PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 5/00, 1/00 | A1

(11) International Publication Number: **WO 99/60099**

(43) International Publication Date: 25 November 1999 (25.11.99)

(21) International Application Number: PCT/US98/10272

(22) International Filing Date: 20 May 1998 (20.05.98)

(71) Applicant: UNITED STATES SURGICAL CORPORATION [US/US]; 150 Glover Avenue, Norwalk, CT 06856 (US).

(72) Inventors: GRUSKIN, Elliott, A.; 23 Beach Tree Ridge, Killingworth, CT 06419 (US). BUECHTER, Douglas, D.; 51 Pierson Drive, Wallingford, CT 06492 (US). ZHANG, Guanghui; 975 Little Meadow Road, Guilford, CT 06473 (US). CONNOLLY, Kevin; 4024 La Salle Avenue, Culver City, CA 90232 (US).

(74) Agents: BARRESE, Rocco, S. et al.; Dilworth & Barrese, 333 Earle Ovington Boulevard, Uniondale, NY 11553 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: AMINO ACID MODIFIED POLYPEPTIDES

(57) Abstract

Incorporation of certain amino acid analogs into polypeptides produced by cells which do not ordinarily provide polypeptides containing such amino acid analogs is accomplished by subjecting the cells to growth media containing such amino acid analogs. The degree of incorporation can be regulated by adjusting the concentration of amino acid analogs in the media and/or by adjusting osmolality of the media. Such incorporation allows the chemical and physical characteristics of polypeptides to be altered and studied.

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.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

AMINO ACID MODIFIED POLYPEPTIDES

BACKGROUND

1. Technical Field

Engineered polypeptides having incorporated amino acids which enhance or otherwise modify properties of such polypeptides.

2. Description of Related Art

Genetic engineering allows polypeptide production to be transferred from one organism to another. In doing so, a portion of the production apparatus indigenous to an original host is transplanted into a recipient. Frequently, the original host has evolved certain unique processing pathways in association with polypeptide production which are not contained in or transferred to the recipient. For example, it is well known that mammalian cells incorporate a complex set of post-translational enzyme systems which impart unique characteristics to protein products of the systems. When a gene encoding a protein normally produced by mammalian cells is transferred into a bacterial or yeast cell, the protein may not be subjected to such post translational modification and the protein may not function as originally intended.

Normally, the process of polypeptide or protein synthesis in living cells involves transcription of DNA into RNA and translation of RNA into protein. Three forms of RNA are involved in protein synthesis: messenger RNA (mRNA) carries genetic information to ribosomes made of ribosomal RNA (rRNA) while transfer RNA (tRNA) links to free amino acids in the cell pool. Amino acid/tRNA complexes line up next to codons of mRNA, with actual recognition and binding being mediated by tRNA. Cells can contain up to twenty amino acids which are combined and incorporated in sequences of varying permutations into proteins. Each amino acid is distinguished from the other nineteen amino acids and charged to tRNA by enzymes known as aminoacyl-tRNA synthetases. As a general rule, amino acid/tRNA complexes are quite specific and normally only a molecule with an exact stereochemical configuration is acted upon by a particular aminoacyl-tRNA synthetase.

In many living cells some amino acids are taken up from the surrounding environment and some are synthesized within the cell from precursors, which in turn have been assimilated from outside the cell. In certain instances, a cell is auxotrophic, i.e., it requires a specific growth substance beyond the minimum required for normal metabolism and reproduction which it must obtain from the surrounding environment. Some auxotrophs depend upon the external environment to supply certain amino acids. This feature allows certain amino acid analogs to be incorporated into proteins produced by auxotrophs by taking advantage of relatively rare exceptions to the above rule regarding stereochemical specificity of aminoacyl-tRNA synthetases. For example, proline is such an exception, i.e., the amino acid activating enzymes responsible for the synthesis of prolyl-tRNA complex are not as specific as others. As a consequence certain proline analogs have been incorporated into bacterial, plant, and animal cell systems. See Tan et al., Proline Analogues Inhibit Human Skin Fibroblast Growth and Collagen Production in Culture, Journal of Investigative Dermatology, 80:261-267(1983).

A method of incorporating unnatural amino acids into proteins is described, e.g., in Noren et al., A General Method For Site-Specific Incorporation of Unnatural Amino Acids Into Proteins, Science, Vol. 244, pp. 182-188 (1989) wherein chemically acylated suppressor tRNA is used to insert an amino acid in response to a stop codon substituted for the codon encoding residue of interest. See also, Dougherty et al., Synthesis of a Genetically Engineered Repetitive Polypeptide Containing Periodic Selenomethionine Residues, Macromolecules, Vol. 26, No. 7, pp. 1779-1781 (1993), which describes subjecting an *E. coli* methionine auxotroph to selenomethionine containing medium and postulates on the basis of experimental data that selenomethionine may completely replace methionine in all proteins produced by the cell.

cis-Hydroxy-L-proline has been used to study its effects on collagen by incorporation into eukaryotic cells such as cultured normal skin fibroblasts (see Tan et al., *supra*) and tendon cells from chick embryos (see e.g., Uitto et al., Procollagen Polypeptides Containing *cis*-4-Hydroxy-L-proline are Overglycosylated and Secreted as Nonhelical Pro-γ-Chains, Archives of Biochemistry and

Biophysics, 185:1:214-221(1978)). However, investigators found that *trans*-4-hydroxyproline would not link with proline specific tRNA of prokaryotic *E. coli*. See Papas et al., Analysis of the Amino Acid Binding to the Proline Transfer Ribonucleic Acid Synthetase of *Escherichia coli*, Journal of Biological Chemistry, 245:7:1588-1595(1970). Another unsuccessful attempt to incorporate *trans*-4-hydroxyproline into prokaryotes is described in Deming et al., In Vitro Incorporation of Proline Analogs into Artificial Proteins, Poly. Mater. Sci. Engin. Proceed., Vol. 71, p. 673-674 (1994). Deming et al. report surveying the potential for incorporation of certain proline analogs, i.e., L-azetidine-2-carboxylic acid, L-γ-thiaproline, 3,4-dehydroproline and L-*trans*-4-hydroxyproline into artificial proteins expressed in *E. coli* cells. Only L-azetidine-2-carboxylic acid, L-γ-thiaproline and 3,4 dehydroproline are reported as being incorporated into proteins in *E. coli* cells in vivo.

Type I collagen is the most abundant form of the fibrillar, interstitial collagens and is the main component of the extracellular matrix. Collagen monomers consist of about 1000 amino acid residues in a repeating array of Gly-X-Y triplets. Approximately 35% of the X and Y positions are occupied by proline and 4-hydroxyproline. Collagen monomers associate into triple helices which consist of one α2 and two α1 chains. The triple helices associate into fibrils which are oriented into tight bundles. The bundles of collagen fibrils are further organized to form the scaffold for extracellular matrix.

In mammalian cells, post-translational modification of collagen contributes to its ultimate chemical and physical properties and includes proteolytic digestion of pro-regions, hydroxylation of lysine and proline, and glycosylation of hydroxylated lysine. The proteolytic digestion of collagen involves the cleavage of pro regions from the N and C termini. It is known that hydroxylation of proline is essential for the mechanical properties of collagen. Collagen with low levels of 4-hydroxyproline has poor mechanical properties, as highlighted by the sequelae associated with scurvy. 4-hydroxyproline adds stability to the triple helix through hydrogen bonding and through restricting rotation about C-N bonds in the polypeptide backbone. In the absence of a stable structure, naturally occurring

cellular enzymes contribute to degrading the collagen polypeptide.

The structural attributes of Type I collagen along with its generally perceived biocompatability make it a desirable surgical implant material. Collagen is purified from bovine skin or tendon and used to fashion a variety of medical devices including hemostats, implantable gels, drug delivery vehicles and bone substitutes. However, when implanted into humans bovine collagen can cause acute and delayed immune responses.

As a consequence, researchers have attempted to produce human recombinant collagen with all of its structural attributes in commercial quantities through genetic engineering. Unfortunately, production of collagen by commercial mass producers of protein such as *E. coli* has not been successful. A major problem is the extensive post-translational modification of collagen by enzymes not present in *E. coli*. Failure of *E. coli* cells to provide proline hydroxylation of unhydroxylated collagen proline prevents manufacture of structurally sound collagen in commercial quantities.

SUMMARY

A method of incorporating an amino acid analog into a polypeptide produced by a cell is provided which includes providing a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, providing growth media containing at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof and contacting the cell with the growth media wherein the at least one amino acid analog is assimilated into the cell and incorporated into at least one polypeptide.

Also provided is a method of substituting an amino acid analog of an amino acid in a polypeptide produced by a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, which includes providing a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, providing growth media containing at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations

thereof and contacting the cell with the growth media wherein the at least one amino acid analog is assimilated into the cell and incorporated as a substitution for at least one naturally occurring amino acid in at least one polypeptide.

A method of controlling the amount of an amino acid analog incorporated into a polypeptide is also provided which includes providing at least a first cell selected from the group consisting of prokaryotic cell and eukaryotic cell, providing a first growth media containing a first predetermined amount of at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof and contacting the first cell with the first growth media wherein a first amount of amino acid analog is assimilated into the first cell and incorporated into at least one polypeptide. At least a second cell selected from the group consisting of prokaryotic cell and eukaryotic cell, is also provided along with a second growth media containing a second predetermined amount of an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof and the at least second cell is contacted with the second growth media wherein a second amount of amino acid analog is assimilated into the second cell and incorporated into at least one polypeptide.

Also provided is a method of increasing stability of a recombinant polypeptide produced by a cell which includes providing a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, and providing growth media containing an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof and contacting the cell with the growth media wherein the amino acid analog is assimilated into the cell and incorporated into a recombinant polypeptide, thereby stabilizing the polypeptide.

A method of increasing uptake of an amino acid analog into a cell and causing formation of an amino acid analog/tRNA complex is also provided which includes providing a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, providing hypertonic growth media containing amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline,

cis-4-fluoro-L-proline and combinations thereof and contacting the cell with the hypertonic growth media wherein the amino acid analog is assimilated into the cell and incorporated into an amino acid analog/tRNA complex. In any of the other above methods, a hypertonic growth media can optionally be incorporated to increase uptake of an amino acid analog into a cell.

A composition is provided which includes a cell selected from the group consisting of prokaryotic cell and eukaryotic cell, and hypertonic media including an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a plasmid map illustrating pMAL-c2.

Figure 2 is a graphical representation of the concentration of intracellular hydroxyproline based upon concentration of *trans*-4-hydroxyproline in growth culture over time.

Figures 3 and 3A depict a DNA sequence encoding human Type 1 (α1) collagen.

Figure 4 is a plasmid map illustrating pHuCol.

Figure 5 depicts a DNA sequence encoding a fragment of human Type 1 (α1) collagen.

Figure 6 is a plasmid map illustrating pHuCol-Fl.

Figure 7 depicts a DNA sequence encoding a collagen-like peptide.

Figure 8 depicts an amino acid sequence of a collagen-like peptide.

Figure 9 is a plasmid map illustrating pCLP.

Figure 10 depicts a DNA sequence encoding mature bone morphogenic protein.

Figure 11 is a plasmid map illustrating pCBC.

Figure 12 is a graphical representation of the percent incorporation of proline and *trans*-4-hydroxyproline into maltose binding protein under various conditions.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Prokaryotic cells and eukaryotic cells can unexpectedly be made to assimilate and incorporate *trans*-4-hydroxyproline into proteins contrary to both Papes et al. and Deming et al., supra. Such assimilation and incorporation is especially useful when the structure and function of a polypeptide depends on post translational hydroxylation of proline not provided by the native protein production system of a recombinant host. Thus, prokaryotic bacteria such as *E. coli* and eukaryotic cells such as *Saccharomyces cerevisiae, Saccharomyces carlsbergensis* and *Schizosaccharomyces pombe* that ordinarily do not hydroxylate proline and additional eukaryotes such as insect cells including lepidopteran cell lines including *Spodoptera frugiperda, Trichoplasia ni, Heliothis virescens, Bombyx mori* infected with a baculovirus; CHO cells, COS cells and NIH 3T3 cells which fail to adequately produce certain polypeptides whose structure and function depend on such hydroxylation can be made to produce polypeptides having hydroxylated prolines. Incorporation includes adding *trans*-4-hydroxyproline to a polypeptide, for example, by first changing an amino acid to proline, creating a new proline position that can in turn be substituted with *trans*-4-hydroxyproline or substituting a naturally occurring proline in a polypeptide with *trans*-4-hydroxyproline as well.

The process of producing recombinant polypeptides in mass producing organisms is well known. Replicable expression vectors such as plasmids, viruses, cosmids and artificial chromosomes are commonly used to transport genes encoding desired proteins from one host to another. It is contemplated that any known method of cloning a gene, ligating the gene into an expression vector and transforming a host cell with such expression vector can be used in furtherance of the present disclosure.

Not only is incorporation of *trans*-4-hydroxyproline into polypeptides which depend upon *trans*-4-hydroxyproline for chemical and physical properties useful in production systems which do not have the appropriate systems for converting proline to *trans*-4-hydroxyproline, but useful as well in studying the structure and function of polypeptides which do not normally contain *trans*-4-hydroxyproline. It is contemplated that the following amino acid analogs may also be incorporated in

accordance with the present disclosure: *trans*-4 hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof (hereinafter referred to as the "amino acid analogs"). Use of prokaryotes and eukaryotes is desirable since they allow relatively inexpensive mass production of such polypeptides. It is contemplated that the amino acid analogs can be incorporated into any desired polypeptide. In a preferred embodiment the prokaryotic cells and eukaryotic cells are starved for proline by decreasing or eliminating the amount of proline in growth media prior to addition of an amino acid analog herein.

Expression vectors containing the gene for maltose binding protein (MBP), e.g., see Figure 1 illustrating plasmid pMAL-c2, commercially available from New England Bio-Labs, are transformed into prokaryotes such as *E. coli* proline auxotrophs or eukaryotes such as *S. cerevisiae* auxotrophs which depend upon externally supplied proline for protein synthesis and anabolism. Other preferred expression vectors for use in prokaryotes are commercially available plasmids which include pKK-223 (Pharmacia), pTRC (Invitrogen), pGEX (Pharmacia), pET (Novagen) and pQE (Quiagen). Substitution of the amino acid analogs for proline in protein synthesis occurs since prolyl tRNA synthetase is sufficiently promiscuous to allow misacylation of proline tRNA with any one of the amino acid analogs. A sufficient quantity, i.e., typically ranging from about .001M to about 0.1M, but more preferably from about .005M to about 0.5M of the amino acid analog(s) is added to the growth medium for the transformed cells to compete with proline in cellular uptake. After sufficient time, generally from about 30 minutes to about 24 hours or more, the amino acid analog(s) is assimilated by the cell and incorporated into protein synthetic pathways. As can be seen from Figure 2, intracellular concentration of *trans*-4-hydroxyproline increases by increasing the concentration of *trans*-4-hydroxyproline in the growth media. In a preferred embodiment the prokaryotic cells and/or eukaryotic cells are starved for proline by decreasing or eliminating the amount of proline in growth media prior to addition of an amino acid analog herein.

Expression vectors containing the gene for human Type I (α1) collagen (DNA sequence illustrated in Figures 3 and 3A; plasmid map illustrated in Figure

4) are transformed into prokaryotic or eukaryotic proline auxotrophs which depend upon externally supplied proline for protein synthesis and anabolism. As above, substitution of the amino acid analog(s) occurs since prolyl tRNA synthetase is sufficiently promiscuous to allow misacylation of proline tRNA with the amino acid analog(s). The quantity of amino acid analog(s) in media given above is again applicable.

Expression vectors containing DNA encoding fragments of human Type 1 (α1) collagen (e.g., DNA sequence illustrated in Figure 5 and plasmid map illustrated in Figure 6) are transformed into prokaryotic or eukaryotic auxotrophs as above. Likewise, expression vectors containing DNA encoding collagen-like polypeptide (e.g., DNA sequence illustrated in Figure 7, amino acid sequence illustration in Figure 8 and plasmid map illustrated in Figure 9) can be used to transform prokaryotic or eukaryotic auxotrophs as above. Collagen-like peptides are those which contain at least partial homology with collagen and exhibit similar chemical and physical characteristics to collagen. Thus, collagen-like peptides consist, e.g., of repeating arrays of Gly-X-Y triplets in which about 35% of the X and Y positions are occupied by proline and 4-hydroxyproline. Certain preferred collagen fragments and collagen-like peptides in accordance herewith are capable of assembling into an extracellular matrix. In both collagen fragments and collagen-like peptides as described above, substitution with amino acid analog(s) occurs since prolyl tRNA synthetase is sufficiently promiscuous to allow misacyclation of proline tRNA with one or more of the amino acid analog(s). The quantity of amino acid analog(s) given above is again applicable.

It is contemplated that any polypeptide having a collagen, collagen fragment or collagen-like peptide domain can be made to incorporate amino acid analog(s) in accordance with the disclosure herein. Such polypeptides include collagen, a collagen fragment or collagen-like peptide domain and a domain having a region incorporating one or more physiologically active substances such as glycoproteins, proteins, peptides and proteoglycans. Physiologically active substances exert control over or modify existing physiologic functions in living things. Physiologically active substances include hormones, growth factors, enzymes,

ligands and receptors. Many active domains of physiologically active substances have been defined and isolated. It is contemplated that polypeptides having a collagen, collagen fragment or collagen-like peptide domain can also have a domain incorporating one or more physiologically active domains which are active fragments of such physiologically active substances. Thus, chimeric proteins are made to incorporate amino acid analog(s) by transforming a prokaryotic proline auxotroph or a eukaryotic proline auxotroph with an appropriate expression vector and contacting the transformed auxotroph with growth media containing at least one of the amino acid analogs. For example, a chimeric collagen/bone morphogenic protein (BMP) construct or various chimeric collagen/growth factor constructs are useful in accordance herein. Such growth factors are well-known and include insulin-like growth factor, transforming growth factor, platelet derived growth factor and the like. Figure 10 illustrates DNA of BMP which can be fused to the 3' terminus of DNA encoding collagen, DNA encoding a collagen fragment or DNA encoding a collagen-like peptide. Figure 11 illustrates a map of plasmid pCBC containing a collagen/BMP construct. In a preferred embodiment, proteins having a collagen, collagen fragment or collagen-like peptide domain assemble to form an extracellular matrix which can be used as a surgical implant.

In another aspect, the amount of amino acid analog(s) transport into a target cell can be regulated by controlling the tonicity of the growth media. A hypertonic growth media increases uptake of *trans*-4-hydroxyproline into *E. coli* as illustrated in Figure 12. All known methods of increasing osmolality of growth media are appropriate for use herein including addition of salts such as sodium chloride, KCl, $MgCl_2$ and the like, and sugars such as sucrose, glucose, maltose, etc. and polymers such as polyethylene glycol (PEG), dextran, cellulose, etc. and amino acids such as glycine. Increasing the osmolality of growth media results in greater intracellular concentration of amino acid analog(s) and a higher degree of complexation of amino acid analog(s) to tRNA. As a consequence, proteins produced by the cell achieve a higher degree of incorporation of amino acid analogs. Figure 12 illustrates percentage of incorporation of proline and hydroxyproline into MBP under isotonic and hypertonic media conditions in comparison to proline in native MBP. Thus,

manipulating osmolality, in addition to adjusting concentration of amino acid analog(s) in growth media allows a dual-faceted approach to regulating their uptake into prokaryotic cells and eukaryotic cells as described above and consequent incorporation into target polypeptides.

Any growth media can be used herein including commercially available growth media such as M9 minimal medium (available from Gibco Life Technologies, Inc.), LB medium, NZCYM medium, terrific broth, SOB medium and others that are well known in the art.

Collagen from different tissues can contain different amounts of *trans*-4-hydroxyproline. For example, tissues that require greater strength such as bone contain a higher number of *trans*-4-hydroxyproline residues than collagen in tissues requiring less strength, e.g., skin. The present system provides a method of adjusting the amount of *trans*-4-hydroxyproline in collagen, collagen fragments, collagen-like peptides, and chimeric polypeptides having a collagen domain, collagen fragment domain or collagen-like peptide domain fused to a physiologically active domain, since by increasing or decreasing the concentration of *trans*-4-hydroxyproline in growth media, the amount of *trans*-4-hydroxyproline incorporated into such polypeptides is increased or decreased accordingly. The collagen, collagen fragments, collagen-like peptides and above-chimeric peptides can be expressed with predetermined levels of *trans*-4-hydroxyproline. In this manner physical characteristics of an extracellular matrix can be adjusted based upon requirements of end use.

Human collagen, collagen fragments, collagen-like peptides and the above chimeric polypeptides produced by recombinant processes have distinct advantages over collagen and its derivatives obtained from non-human animals. Since the human gene is used, the collagen will not act as a xenograft in the context of a medical implant. Moreover, unlike naturally occurring collagen, the extent of proline hydroxylation can be predetermined. This unprecedented degree of control permits detailed investigation of the contribution of *trans*-4-hydroxyproline to triple helix stabilization, fibril formation and biological activity. In addition, design of medical implants based upon the desired strength of collagen fibrils is enabled.

The following examples are included for purposes of illustration and are not to be construed as limitations herein.

EXAMPLE 1

Trans-membrane Transport

A 5 mL culture of *E. coli* strain DH5α (*supE*44 Δ*lac*U169 (ϕ80*lacZ* ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1) containing a plasmid conferring resistance to ampicillin (pMAL-c2, Fig. 1) was grown in Luria Broth to confluency (~16 hours from inoculation). These cells were used to inoculate a 1 L shaker flask containing 500 mL of M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 μg/mL ampicillin supplemented with all amino acids at 20 μg/mL) which was grown to an AU_{600} of 1.0 (18-20 hours). The culture was divided in half and the cells harvested by centrifugation. The cells from one culture, were resuspended in 250 mL M9 media and those from the other in 250 mL of M9 media containing 0.5M NaCl. The cultures were equilibrated in an air shaker for 20 minutes at 37 °C (225 rpm) and divided into ten 25 mL aliquots. The cultures were returned to the shaker and 125 μL of 1M hydroxyproline in distilled H_20 was added to each tube. At 2, 4, 8, 12, and 20 minutes, 4 culture tubes (2 isotonic, 2 hypertonic) were vacuum filtered onto 1 μm polycarbonate filters that were immediately placed into 2 mL microfuge tubes containing 1.2 mL of 0.2M NaOH/2% SDS in distilled H_20. After overnight lysis, the filters were carefully removed from the tubes, and the supernatant buffer was assayed for hydroxyproline according to the method of Grant, Journal of Clinical Pathology, 17:685 (1964). The intracellular concentration of *trans*-4-hydroxyproline versus time is illustrated graphically in Figure 2.

EXAMPLE 2

Effects of Salt Concentration on Transmembrane Transport

To determine the effects of salt concentration on transmembrane transport, an approach similar to Example 1 was taken. A 5 mL culture of *E. coli* strain DH5α (*supE*44 Δ*lac*U169 (ϕ80*lacZ* ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1) containing a plasmid conferring resistance to ampicillin (pMAL-c2, Fig. 1)

was grown in Luria Broth to confluency (~16 hours from inoculation). These cells were used to inoculate a 1 L shaker flask containing 500 mL of M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 μg/mL ampicillin supplemented with all amino acids at 20 μg/mL) that was then grown to an AU_{600} of 0.6. The culture was divided into three equal parts, the cells in each collected by centrifugation and resuspended in 150 mL M9 media, 150 mL M9 media containing 0.5M NaCl, and 150 mL M9 media containing 1.0M NaCl, respectively. The cultures were equilibrated for 20 minutes on a shaker at 37° C (225rpm) and then divided into six 25 mL aliquots. The cultures were returned to the shaker and 125 μL of 1M hydroxyproline in distilled H_20 was added to each tube. At 5 and 15 minutes, 9 culture tubes (3 isotonic, 3 x 0.5M NaCl, and 3 x 1.0M NaCl) were vacuum filtered onto 1 μm polycarbonate filters that were immediately placed into 2 mL microfuge tubes containing 1.2 mL of 0.2M NaOH/2% SDS in distilled H_20. After overnight lysis, the filters were removed from the tubes and the supernatant buffer assayed for hydroxyproline according to the method of Grant, *supra*.

EXAMPLE 3

Determination Of Proline Starvation Conditions in *E. Coli*

Proline auxotrophic *E. coli* strain NM519 (*pro*$^-$) including plasmid pMAL-c which confers ampicillin resistance was grown in M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 microgram/mL ampicillin supplemented with all amino acids at 20 μg/mL except proline which was supplemented at 12.5 mg/L) to a constant AU_{600} of 0.53 AU (17 hours post-inoculation). Hydroxyproline was added to 0.08M and hydroxyproline-dependent growth was demonstrated by the increase in the OD_{600} to 0.61 AU over a one hour period.

EXAMPLE 4

Hydroxyproline Incorporation Into Protein in *E. coli* Under Proline Starvation Conditions

Plasmid pMAL-c2 (commercially available from New England Biolabs)

containing DNA encoding for maltose-binding protein (MBP) was used to transform proline auxotrophic *E. coli* strain NM519 (*pro*$^-$). Two 1 L cultures of transformed NM519 (*pro*$^-$) in M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 μg/mL ampicillin supplemented with all amino acids at 20 μg/mL except proline which was supplemented at 12.5 mg/L) were grown to an AU_{600} Of 0.53 (~17 hours post-inoculation). The cells were harvested by centrifugation, the media in one culture was replaced with an equal volume of M9 media containing 0.08M hydroxyproline and the media in the second culture was replaced with an equal volume of M9 media containing 0.08M hydroxyproline and 0.5M NaCl. After a one hour equilibration, the cultures were induced with 1mM isopropyl-β-D-thiogalactopyranoside. After growing for an additional 3.25 hours, cells were harvested by centrifugation, resuspended in 10 mL of 10mM Tris-HCl (pH 8), 1mM EDTA, 100mM NaCl (TEN buffer), and lysed by freezing and sonication. MBP was purified by passing the lysates over 4 mL amylose resin spin columns, washing the columns with 10 mL of TEN buffer, followed by elution of bound MBP with 2 mL of TEN buffer containing 10mM maltose. Eluted samples were sealed in ampules under nitrogen with an equal volume of concentrated HCl (11.7M) and hydrolysed for 12 hours at 120 °C. After clarification with activated charcoal, hydroxyproline content in the samples was determined by HPLC and the method of Grant, *supra*. The percent incorporation of *trans*-4-hydroxyproline compared to proline into MBP is shown graphically in Figure 12.

EXAMPLE 5

Hydroxyproline Incorporation Into Protein in *S. cerevisiae* via Integrating Vectors Under Proline Starvation Conditions

The procedure described in Example 4 above is performed in yeast using an integrating vector which disrupts the proline biosynthetic pathway. A gene encoding human Type 1 (α_1) collagen is inserted into a unique shuttle vector behind the inducible *GAL10* promoter. This promoter/gene cassette is flanked by a 5' and 3' terminal sequence derived from a *S. cerevisiae* proline synthetase gene. The plasmid is linearized by restriction digestion in both the 5' and 3' terminal regions and used to transform a proline-prototrophic *S. cerevisiae* strain. The

transformation mixture is plated onto selectable media and transformants are selected. By homologous recombination and gene disruption, the construct simultaneously forms a stable integration and converts the *S. cerevisiae* strain into a proline auxotroph. A single transformant is selected and grown at 30 °C in YPD media to an OD_{600} of 2 AU. The culture is centrifuged and the cells resuspended in yeast dropout media supplemented with all amino acids except proline and grown to a constant OD_{600} indicating proline starvation conditions. 0.08M L-hydroxyproline and 2% (w/v) galactose is then added. Cultures are grown for an additional 6-48 hours. Cells are harvested by centrifugation (5000 rpm, 10 minutes) and lysed by mechanical disruption. Hydroxyproline-containing human Type 1 (α_1) collagen is purified by ammonium sulfate fractionation and column chromatography.

EXAMPLE 6

Hydroxyproline Incorporation Into Protein in *S. cerevisiae* via Non-Integrating Vectors Under Proline Starvation Conditions

The procedure described above in Example 4 is performed in a yeast proline auxotroph using a non-integrating vector. A gene encoding human Type 1 (α_1) collagen is inserted behind the inducible *GAL10* promoter in the YEp24 shuttle vector that contains the selectable Ura^+ marker. The resulting plasmid is transformed into proline auxotrophic *S. cerevisiae* by spheroplast transformation. The transformation mixture is plated on selectable media and transformants are selected. A single transformant is grown at 30 °C in YPD media to an OD_{600} of 2 AU. The culture is centrifuged and the cells resuspended in yeast dropout media supplemented with all amino acids except proline and grown to a constant OD_{600} indicating proline starvation conditions. 0.08M L-hydroxyproline and 2% (w/v) galactose is then added. Cultures are grown for an additional 6-48 hours. Cells are harvested by centrifugation (5000 rpm, 10 minutes) and lysed by mechanical disruption. Hydroxyproline-containing human Type 1 (α_1) collagen is purified by ammonium sulfate fractionation and column chromatography.

EXAMPLE 7

Hydroxyproline Incorporation Into Protein in a Baculovirus Expression System

A gene encoding human Type 1 (α_1) collagen is inserted into the pBacPAK8 baculovirus expression vector behind the AcMNPV polyhedron promoter. This construct is co-transfected into SF9 cells along with linearized AcMNPV DNA by standard calcium phosphate co-precipitation. Transfectants are cultured for 4 days at 27 °C in TNM-FH media supplemented with 10 % FBS. The media is harvested and recombinant virus particles are isolated by a plaque assay. Recombinant virus is used to infect 1 liter of SF9 cells growing in Grace's media minus proline supplemented with 10% FBS and 0.08 M hydroxyproline. After growth at 27 °C for 2-10 days, cells are harvested by centrifugation and lysed by mechanical disruption. Hydroxyproline-containing human Type 1 (α_1) collagen is purified by ammonium sulfate fractionation and column chromatography.

EXAMPLE 8

Hydroxyproline Incorporation Into Human Collagen Protein in *Escherichia coli* Under Proline Starvation Conditions

A plasmid (pHuCol, Fig. 2) encoding the gene sequence of human Type I (α_1) collagen (Figures 3 and 3A) placed behind the isopropyl-β-D--thiogalactopyranoside (IPTG)-inducible *tac* promotor and also encoding β-lactamase is transformed into *Escherichia coli* proline auxotrophic strain NM519 (*pro*$^-$) by standard heat shock transformation. Transformation cultures are plated on Luria Broth (LB) containing 100 μg/ml ampicillin and after overnight growth a single ampicillin-resistant colony is used to inoculate 5 ml of LB containing 100 μg/ml ampicillin. After growth for 10-16 hours with shaking (225 rpm) at 37 °C, this culture is used to inoculate 1 L of M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 μg/ml ampicillin, supplemented with all amino acids at 20 μg/mL except proline which is supplemented at 12.5 mg/L) in a 1.5 L shaker flask. After growth at 37 °C, 225 rpm, for 15-20 hours post-inoculation, the optical density at 600 nm is constant at approximately 0.5 OD/ml. The cells are harvested by centrifugation (5000 rpm, 5 minutes), the media decanted, and the cells resuspended in 1 L of M9 minimal media containing 100 μg/ml ampicillin, 0.08M L-hydroxyproline, and 0.5M NaCI. Following growth for 1 hour at 37 °C, 225 rpm, IPTG is added to 1mM and the cultures allowed to grow for an additional 5-15 hours. Cells are harvested by centrifugation (5000 rpm, 10 minutes) and lysed by mechanical disruption. Hydroxyproline-containing collagen is purified by ammonium sulfate fractionation and column chromatography.

EXAMPLE 9

Hydroxyproline Incorporation Into Fragments of Human Collagen Protein in *Escherichia coli* Under Proline Starvation Conditions

A plasmid (pHuCol-Fl, Figure 6) encoding the gene sequence of the first 80 amino acids of human Type 1 (α_1) collagen (Figure 5) placed behind the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *tac* promotor and also encoding β-lactamase is transformed into *Escherichia coli* proline auxotrophic strain NM519 (*pro*$^-$) by standard heat shock transformation. Transformation cultures are plated on

Luria Broth (LB) containing 100 μg/ml ampicillin and after overnight growth a single ampicillin-resistant colony is used to inoculate 5 ml of LB containing 100 μg/ml ampicillin. After growth for 10-16 hours with shaking (225 rpm) at 37 °C, this culture is used to inoculate 1 L of M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 100 μg/ml ampicillin, supplemented with all amino acids at 20 μg/mL except proline which is supplemented at 12.5 mg/L) in a 1.5 L shaker flask. After growth at 37 °C, 225 rpm, for 15-20 hours post-inoculation, the optical density at 600 nm is constant at approximately 0.5 OD/ml. The cells are harvested by centrifugation (5000 rpm, 5 minutes), the media decanted, and the cells resuspended in 1 L of M9 minimal media containing 100 μg/ml ampicillin, 0.08M L-hydroxyproline, and 0.5M NaCl. Following growth for 1 hour at 37 °C, 225 rpm, IPTG is added to 1mM and the cultures allowed to grow for an additional 5-15 hours. Cells are harvested by centrifugation (5000 rpm, 10 minutes) and lysed by mechanical disruption. The hydroxyproline-containing collagen fragment is purified by ammonium sulfate fractionation and column chromatography.

It will be understood that various modifications may be made to the embodiments disclosed herein. For example, it is contemplated that any protein produced by prokaryotes and eukaryotes can be made to incorporate one or more amino acid analogs in accordance with the present disclosure. Therefore, the above description should not be construed as limiting, but merely as exemplifications of preferred embodiments. Those skilled in art will envision other modifications within the scope and spirit of the claims appended hereto.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: GRUSKIN, ELLIOTT A.
BUECHTER, DOUGLAS
ZHANG, GUANGHUI
CONNOLLY, KEVIN

(II) TITLE OF INVENTION: AMINO ACID MODIFIED POLYPEPTIDES

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: DILWORTH & BARRESE
(B) STREET: 333 EARLE OVINGTON BOULEVARD
(C) CITY: UNIONDALE
(D) STATE: NY
(E) COUNTRY: US
(F) ZIP: 11553

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/655,086
(B) FILING DATE: 03-JUN-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: STEEN, JEFFREY S.
(B) REGISTRATION NUMBER: 32,063
(C) REFERENCE/DOCKET NUMBER: 203-1632

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 516-228-8484
(B) TELEFAX: 516-228-8516

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CAGCTGTCTT ATGGCTATGA TGAGAAATCA ACCGGAGGAA TTTCCGTGCC TGGCCCCATG      60
GGTCCCTCTG GTCCTCGTGG TCTCCCTGGC CCCCCTGGTG CACCTGGTCC CCAAGGCTTC     120
CCAAGGCTTC CAAGGTCCCC CTGGTGAGCC TGGCGAGCCT GGAGCTTCAG GTCCCATGGG     180
TCCCCGAGGT CCCCCAGGTC CCCCTGGAAA GAATGGAGAT GATGGGGAAG CTGGAAAACC     240
TGGTCGTCCT GGTGAGCGTG GGCCTCCTGG GCCTCAGGGT GCTCGAGGAT TGCCCGGAAC     300
AGCTGGCCTC CCTGGAATGA AGGGACACAG AGGTTTCAGT GGTTTGGATG GTGCCAAGGG     360
AGATGCTGGT CCTGCTGGTC CTAAGGGTGA GCCTGGCAGC CCTGGTGAAA ATGGAGCTCC     420
TGGTCAGATG GGCCCCCGTG GCCTGCCTGG TGAGAGAGGT CGCCCTGGAG CCCCTGGCCC     480
TGCTGGTGCT CGTGGAAATG ATGGTGCTAC TGGTGCTGCC GGGCCCCCTG GTCCCACCGG     540
CCCCGCTGGT CCTCCTGGCT TCCCTGGTGC TGTTGGTGCT AAGGGTGAAG CTGGTCCCCA     600
AGGGCCCCGA GGCTCTGAAG GTCCCCAGGG TGTGCGTGGT GAGCCTGGCC CCCCTGGCCC     660
TGCTGGTGCT GCTGGCCCTG CTGGAAACCC TGGTGCTGAT GGACAGCCTG GTGCTAAAGG     720
TGCCAATGGT GCTCCTGGTA TTGCTGGTGC TCCTGGCTTC CCTGGTGCCC GAGGCCCCTC     780
TGGACCCCAG GGCCCCGGCG GCCCTCCTGG TCCCAAGGGT AACAGCGGTG AACCTGGTGC     840
TCCTGGCAGC AAAGGAGACA CTGGTGCTAA GGGAGAGCCT GGCCCTGTTG GTGTTCAAGG     900
ACCCCCTGGC CCTGCTGGAG AGGAAGGAAA GCGAGGAGCT CGAGGTGAAC CCGGACCCAC     960
TGGCCTGCCC GGACCCCCTG GCGAGCGTGG TGGACCTGGT AGCCGTGGTT TCCCTGGCGC    1020
AGATGGTGTT GCTGGTCCCA AGGGTCCCGC TGGTGAACGT GGTTCTCCTG GCCCCGCTGG    1080
CCCCAAAGGA TCTCCTGGTG AAGCTGGTCG TCCCGGTGAA GCTGGTCTGC CTGGTGCCAA    1140
GGGTCTGACT GGAAGCCCTG GCAGCCCTGG TCCTGATGGC AAAACTGGCC CCCCTGGTCC    1200
CGCCGGTCAA GATGGTCGCC CCGGACCCCC AGGCCCACCT GGTGCCCGTG GTCAGGCTGG    1260
TGTGATGGGA TTCCCTGGAC CTAAAGGTGC TGCTGGAGAG CCCGGCAAGG CTGGAGAGCG    1320
AGGTGTTCCC GGACCCCCTG GCGCTGTCGG TCCTGCTGGC AAAGATGGAG AGGCTGGAGC    1380
TCAGGGACCC CCTGGCCCTG CTGGTCCCGC TGGCGAGAGA GGTGAACAAG GCCCTGCTGG    1440
CTCCCCCGGA TTCCAGGGTC TCCCTGGTCC TGCTGGTCCT CCAGGTGAAG CAGGCAAACC    1500
TGGTGAACAG GGTGTTCCTG GAGACCTTGG CGCCCCTGGC CCCTCTGGAG CAAGAGGCGA    1560
GAGAGGTTTC CCTGGCGAGC GTGGTGTGCA AGGTCCCCCT GGTCCTGCTG GACCCCGAGG    1620
GGCCAACGGT GCTCCCGGCA ACGATGGTGC TAAGGGTGAT GCTGGTGCCC CTGGAGCTCC    1680
CGGTAGCCAG GGCGCCCCTG GCCTTCAGGG AATGCCTGGT GAACGTGGTG CAGCTGGTCT    1740
TCCAGGGCCT AAGGGTGACA GAGGTGATGC TGGTCCCAAA GGTGCTGATG GCTCTCCTGG    1800
CAAAGATGGC GTCCGTGGTC TGACCGGCCC CATTGGTCCT CCTGGCCCTG CTGGTGCCCC    1860
```

TGGTGACAAG GGTGAAAGTG GTCCCAGCGG CCCTGCTGGT CCCACTGGAG CTCGTGGTGC 1920

CCCCGGAGAC CGTGGTGAGC CTGGTCCCCC CGGCCCTGCT GGCTTTGCTG GCCCCCCTGG 1980

TGCTGACGGC CAACCTGGTG CTAAAGGCGA ACCTGGTGAT GCTGGTGCCA AAGGCGATGC 2040

TGGTCCCCCT GGGCCTGCCG GACCCGCTGG ACCCCCTGGC CCCATTGGTA ATGTTGGTGC 2100

TCCTGGAGCC AAAGGTGCTC GCGGCAGCGC TGGTCCCCCT GGTGCTACTG GTTTCCCTGG 2160

TGCTGCTGGC CGAGTCGGTC CTCCTGGCCC CTCTGGAAAT GCTGGACCCC CTGGCCCTCC 2220

TGGTCCTGCT GGCAAAGAAG GCGGCAAAGG TCCCCGTGGT GAGACTGGCC CTGCTGGACG 2280

TCCTGGTGAA GTTGGTCCCC CTGGTCCCCC TGGCCCTGCT GGCGAGAAAG GATCCCCTGG 2340

TGCTGATGGT CCTGCTGGTG CTCCTGGTAC TCCCGGGCCT CAAGGTATTG CTGGACAGCG 2400

TGGTGTGGTC GGCCTGCCTG GTCAGAGAGG AGAGAGAGGC TTCCCTGGTC TTCCTGGCCC 2460

CTCTGGTGAA CCTGGCAAAC AAGGTCCCTC TGGAGCAAGT GGTGAACGTG GTCCCCCCGG 2520

TCCCATGGGC CCCCCTGGAT TGGCTGGACC CCCTGGTGAA TCTGGACGTG AGGGGGCTCC 2580

TGCTGCCGAA GGTTCCCCTG GACGAGACGG TTCTCCTGGC GCCAAGGGTG ACCGTGGTGA 2640

GACCGGCCCC GCTGGACCCC CTGGTGCTCC TGGTGCTCCT GGTGCCCCTG GCCCCGTTGG 2700

CCCTGCTGGC AAGAGTGGTG ATCGTGGTGA GACTGGTCCT GCTGGTCCCG CCGGTCCCGT 2760

CGGCCCCGCT GGCGCCCGTG GCCCCGCCGG ACCCCAAGGC CCCCGTGGTG ACAAGGGTGA 2820

GACAGGCGAA CAGGGCGACA GAGGCATAAA GGGTCACCGT GGCTTCTCTG GCCTCCAGGG 2880

TCCCCCTGGC CCTCCTGGCT CTCCTGGTGA ACAAGGTCCC TCTGGAGCCT CTGGTCCTGC 2940

TGGTCCCCGA GGTCCCCCTG GCTCTGCTGG TGCTCCTGGC AAAGATGGAC TCAACGGTCT 3000

CCCTGGCCCC ATTGGGCCCC CTGGTCCTCG CGGTCGCACT GGTGATGCTG GTCCTGTTGG 3060

TCCCCCCGGC CCTCCTGGAC CTCCTGGTCC CCCTGGTCCT CCCAGCGCTG GTTTCGACTT 3120

CAGCTTCCTC CCCCAGCCAC CTCAAGAGAA CGCTCACGAT GGTGGCCGCT ACTACCGGGC 3180

T 3181

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAGCTGTCTT ATGGCTATGA TGAGAAATCA ACCGGAGGAA TTTCCGTGCC TGGCCCCATG 60

GGTCCCTCTG GTCCTCGTGG TCTCCCTGGC CCCCCTGGTG CACCTGGTCC CCAAGGCTTC 120

CAAGGTCCCC CTGGTGAGCC TGGCGAGCCT GGAGCTTCAG GTCCCATGGG TCCCCGAGGT 180

CCCCCAGGTC CCCCTGGAAA GAATGGAGAT GATGGGGAAG CTGGAAAACC TGGTCGTCCT 240

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGATCCATGG GGCTCGCTGG CCCACCGGGC GAACCGGGTC CGCCAGGCCC GAAAGGTCCG 60

CGTGGCGATA GCGGGCTCCC GGGCGATTCC TAATGGATCC 100

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly Leu Ala Gly Pro Pro Gly Glu Pro Gly Pro Pro Gly Pro Lys Gly
1 5 10 15

Pro Arg Gly Asp Ser
20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGCGGGCCA GGAAGAAGAA TAAGAACTGC CGGCGCCACT CGCTCTATGT GGACTTCAGC 60

GATGTGGGCT GGAATGACTG GATTGTGGCC CCACCAGGCT ACCAGGCCTT CTACTGCCAT 120

GGGGACTGCC CCTTTCCACT GGCTGACCAC CTCAACTCAA CCAACCATGC CATTGTGCAG 180

ACCCTGGTCA ATTCTGTCAA TTCCAGTATC CCCAAAGCCT GTTGTGTGCC CACTGAACTG 240

AGTGCCATCT CCATGCTGTA CCTGGATGAG TATGATAAGG TGGTACTGAA AAATTATCAG 300

GAGATGGTAG TAGAGGGATG TGGGTGCCGC 330

WHAT IS CLAIMED IS:

1. A method of incorporating an amino acid analog into a polypeptide produced by a cell selected from the group consisting of prokaryotic cell and eukaryotic cell comprising:

providing a cell selected from the group consisting of a prokaryotic cell and eukaryotic cell;

providing growth media containing at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the cell with the growth media wherein the at least one amino acid analog is assimilated into the cell and incorporated into at least one polypeptide.

2. A method according to claim 1 wherein the cell is a proline auxotroph.

3. A method according to claim 2 wherein the cell is selected from the group consisting of bacterial cell, yeast cell and insect cell.

4. A method according to claim 1 wherein the at least one polypeptide is at least a portion of a collagen molecule.

5. A method according to claim 4 wherein the polypeptide is encoded for by the nucleic acid sequence shown in Figure 3.

6. A method according to claim 4 wherein the polypeptide is a fragment encoded for by the nucleic acid sequence shown in Figure 5.

7. A method according to claim 4 wherein the polypeptide is encoded for by the nucleic acid sequence shown in Figure 7.

8. A method according to claim 4 wherein the at least a portion of a collagen is fused to a physiologically active substance.

9. A method according to claim 8 wherein the physiologically active substance is a BMP.

10. A method according to claim 1 wherein nucleic acid encoding the at least one polypeptide is carried on a replicable expression vector.

11. A method according to claim 10 wherein the replicable expression vector is a plasmid.

12. A method according to claim 3 wherein the proline auxotroph is selected from the group consisting of *E. coli* and *S. cerevisiae*.

13. A method according to claim 1 wherein the at least one polypeptide is at least a portion of a maltose binding protein molecule.

14. A method according to claim 1 wherein the growth media is hypertonic.

15. A method according to claim 14 wherein an osmolality increasing agent selected from a group consisting of NaCl, KCl, $MgCl_2$, sucrose, glucose, maltose, PEG, dextran, cellulose and glycine is added to the growth media.

16. A method according to claim 14 wherein NaCl ranges from about 0.5M to about 1M.

17. A method according to claim 1 wherein the growth media contains an amount of proline which causes proline starvation of the cell.

18. A method of substituting an amino acid analog in a polypeptide

manufactured by a cell selected from the group consisting of prokaryotic cell and eukaryotic cell comprising:

providing a cell selected from the group consisting of prokaryotic cell and eukaryotic cell;

providing growth media containing at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the cell with the growth media wherein the at least one amino acid analog is assimilated into the cell and incorporated as a substitution for at least one naturally occurring amino acid in at least one polypeptide.

19. A method of controlling the amount of an amino acid analog incorporated into a polypeptide comprising:

providing a first cell selected from the group consisting of prokaryotic cell and eukaryotic cell comprising:

providing a first growth media containing a first predetermined amount of at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the cell with the first growth media wherein a first amount of amino acid analog is assimilated into the first cell and incorporated into at least one polypeptide;

providing at least a second cell selected from the group consisting of prokaryotic cell and eukaryotic cell;

providing a second growth media containing a second predetermined amount of an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the at least second cell with the second growth media wherein a second amount of amino acid analog is assimilated into the second cell and incorporated into at least one polypeptide.

20. A method of increasing stability of a recombinant polypeptide produced by a cell comprising:

providing a selected from the group consisting of prokaryotic cell and eukaryotic cell;

providing growth media containing an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the cell with the growth media wherein the amino acid analog is assimilated into the cell and incorporated into a recombinant polypeptide, thereby stabilizing the polypeptide.

21. A method of increasing uptake of an amino acid analog into a cell and causing formation of an amino acid analog/tRNA complex comprising:

providing a prokaryotic cell selected from the group consisting of prokaryotic cell and eukaryotic cell;

providing hypertonic growth media containing an amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof; and

contacting the cell with the hypertonic growth media wherein the amino acid analog is assimilated into the cell and incorporated into an amino acid analog/tRNA complex.

22. A composition comprising a cell selected from group consisting of prokaryotic cell and eukaryotic cell and hypertonic growth media including at least one amino acid analog selected from the group consisting of *trans*-4-hydroxyproline, *cis*-4-hydroxyproline, 3-hydroxyproline, *cis*-4-fluoro-L-proline and combinations thereof wherein the hypertonic growth media increases cell uptake of the at least one amino acid analog.

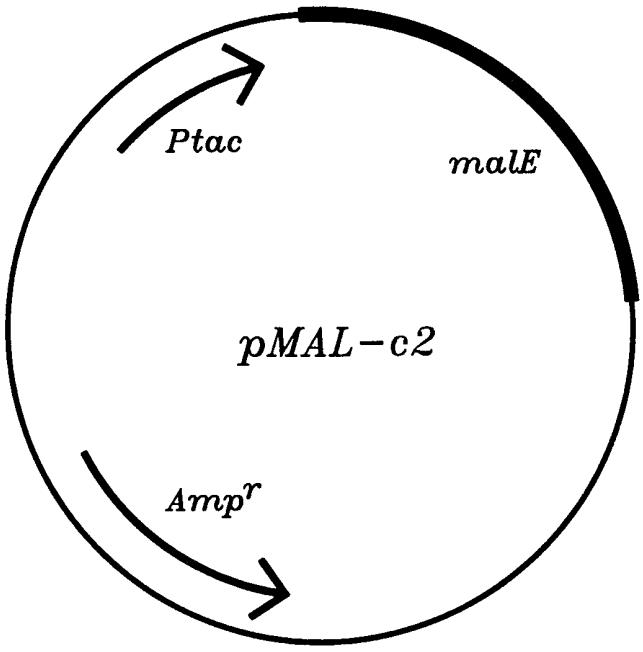


FIG. 1

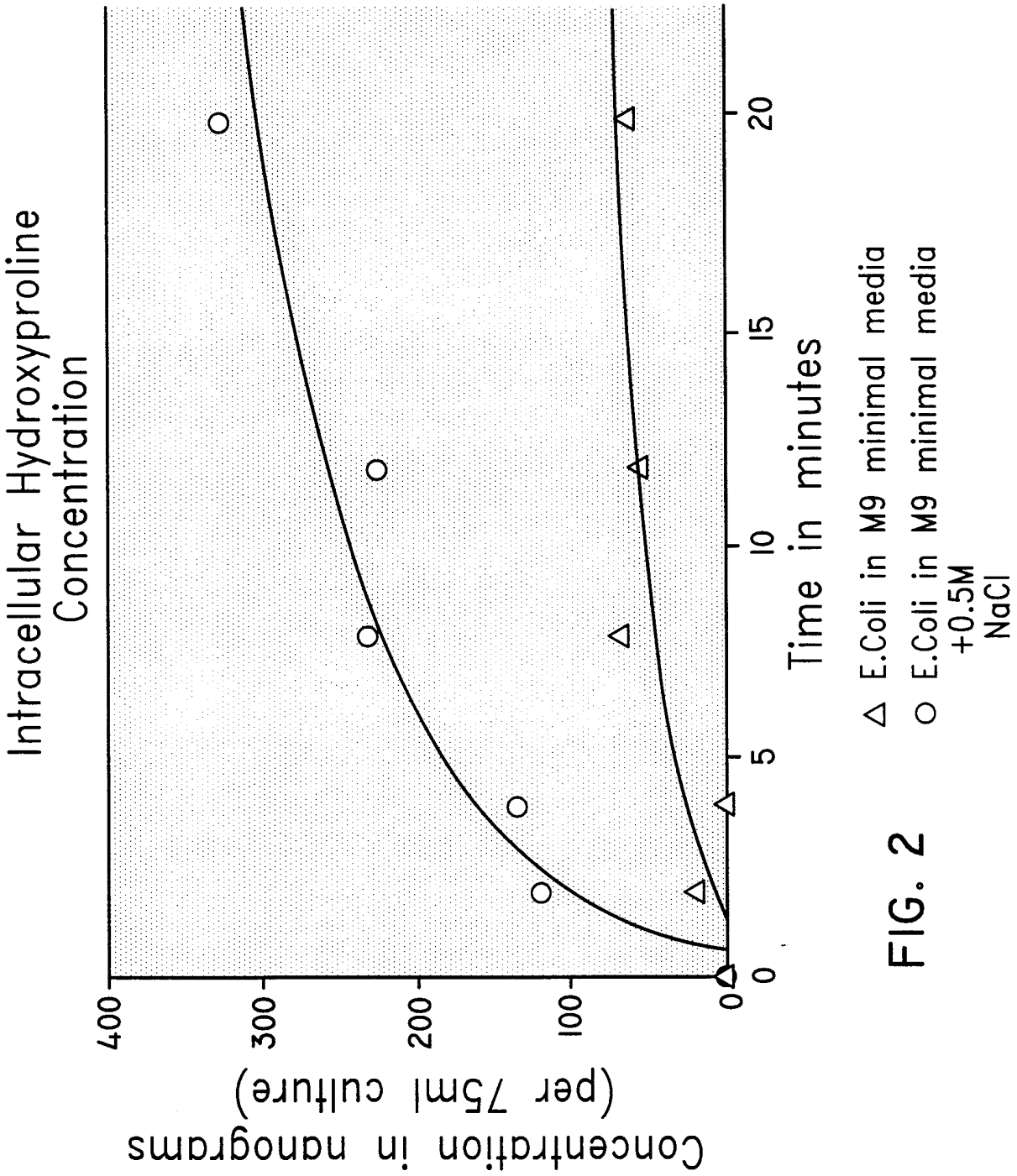

Intracellular Hydroxyproline Concentration
Concentration in nanograms (per 75ml culture)
400
300
200
100
0
0
5
10
15
20
Time in minutes
△ E.Coli in M9 minimal media
○ E.Coli in M9 minimal media +0.5M NaCl

FIG. 2

5'- CAGCTGTCTT ATGGCTATGA TGAGAAATCA ACCGGAGGAA TTTCCGTGCC
TGGCCCCATG GGTCCCTCTG GTCCTCGTGG TCTCCCTGGC CCCCCTGGTG
CACCTGGTCC CCAAGGCTTC CAAGGTCCCC CTGGTGAGCC TGGCGAGCCT
GGAGCTTCAG GTCCCATGGG TCCCCGAGGT CCCCCAGGTC CCCCTGGAAA
GAATGGAGAT GATGGGGAAG CTGGAAAACC TGGTCGTCCT GGTGAGCGTG
GGCCTCCTGG GCCTCAGGGT GCTCGAGGAT TGCCCGGAAC AGCTGGCCTC
CCTGGAATGA AGGGACACAG AGGTTTCAGT GGTTTGGATG GTGCCAAGGG
AGATGCTGGT CCTGCTGGTC CTAAGGGTGA GCCTGGCAGC CCTGGTGAAA
ATGGAGCTCC TGGTCAGATG GGCCCCCGTG GCCTGCCTGG TGAGAGAGGT
CGCCCTGGAG CCCCTGGCCC TGCTGGTGCT CGTGGAAATG ATGGTGCTAC
TGGTGCTGCC GGGCCCCCTG GTCCCACCGG CCCCGCTGGT CCTCCTGGCT
TCCCTGGTGC TGTTGGTGCT AAGGGTGAAG CTGGTCCCCA AGGGCCCCGA
GGCTCTGAAG GTCCCCAGGG TGTGCGTGGT GAGCCTGGCC CCCCTGGCCC
TGCTGGTGCT GCTGGCCCTG CTGGAAACCC TGGTGCTGAT GGACAGCCTG
GTGCTAAAGG TGCCAATGGT GCTCCTGGTA TTGCTGGTGC TCCTGGCTTC
CCTGGTGCCC GAGGCCCCTC TGGACCCCAG GGCCCCGGCG GCCCTCCTGG
TCCCAAGGGT AACAGCGGTG AACCTGGTGC TCCTGGCAGC AAAGGAGACA
CTGGTGCTAA GGGAGAGCCT GGCCCTGTTG GTGTTCAAGG ACCCCCTGGC
CCTGCTGGAG AGGAAGGAAA GCGAGGAGCT CGAGGTGAAC CCGGACCCAC
TGGCCTGCCC GGACCCCCTG GCGAGCGTGG TGGACCTGGT AGCCGTGGTT
TCCCTGGCCC AGATGGTGTT GCTGGTCCCA AGGGTCCCGC TGGTGAACGT
GGTTCTCCTG GCCCCGCTGG CCCCAAAGGA TCTCCTGGTG AAGCTGGTCG
TCCCGGTGAA GCTGGTCTGC CTGGTGCCAA GGGTCTGACT GGAAGCCCTG
GCAGCCCTGG TCCTGATGGC AAAACTGGCC CCCCTGGTCC CGCCGGTCAA
GATGGTCGCC CCGGACCCCC AGGCCCACCT GGTGCCCGTG GTCAGGCTGG
TGTGATGGGA TTCCCTGGAC CTAAAGGTGC TGCTGGAGAG CCCGGCAAGG
CTGGAGAGCG AGGTGTTCCC GGACCCCCTG GCGCTGTCGG TCCTGCTGGC
AAAGATGGAG AGGCTGGAGC TCAGGGACCC CCTGGCCCTG CTGGTCCCGC
TGGCGAGAGA GGTGAACAAG GCCCTGCTGG CTCCCCCGGA TTCCAGGGTC
TCCCTGGTCC TGCTGGTCCT CCAGGTGAAG CAGGCAAACC TGGTGAACAG
GGTGTTCCTG GAGACCTTGG CGCCCCTGGC CCCTCTGGAG CAAGAGGCGA
GAGAGGTTTC CCTGGCGAGC GTGGTGTGCA AGGTCCCCCT GGTCCTGCTG
GACCCCGAGG GGCCAACGGT GCTCCCGGCA ACGATGGTGC TAAGGGTGAT
GCTGGTGCCC CTGGAGCTCC CGGTAGCCAG GGCGCCCCTG GCCTTCAGGG
AATGCCTGGT GAACGTGGTG CAGCTGGTCT TCCAGGGCCT AAGGGTGACA
GAGGTGATGC TGGTCCCAAA GGTGCTGATG GCTCTCCTGG CAAAGATGGC

GTCCGTGGTC TGACCGGCCC CATTGGTCCT CCTGGCCCTG CTGGTGCCCC
TGGTGACAAG GGTGAAAGTG GTCCCAGCGG CCCTGCTGGT CCCACTGGAG
CTCGTGGTGC CCCCGGAGAC CGTGGTGAGC CTGGTCCCCC CGGCCCTGCT
GGCTTTGCTG GCCCCCCTGG TGCTGACGGC CAACCTGGTG CTAAAGGCGA
ACCTGGTGAT GCTGGTGCCA AAGGCGATGC TGGTCCCCCT GGGCCTGCCG
GACCCGCTGG ACCCCCTGGC CCCATTGGTA ATGTTGGTGC TCCTGGAGCC
AAAGGTGCTC GCGGCAGCGC TGGTCCCCCT GGTGCTACTG GTTTCCCTGG
TGCTGCTGGC CGAGTCGGTC CTCCTGGCCC CTCTGGAAAT GCTGGACCCC
CTGGCCCTCC TGGTCCTGCT GGCAAAGAAG GCGGCAAAGG TCCCCGTGGT
GAGACTGGCC CTGCTGGACG TCCTGGTGAA GTTGGTCCCC CTGGTCCCCC
TGGCCCTGCT GGCGAGAAAG GATCCCCTGG TGCTGATGGT CCTGCTGGTG
CTCCTGGTAC TCCCGGGCCT CAAGGTATTG CTGGACAGCG TGGTGTGGTC
GGCCTGCCTG GTCAGAGAGG AGAGAGAGGC TTCCCTGGTC TTCCTGGCCC
CTCTGGTGAA CCTGGCAAAC AAGGTCCCTC TGGAGCAAGT GGTGAACGTG
GTCCCCCCGG TCCCATGGGC CCCCCTGGAT TGGCTGGACC CCCTGGTGAA
TCTGGACGTG AGGGGGCTCC TGCTGCCGAA GGTTCCCCTG GACGAGACGG
TTCTCCTGGC GCCAAGGGTG ACCGTGGTGA GACCGGCCCC GCTGGACCCC
CTGGTGCTCC TGGTGCTCCT GGTGCCCCTG GCCCCGTTGG CCCTGCTGGC
AAGAGTGGTG ATGGTGGTGA GACTGGTCCT GCTGGTCCCG CCGGTCCCGT
CGGCCCCGCT GGCGCCCGTG GCCCCGCCGG ACCCCAAGGC CCCCGTGGTG
ACAAGGGTGA GACAGGCGAA CAGGGCGACA GAGGCATAAA GGGTCACCGT
GGCTTCTCTG GCCTCCAGGG TCCCCCTGGC CCTCCTGGCT CTCCTGGTGA
ACAAGGTCCC TCTGGAGCCT CTGGTCCTGC TGGTCCCCGA GGTCCCCCTG
GCTCTGCTGG TGCTCCTGGC AAAGATGGAC TCAACGGTCT CCCTGGCCCC
ATTGGGCCCC CTGGTCCTCG CGGTCGCACT GGTGATGCTG GTCCTGTTGG
TCCCCCCGGC CCTCCTGGAC CTCCTGGTCC CCCTGGTCCT CCCAGCGCTG
GTTTCGACTT CAGCTTCCTC CCCCAGCCAC CTCAAGAGAA GGCTCACGAT
GGTGGCCGCT ACTACCGGGC T-3'

FIG. 3A

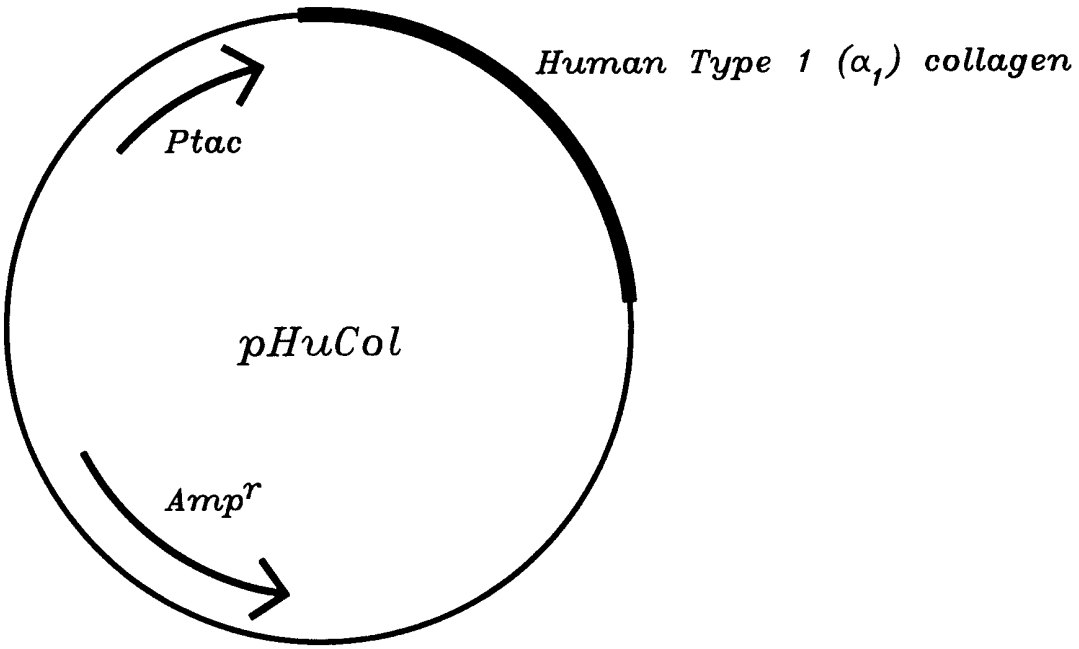


FIG. 4

5'- CAGCTGTCTT ATGGCTATGA TGAGAAATCA ACCGGAGGAA TTTCCGTGCC
TGGCCCCATG GGTCCCTCTG GTCCTCGTGG TCTCCCTGGC CCCCCTGGTG
CACCTGGTCC CCAAGGCTTC CAAGGTCCCC CTGGTGAGCC TGGCGAGCCT
GGAGCTTCAG GTCCCATGGG TCCCCGAGGT CCCCCAGGTC CCCCTGGAAA
GAATGGAGAT GATGGGGAAG CTGGAAAACC TGGTCGTCCT-3'

FIG. 5

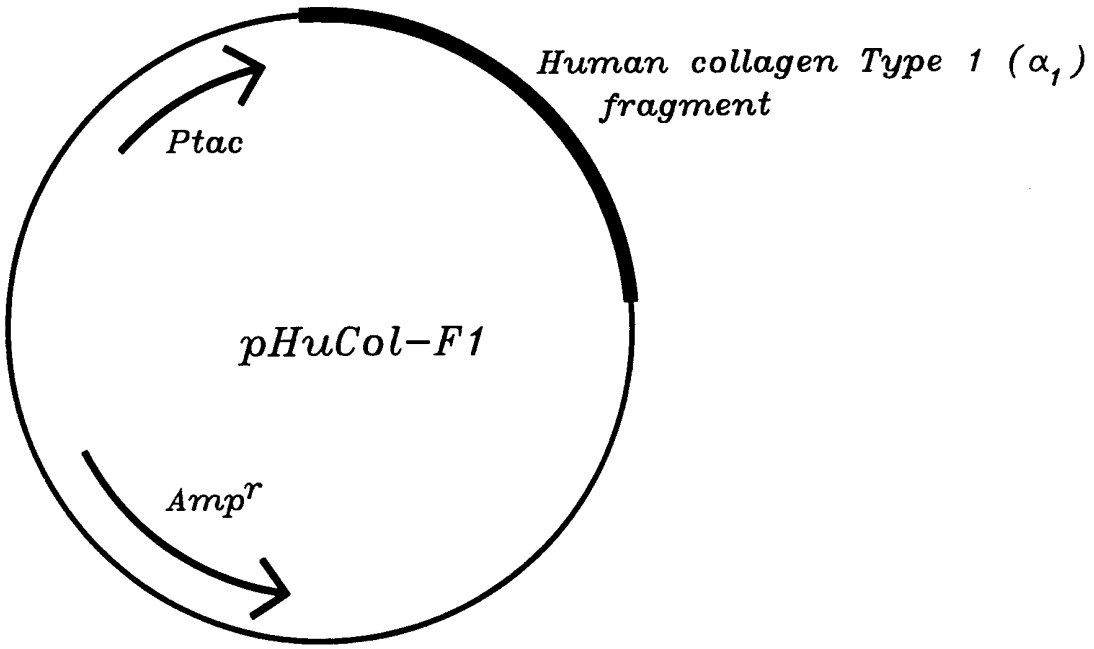


FIG. 6

GGA TCC ATG GGG CTC GCT GGC CCA CCG GGC GAA CCG GGT CCG CCA GGC CCG AAA GGT CCG CGT GGC GAT AGC GGG CTC CCG GGC GAT TCC TAA TGG ATC C

FIG. 7

Gly-Leu-Ala-Gly-Pro-Pro-Gly-Glu-Pro-Gly-Pro-Pro-Gly-Pro-Lys-Gly-Pro-Arg-Gly-Asp-Ser

FIG. 8

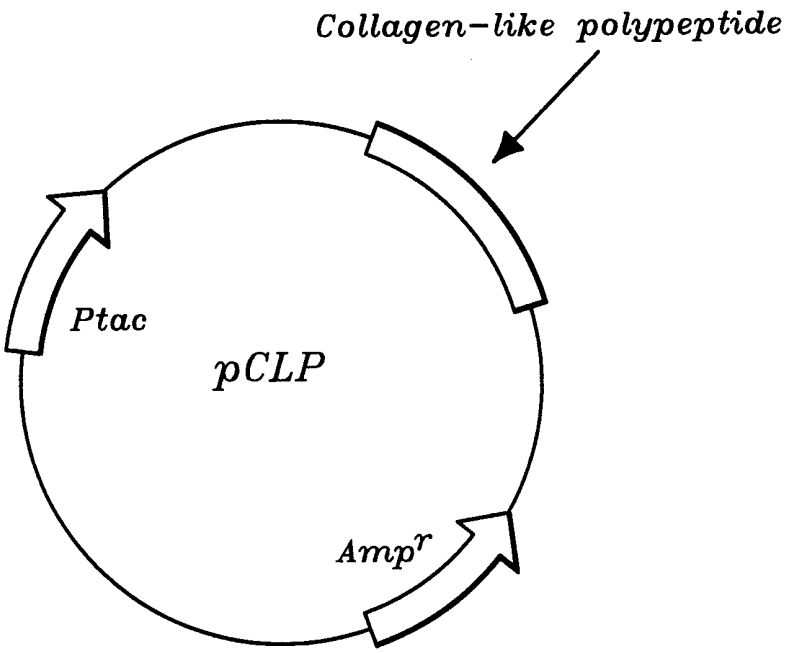


FIG. 9

```
5'- CAGCGGGCCA GGAAGAAGAA TAAGAACTGC CGGCGCCACT CGCTCTATGT
    GGACTTCAGC GATGTGGGCT GGAATGACTG GATTGTGGCC CCACCAGGCT
    ACCAGGCCTT CTACTGCCAT GGGGACTGCC CCTTTCCACT GGCTGACCAC
    CTCAACTCAA CCAACCATGC CATTGTGCAG ACCCTGGTCA ATTCTGTCAA
    TTCCAGTATC CCCAAAGCCT GTTGTGTGCC CACTGAACTG AGTGCCATCT
    CCATGCTGTA CCTGGATGAG TATGATAAGG TGGTACTGAA AAATTATCAG
    GAGATGGTAG TAGAGGGATG TGGGTGCCGC      -3'
```

FIG.10

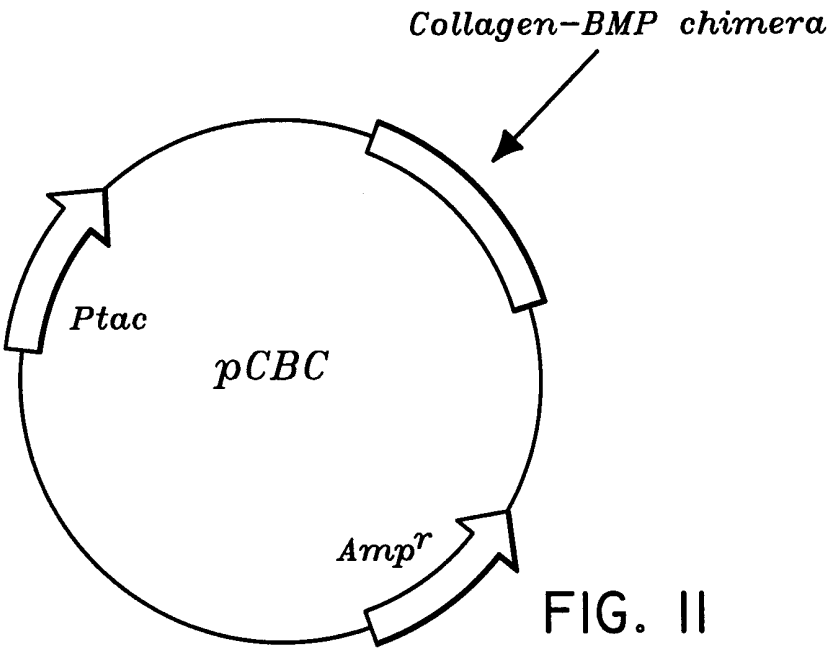

Collagen–BMP chimera
Ptac
pCBC
Amp^r

FIG. 11

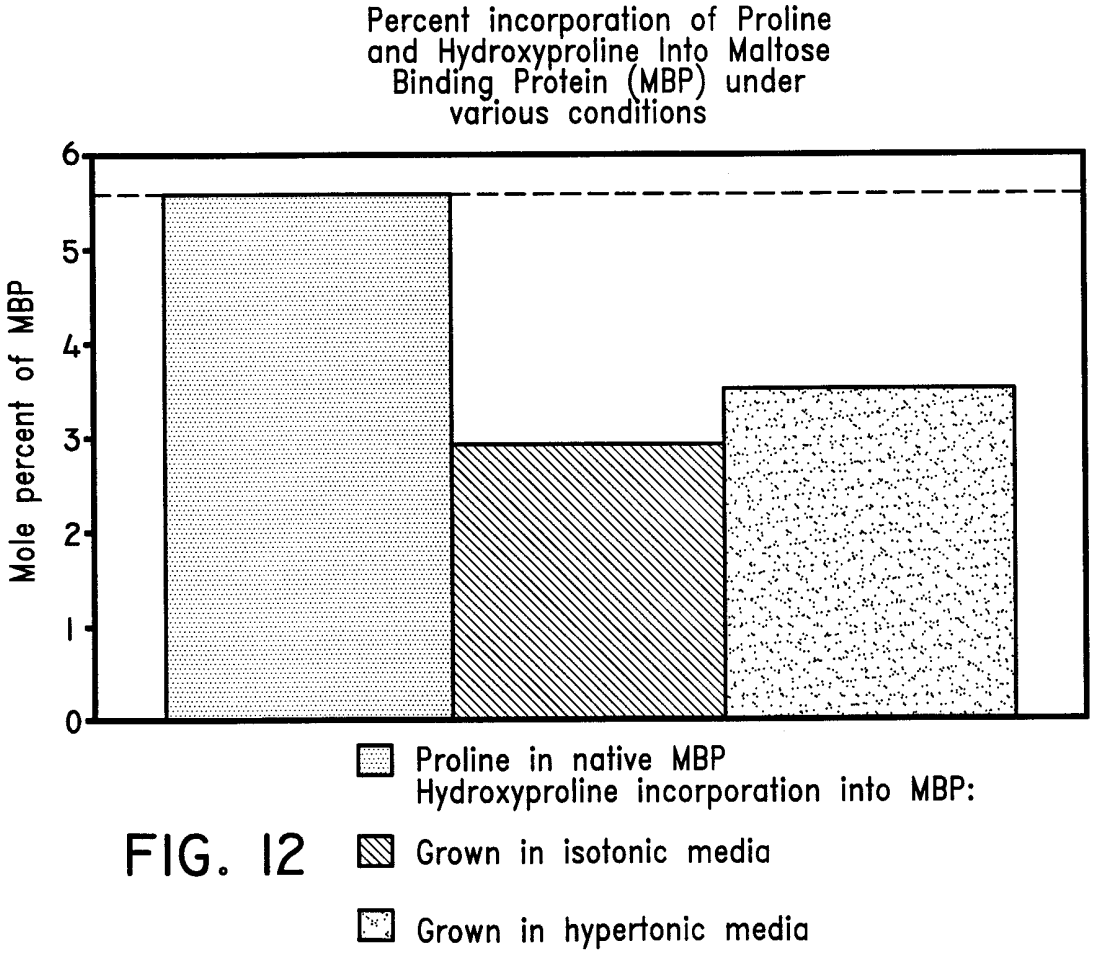

Percent incorporation of Proline and Hydroxyproline Into Maltose Binding Protein (MBP) under various conditions
Mole percent of MBP
0
1
2
3
4
5
6
Proline in native MBP
Hydroxyproline incorporation into MBP:
Grown in isotonic media
Grown in hypertonic media

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10272

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 5/00, 1/00
US CL :435/70.1, 243, 404

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/70.1, 243, 404

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 4,066,658 A (FELIX) 03 January 1978, see entire document.	1-22
Y	US 4,894,441 A (MENICAGLI) 16 January 1990, see entire document.	1-22

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"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
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"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
"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)	"&"	document member of the same patent family
"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		

Date of the actual completion of the international search 19 JUNE 1998	Date of mailing of the international search report 16 JUL 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer L. BLAINE LANKFORD Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★