(54) Title: EXPRESSION OF HUMAN CONNECTIVE TISSUE ACTIVATOR AND RELATED PEPTIDES

(57) Abstract

Methods and compositions for the efficient production of a polypeptide product which has biological activity corresponding to that of natural human connective tissue activator peptide-III (CTAP-III) and degradation products thereof, such as β-thromboglobulin. A synthetic gene for the CTAP-III polypeptide is joined to a secretory leader and processing signal sequence derived from yeast α-factor. The resulting DNA construct provides for expression and secretion of the gene product from yeast. Enhanced yields of the product may then be recovered from the nutrient medium.
FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AT | Austria       | AU | Australia   | BB | Barbados  | BE | Belgium    | BG | Bulgaria   | BR | Brazil    | CF | Central African Republic | CG | Congo      | CH | Switzerland | CM | Cameroon   | DE | Germany, Federal Republic of | DK | Denmark    | FI | Finland    | FR | France     |
|----|---------------|----|-------------|----|-----------|----|------------|----|------------|----|-----------|----|-------------------------|----|------------|----|-------------|----|------------|----|-------------|----|-------------|    |            |    |            |
| MR | Mauritania,   | MW | Malawi      | NL | Netherlands | NO | Norway     | RO | Romania    | SD | Sudan      | SE | Sweden     | SN | Senegal    | SU | Soviet Union | TD | Chad       | TG | Togo       | US | United States of America |
EXPRESSION OF HUMAN CONNECTIVE
TISSUE ACTIVATOR AND RELATED PEPTIDES

Background of the Invention

1. Field of the Invention

The phrase "connective tissue activation" embraces a number of metabolic events which occur during the regeneration phase of cellular inflammation, specifically including the stimulation of cellular energy metabolism such as glycolysis, the stimulation of glycosaminoglycan synthesis, and the stimulation of DNA synthesis. Proteins capable of imparting connective tissue activation are referred to as connective tissue activator peptides (CTAP's).

Particular CTAP's are designated based on the cell or tissue of origin. CTAP-I is of lymphoid origin, CTAP-II is of tumor cell origin, CTAP-PMN is of polymorphonuclear leukocyte origin, while CTAP-III and CTAP-P₂ are of human platelet origin. CTAP-III exists in multiple forms having a number of biological activities, including the stimulation of synthesis of DNA, hyaluronic acid, and sulphated glycosaminoglycan; stimulation of glycolysis and prostaglandin E₂ secretion, and intracellular cAMP accumulation. β-thromboglobulin (β-TG), a derivative of CTAP-III generated in vivo by proteolytic removal of the four amino-terminal amino acids, has no presently established physiological activity, but is useful as a serum marker which is indicative of platelet status, particularly for patients suffering from diseases and conditions which result in platelet destruction or activation, such as acute myocardial infarction, prosthetic heart valve implantation, and the like.
Based on its demonstrated activity, CTAP-III promises to have substantial therapeutic value. For example, patients suffering from rheumatoid arthritis are found to have a deficiency of CTAP-III activity.

Administration of exogenous CTAP-III may be a useful treatment. CTAP-III is also believed to be useful as a supplementary growth factor for mammalian cell culture, in particular synovium, cartilage, skin, thyroid, and arterial endothelium. Finally, CTAP-III is believed to have wide application when administered to patients for promoting connective tissue healing. β-TG can find use as a diagnostic reagent in the monitoring of plasma β-TG, particularly as an indicator of platelet status.

At present, only small quantities of CTAP-III and its derivatives, including β-TG, may be obtained by separation from human blood platelets. It would therefore be of substantial value to provide greater supplies of CTAP-III peptides for therapeutic use in vivo and as a reagent for in vitro use, particularly in immunological assays for the detection of CTAP-III and β-TG. By employing recombinant DNA techniques, there is an opportunity to produce CTAP-III and related peptides efficiently, where the products will have the biological activity of a naturally-occurring peptide, in particular, the mitogenic activity and/or the immunological reactivity.

2. Description of the Prior Art

of CTAP-III and β-TG, respectively. Low-affinity platelet factor 4, a growth factor derived from thrombin-treated platelets, is thought to be identical to CTAP-III. See, e.g., Castor et al. (1983) ibid., and Brown et al. (1980) Clinica Chimica Acta 101:225-233. Castor et al. (1984) Abstracts: Annual Meeting, American Rheumatism Association (48th) and Arthritis Health Professions Association (19th): S33, indicate that there are multiple forms of biologically-active CTAP-III. Other references of interest include Doolittle et al. (1983) Science 221:275-277 which suggests that CTAP-III may be derived from the same or a closely related protein as the transforming protein of a primate sarcoma; and Castor et al. (1981) In Vitro 17:777-785 which discusses the use of CTAP-III as a growth factor.

Summary of the Invention

Novel methods and DNA constructs are utilized for the production of polypeptides having biological activities, in particular mitogenic activity and/or immunological reactivity, which are analogous to that of human connective tissue activator protein-III (CTAP-III) and related peptides, particularly proteolytic fragments such as β-thromboglobulin (β-TG).

The DNA constructs include a sequence coding for an amino acid sequence which is substantially the same as that of natural human CTAP-III or biologically-active fragments thereof. The sequence may be derived from cDNA, synthetic DNA, or a combination of the two. By providing a secretory leader and processing signal sequence in proper reading frame with the coding sequence, and introducing the resulting construct into a suitable unicellular host, the mature polypeptide is secreted into the growth medium, allowing efficient recovery. After purification, the polypeptide will
usually contain only trace amounts of natural gene products from the host which do not interfere with its biological activity or have undesirable side effects when administered to a mammalian, e.g., human host. In the specific embodiment, coding sequences for CTAP-III and β-TG are chemically synthesized in vitro and utilize codons preferentially recognized by a yeast host. The constructs include a replication system for yeast, and the resulting yeast transformants provide for efficient and economic production of a product having biological activities characteristic of human CTAP-III and β-TG.

**Brief Description of the Drawings**

Fig. 1 illustrates the construction of an extrachromosomal element capable of expressing CTAP-III according to the present invention, as described in the Experimental section hereinafter.

Fig. 2 illustrates the construction of an extrachromosomal element capable of expressing β-TG according to the present invention as described in the Experimental section hereinafter.

**Description of the Specific Embodiments**

DNA constructs are provided which are capable of expressing polypeptides having biological activities, e.g., mitogenic activity and immunological activity, corresponding to those of human connective tissue activator peptide-III (CTAP-III) and related peptides, such as β-thromboglobulin (β-TG). These polypeptides are derived from human platelets. The DNA constructs are usually joined to replication systems which allow for autonomous replication, and the resulting extrachromosomal elements are used to transform susceptible unicellular hosts. By including
a secretory leader and processing signal sequence in proper reading frame with the sequence coding for CTAP-III or related peptides, transformation of the host results in the expression of a precursor polypeptide which may be processed by the host into the mature polypeptide which is then secreted into the growth medium.

As used hereinafter and in the claims, CTAP-III and related polypeptides will refer to the polypeptide products which are produced according to the present invention and which display biological activities, e.g., mitogenic activity and/or immunological activity, similar to the various forms of natural human connective tissue activator peptide-III and related peptides, such as β-thromboglobulin, as measured in recognized bioassays. The polypeptide products will have amino acid sequences, which are the same or substantially the same as the natural proteins, usually differing by no more than 5 amino acids, more usually differing by 3 or fewer amino acids. For the most part, the CTAP-III or β-TG sequence will differ, if at all, by substitutions among the non-polar amino acids, i.e., aliphatic and aromatic, or by truncation of up to 5 amino acids at the carboxy-terminus. The deviations from the natural amino acid sequence will not adversely affect the biological activity of interest.

The extrachromosomal elements of the present invention will include a replication system recognized by the desired host, typically yeast, a structural gene coding for the polypeptide product, and a secretory leader and processing signal sequence in reading frame with the structural gene and 5' to the structural gene in the direction of transcription. The precursor polypeptide which is expressed results in efficient processing and secretion of the precursor polypeptide
by the host. Such constructs provide for a high yield of a mature polypeptide product.

The structural gene utilized in the present invention may be derived from cDNA, synthetic DNA, or a combination of both, e.g., synthetic DNA may be combined with cDNA to complete the structural gene. The term structural gene intends the DNA sequence which encodes for the mature polypeptide product, i.e., the polypeptide which is secreted by the host after processing. The term structural gene does not refer to regulatory or other sequences associated with transcription, expression, or processing.

The subject invention will be illustrated with extrachromosomal elements employing synthetic structural genes having nucleotide sequences based on the amino acid sequences of naturally-occurring human connective tissue activator peptide-III and β-TG. The amino acid sequences are reported by Castor et al. (1983), supra., which sequences are reproduced in the Experimental section hereinafter. The amino acid sequence will usually be identical to that of the natural polypeptide, although some deviation may be acceptable so long as the polypeptide product retains the desired biological activity, i.e., mitogenic activity and/or immunogenic activity.

In devising the nucleotide sequence, it is preferred to use codons which are preferentially utilized by the intended yeast host. Specifically, such codons are those which appear at high frequency in the structural genes encoding for the yeast glycolytic enzymes. The nucleotide sequence will comprise at least 50% preferred yeast codons, usually at least 60%, more usually at least 75%. A specific nucleotide sequence which employs the preferred yeast codons is also set forth in the Experimental section.

The synthetic structural gene is conveniently prepared by in vitro synthesis of a series of single
stranded DNA fragments including from about 10 to 125 bases, usually from about 25 to 100 bases. The single stranded fragments are selected so that when brought together under annealing conditions, a double stranded fragment having the desired nucleotide sequence is produced.

When devising the synthetic nucleotide sequence, the sequences of the individual single stranded DNA segments are examined to assure that annealing between unmatched segments is avoided. In this way, base pairing occurs only between those segments which are intended to be annealed in the resulting double stranded DNA fragment. Undesired base pairing is avoided by altering one nucleotide (usually the third) in certain of the codons, while maintaining the desired yeast bias to the extent possible.

The double stranded synthetic DNA fragment encoding CTAP-III or β-TG will usually have cohesive ends which facilitate insertion of the fragment into cloning and expression vectors. The fragment will typically include stop codon(s) at the end of the structural gene, and may include a portion of the secretory leader and processing signal sequence at the 5'-end of the fragment. When devising the sequence of the structural gene, unique restriction sites may be incorporated to allow for subsequent modification of the gene, e.g., hybrid constructions, repairs, mutations, truncations, and the like.

Secretory leader and processing signal sequences will normally be derived from naturally-occurring DNA sequences in yeast which provide for secretion of a polypeptide. Such polypeptides which are naturally secreted by yeast include α-factor, a-factor, acid phosphatase, and the like. If desired, the naturally-occurring sequence may be modified, for example, by reducing the number of or eliminating the glu-ala pairs which are located at the
processing site of the α-factor precursor, or by reducing the length of the mRNA leader (while retaining sufficient length to provide for secretion), or by introducing point mutations, deletions or other modifications which facilitate manipulation, e.g., restriction recognition sites. Conveniently, the secretory leader and processing signal sequence may be joined to the structural gene by providing appropriate cohesive ends on the synthetic structural gene, by means of adaptor molecules, or a combination of both.

In addition to the structural gene and secretory leader and processing signal sequence, the extrachromosomal elements of the present invention will include transcriptional initiation regulatory and terminator sequences, optionally, replication systems capable of autonomous replication in the desired yeast host, and replication systems capable of autonomous replication in bacterial hosts. Conveniently, the promoter associated with the secretory leader and processing signal sequence may be used, although a wide variety of other promoters would also be suitable. Promoters of particular interest include those involved with enzymes in the yeast glycolytic pathway, such as the promoters for alcohol dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, pyruvate kinase, triose phosphate isomerase, phosphoglucoisomerase, phosphofructokinase, and the like. By employing these promoters with regulatory sequences, such as enhancers, operators, and the like, and using a host having an intact regulatory system, one can regulate the expression of the CTAP-III or β-TG product by varying the concentration of certain small organic molecules, such as glucose. Alternatively, temperature-sensitive regulatory mutants may be employed which allow for modulation of transcription by varying the temperature.
The terminator utilized in the DNA constructs of the present invention should be balanced with the promoter to provide for proper transcription. Conveniently, the terminator which is naturally utilized by the promoter may be employed. The remaining sequences in the construct, including the replication systems for both yeast and bacteria, are well known and amply described in the literature.

Other capabilities may also be introduced into the DNA construct. The DNA construct may include segment(s) homologous to the host genome to provide for integration. Integration of the structural gene and associated regulatory sequences (excluding the remaining portions of the extrachromosomal elements) into the host chromosomes can be achieved by placing a homologous sequence on either side of the structural gene and regulatory sequences. This allows for integration by a double cross-over event which may limit the insertion to the precursor CTAP-III or β-TG cistron. In addition, certain integrated genes together with their flanking regions, are amplified upon stress to the host. By placing such a gene upstream from the structural gene and associated regulatory regions, and stressing the host, repeating sequence may be obtained where each of the tandem repeats includes the structural gene and associated regulatory sequences and the amplifying gene with associated regulatory sequences. Illustrative genes which provide for such amplification include those for metallothioneins and dihydrofolate reductase.

Conveniently, the DNA construct may be prepared using preexisting extrachromosomal elements, e.g., plasmid and phage vectors, which carry DNA sequences of interest. Usually, a bacterial cloning vector will be used for joining the CTAP-III or related gene coding region to the secretory leader and processing signal sequences. The CTAP-III or related
gene and secretory sequence may be prepared separately, as described above, and introduced into any convenient bacterial vector, such as pBR322, pBR325, and the like, in a manner which assures their proper relative positioning. In the Experimental section hereinafter, advantage was taken of a bacterial cloning vector (pωEGF-24) which already carried a modified ω-factor secretory leader and processing signal sequence. By properly coding the 5'-end of the synthetic CTAP-III or β-TG gene, it was possible to cleave within the modified ω-factor sequence, insert the synthetic gene, and substantially regenerate the ω-factor sequence.

After inserting the desired sequences, the bacterial vector may be cloned to expand the fragment carrying the secretory leader and processing signal sequence (usually including the necessary transcriptional control signals) and the CTAP-III or related gene. The fragment may be excised and inserted into an appropriate expression vector, e.g., a yeast expression vector. The yeast expression vector may include a yeast replication system and, usually, at least one selective marker. In the Experimental section, the utilized yeast expression vector (pCl/1) carries a replication system derived from the 2μm plasmid and a LEU2 gene which allows for selection in leu" yeast hosts.

The plasmids may be introduced into the yeast host by any convenient means, employing yeast host cells or spheroplasts and using calcium-precipitated DNA for transformation, or liposomes, or other conventional techniques. The modified yeast hosts may be selected in accordance with genetic markers which are usually provided in the vector used to construct the expression plasmid. An auxotrophic host may be employed, where the plasmid includes a gene which complements the host and provides for prototrophy. Alternatively, resistance to an appropriate biocide, e.g.,
antibiotic, heavy metal, toxin, or the like, may be included as a marker in the plasmid. Selection may then be achieved by employing a nutrient medium which stresses the host cells, so as to select for the cells containing the plasmid. The plasmid-containing cells may then be grown in an appropriate nutrient medium, and the desired secreted polypeptide isolated in accordance with conventional techniques. The polypeptide may be purified by chromatography, filtration, extraction, and the like. Since the polypeptide will be present in mature form in the nutrient medium, one can cycle nutrient medium continuously while removing the desired polypeptide. After purification, the polypeptide product will be substantially free from contaminants from the host, but will usually still contain at least trace amounts of natural gene products from the unicellular host, e.g., yeast.

The following experiments are offered by way of illustration and not by way of limitation.

Experimental

1. Connective Tissue Activator Peptide-III (CTAP-III)

A nucleotide sequence for CTAP-III based on the amino acid sequence reported by Castor et al. (1983) supra. and employing preferred yeast codons was devised. The sequence, with the coding strand shown 5' to 3', was as follows:
Processing Site
LeuLeuAspLysArgAsnLeuAlaLysGlyLysGluGluSerLeuAspSerAsp
5'-TATTGGATAAAAGAACAACCTGGCCCAAGGTTAAGGAAAGAATCCTTGAGACTGAGC
3'-CATGATAACCTTTTCCTTTGCAAGGTCCTCTCTATAGGAAACCTGAGACTG

5
FokI
LeuTyrAlaGluLeuArgCysMetCysIleLeuThrThrThrSerGlyIleHisProLysAsn
TTGTACCGCTGATTGAGATGTATGTTATGATCACAGACCTGCTGATCACCACCCAAAAGAC
AACATGCCGACTTAACTCCTACATAACATAGTTCTCTGTGGAGACCATAAGTTGAGTTCTTG

10
KpnI
IleGlnSerLeuGluValIleGlyLysGlyThrHisCysAsnGlnValGluValIleAla
ATCCCAATCTTTGGAAGTGATTTGATGGGTAAGGTCTACCTACGATCGCT
TAGGTAGAAGACCTCGAATCGACATCCCATTGGTGAGTGGTTGCTCAAACTCTCAATAGGA

15
BglII
ThrLeuLysAspGlyArgLysIleCysLeuAspProAspAlaProArgIleLysLysIle
ACCTGGAAAGGACGATGAGAAAGATCTCTTTGGACCCAGATGCTCGAACCAGAATCAGAGATC
TGGAAACCTTCTCGCCATCTCTTCAGAACAAACCTGGCTACGAGTCTTCTATTAGGTCTCTTAG

20
ValGlnLysLysLeuAlaGlyAspGluSerAlaAspOC AM
GTTCAAAAAAGAAGTTGGCGTGGTGACCAGATCGTCTGACTAATAGGCTCG-3'
CAAGTTTTCTTTCAACCAGACCCACTGCTCATGAGACGACTGATTACGACGAGCACGACTGGTCT-5'

The sequence includes a KpnI cohesive end at
its 5'-end (with reference to the coding strand) and a
SalI cohesive end at its 3'-end. Coding for the mature
polypeptide begins after the LysArg processing site.
Unique KpnI and BglII restriction sites are provided
internal to the coding region for the mature
polypeptide to allow for subsequent repairs,
modifications, mutations and truncations. When making
the final coding selection, undesirable annealing
combinations, i.e., those segments which could anneal
at more than one location, were altered by substituting
the nucleotide at the third position of appropriate
codons in a manner consistent with the yeast codon bias
when possible.

The 5'-end of the sequence up to the process-
ing site is a modification of the 3'-end of the
naturally-occurring g-factor secretory leader and
processing signal sequence, where three glu-alanine pairs
have been deleted and a leu (rather than ser) is
regenerated when the sequence is inserted into the
cloning vector which carries the remaining portion of the g-factor gene (as described below).

A synthetic DNA fragment for CTAP-III having the sequence just described was prepared by synthesizing 20 overlapping ssDNA segments using a modification of the phosphoramidite method described by Beaucage and Carruthers (1981) Tetrahedron Lett. 22:1859-1862 as reported in Urdea et al. (1983) Proc. Natl. Acad. Sci. USA 80:7461-7465. The sequences of the ssDNA segments were as follows:

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-1</td>
<td>TATTGGAATAAAAAGA</td>
</tr>
<tr>
<td>C-2</td>
<td>AACTTGCCAAAGGTAAGGAAGAATCTTT</td>
</tr>
<tr>
<td>C-3</td>
<td>GGACTCTGACTTGTACGCTGAATTTGAGAT</td>
</tr>
<tr>
<td>C-4</td>
<td>GTATGTGTAACAGAAAGACCTCCTGGAT</td>
</tr>
<tr>
<td>C-5</td>
<td>CCACCCAAAGAACATCAATCGTTTGGAAG</td>
</tr>
<tr>
<td>C-6</td>
<td>TCATTGGTAAGGTACCCACTGTAACCAAG</td>
</tr>
<tr>
<td>C-7</td>
<td>TTGAAGTTATCCTACCTTGGAAGACGGT</td>
</tr>
<tr>
<td>C-8</td>
<td>AGAAAGATCTGTTTGGACCCAGATGCTCC</td>
</tr>
<tr>
<td>C-9</td>
<td>AAGAATCAAAGAGATCGTTTTCAAAAAAGATTTG</td>
</tr>
<tr>
<td>C-10</td>
<td>GCTGGTGACGAATCTGCTGACTAATAGCGTG</td>
</tr>
<tr>
<td>C-11</td>
<td>CTTGGCCAAAGTTTTTCTTATCAAATAGTAC</td>
</tr>
<tr>
<td>C-12</td>
<td>ACAAGTCAGAGTCAGTTACCTTCTCCCTTAC</td>
</tr>
<tr>
<td>C-13</td>
<td>CTGATACACACATCATCAGTTACACAGCTC</td>
</tr>
<tr>
<td>C-14</td>
<td>ATGTCTTCTGGTTGAGATACACAGGTTGTT</td>
</tr>
<tr>
<td>C-15</td>
<td>ACCCTTACAAATGACTTCTAAAGAAGATTG</td>
</tr>
<tr>
<td>C-16</td>
<td>AGCGATAACCTTCAACTTGTACAGTGTTGG</td>
</tr>
<tr>
<td>C-17</td>
<td>AAACAGATCTTTCTCTCGCTTTGGAAGGT</td>
</tr>
<tr>
<td>C-18</td>
<td>ATCTCCTGATCTTGGAGCATCTGGGCTCC</td>
</tr>
<tr>
<td>C-19</td>
<td>CAGATTCGGTCAACGCGAATCTTTGGAACG</td>
</tr>
<tr>
<td>C-20</td>
<td>TCGACGACGCTATTAGTCAG</td>
</tr>
</tbody>
</table>

The ssDNA fragments were joined as follows: 500 pmole of each segment except C-1 and C-20 were individually 5'-phosphorylated with 5.6 units of T4
polynucleotide kinase (New England Nuclear) in 10mM dithiothreitol (DTT), 1mM ATP, 10mM MgCl₂, 100ng/ml spermidine, 50mM Tris-HCl, pH 7.8 (total volume: 20μl) for 1 hour at 37°C. Additional T4 kinase (5.6 units) was then added and the reaction continued for 2.5 hours at 37°C. Each reaction mixture was then diluted to 50μl with water, and heated to 92°C for 4 minutes. To the pooled phosphorylated segments (5μl, 50pmoles of each segment) in their kinase reaction mixtures were added the C-1 and C-20 segments (50 pmoles) and poly ra (10μg) to a final volume of 115μl. The segments were coprecipitated by adjusting the mixture to 0.6M NaOAc with 2M NaOAc followed by addition of 3 volumes of ethanol (100%) and chilling at -80°C for 7 hours. After centrifugation, the pellet was washed once with aqueous ethanol (95%) and dried under vacuum. The pellet was redisolved in water (18μl) and heated to 90°C for 3 minutes and then cooled slowly over 1.5 hours to 25°C in a water bath.

The annealed fragment pool was ligated in a reaction mixture containing T4 DNA ligase (New England Biolabs 1200 units) 1mM ATP, 10mM DTT, 10mM MgCl₂, 100 ng/ml spermidine, and 50mM Tris-HCl, pH 7.8 (30μl). After incubation for 14 hours at 14°C, the full length, double-stranded fragment was partially purified by preparative polyacrylamide gel (7%, native) electrophoresis. The DNA was removed from the appropriate gel slice by electroelution and ethanol coprecipitated with poly ra (5μg). The resulting double-stranded fragment was assembled from the synthetic, single-stranded segments as follows:

C-1  C-2  C-3  C-4  C-5  C-6  C-7  C-8  C-9  C-10
C-11  C-12  C-13  C-14  C-15  C-16  C-17  C-18  C-19  C-20

After assembly, the synthetic CTAPIII sequence was inserted into bacterial cloning plasmid pEGF-24 which had been previously digested with restriction endonucleases KpnI and SalI to remove a
KpnI/SalI fragment. Plasmid pωEGF-24 carries the 5'-end of the α-factor secretory leader and processing signal sequence so that when the synthetic CTAP-III sequence is inserted, a functional leader and processing sequence is regenerated in proper reading frame with the CTAP-III gene. The plasmid also carries the α-factor promoter and terminator upstream and downstream, respectively, from the insertion site.

Plasmid pωEGF-24 was prepared as described in copending application Serial No. 522,909 filed August 12, 1983. The preparation was as follows. A synthetic sequence for human epidermal growth factor (EGF) based on the amino acid sequence of EGF reported by H. Gregory and B. M. Preston (1977) Int. J. Peptide Protein Res. 9:107-118 was prepared. The sequence was inserted into the EcoRI site of pBR328 to produce a plasmid p328EGF-1 and cloned. Approximately 30μg of p328EGF-1 was digested with EcoRI and approximately 1μg of the expected 190 base pair EcoRI fragment was isolated. This was followed by digestion with the restriction enzyme HgaI. Two synthetic oligonucleotide connectors HindIII-HgaI and HgaI-SalI were then ligated to the 159 base pair HgaI fragment. The HgaI-HindIII linker had the following sequence:

```
AGCTGAAGCT
CTTCGATTGAG
```

This linker restores the α-factor processing signals interrupted by the HindIII digestion and joins the HgaI end at the 5'-end of the EGF gene to the HindIII end of pAB112. The HgaI-SalI linker had the following sequence:

```
TGAGATGATAAG
ACTATTAGCT
```

This linker has two stop codons and joins the HgaI end at the 3'-end of the EGF gene to the SalI end of pAB112. The resulting 181 base pair fragment was purified by preparative gel electrophoresis and ligated
to 100 ng of pAB112 which had been previously completely digested with the enzymes HindIII and SalI. The resulting mixture was used to transform E. coli HB101 cells and plasmid pAB201 obtained.

pAB112 is a plasmid containing a 1.75 kb EcoRI fragment with the yeast $\alpha$-factor gene cloned in the EcoI site of pAB11 (pBR322 from which the HindIII and SalI sites had been deleted). pAB112 was derived from plasmid pAB101 which contains the yeast $\alpha$-factor gene as a partial Sau3A fragment cloned in the BamHI site of plasmid YEp24. pAB101 was obtained by screening a yeast genomic library in YEp24 using a synthetic 20-mer oligonucleotide probe (3'-GGCCGGTTGTTACATGATT-5') homologous to the published $\alpha$-factor coding region (Kurjan and Herskowitz, Abstracts 1981 Cold Spring Harbor meeting on the Molecular Biology of Yeasts, page 242).

A PstI-SalI fragment of pAB201 containing the $\alpha$-factor leader-hEGF fusion was cloned in phage M13 and isolated in a single-stranded form. A synthetic 36-mer of sequence 5'-GGGTACCTTTGGATAAAAAGAAACTCCGACTCCGAAT-3' was synthesized and 70 picomoles were used as a primer for the synthesis of the second strand from 1 picomole of the above template by the Klenow fragment of DNA polymerase. After fill-in and ligation at 14° for 18 hrs. the mixture was treated with S1 nuclease (5 units for 15 min) and used to transfec E. coli JM101 cells. Bacteriophage containing DNA sequences in which the region coding for (glu-alal)3 at the processing site was removed were located by filter plaque hybridization using the 32P-labelled primer as probe. RF DNA from positive plaques was isolated, digested with PstI and SalI and the resulting fragment inserted in pAB114 which had been previously digested to completion with SalI, partially with PstI, and treated with alkaline phosphatase. The resulting plasmid (pαEGF-24) included
a KpnI site where the yeast α-factor sequence joined the hEGF sequence.

The plasmid pAB114 was derived as follows: plasmid pAB112 was digested to completion with HindIII and then religated at low (4μg/ml) DNA concentration and plasmid pAB113 was obtained in which three 63bp HindIII fragments have been deleted from the α-factor structural gene, leaving only a single copy of mature α-factor coding region. A BamHI site was added to plasmid pAB11 by cleavage with EcoRI, filling in of the overhanging ends by the Klenow fragment of DNA polymerase, ligation of BamHI linkers, cleavage with BamHI and religation to obtain pAB12. Plasmid pAB113 was digested with EcoRI, the overhanging ends filled in, and ligated to BamHI linkers. After digestion with BamHI, the 1500bp fragment was gel-purified and ligated to pAB12 which had been digested with BamHI and treated with alkaline phosphatase, resulting in plasmid pAB114, which contains the 1500bp BamHI fragment carrying the α-factor gene.

Referring now to Fig. 1, the synthetic CTAP-III gene sequence was dissolved in a 20μl reaction pool containing 10mM dithiothreitol, 1mM ATP, 10mM MgCl₂, 100ng/ml spermidine, 50mM Tris-HCl, pH 7.8 and 8 units of T4 polynucleotide kinase, and the mixture was incubated for 50 minutes at 37°C. Plasmid pαEGF-24 (1.0 μl, 100 ng/μl) was digested with SalI and KpnI and the larger fragment gel isolated and added to the reaction mixture together with T4 ligase (400 units). The resulting mixture was incubated at 14°C for 22 hours, after which the DNA was ethanol precipitated and the pellet washed once with ethanol.

The DNA was resuspended and used to transform E. coli HB101 cells. Of 8 transformants, alkaline SDS plasmid mini-preparations followed by restriction site screening yielded 3 putative recombinants which were sequence-screened by the M13 dideoxy-sequencing method.
All 3 transformants possessed mutant sequences, and one contained a single silent mutation in the third position of the 39th codon (codon for valine) where the cytosine in the coding strand was replaced by a thymine. Because this new codon was even more preferred by yeast (according to the glycolytic yeast codon usage bias observed for yeast) than the codon initially chosen, no further screening was performed. The plasmid was designated pωCTAP.

A yeast shuttle vector carrying the CTAP-III gene and associated α-factor leader and processing signal sequence was prepared as follows. Plasmid pωCTAP was digested with BamHI and an approximately 1.7 kbp fragment isolated by preparative gel electrophoresis. Approximately 4 µg of this fragment was ligated with 0.1 µg of a vector prepared from BamHI digestion and alkaline phosphatase treatment of pCl/l in a 20 µl reaction mixture containing 1mM ATP, 10mM DDT, 10mM MgCl₂, 100ng/ml spermidine, 50mM Tris-HCl, pH 7.8, and 400 units of T4 ligase for 21.5 hours at 14°C. Plasmid pCl/l is a derivative of pJDB219 (Beggs (1978) Nature 275:104) in which the region corresponding to bacterial plasmid pMB9 in pJDB219 has been replaced by pBR322 in pCl/l. E. coli transformants are selected by ampicillin resistance.

Transformation of E. coli HB101 cells yielded approximately 1200 ampicillin resistant transformants of which 22 were restriction site screened. Of these, 8 were putative CTAP-III recombinants. Further restriction site screening demonstrated that one of the recombinant plasmids (designated pYαCTAP3-1) carried a single copy of the CTAP-III gene and accompanying α-factor segments, while a second recombinant (designated pYαCTAP3-2) carried a tandem repeat of the fragment in a head-to-tail arrangement, as illustrated in Fig. 1.
Yeast (Saccharomyces cerevisiae strain AB103) spheroplasts were transformed with pYocTAP3-1 and pYocTAP3-2, as described by Hinnen et al. (1978) Proc. Nat. Acad. Sci. USA 75:1929-1933, and leucine prototrophs were selected. Cultures of these transformants were grown to stationary phase in leucine-selective media. After removal of the cells by centrifugation, 45ml of supernatants, adjusted to pH 3.0 with 0.1N HCl, were passed over a 0.8ml Bio-Rex 70 (50-100 mesh) cation exchange column pre-equilibrated with 0.1M acetic acid. The columns were washed with two bed volumes of 0.1M acetic acid and two volumes of 50% ethanol. Protein was eluted with 3.0 bed volumes of a solution of 2 parts 5mM HCl and 8 parts ethanol and then evaporated to dryness, resuspended in water and stored at 4°C.

The proteins produced by both pYocTAP3-1 and pYocTAP3-2 were characterized (as described below) and found to have properties substantially identical to human CTAP-III isolated from a natural source. The proteins were electrophoresed on an SDS-polyacrylamide gel and Coomassie Blue staining revealed a strong band slightly smaller in size than the cytochrome C marker, thus representing a protein of about 10,000 daltons. CTAP-III has a calculated molecular weight of 9278 daltons. Comparison of the staining intensity to known standards indicated a yield of approximately 2mg protein per liter of culture for both transformants. Protein sequencing established that the first 15 N-terminal amino acids of both proteins were identical to those reported for human connective tissue activator peptide III.

The purified proteins were shown to display biological activity. Human connective tissue activator peptide III isolated from human platelets is a potent mitogen for human connective tissue cells in culture. Marked stimulation of human dermal fibroblasts by the
protein purified from human platelets occurs at concentrations ranging from 1-20 μg/ml, depending upon the source of cells used for assay and variation in individual preparations. The mitogenic activity of the material produced by both transformants was assayed using human dermal fibroblasts, as described by Castor et al. (1983), supra. Incorporation of $^3$H-labelled thymidine in the cellular DNA in the presence of different concentrations of the gene product of both pYαCTAP3-1 and pYαCTAP3-2 transformants was measured. These results are summarized in Table 1 below. The addition of 10μg/ml of protein from either transformant resulted in marked stimulation of proliferation. Stimulation was observed with as little as 1μg/ml of either protein. The mitogenic activity of the proteins from both transformants falls within the range of that reported in the literature for the protein purified from platelets.

Table 1: Assay for mitogenic activity using contact inhibited human dermal fibroblasts.

<table>
<thead>
<tr>
<th>Substance tested</th>
<th>$^3$H-Thymidine Incorporation (CPM)</th>
<th>Ratio of test sample to control</th>
</tr>
</thead>
<tbody>
<tr>
<td>no addition (control)</td>
<td>9,603</td>
<td>1.0</td>
</tr>
<tr>
<td>20% human serum</td>
<td>25,939</td>
<td>2.7</td>
</tr>
<tr>
<td>pYαCTAP3-1</td>
<td>36,183</td>
<td>3.7</td>
</tr>
<tr>
<td>10 μg/ml</td>
<td>16,345</td>
<td>1.7</td>
</tr>
<tr>
<td>1 μg/ml</td>
<td>7,930</td>
<td>0.8</td>
</tr>
<tr>
<td>pYαCTAP3-2</td>
<td>39,158</td>
<td>4.1</td>
</tr>
<tr>
<td>10 μg/ml</td>
<td>17,348</td>
<td>1.8</td>
</tr>
<tr>
<td>0.1 μg/ml</td>
<td>8,364</td>
<td>0.9</td>
</tr>
</tbody>
</table>
2. **β-Thromboglobulin (β-TG) Expression**

The nucleotide sequence devised for β-TG is identical to that shown above for CTAP-III except that it lacks the 12bp encoding the four N-terminal amino acids (AsnLeuAlaLys) of CTAP-III and utilizes the codon GTT for valine at position 35 of β-TG, which corresponds to the 39th codon of CTAP-III. The sequence, with the coding strand indicated 5' to 3', is as follows:

**Processing Site**

LeuLeuAspLysArgGlyLysGluGluSerLeuAspSerAsp
5'-TATTGATAAAAAGAGTAAGGAAAGAATCTTGGACTCTGAC
3'-CATGATAACCTATTTTCTACTCTCTTTAGGAACCTGAGCTG

**FokI**

LeuTyrAlaGluLeuArgCysMetCysIleLysThrThrSerGlyIleHisProLysAsn
TTGTAACCTCTGAAAGTAGTATGTATGATCAAGACACCCCTCTGATACCTCAACAAAAGAAC
AATGCTGACTGAACTCTACATACACAGTAGTCTGAGAGGCACATAGGTTGGTTCTCTTG

**KpnI**

IleGlnSerLeuGluValIleGlyLysGlyThrHisCysAsnGlnValGluValIleAla
ATCCAAATTTTTGGAAAGTTATGTAAGGTAACGACTGTAAACCACGAGTTTATCGCT
TAGTTAGAAACCTCAAACCACCTACGGGTAGATGTTGTTCAACTCTCAATAGCGA

**BglII**

ThrLeuLysAspGlyArgLysIleCysLeuAspProAspAlaProArgIleLysIle
ACCTTGAAGCGGCTAGAAAAGATCTGTTTGTGACCAGATGCTCACAAGATCAAGAGATC
25

**ValGlnLysLeuAlaGlyAspGluSerAlaAspOC AM**

GTCCAAGAAAGTTGCTGTTGACGAAATCCTGCTGAATATCGCGTCG-3'
CAAGTTTCTTCCACCCAGACCACGCTGCTAGACGTATATGTCACGACGACGCT-5'

As described above for CTAP-III, the sequence includes KpnI and SalI cohesive ends at its 5' and 3' termini, respectively, utilizes preferred yeast codons, and contains the yeast g-factor secretory leader (3'-portion) and modified processing signal sequence.

A DNA fragment for β-TG having the sequence just described is prepared as follows (See Fig. 2). The two single-stranded oligonucleotides shown below of length 87 and 95 bases, designated B-1 and B-2, respectively, are chemically synthesized using automated phosphoramidite methods (Urdea et al. (1983) *supra*., Beaucage and Carruthers (1981) *supra*.)
Designation | Sequence (5' to 3')
---|---
B-1 | 5'-TATTGGATAAAAGAGGCTAGGAAGAAATCCTTGGACTCTGACCTGAGTCTGAGTCTGAGATCTGATGATATCAAGACCACCTCTCTGTA-3'  
5 | 5'-TGGATACCCAGGGTTGTCTCTTGATAACACATACATCTCAATTCACGGCAGTCAGAAGTCGCTGCCAGAGATTCTCTCCCTACCTCTTTTATCCAAATAGTAC-3'

After purification of the oligonucleotides by gel electrophoresis, segment B-2 is treated with T4 polynucleotide kinase, the enzyme inactivated (see above) and segment B-1 added. The two segments are coprecipitated with NaOAc and ethanol, recovered and washed by centrifugation, and dried as described above.

The pellet is then redissolved in water (18μl), heated to 95°C for three minutes and cooled slowly (1.5 hours) to room temperature in a water bath. The resultant double-stranded DNA fragment contains the α-factor secretory leader (3'-part) and processing signal sequences, and encodes the N-terminal portion (about 1/3) of β-TG.

To obtain the C-terminal portion of the β-TG gene, plasmid pωCTAP (described above) is digested with restriction enzymes FokI and SalI, and the 177bp fragment containing the C-terminal sequence (approximately 2/3) common to both the β-TG and CTAP-III structural genes, is gel isolated. The C-terminal sequence is then ligated with a 50-fold excess of the synthetic N-terminal fragment (prepared above), and the ligation mixture treated with SalI, extracted with phenol, and precipitated with ethanol. The fusion product containing the β-TG structural gene is kinased (see above) and inserted into the bacterial cloning plasmid pωEGF-24 (described above) previously digested with KpnI and SalI to remove a KpnI/SalI fragment.

The resulting DNA construct is used to transform E. coli HB101 cells, and the desired recombinant plasmid identified in alkaline SDS plasmid
mini-preparations by restriction site screening. The plasmid carrying the \( \beta \)-TG gene (designated p\( \alpha \)\( \beta \)-TG in Fig. 2) is isolated and its structure confirmed by sequencing.

Alternatively, a synthetic fragment for \( \beta \)-thromboglobulin may be prepared in precisely the same fashion as the synthesis of the fragment for CTAP-III above, except that the single-stranded fragments C-2, C-11 and C-12 are replaced with equimolar amounts of fragments C-21, C-22 and C-23, respectively. The sequence of these ssDNA fragments are as follows:

5' to 3' direction

C-21 (17mer) GGTAAGGAAAGAATCTTT
C-22 (25mer) TCCTACTCTTTTTATCCAATAGTAC
C-23 (23mer) ACAAGTCAAGTGACCAAGGTCT

As a second alternative, a synthetic gene fragment coding for \( \beta \)-thromboglobulin may be prepared as follows. A PstI/SalI fragment containing the \( \alpha \)-factor leader/CTAP-III fusion is gel isolated from a PstI/SalI digest of pY\( \alpha \)CTAP3-1 (described above), cloned in phage M13, and isolated in a single stranded form. A synthetic 30-mer of sequence

5'-GGATTCTTCTCTACCTCTTTATCCAATAG-3'

is synthesized, and 70 picomoles are used as a primer for the synthesis of the second strand from 1.0 picomole of the \( \alpha \)-factor leader/CTAP-III template using the Klenow fragment of DNA polymerase. After fill-in and ligation at 14°C for 18 hours, the mixture is treated with \( S_1 \) nuclease (5 units for 15 min) and then used to transfect E. coli JM101 cells. Bacteriophage containing DNA sequences in which the region coding for the four N-terminal amino acids (AsnLeuAlaLys) have been removed are located by filter plaque hybridization using the \( ^{32}P \)-labelled primer as probe. RF (replicative form) DNA from positive plaques is isolated and digested with PstI and SalI, and the resulting fragment is inserted into p\( \alpha \)EGF-24 which has
been previously digested to completion with SalI and partially with PstI and treated with alkaline phosphatase.

The resulting DNA construct is used to transform E. coli HB101 cells, and the desired recombinant plasmid identified in alkaline SDS mini-preparations by restriction site screening. Plasmids carrying the β-TG gene may be isolated and their structure confirmed by sequencing.

A yeast shuttle vector carrying the β-TG gene and associated α-factor leader and processing signals may be prepared from any of the plasmids described by digestion of the plasmid (e.g., pαβ-TG) with BamHI and isolation of an about 1.7kb fragment by gel electrophoresis. This isolated fragment is then inserted into plasmid pCl/1 (described above) previously digested with BamHI and treated with alkaline phosphatase. Ampicillin resistant transformants of E. coli HB101 are restriction site screened, and the resultant plasmid (designated pYαβ-TG in Fig. 2) isolated and used to transform S. cerevisiae strain AB103.1. The presence and structure of secreted β-thromboglobulin in supernatant medium from cultures of selected leucine prototrophs are verified by SDS/polyacrylamide gel electrophoresis, amino acid sequencing, and immunological assay.

In accordance with the subject invention, novel DNA constructs are provided which may be inserted into appropriate vectors to provide for expression of precursor polypeptides for both CTAP-III and β-TG and subsequent intracellular processing and secretion of the mature polypeptides. The mature CTAP-III polypeptide has been found to display biological activity which closely corresponds to that of natural human CTAP-III in a recognized bioassay based on the stimulation of tissue regeneration in cell culture of human dermal fibroblasts. By providing for secretion,
greatly enhanced yields of the desired polypeptide product are obtained and subsequent isolation and purification of the polypeptides simplified.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.
WHAT IS CLAIMED IS:

1. A polypeptide having substantially the same amino acid sequence and providing a substantially equivalent biological activity as human connective tissue activator peptide-III and degradation products thereof isolated from a human host, said polypeptide having been produced in a unicellular host and being present in combination with at least trace amounts of natural gene products of the unicellular host.

2. A polypeptide as in claim 1, wherein the amount of natural gene product present does not interfere with the biological activity of the polypeptide or have an adverse physiological effect when administered to a mammalian host.

3. A polypeptide as in claim 1, wherein the polypeptide displays substantially the same mitogenic activity as that displayed by human connective tissue activator peptide isolated from human platelets.

4. A polypeptide as in claim 1, wherein the polypeptide displays substantially the same immunological activity as that displayed by human β-thromboglobulin isolated from platelets.

5. A polypeptide as in claim 1, wherein the unicellular host is a yeast.

6. A polypeptide as in claim 5, wherein the yeast is Saccharomyces cerevisiae.

7. A polypeptide as in claim 1, wherein the polypeptide is encoded on an extrachromosomal element comprising a synthetic structural gene in proper
reading frame with a secretory leader and processing signal sequence recognized by the host.

8. A polypeptide as in claim 7, wherein the secretory leader and processing signal sequence is derived from the yeast α-factor gene.

9. A method for producing a polypeptide having substantially the same amino acid sequence as human connective tissue activator peptide-III and degradation products thereof in a suitable unicellular host, said method comprising growing host cells containing an extrachromosomal element having a gene encoding the peptide in proper reading frame with a secretory leader and processing signal sequences recognized by said host so that the host secretes mature polypeptide, and isolating the mature polypeptide.

10. A method as in claim 9, wherein the peptide gene is a synthetic gene having codons preferentially utilized by the host.

11. A method as in claim 10, wherein the host is a yeast.

12. A method as in claim 11, wherein the secretory leader and processing signal sequence is derived from the yeast α-factor gene.

13. A method as in claim 9, wherein the polypeptide displays substantially the same mitogenic activity as that displayed by naturally-occurring human connective tissue activator peptide-III.

14. A method as in claim 13, wherein the peptide gene has the following nucleotide sequence:
5'-AACTTGGCCAAGGTAAAGGGATCTGTGACTCTTGACTG
3'-TTGAAACCGGTTCACATCTCTCTTCTTAGGAAACCTGAGCTG

TTGTACGCTGAAATGGAGATGTATGCTATCAAGACACCCTCTCTG
AACATGGCACTTAAACTCTACTACATAACATAGTCTTCTTGGGAGACATAGGATTGGTTTTCTTG

5'
ATCCCAATCTTTGGAGATTTATTGTAGTTTGAAGGGTACCTGCTTAAGCAGGTTAATGGCT
TGGGTTAGAAAATTTCAATATTCCATCCCAGGTGAGATGTGGATTCTCAGGAGGTTC

ACCTGGAAGGACGTTGAAGGAGAGATCTGTTGGAGCCAGATGCTCCAAGAATGCAAGAGATCG
TGGAACTTCTTGGCCTTCTTTCTAGAGACAACAAACTGTTGCTTCTTAGGTTCTTCTTAG

GTTCAAAAAGAAGTTGGCTTTGGTACGAGATCCTGCTGAC-3'
CAAGTTTTTCTTCAACCAGACCCTCTGCTTAGAGACGTG-5'

15. A method as in claim 14, wherein the host is a yeast and the extrachromosomal element is pYCTAP3-1 or pYCTAP3-2.


17. A method as in claim 9, wherein the polypeptide displays substantially the same immunological activity as that displayed by human \( \beta \)-thromboglobulin.

18. A method as in claim 17, wherein the polypeptide gene has the following nucleotide sequence:
5'-GGTAAAGGAAGAATCCTTTGGACTCTGAC
3'-CCAATTCTTCTTTAGGAAACCTGAGCTG

TTGTACGCTGAAATGGAGATGTATGCTATCAAGACACCCTCTCTG
AACATGGCACTTAAACTCTACTACATAACATAGTCTTCTTGGGAGACATAGGATTGGTTTTCTTG

5'
ATCCCAATCTTTGGAGATTTATTGTAGTTTGAAGGGTACCTGCTTAAGCAGGTTAATGGCT
TGGGTTAGAAAATTTCAATATTCCATCCCAGGTGAGATGTGGATTCTCAGGAGGTTC

ACCTGGAAGGACGTTGAAGGAGAGATCTGTTGGAGCCAGATGCTCCAAGAATGCAAGAGATCG
TGGAACTTCTTGGCCTTCTTTCTAGAGACAACAAACTGTTGCTTCTTAGGTTCTTCTTAG

GTTCAAAAAGAAGTTGGCTTTGGTACGAGATCCTGCTGAC-3'
CAAGTTTTTCTTCAACCAGACCCTCTGCTTAGAGACGTG-5'

19. A DNA construct comprising a coding sequence for a polypeptide having substantially the same amino acid sequence as human connective tissue
activator peptide-III and degradation products thereof in proper reading frame with a secretory leader and processing signal sequence recognized by yeast, and a replication system recognized by yeast.

20. A DNA construct as in claim 19, wherein the coding sequence is synthesized in vitro and includes codons preferentially utilized by yeast.

21. A DNA construct as in claim 19, further comprising a replication system recognized by bacteria.

22. A DNA construct as in claim 19, wherein the yeast replication system is derived from the 2μm plasmid.

23. A DNA construct as in claim 19, wherein the polypeptide displays substantially the same mitogenic activity as that displayed by naturally-occurring human connective tissue activator peptide-III.

24. A DNA construct as in claim 23, wherein the coding sequence is as follows:

```
5'-AACCTTGGCGGCAAGGGTAAAGGAGAATCTTGGACTCTGAC
3'-TTGAACCGGGTTCCCCATTCCTCTTCTTGAACCTGAGACTG
TTGTACGCTGAATTTGATGTATGCTATCAAGGACCACCTCCTGATCCACCCAAAGAAAC
AACATGCAGCTTTAACTCTACATACATAGTTTCTGTGTGAACCATAGGTTGCTTCTTG
ATCCAAATCTTTGGAAGTTATTGGAAGGCTACCACACTGTAACCAAGTTAGTTATCGCT
TTAGTTTAGAAACCTTTCAATAACCATTCCCATGGTGACATTGTTTCAACTTCAAATAGCGA
ACCTTGAAGGACGGTAGAAAGATCTGTGTTGGACCCAGATGCTCACAAGAATCAAGAAGATC
TGAAACTCTCTCCTCTCTTAGACAAACCTCTGCTAGGATCTCTTAGTTCTCTCTCTTAG
GTTCAAAAAGAAAGTTGGCTGGTACGGAATCTGCTGAC-3'
CAAGTTTTCTTCAACCGAGCGCTTGAACGCTG-5'
```

25. Plasmid pYαCTAP3-1 or pYαCTAP3-2.
26. A yeast cell transformed with a DNA construct as in claim 19.

27. A DNA construct as in claim 19, wherein the polypeptide displays substantially the same immunological activity as that displayed by human β-thromboglobulin.

28. A DNA construct as in claim 27, wherein the polypeptide gene has the following nucleotide sequence:

```
5'-GGTAAGGAAAGAATCTTTGGAATCCTGAC
3'-CCATTCCCCTTAGGAACCTTGAC
TTGTAGCCTGAAATTGAGATGTTATGTGTATCAAGACACACCTCCTGGTATCCACCACCACAGAC
AACATGCGACTTAACTCTACATACATACATAGTTCTGGGAGACCATAGTGGGTTCTTT
ATCCAATCTTTGGGAAGTTATGTGGTAAGGGTTACCCACTGTAACCAAGTTGAATTTATCGCT
TAGTTTAGAACAACCCTTACAATTTACCATCCCATGGGTGACATTTGTTTCAACTTCAATAGCGA
ACCTTGAGGACGGTAGAAAGATCCTGTATGGACCAGATGCTCCAAGAATCAAGAGATCC
TGGACTCTCCTGCCATTTTCTAGACAAACCTGGGTCTAGGAGTTTTCTATTTCTTAG
GGTTAAAAGAAGTTGGCTGGTGACGAATCTGCTGAC-3'
CAAGTTTTCTTCAACCGACCACCTGCTTAGACCGACTG-5'
```
FIG. 2.
**INTERNATIONAL SEARCH REPORT**

**International Application No** PCT/US85/01657

**I. CLASSIFICATION OF SUBJECT MATTER**

(If several classification symbols apply, indicate all) 3

According to International Patent Classification (IPC) or to both National Classification and IPC

**I.P.C.** C07K 7/04; C07K 13/00; C07H 21/04; C12P 21/02; C12P 21/00; C12N 1/00

**II. FIELDS SEARCHED**

<table>
<thead>
<tr>
<th>Classification System</th>
<th>Classification Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>435/68, 70, 71, 91, 172.3, 243, 253, 255, 256, 317; 536/27; 260/112R, 112.5R; 514/2; 935/11, 12, 13, 14, 28, 29, 37, 38; 69, 72, 73, 74, 75</td>
</tr>
</tbody>
</table>

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched 6

**CA SEARCH DATABASE 1972-1985**

**III. DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of Document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to Claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>US.A, 4, 418, 149, Published 29 November 1983, Ptashne et al.</td>
<td>1-28</td>
</tr>
</tbody>
</table>

* Special categories of cited documents: 15

**A** - document defining the general state of the art which is not considered to be of particular relevance

**E** - earlier document but published on or after the international filing date

**L** - document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

**O** - document referring to an oral disclosure, use, exhibition or other means

**P** - document published prior to the international filing date but later than the priority date claimed

"T" - later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" - document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" - document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"A" - document member of the same patent family

**IV. CERTIFICATION**

Date of the Actual Completion of the International Search 1 31 October 1985

Date of Mailing of this International Search Report 2 15 NOV 1985

International Searching Authority 1 ISA/US

Signature of Authorized Officer 19

James Martinell
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of Document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to Claim No</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>N, Begg et al, Biochemistry, Vol. 17, 1978, pages 1739-1744</td>
<td>1-6 and 16</td>
</tr>
<tr>
<td>Y</td>
<td>EP,A, 0068375, Published 05 January 1983, Stewart et al.</td>
<td>1-28</td>
</tr>
</tbody>
</table>
### FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

|---|---|---|

### OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. Claim numbers ......... because they relate to subject matter not required to be searched by this Authority, namely:

2. Claim numbers ......... because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

### OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This international Searching Authority found multiple inventions in this international application as follows:

1. Claims 1-8 and 16 drawn to proteins.
2. Claims 9-15 and 17-28 drawn to DNAs, vectors, and cells.

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.
2. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:
3. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:
4. As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest
- The additional search fees were accompanied by applicant’s protest.
- No protest accompanied the payment of additional search fees.