequences duplicating a sequence putatively present in the "target" cDNA may be employed in DNA/DNA hybridization procedures carried out on cloned copies of the cDNA which have been denatured to single stranded form. [See, generally, the disclosure and discussions of the art provided in U.S. Pat. No. 4,394,443 to Weissman, et al. and the recent demonstrations of the use of long oligonucleotide hybridization probes reported in Wallace, et al., *Nuc.Acids Res.*, 6, pp. 3543–3557 (1979), and Reyes, et al., *P.N.A.S.* (*U.S.A.*), 79, pp. 3270–3274 (1982), and Jaye, et al., *Nuc.Acids Res.*, 11, pp. 2325–2335 (1983). See also, U.S. Pat. No. 4,358,535 to Falkow, et al., relating to DNA/DNA hybridization procedures in effecting diagnosis; published European Patent Application Nos. 0070685 and 0070687 relating to light-emitting labels on single stranded polynucleotide probes; Davis, et al,, "A Manual for Genetic Engineering, Advanced Bacterial Genetics", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1980) at pp. 55–58 and 174–176, relating to colony and plaque hybridization techniques; and, New England Nuclear (Boston, Mass.) brochures for "Gene Screen" Hybridization Transfer Membrane materials providing instruction manuals for the transfer and hybridization of DNA and RNA, Catalog No. NEF-972.]

Among the more signficant recent advances in hybridization procedures for the screening of recombinant clones is the use of labelled mixed synthetic oligonucleotide probes, each of which is potentially the complete complement of a specific DNA sequence in the hybridization sample including a heterogenous mixture of single stranded DNAs or RNAs. These procedures are acknowledged to be especially useful in the detection of cDNA clones derived from sources which provide extremely low amounts of mRNA sequences for the polypeptide of interest. Briefly put, use of stringent hybridization conditions directed toward avoidance of non-specific binding can allow, e.g., for the autoradiographic visualization of a specific cDNA clone upon the event of hybridization of the target DNA to that single probe within the mixture which is its complete complement. See generally, Wallace, et al, *Nuc Acids Res.*, 9, pp. 879–897 (1981); Suggs, et al. *P.N.A.S.* (*U.S.A.*), 78, pp. 6613–6617 (1981); Choo, et al., *Nature*, 299, pp. 178–180 (1982); Kurachi, et al., *P.N.A.S.* (*U.S.A.*), 79, pp. 6461–6464 (1982); Ohkubo, et al., *P.N.A.S.* (*U.S.A.*), 80, pp. 2196–2200 (1983); and Kornblihtt, et al. *P.N.A.S.* (*U.S.A.*), 80, pp. 3218–3222 (1983). In general, the mixed probe procedures of Wallace, et al. (1981), supra, have been expanded upon by various workers to the point where reliable results have reportedly been obtained in a cDNA clone isolation using a 32-member mixed "pool" of 16-base-long (16-mer) oligonucleotide probes of uniformly, varying DNA sequences together with a single 11-mer to effect a two-site "positive" confirmation of the presence of cDNA of interest. See, Singer-Sam, et al., *P.N.A.S.* (*U.S.A.*), 80, pp. 802–806 (1983).

The use of genomic DNA isolates is the least common of the three above-noted methods for developing specific DNA sequences for use in recombinant procedures. This is especially true in the area of recombinant procedures directed to securing microbial expression of mammalian polypeptides and is due, principally to the complexity of mammalian genomic DNA. Thus, while reliable procedures exist for developing phage-borne libraries of genomic DNA of human and other mammalian species origins [See, e.g., Lawn, et al. *Cell*, 15, pp. 1157–1174 (1978) relating to procedures for generating a human genomic library commonly referred to as the "Maniatis Library"; Karn, et al., *P.N.A.S.* (*U.S.A.*), 77, pp. 5172–5176 (1980) relating to a human genomic library based on alternative restriction endonuclease fragmentation procedure; and Blattner, et al., *Science*, 196, pp. 161–169 (1977) describing construction of a bovine genomic library] there have been relatively few successful attempts at use of hybridization procedures in isolating genomic DNA in the absence of extensive foreknowledge of amino acid or DNA sequences. As one example, Fiddes, et al., *J.Mol. and App.Genetics*, 1, pp. 3–18 (1981) report the successful isolation of a gene coding for the alpha subunit of human pituitary glycoprotein hormones from the Maniatis Library through use of a "full length"