(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 August 2012 (30.08.2012)

(10) International Publication Number
WO 2012/115641 A1

(51) International Patent Classification:
C07K 14/62 (2006.01) *C12P 21/02* (2006.01)
C12N 15/00 (2006.01) *A61P 3/10* (2006.01)

(21) International Application Number:
PCT/US2011/025934

(22) International Filing Date:
23 February 2011 (23.02.2011)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **ELONA BIOTECHNOLOGIES** [US/US]; 1040 Sierra Drive, Suite 1000, Greenwood, Indiana 46143 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZIMMERMAN, Ronald, E.** [US/US]; 4465 W. Stones Crossing Road, Greenwood, Indiana 46143 (US). **STOKELL, David, John** [US/US]; 5513 Floating Leaf Drive, Indianapolis, Indiana 46237 (US). **AKERS, Michael, Patrick** [US/US]; 517 Kessler Blvd. E Dr., Indianapolis, Indiana 46220 (US).

(74) Agents: **MAYFIELD, Denise, L.** et al.; Katten Muchin Rosenman LLP, 2900 K Street, NW, Suite 200, Washington, DC 20007-5118 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: LIS-PRO PROINSULIN COMPOSITIONS AND METHODS OF PRODUCING LIS-PRO INSULIN ANALOGS THEREFROM

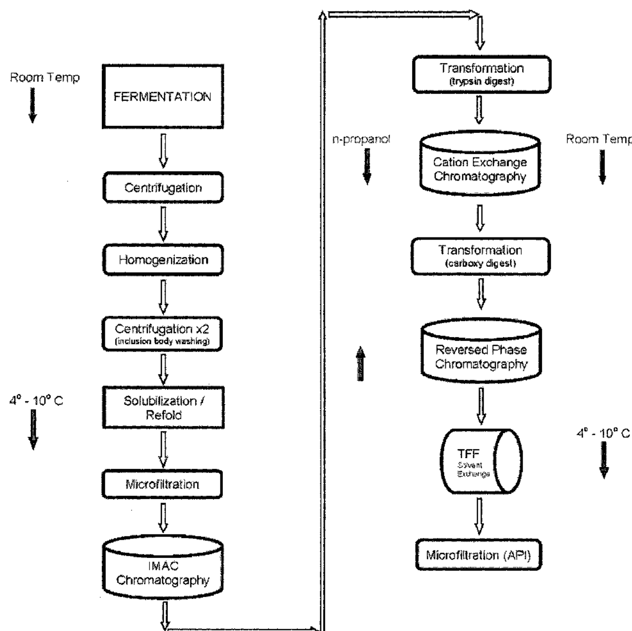


Figure 2a: Production Block Flow Diagram for Insulin Lispro

(57) Abstract: Lis-Pro modified proinsulin sequences that have a modified C-peptide amino acid and/or nucleic acid modification are presented. Methods for producing Lis-Pro insulin analogs are also disclosed. Highly efficient processes for preparing the Lis-Pro insulin analogs and improved preparations containing the Lis-Pro insulin analogs according to the methods described herein are also provided.

WO 2012/115641 A1

Published:

— *with international search report (Art. 21(3))*

LIS-PRO PROINSULIN COMPOSITIONS AND METHODS OF PRODUCING LIS-PRO INSULIN ANALOGS THEREFROM

Field of the Invention

[0001] The invention relates to compositions and preparations that comprise Lis-Pro proinsulin, and in particular Lis-Pro proinsulin with modified C-peptide sequences. The invention also relates to methods of manufacture for manufacturing Lis-Pro insulin analogs from modified proinsulin sequences.

Background

[0002] Insulin is a hormone that regulates glucose metabolism in animals. Insulin is a polypeptide hormone secreted by beta-cells of the pancreas. This hormone is made up of two polypeptide chains, an A-chain of 21 amino acids, and a B-chain of 30 amino acids. These two chains are linked to one another in the mature form of the hormone by two interchain disulphide bridges. The A-chain also features one intra-chain disulphide bridge.

[0003] Insulin analogs are altered forms of native insulin that are available to the body for performing the same action as native insulin. A specific insulin analog known as Lis-Pro insulin has also been described in US Pat. Nos. 5,474,978 and 5,504,188. This analog is used in the treatment of diabetes. Lis-Pro insulin is characterized as a short acting insulin analog, which, when combined with an insulin pump, allows for better blood glucose stability without the risk of hyperglycemia. This Lis-Pro insulin analog has been available commercially as HUMALOG® (Eli Lilly). HUMALOG® is an insulin analog wherein the molecule includes a Lys(B_{28})-Pro(B_{29}) amino acid sequence in place of the native insulin Pro(B_{28})-Lys(B_{29}). HUMALOG® is an injectable, fast-acting insulin.

[0004] Native insulin is a hormone that is synthesized in the body in the form of a single-chain precursor molecule, proinsulin. Proinsulin is a molecule comprised of a prepeptide of 24 amino acids, followed by the B-chain peptide, a C-peptide of 35 amino acids, and the A-chain peptide. The C-peptide of this precursor insulin molecule ("proinsulin") contains the two amino acids, lysine-arginine (KR) at its carboxy end (where it attaches to the A-chain), and the two amino acids, arginine-arginine (RR) at its amino end (where it attaches to the B-chain). In the mature

insulin molecule, the C-peptide is cleaved away from the peptide so as to leave the A-chain and the B-chain connected directly to one another in its active form.

[0005] Molecular biology techniques have been used to produce human proinsulin. In this regard, three major methods have been used for the production of this molecule. Two of these methods involve *Escherichia coli,* with either the expression of a large fusion protein in the cytoplasm (Chance *et al.* (1981), and Frank *et al.* (1981) in Peptides: Proceedings of the 7th American Peptide Chemistry Symposium (Rich, D. and Gross, E., eds.), pp. 721-728, 729-739, respectively, Pierce Chemical Company, Rockford, IL), or the use of a signal peptide to enable secretion into the periplasmic space (Chan *et al.* (1981) P.N.A.S., USA., 78:5401-5404). A third method utilizes yeast, especially *Saccharomyces cerevisiae*, to secrete the insulin precursor into the medium (Thim, *et al.* (1986), P.N.A.S., USA., 83: 6766-6770).

[0006] Chance *et al.* reported a process for preparing insulin by producing each of the A and B chains of insulin in the form of a fusion protein by culturing *E. coli* that carries a vector compromising a DNA encoding the fusion protein, cleaving the fusion protein with cyanogen bromide to obtain the A and the B chains, sulfonating the A and B chains to obtain sulfonated chains, reacting the sulfonated B chain with an excess amount of the sulfonated A chain; and then purifying the resultant products to obtain insulin. US Pat. No. 5,700,662 describes recombinant processes for producing various insulin analogs, including using the process of Chance *et al.* to produce Lis-Pro insulin analogs.

[0007] Drawbacks associated with this process are that it requires two fermentation processes and the requirement of a reaction step for preparing the sulfonated A chain and the sulfonated B chain. This results in a low insulin yield.

[0008] Frank *et al.* described a process for preparing insulin in the form of a fusion protein in *E. coli.* In this process, proinsulin is produced in the form of a fusion protein by culturing *E. coli* which carries a vector comprising a nucleic acid sequence (DNA) encoding for the fusion protein, cutting the fusion protein with cyanogen bromide to obtain proinsulin, sulfonating the proinsulin and separation of the sulfonated proinsulin, refolding the sulfonated proinsulin to form correct disulfide bonds, treating the refolded proinsulin with trypsin and carboxypeptidase B, and then purifying the resultant product to obtain insulin. However, the yield of the refolded proinsulin having correctly folded disulfide bonds is reported to sharply decrease as the concentration of the proinsulin increases. This is allegedly due to, at least among other reasons,

misfolding of the protein, and some degree of polymerization being involved. Hence, the process entails the inconvenience of using laborious purification steps during the recovery of proinsulin.

[0009] Thim *et al.* reported a process for producing insulin in yeast, *Saccharomyces cerevisiae*. This process has the steps of producing a single chain insulin analog having a certain amino acid sequence by culturing *Saccharomyces cerevisiae* cells, and isolating insulin therefrom through the steps of: purification, enzyme reaction, acid hydrolysis and a second purification. This process, however, results in an unacceptably low yield of insulin.

[0010] The role of the native C-peptide in the folding of proinsulin is not precisely known. The dibasic terminal amino acid sequence at both ends of the C-peptide sequence has been considered necessary to preserve the proper processing and/or folding of the proinsulin molecule to insulin.

[0011] Other amino acids within the C-peptide sequence, however, have been modified with some success. For example, Chang *et al.* (1998) (Biochem. J., 329:631-635) described a shortened C-peptide of a five (5) amino acid length, -YPGDV- (SEQ ID NO: 1), that includes a preserved terminal di-basic amino acid sequence, RR at one terminal end, and KR at the other terminal end, of the peptide. Preservation of the dibasic amino acid residues at the B-chain-C peptide and C-peptide-A-chain junctures is taught as being a minimal requirement for retaining the capacity for converting the proinsulin molecule into a properly folded mature insulin protein. The production of the recombinant human insulin was described using *E. coli* with a shortened C-peptide having a dibasic amino acid terminal sequence. US Pat. No. 5,962,267 also describes dibasic terminal amino acid sequences at both ends of the C-peptide.

[0012] One of the difficulties and/or inefficiencies associated with the production of recombinant insulin employing a proinsulin construct having the conserved, terminal di-basic amino acid sequence in the C-peptide region is the presence of impurities, such as Arg(A_0)-insulin, in the reaction mixture, once enzymatic cleavage to remove the C-peptide is performed. This occurs as a result of misdirected cleavage of the proinsulin molecule so as to cleave the C-peptide sequence away from the A-chain at this juncture, by the action of trypsin. Trypsin is a typical serine protease, and hydrolyses a protein or peptide at the carboxyl terminal of an arginine or lysine residue (*Enzymes*, pp. 261-262 (1979), ed. Dixon, M. & Webb, E. C. Longman Group Ltd., London). This unwanted hydrolysis results in the unwanted Arg(A_0)-insulin by-product, and typically constitutes about 10% of the reaction yield. Hence, an additional

purification step is required. The necessity of an additional purification step makes the process much more time consuming, and thus expensive, to use. Moreover, an additional loss of yield may be expected from the necessity of this additional purification step.

[0013] Others have described the use of proinsulin constructs that do not have a conserved terminal dibasic amino acid sequence of the C-peptide region. For example, US Pat. No. 6,777,207 (Kjeldsen *et al.*) relates to a novel proinsulin peptide construct containing a shortened C-peptide that includes the two terminal amino acids, glycine-arginine or glycine-lysine at the carboxyl terminal end that connects to the A-chain of the peptide. The B-chain of the proinsulin construct described therein has a length of 29 amino acids, in contrast to the native 30 amino acid length of the native B-chain in human insulin. The potential effects of this change to the native amino acid sequence of the B-chain in the human insulin produced are yet unknown. Methods of producing insulin using these proinsulin constructs in yeast are also described. Inefficiencies associated with correct folding of the mature insulin molecule when yeast is used as the expression host, render this process, among other things, inefficient and more expensive and time consuming to use. In addition, yeast provides a relatively low insulin yield, due to the intrinsically low expression levels of a yeast system as compared to *E. coli.*

[0014] As evidenced from the above review, a present need exists for a more efficient process for production of Lis-Pro insulin analogs that is efficient and that may also improve and/or preserve acceptable production yield requirements of the pharmaceutical industry.

[0015] The above references are incorporated by reference herein where appropriate for appropriate teachings of additional or alternative details, features and/or technical background.

Summary of the Invention

[0016] The present invention provides processes for using a modified Lis-Pro proinsulin sequence to produce Lis-Pro insulin analogs. The modified Lis-Pro proinsulin sequence has the formula

$$R_1\text{-}(B_1\text{-}B_{27})\text{-}B_{28}\text{-}B_{29}\text{-}B_{30}\text{-}R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5\text{-}(A_1\text{-}A_{21})\text{-}R_6 \qquad \text{Formula I}$$

[0017] wherein

[0018] R_1 is a tag sequence containing one or more amino acids, preferably with a C-terminal Arg or Lys, or R_1 is absent with an Arg or Lys present prior to the start of the B chain;

[0019] $(B_1\text{-}B_{27})$ and $(A_1\text{-}A_{21})$ comprise amino acid sequences of native human insulin;

[0020] B_{28} is any amino acid other than Pro, preferably B_{28} is Lys;

[0021] B_{29} is any amino acid other than Lys or Arg, preferably B_{29} is Pro;

[0022] B_{30} is Thr;

[0023] R_2, R_3 and R_5 are Arg;

[0024] R_4 is any amino acid other than Gly, Lys or Arg or is absent, preferably Ala;

[0025] X is a sequence comprises one or more amino acids or is absent, provided that X is not EAEALQVGQVELGGGPGAGSLQPLALEGSLQ (SEQ ID NO: 2) and X does not comprise a C-terminal Gly, Lys, or Arg when R_4 is absent; and

[0026] R_6 is a tag sequence containing one or more amino acids, preferably with a N-terminal Arg or Lys, or R_6 is absent.

[0027] In one aspect of the present invention is related to a process for producing Lis-Pro insulin analogs comprising the steps of culturing *E. coli* cells under conditions suitable for expression of a modified Lis-Pro proinsulin sequence of Formula I, disrupting the cultured *E. coli* cells to provide a composition comprising inclusion bodies containing the modified Lis-Pro proinsulin sequence, solubilizing the composition of inclusion bodies, and recovering Lis-Pro insulin analogs from the solubilized composition.

[0028] Another aspect of the present invention is related to a process for producing Lis-Pro insulin analogs comprising the steps of providing a modified Lis-Pro proinsulin sequence of Formula I, folding the modified Lis-Pro proinsulin sequence to provide a Lis-Pro proinsulin derivative peptide, enzymatically cleaving the Lis-Pro proinsulin derivative peptide to remove a connecting peptide and provide an intermediate solution comprising Lis-Pro insulin analog, and

purifying the intermediate solution in a chromatography column wherein the Lis-Pro insulin analog is eluted using a buffer comprising n-propanol.

[0029] Another aspect of the present invention is related to a process for producing Lis-Pro insulin analogs comprising the steps of culturing transformed *E. coli* cells having modified Lis-Pro proinsulin sequence of Formula I under conditions suitable for expression of the modified Lis-Pro proinsulin sequence, disrupting said cultured *E. coli* cells to provide a composition comprising inclusion bodies containing the modified Lis-Pro proinsulin sequence, solubilizing said composition of inclusion bodies by adjusting the pH to at least 10.5, folding the modified Lis-Pro proinsulin sequence to provide a Lis-Pro proinsulin derivative peptide, enzymatically cleaving the Lis-Pro proinsulin derivative peptide to remove a connecting peptide and provide an intermediate solution comprising Lis-Pro insulin analog, and purifying the intermediate solution in a chromatography column. Preferably, the Lis-Pro insulin analog is eluted using a buffer comprising n-propanol.

[0030] Additional advantages, objects, and features of the invention will be set forth in part in the description which follows, and in part will become apparent to those having ordinary skill in the art upon examination of the following or may be learned from practice of the invention. The objects and advantages of the invention may be realized and attained as particularly pointed out in the appended claims.

Brief Description of the Figures

[0031] The invention will be described in detail with reference to the following drawings in which like reference numerals refer to like elements wherein.

[0032] FIG. 1, according to one aspect of the invention, is a vector map of plasmid pTrcHis2A (Kan) with a Lis-Pro proinsulin gene insert.

[0033] FIG. 2a and FIG 2b, according to some aspects to the invention, present process flow schemes for the purification of Lis-Pro insulin analogs.

Detailed Description

[0034] The present invention generally relates to the preparation of insulin analogs, specifically Lis-Pro insulin analog, from modified Lis-Pro proinsulin sequences. Lis-Pro insulin analog comprises a modified B-chain having Lys(B_{28}) and Pro(B_{29}). Modified Lis-Pro proinsulin

sequences refer to a single-chain polypeptide that may be converted into human insulin or insulin analogs and comprise a connecting peptide (C-peptide) having at least one non-dibasic terminal amino acid sequence. In one embodiment, non-dibasic terminal amino acid sequences may comprise any amino acid except Lys or Arg-Arg ((any except R or K)R), and more preferably any amino acid except Gly, Lys, or Arg-Arg ((any except G, R, or K)R). In one embodiment, the terminal amino acid sequence may comprise Ala-Arg. Advantageously, the positioning of these particular terminal amino acids in the C-peptide provides for an improved method for producing recombinant Lis-Pro insulin analog, having fewer steps, improved yields of the recombinant Lis-Pro insulin analog and less contaminating byproducts.

[0035] The process for producing Lis-Pro insulin analogs of the invention presents many advantages, among them the advantage of reducing and/or eliminating the presence of unwanted and contaminating cleavage by-products characteristic of conventional manufacturing processes for producing recombinant human Lis-Pro insulin in *E. coli.* Previously undesirable by-products evident in yield mixtures using conventional methods of producing recombinant human insulin analogs included, by way of example, the production of an unwanted cleavage product, Arg(A_0)-insulin analogs. A highly efficient process for the production of recombinant human insulin analogs is presented that reduces and/or eliminates the presence of this and other unwanted and undesirable cleavage by-products, and that further presents the advantages of eliminating several time consuming, expensive, purification steps. A process having fewer technician-assisted steps is thus devised, and illustrates the additional advantage of eliminating the degree of inconsistency and/or error associated with technician assisted steps in the manufacturing process.

[0036] In one embodiment, the modified Lis-Pro proinsulin sequence of the present invention has a formula I:

$$R_1\text{-}(B_1\text{-}B_{27})\text{-}B_{28}\text{-}B_{29}\text{-}B_{30}\text{-}R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5\text{-}(A_1\text{-}A_{21})\text{-}R_6 \qquad \text{Formula I}$$

[0037] wherein

[0038] R_1 is a tag sequence containing one or more amino acids, preferably with a C-terminal Arg or Lys, or R_1 is absent with an Arg or Lys present prior to the start of the B chain;

[0039] (B_1-B_{27}) and (A_1-A_{21}) comprise amino acid sequences of native human insulin;

[0040] B_{28} is any amino acid other than Pro, preferably B_{28} is Lys;

[0041] B_{29} is any amino acid other than Lys or Arg, preferably B_{29} is Pro;

[0042] B_{30} is Thr;

[0043] R_2, R_3 and R_5 are Arg;

[0044] R_4 is any amino acid other than Gly, Lys or Arg or is absent, preferably Ala;

[0045] X is a sequence comprising one or more amino acids, or is absent, provided that X is not EAEALQVGQVELGGGPGAGSLQPLALEGSLQ (SEQ ID NO: 2), and X does not comprise a C-terminal Gly, Lys, or Arg when R_4 is absent; and

[0046] R_6 is a tag sequence containing one or more amino acids, preferably with a N-terminal Arg or Lys, or R_6 is absent.

[0047] R_1 or R_6 in the modified proinsulin of Formula I comprises a pre or post-peptide that may be a native pre-peptide or an N-terminal multiple His-tag sequence, or any other commercially available tag utilized for protein purification, e.g. DSBC, Sumo, Thioredein, T7, S tag, Flag Tag, HA tag, VS epitope, Pel B tag, Xpress epitope, GST, MBP, NusA, CBP, or GFP. In one embodiment, at least one of R_1 or R_6 is present in Formula I. In some embodiments, the terminal amino acid of the pre or post-peptide that connects to the B-chain or A-chain comprises Arg or Lys. Native pre-peptide has the sequence of MALWMRLLPLLALLALWGPDPAAA (SEQ ID NO: 3). In some embodiments, the N-terminal multiple His-tagged Lis-Pro proinsulin construct comprises a 6-histidine (SEQ ID NO: 20) N-terminal tag and may have the sequence of MHHHHHHGGR (SEQ ID NO: 4). The modified Lis-Pro proinsulin sequence may replace the native 24 amino acid pre-peptide with the 6-histidine (SEQ ID NO: 20) N-terminal tag sequence. In some embodiments, R_1 and/or R_6 may be a sequence of one or more amino acids, e.g., preferably from 1 to 30 and more preferably from 6 to 10.

[0048] Native insulin comprises an A-chain having the sequence GIVEQCCTSICSLYQLENYCN (SEQ ID NO: 5) and a B-chain having the sequence FVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO: 6). According to the invention, the B-chain of Formula I is modified from native insulin and contains at least one amino acid mutation, substitution, deletion, insertion, and/or addition. For Lis-Pro insulin analogs, preferably B_{28} and B_{29} of the B-chain are modified. The lysine B_{29} of native insulin is substituted with proline and the proline B_{28} of native insulin is substituted with lysine. In some embodiments, the B-chain that is modified is human insulin B-chain. In another embodiment, the B-chain that is modified is porcine insulin B-chain.

[0049] As used in the description of the present invention, the term "connecting peptide" or "C-peptide" is meant the connecting moiety "C" of the B-C-A polypeptide sequence of a single chain Lis-Pro proinsulin molecule. As in the native human proinsulin, the N-terminus of the C-peptide connects to C-terminus of the modified B-chain, e.g., position 30 of the B-chain, and the C-terminus of the C-peptide connects to N-terminus of the A-chain, e.g., position 1 of the A-chain.

[0050] In one embodiment, the C-peptide may have a sequence of formula II:

R_2-R_3-X-R_4-R_5 Formula II

[0051] wherein R_2, R_3, R_4, R_5, and X have the same meaning as in Formula I. In one embodiment, X may be a sequence having up to 40 amino acids, preferably up to 35 amino acids or more preferably up to 30 amino acids.

[0052] The C-peptide sequences of the present invention may include:

[0053] (1) RREAEDLQVGQVELGGGPGAGSLQPLALEGSLQAR (SEQ ID NO: 7);

[0054] (2) RREAEDLQVGQVGLGGGPGAGSLQPLALEGSLQAR (SEQ ID NO:8);

[0055] (3) RREAEALQVGQVGLGGGPGAGSLQPLALEGSLQAR (SEQ ID NO: 9);

[0056] (4) RREAEDLQVGQVELGGGPGAGSLQPLAIEGSLQAR (SEQ ID NO: 10);

[0057] (5) RREAEDLQVGQVGLGGGPGAGSLQPLAIEGSLQAR (SEQ ID NO: 11); or

[0058] (6) RREAEALQVGQVGLGGGPGAGSLQPLAIEGSLQAR (SEQ ID NO: 12).

[0059] In the above embodiments, where the designation A appears at the terminal end of the C-peptide sequence, AR cannot be replaced with KR or RR.

[0060] Preferred modified Lis-Pro proinsulin sequences of the present invention may include:

[0061] (1)FVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 13);

[0062] (2)MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 14);

[0063] (3)MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 15);

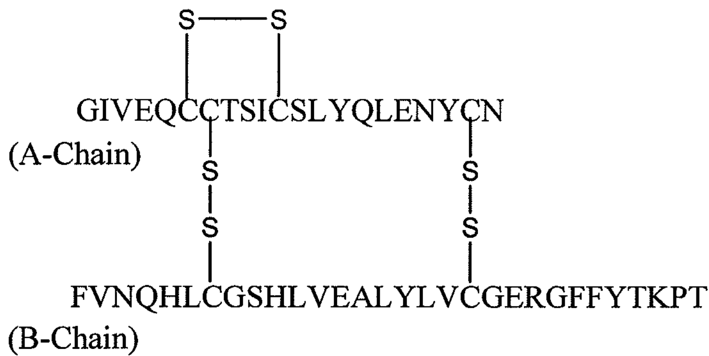
[0064] (4)MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNRHHHHHH (SEQ ID NO: 16);

[0065] (5)MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNKHHHHHH (SEQ ID NO: 21);

[0066] (6)MRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNRHHHHHH (SEQ ID NO: 17); or

[0067] (7)MRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNKHHHHHH (SEQ ID NO: 22).

[0068] The single chain Lis-Pro insulin analogs of the invention will include three (3) correctly positioned, disulphide bridges, as is characteristic of the native human insulin. In some embodiments, the folded modified Lis-Pro proinsulin, or Lis-Pro proinsulin derivative peptide, may include three (3) correctly positioned, disulphide bridges. In production, the C-peptide of the Lis-Pro proinsulin derivative peptide is removed to produce the Lis-Pro insulin analog. Lis-Pro insulin analogs of the invention have a sequence (SEQ ID NOS 5 and 18, respectively, in order of appearance) of formula III, where the disulphide bridges are represented as –S–S–:



Formula III

[0069] The present invention provides modified Lis-Pro proinsulin sequences having the modified C-peptide and methods for using these in a process to provide high yields of mature recombinant Lis-Pro insulin analog. Advantageously, the positioning of these particular terminal amino acids in the C-peptide may provide for an improved method for producing recombinant Lis-Pro insulin analog, having fewer steps, improved yields of the recombinant Lis-Pro insulin analog and less contaminating byproducts.

[0070] As used in the description of the present invention, the terms "insulin precursor" or "proinsulin" are described as a single-chain polypeptide in which, by one or more subsequent chemical and/or enzymatic processes, may be converted into human insulin or insulin analog.

[0071] A proinsulin analog or modified proinsulin is defined as a proinsulin molecule having one or more mutations, substitutions, deletions, and or additions, of the A, B and/or C chains relative to the native human proinsulin nucleic acid sequence. The Lis-Pro proinsulin analogs are preferably such wherein one or more of the naturally occurring nucleic acids have been substituted with another nucleic acid within a triplet encoding for a particular amino acid. For purposes of convenience, proinsulin analog is understood to refer to Lis-Pro proinsulin analog, unless otherwise specified.

[0072] The term "insulin analog" includes insulin molecules having one or more mutations, substitutions, deletions, additions, or modifications to one or more amino acids of a native insulin sequence. For example, in one embodiment, the native insulin sequence is porcine insulin, while in another embodiment, the native insulin sequence is human. For purposes of convenience insulin analog is understood to refer to Lis-Pro insulin analog, unless otherwise specified.

[0073] The term "a" as used in the description of the present invention is intended to mean "one or more", and is used to define both the singular and plural forms of the item or items to which it

references, or to a feature or characteristic to which it refers. The use of the singular or plural in the claims or specification is not intended to be limiting in any way and also includes the alternative form.

[0074] The term "about" is intended to be inclusive of and to encompass both an exact amount as well as an approximate amount or range of values or levels of the item, ingredient, element, activity, or other feature or characteristic to which it references. Generally, and in some embodiments, the term "about" is intended to reference a range of values relatively close to the specific numerical value specifically identified. For example, "about 3 grams to about 5 grams" is intended to encompass a measure of in or around a value of 3 grams, concentration values between 3 grams and 5 grams, concentration values in and around 5 grams, as well as concentration values that are exactly 3 grams and exactly 5 grams.

[0075] As used in the description of the present process, a high protein concentration of the Lis-Pro proinsulin or insulin analog product is defined as a protein yield concentration of at least about 3 grams/liter, or between about 3 grams to about 5 grams per liter. The expression yield to be expected may be defined as a protein/peptide yield that is sufficient to detect via polyacrylamide gel electrophoreses (PAGE).

[0076] The invention provides a processes for producing highly purified Lis-Pro insulin analog that is more efficient than current techniques. The invention in a general and overall sense relates to an improved process for preparing a heterologous recombinant protein in a prokaryotic host cell. This process is characterized in that it employs a recombinant Lis-Pro insulin protein that provides a useful and efficiently processed modified proinsulin sequence analog as described above.

[0077] The term "heterologous protein" is intended to mean that the protein in the prokaryotic host cell is not native, *i.e.*, it occurs as a protein in peculiar or foreign (i.e., not native to) the host prokaryotic cell.

[0078] The term "recombinant" is intended to mean produced or modified by molecular-biological methods. For example, according to one embodiment, a recombinant protein is made using genetic engineering techniques and is not found in nature.

[0079] As used in the description of the present invention, the term "heterologous recombinant protein" is defined as any protein known to the skilled person in the molecular biological arts,

such as, for example, insulin, insulin analog, proinsulin, proinsulin analog, C-peptide, and proteins containing these together with any other protein or peptide fragment.

[0080] Prokaryotic host cells may be any host cells known to the skilled artisan in the molecular biological arts, and by way of example, *Escherichia coli*. Such types of cells available form public collections and useful in the practice of the present invention include, by way of example, the Deutsche Sammlung von Mikrooganismen and Zellkulturen GmbH, raunschweig, Germany, e.g., *E. coli* Strain K12 JM107 (DSM 3950).

[0081] The following reference table, Table 1, provides the triplet codons corresponding to each of the various amino acids that are used in the description of the present invention. As will be understood by those of skill in the art, the amino acid that may be used in any particularly defined position as part of any of the peptide, protein, or constructs otherwise defined herein by reference to a particular nucleotide triplet base pair may be encoded by a number of different nucleotide triplets that function to encode the same amino acid. For example, where the amino acid of the sequence defined herein is alanine (Ala, or A), the triplet codon of nucleic acids that may encode for this amino acid are: GCT, GCC, GCA, or GCG. The following table illustrates this definition of variables at and substitutions as can be applied to all of the naturally occurring amino acids sequences of the disclosure.

Table 1.

	U	C	A	G	
U	UUU, UUC – Phe UUA, UUG – Leu	UCU, UCC, UCA, UCG – Ser	UAU, UAC – Tyr UAA Stop UAG Stop	UGU, UGG – Cys UGA Stop UGG Trp	U C A G
C	CUU, CUC, CUA, CUG – Leu	CCU, CCC, CCA, CCG – Pro	CAU, CAC – His CAA, CAG – Gln	CGU, CGC, CGA, CGG – Arg	U C A G
A	AUU, AUC, AUA – Ile AUG Met	ACU, ACC, ACA, ACG – Thr	AAU, AAC – Asn AAA, AAG – Lys	AGU, AGC – Ser AGA, AGG – Arg	U C A G
G	GUU, GUC, GUA, GUG – Val	GCU, GCC, GCA, GCG – Ala	GAU, GAC – Asp GAA, GAG – Glu	GGU, GGC, GGA, GGG – Gly	U C A G

[0082] It should be understood that process steps within the following description of the method may be modified, changed and/or eliminated, depending on the particular preferences of the processor and/or the particular mechanical apparatus available to the processor, as well as the specific reagents and/or materials available and/or convenience and/or economics of use.

[0083] The Lis-Pro insulin analog prepared by the present invention may be formulated as liquid Lis-Pro insulin analog or crystalline Lis-Pro insulin analog. According to an embodiment of the invention, a preparation of recombinant liquid Lis-Pro insulin analog is in a substantially liquid form and that has not been through a crystallization process. Eliminating these steps has no

negative impact on the purity of the liquid Lis-Pro insulin analog produced, but has the added advantage of reducing the amount of inactive insulin multimers in the liquid Lis-Pro insulin analog of the invention. Lis-Pro insulin analog reconstituted from lyophilized and crystallized insulin may be contaminated with inactive insulin multimers and is less preferred.

[0084] According to one embodiment, the methods of producing Lis-Pro insulin analog described herein generally include the following steps: fermentation/expression, Inclusion body isolation, solubilization of Lis-Pro proinsulin analog; refolding processing and transformation of Lis-Pro proinsulin analog to Lis-Pro insulin analog; and purification of Lis-Pro insulin analog. FIGS. 2A and 2B illustrate flow charts of preferred process steps in producing Lis-Pro insulin analog according to embodiments of the present invention.

[0085] Expression of Lis-Pro proinsulin analog may occur in a recombinant expression system. According to one embodiment, the recombinant expression system is a transformed *E. coli* containing a Lis-Pro proinsulin analog expression vector. For example, the transformed cells may be vertebrate or invertebrate cells, such as prokaryote or eukaryote cells, and most preferably the cells may be mammalian, bacterial, insect, or yeast cells. In one embodiment, the cell is a bacterial cell and in a further embodiment, the bacteria is *E. coli*. In another embodiment, the cell is a yeast cell and in a further embodiment, the yeast cell is *S. cerevisiae or S. pombe.*

[0086] In one embodiment, *E. coli* cells may be cultured and disrupted to provide a composition comprising inclusion bodies. The inclusion bodies contain the modified Lis-Pro proinsulin sequence. The Lis-Pro proinsulin analogs expressed by transformed *E. coli* cells according to the method of the invention may be secreted from the cells and include a secretory sequence. In other embodiments, Lis-Pro proinsulin analogs expressed by transformed cells are not secreted from the cells, and thus do not include a secretory sequence.

[0087] The step of solubilizing of the composition of inclusion bodies may involve adjusting the pH to achieve complete solubilization of the modified Lis-Pro proinsulin sequences. In one embodiment, the inclusion bodies may be solubilized by adjusting the pH to at least 10.5, preferably from 10.5 to 12.5, preferably from 11.8-12. The pH may be adjusted by adding an alkali hydroxide such as NaOH or KOH to the composition of inclusion bodies. In addition, the step of solubilization may use one or more reducing agents and/or chaotropic agent. Suitable reducing agents may include those selected from the group consisting of 2-mercaptoethanol, L-

cysteine hydrochloride monohydrate, dithiothreitol, dithierythritol, and mixtures thereof. Suitable chaotropic agents include those selected from the group consisting of urea, thiourea, lithium perchlorate or guanidine hydrochloride, and mixtures thereof.

[0088] The solubilized inclusion bodies may be mixed in a refolding buffer, such as glycine or sodium carbonate, at a pH of 7-12, preferably from 10-11, preferably from 10.5-11, to refold the modified proinsulin sequences to a proinsulin derivative peptide, e.g., Lis-Pro proinsulin derivative peptide. The solution with refolded material should be pH adjusted to 7-9, preferably 7.8-8.2, with or without the addition of an alkaline salt, preferably sodium chloride to a final concentration of 100mM to 1M final concentration, preferably 500mM – 1M, preferably 700mM, and may be filtered and loaded onto a column, such as an immobilized metal-ion affinity chromatography (IMAC) column. Commercially available resins suitable for embodiments of the present invention include Nickle Sepharose 6 Fast Flow (GE Healthcare), Nickle NTA Agarose (GE Healthcare), Chelating Sepharose Fast flow(GE Healthcare), IMAC Fast Flow (GE Healthcare).

[0089] Lis-Pro proinsulin derivative peptide is subject to concentration by tangential flow filtration or diafiltration. Next, Lis-Pro proinsulin derivative peptide is enzymatically cleaved, preferred by subjecting the proinsulin derivative peptide to trypsin digestion. Although embodiments of the present invention may use commercially available rat, bovine, porcine or human trypsins or other isoenzymes or derivatives or variants thereof, it is also possible to use the following enzymes: recombinant trypsin, trypzene, trypsin from Fusarium oxysporum and from Streptomyces (*S. griseus, S. exfoliatus, S. erythraeus, S. fradiae* and *S. albidoflavus*), tryptase, mastin, acrosin, kallikrein, hepsin, prostasin I, lysyl endopeptidase (Lysin-C) and endoproteinase Arg-C (clostripain). In one embodiment, trypsin digestion occurs at a pH from about 7 to 10, and more preferably from 8.1 to 8.3. In a further embodiment, the trypsin digest is quenched by adding an organic acid, preferably glacial acetic acid. While it is contemplated that other additives may be employed, acetic acid appears to be most preferred and stable for this purpose.

[0090] Trypsin is an enzyme that has specific cleavage activity at the terminal arginine residues, and to a lesser extent, lysine residues, of the C-peptide. In the transformation reaction, it is required that the terminal arginine or lysine residues of the C-peptide be removed. In native human proinsulin, when trypsin cleaves at the lysine in position 64, it will be unable to remove

the arginine at position 65, due to the fact that it requires at least one residue on both sides of a cleavage site. What results is the production of an unwanted by-product, arg(A_0)-insulin. This by-product constitutes a small loss in yield and generates an undesired contaminant. By converting lysine, such at position 64 of native C-peptide, into another non basic amino acid, particularly alanine, the level of arg(A_0)-insulin byproduct is preferentially not formed. When formed is less than 10%, and more preferably is less than 0.3% of total byproducts from the trypsin transformation reaction may be arg(A_0). This is because the trypsin no longer acts to cleave at this particular site of the proinsulin derivative peptide.

[0091] The proinsulin derivative peptide, may also be subjected to carboxypeptidase B digestion. In one embodiment the Lis-Pro-proinsulin is double digested with trypsin and carboxypeptidase B in a glycine buffer at pH 9.6±0.1. In one embodiment, a trypsin inhibitor may be added to the insulin prior to addition of carboxypeptidase B. Trypsin inhibitor is added in an equal amount to the amount of trypsin added for the trypsin digest step. In another embodiment, a glycine solution is added to the DiR-Lis-Pro insulin analog prior to addition of carboxypeptidase B. For example, in some embodiments, glycine is added to adjust the pH of the insulin solution to about 9.6±0.1. The target concentration of glycine is 50mM using a 1M glycine stock. In some embodiments, the carboxypeptidase B is permitted to digest for at least 1-16 hours, preferably at least 8 hours. A minimum of 10 hours is preferred, but overdigestion is rarely a significant issue so there is no maximum time limit.

[0092] In one embodiment, after typsin digest and pre-carboxy digestion, the intermediate DiR–Lis-Pro insulin is purified on a chromatography column, such as an ion exchange column or reverse phase chromatography column, prior to carboxypeptidase B digestion. Following carboxypeptidase B digestion, the Lis-Pro insulin material may be further purified using ion exchange or reverse phase chromatography. In one embodiment, after trypsin and carboxypeptidase B double digestion, the Lis-Pro insulin solution is preferably purified in a chromatography column, such as an ion exchange chromatography column or reverse phase chromatography column. In one embodiment, the intermediate solution may be purified in a chromatography column by eluting the Lis-Pro insulin analog using a buffer comprising an alcohol or organic solvent, preferably propanol, such as n-propanol. The buffer may also further comprise an alkali metal salt, suchas sodium sulfate. The buffer may also further comprise an organic acid, suchas phosphoric acid.

[0093] The manufacturing process described herein results in a preparation of Lis-Pro insulin analog in liquid active pharmaceutical ingredient (API) form. The process eliminates the need to prepare a crystallized insulin that is later reconstituted. As a result of eliminating the crystallization and drying steps, the amount of inactive insulin multimers present in the liquid formulation is reduced in comparison to the amounts otherwise present in crystallized forms of insulin and reconstituted crystallized insulin. Although crystallization is less preferred, in some embodiments, a crystallization step may be included to produce Lis-Pro insulin analog API crystals. The Lis-Pro insulin analog may be crystallized to allow for increased shelf life to the API material. However, the crystallization process may lead to increased levels of multimers and in turn an overall lower purity.

[0094] Lis-Pro insulin analog may prevent the formation of multimeric (non-monomeric) insulin, such as dimers and hexamers. Accordingly, upon administration of the Lis-Pro insulin analog to a patient, larger amounts of active monomeric insulin are available to act in the patient. In particular, Lis-Pro insulin analog is particularly suitable for postprandial, i.e., after eating, injection as it is available immediately for use by the patient to control glucose levels. Accordingly, this analog has the advantage over native insulin in that its short delay of onset allows more flexibility with eating schedules for diabetic patients than regular insulin which requires a longer waiting period between injection and eating. According to one embodiment of the invention, the Lis-Pro insulin analog is provided to a patient in combination with a longer acting insulin to provide optimal glycemic control.

[0095] In some embodiments, the preparations comprise a pharmaceutically acceptable preparation comprising recombinant Lis-Pro insulin analog and being essentially free of modified proinsulin sequences and/or non-monomeric Lis-Pro insulin molecules.

[0096] It should be understood that process steps within the following description of the method may be modified, changed and/or eliminated, depending on the particular preferences of the processor and/or the particular mechanical apparatus available to the processor, as well as the specific reagents and/or materials available and/or convenience and/or economics of use.

Example 1—Preparation of an *E. coli* clone expressing Lis-Pro proinsulin

[0097] The preparation of a transformed *E. coli* cells capable of expressing recombinant Lis-Pro proinsulin is carried out according to the following processes.

[0098] Step 1: Construction of a purified Lis-Pro proinsulin gene segment for insertion into the vector. The initial gene construct was synthesized in a basic cloning vector (ptrcKis2a(Kan)). The gene construct included the N-terminal histidine tag, MHHHHHHGGR (SEQ ID NO: 4), modified B-chain, and modified C-peptide with the alanine codon in place of the native lysine and having the amino acid sequence MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGG PGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 14). The gene was flanked by Nde1 and EcoR1 restriction sites, for subsequent subcloning into the desired expression vector. The codons selected were optimized for expression in *E.coli*. The following sequence represents the pTrcHis2a(Kan) vector with a Lis-Pro proinsulin insert. The IPTG inducible promoter region which regulates the transcription rate is shown by the dotted underline, while the Lis-Pro proinsulin insert, adjacent the promoter region, is shown by the solid underlined. The sequence shown by the bold and italicized is the Kanamycin gene, which provides the antibiotic selection marker for the vector.

pTrcHis2A(Kan)/Lis-Pro DNA Sequence:

5′GTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGG
AAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTC
CCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGC
TGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACAC
AGGAAACAGCGCCGCTGAGAAAAAGCGAAGCGGCACTGCTCTTTAACAATTTATCAGACAATCT
GTGTGGGCACTCGACCGGAATTATCGATTAACTTTATTATTAAAAATTAAAGAGGTATATATTA
ATGTATCGATTAAATAAGGAGGAATAAACCATGATGCATCATCATCATCATCATGGTGGCCGCT
TTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACG
AGGCTTCTTCTACACAAAGCCGACCCGCCGGGAGGCAGAGGACCTGCAGGTGGGGCAGGTGGAG
CTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCTGGAGGGGTCCCTGCAGAAGC
GTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCGG
CTAGGAATTCGAAGCTTGGGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGT
CGACCATCATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAG
AGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTT
GCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGT
AGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAA
CGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCC
TGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCG
GGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGC
CTTTTTGCGTTTCTACAAACTCTTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT
CATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAA
CATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAG
AAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACT
GGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGC

ACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCG
GTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTCCTGAATCGCCCCATCATCCAGCCA
GAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGACCAGTTGGTGATTTTGAAC
TTTTGCTTTGCCACGGAACGGTCTGCGTTGTCGGGAAGATGCGTGATCTGATCCTTCAACTCAG
CAAAAGTTCGATTTATTCAACAAAGCCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGT
TACAACCAATTAACCAATTCTGA***TTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTAT***
TCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTC
ACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACA
TCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAG
TGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGG
CCAGCCATTACGCTCGTCATCAAAATCACTGCATCAACCAAACCGTTATTCATTCGTGATTGCG
CCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAA
CCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAAT
ACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGA
TAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATC
TGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTC
CCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCAT
ATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCGAGCAAGACGTTTCCCGTTGAATATG
GCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATA
TTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGAATAAATC
GAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGC
AAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCC
CTCACTTTCTGGCTGGATGATGGGGCGATTCAGGACTCACCAGTCACAGAAAAGCATCTTACGG
ATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAA
CTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT
CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTAC
TCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTG
CGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTC
GCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC
GGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATT
AAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATT
TTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACG
TGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCT
TTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTT
TGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACC
AAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTA
CCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTC
GTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA
TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGG
GTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGG
AAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT
TCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC
CGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTG
ATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTA

CAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTC
ATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGG
CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTC
ATCACCGAAACGCGCGAGGCAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATTTACG
TTGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA
ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCT
TATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAG
TGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAA
ACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTC
GCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAA
GCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGAT
CATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCG
GCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACG
GTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGG
CCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAAT
CAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCA
TGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCT
GGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGA
TACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTCAACCACCATCAAACAGGATTTTC
GCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGG
CAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACC
GCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAA
GCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCGCGAATTGATCTG 3′ (SEQ ID
NO: 19)

[0099] The modified proinsulin sequence without the tag is as follows:

TTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAAC
GAGGCTTCTTCTACACAAAGCCGACCCGCCGGGAGGCAGAGGACCTGCAGGTGGGGCAGGTGGA
GCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCTGGAGGGGTCCCTGCAGAAG
CGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCG
GCTAG (SEQ ID NO: 23)

[0100] Step 2: Generation of the pTrcHis2A(Kan) vector containing Lis-Pro proinsulin. Commercially available pTrcHis2A(Kan) vector was modified to include a Kanamycin resistance gene in the middle of the Ampicillin resistance gene to negate the Ampicillin resistance prior to insertion of the proinsulin sequence into the vector. Ampicillin resistance heightens the potential for allergic reactions to preparations made using vector constructs that include the Ampicillin resistance gene. Therefore it is preferable to eliminate the Ampicillin resistance in the constructs that are prepared and used.

[0101] The pTrcHis2A(Kan) vector was modified at the start codon in the multiple cloning site by replacing the Nco1 restriction site with an Nde1 site to simplify subsequent subcloning work.

Nco1 = CCATGG - Nde1 = CATATG

[0102] The proinsulin gene was isolated from the DNA 2.0 plasmid using Nde1 to cleave at the N-terminal side of the gene and EcoR1 to cleave at the C-terminal side of the gene. The Digested DNA was run over a 2% agarose gel to separate the plasmid DNA from the Lis-Pro proinsulin gene. A QIAquick™ (Qiagen) gel purification kit was then used to purify the gene construct.

[0103] Accordingly, a sequential digest of the vector with Nde1 and EcoR1, respectively, was performed. The vector DNA was also purified using a QIAquick gel purification kit. Following purification of the vector and the gene, a 5' Nde1 and a 3' EcoR1 ligation reaction was utilized to insert the proinsulin gene into the pTrcHis2A(Kan) vector.

[0104] Step 3: Transformation. One microliter of the ligation reaction was used to transform competent *E. coli* cells BL21 with the pTrcHis2A(Kan) plasmid containing the proinsulin gene. The transformed *E. coli* BL21 cells were plated on LB-Kan agar plates and incubated overnight at 37°C. Several clones were selected and sequenced. Clones with the correct sequence were then screened for expression.

[0105] The resulting vector is referred to as the His Tagged Lis-Pro proinsulin pTrcHis2A(Kan) vector.

[0106] Step 4: Preparation of the working cell bank (WCB). To establish the WCB, sterile growth medium was inoculated with the recombinant BL21 *E. coli* containing the His Tagged Lis-Pro proinsulin pTrcHis2A(Kan) vector and incubated to allow cell growth. The cells were harvested in an ISO5 (class 100) environment under a biosafety cabinet via centrifugation. Sterile medium and glycerol were added to the cells. 1 mL aliquots of the cells were then dispensed into sterile ampoules and stored at -80°C. Aseptic techniques were utilized to generate a WCB.

Example 2—Product Manufacture of Lis-Pro Insulin Analog from Modified Proinsulin Sequence Carrying Transformed *E. coli*

[0107] Step 1 – Culturing of *E. coli* transformed with Lis-Pro modified proinsulin sequence as described in Example 1. Seed an inoculum preparation of the transformed (WCB) in a sterile growth medium that includes yeastolate (purchased from VWR, Prod. # 90004-426 or – 488), select phytone, sodium chloride, purified water, sterile Kanamycin solution), and incubate until growth to an Optical density (OD_{600nm}) of 2 to 4. Prepare a fermentation media (containing

select phytone, yeastolate, glycerin, BioSpumex 153K (Cognis, Inc.) in a fermentor. Add the following sterilized phosphate solutions to the Fermentor. Prepare a Phosphate flask 1 – potassium phosphate monobasic and potassium phosphate dibasic containing Kanamycin solution. Prepare a Phosphate flask 2 – potassium phosphate monobasic and potassium phosphate dibasic. Add seed inoculate of *E. coli* to the Fermentor – growth to O.D. (optical density) 600nm of 8 to 10 (mid log phase). Add a dioxane free IPTG (purchased from Promega, Catalog No. #PA V3953 (VWR Catalog #PAV3953) solution to the fermentor (to induce transcription of the K64A Lis-Pro proinsulin gene). Incubate for 4 hours. This results in the production of a concentrated cell suspension containing His-tagged Lis-Pro proinsulin inclusion bodies. The cell suspension is then centrifuged to provide a cell paste for the subsequent inclusion body isolation step.

[0108] Step 2 – Disruption —Cells containing inclusion bodies expressing Lis-Pro modified proinsulin sequence are lysed in a basic Tris/salt buffer, using a Niro Soavi homogenizer (1100-1200 bar).

[0109] Step 3 - Inclusion Body Washing—Contaminant protein removal is accomplished via two sequential washes with a Tris/Triton X-100 buffer, followed by two sequential washes with a Tris/Tween-20 buffer, and finally a single wash with a Tris/NaCl buffer.

[0110] Step 4 - Solubilization—Inclusion bodies enriched with the modified proinsulin peptide are solubilized in 4-8M urea, preferably 6-8M urea, containing reducing agents (2-mercaptoethanol, L-cysteine hydrochloride monohydrate). Complete solubilization is achieved by adjusting the pH to 10.5-12, preferably 11.8-12 with NaOH.

[0111] Step 5 - Dilution refolding—The solubilized protein is then diluted into refolding buffer (20 mM Glycine, pH 10-11 at 6-10°C.) to a final concentration of 1 mg/ml and permitted to refold for 24 to 72 hours, preferentially about 48 hours, at 6-10°C. Higher protein concentration may be used in the refold if desired. However, overall refold efficiency will decrease. Sodium Chloride and Phosphate are then added to final concentrations of 700 mM and 25 mM respectively, followed by pH adjustment to 7.0 to 9.0, preferably 7.9-8.0 with 6M HCl.

[0112] Step 6 - IMAC Chromatography—The dilute proinsulin derivative is loaded onto an IMAC column to a maximum capacity of ≤ 26.5 mg main peak protein per ml of resin. A 75mM imidizole buffer is used to isocratically strip the majority of impurities from the column. Lis-Pro proinsulin is eluted isocratically using ≤300 mM imidizole.

[0113] Step 7 - Buffer exchange—To the IMAC main peak pool material, add EDTA to a final concentration of 20 mM. Exchange the buffer using a membrane with a suitable molecular weight cutoff (e.g. 3000 Da). The final buffer should be at least 97% exchanged to a 20 mM Tris-Cl, pH 7.0-10.0, preferably 8.1 at 8-10°C. A protein concentration of approximately 5 mg/ml is desirable.

[0114] Step 8 Trypsin/Carboxypeptidase B Enzymatic Transformation/Proteolysis—The buffer exchanged sample is digested with a 1500:1 mass ratio of main peak protein to trypsin and 1000:1 mass ratio of main peak protein to carboxypeptidase B, in the presence of 5mM CaCl. The ratios of trypsin and carboxypeptidase may be increased or decreased depending on the desired length of time for the reaction. Once complete, based on HPLC, the digest is then quenched by the addition of acetic acid to ≥ 700 mM, to a pH of approximately 3-3.5.

Example 3A Final Purification

[0115] After step 8 in Example 2, the final purification may proceed using alternative processes in Examples 3A or 3B.

[0116] Step 9a - Ion Exchange Chromatography—The digested material is loaded onto a cation exchange column and eluted with a NaCl gradient, in the presence of 20% n-propanol or acetonitrile at pH 2-5, preferably 4.0. Fractions are diluted 1:4 if n-propanol is used for elution or 1:2 with cold purified water if acetonitrile is used for elution, or no dilution if acetonitrile is used for elution. RP-HPLC is used to pool the appropriate fractions containing the Lis-Pro insulin peak of interest at the desired purity level.

[0117] Step 10a - Reverse Phase Chromatography—The S-column pool containing the Lis-Pro insulin is loaded onto an RPC30 or C18 reverse phase column and eluted using an n-propanol or acetonitrile gradient in the presence of 200mM sodium sulfate and 0.136% phosphoric acid. Fractions are immediately diluted 1:4 with 100mM phosphate buffer at pH 7.0-9.0, preferably 7.5-8, as they are collected. RP-HPLC is used to pool the appropriate fractions containing the Lis-Pro Insulin peak of interest at the desired purity level.

[0118] Step 11a - Buffer Exchange—Exchange the sample into WFI (water for injection) using a membrane with a suitable molecular weight cutoff (e.g. 3000 Da). The pH of the solution should be monitored and maintained at 7.0-9.0, preferably 7.5-8.0. The final sample is concentrated to 5.5-8 mg/ml, with an adjusted pH of 7.0-9.0, preferably 7.5-8.0at 6-10°C. This

material represents the liquid API form of the presently disclosed preparations of Lis-Pro Insulin Analog. The API should be stored in the dark at 6-10°C.

Example 3B Final Purification

[0119] Step 10b - Reverse Phase Chromatography—The digested material containing the Lis-Pro insulin is loaded onto an RPC30 or C18 reverse phase column and eluted using a n-propanol or acetonitrile gradient in the presence of 200mM sodium sulfate and 0.136% phosphoric acid. Fractions are immediately diluted 1:4 with 100mM phosphate buffer at pH 7.0-9.0, preferably 7.5-8 as they are collected. RP-HPLC is used to pool the appropriate fractions containing the Lis-Pro insulin peak of interest at the desired purity level.

[0120] Step 11b - Buffer Exchange—Exchange the sample into WFI using a membrane with a suitable molecular weight cutoff (e.g. 3000 Da). The pH of the solution should be monitored and maintained at 7.0-9.0, preferably 7.5-8.0. The final sample is concentrated to 5.5-5.8 mg/ml, with an adjusted pH of 7.5-8.0 at 6-10°C. This material represents the liquid API form of the presently disclosed preparations and formulations of Lis-Pro Insulin Analog.

Example 4 – API Formulation

[0121] The Lis-Pro Insulin Analog purified by Examples 3A or 3B is formulated by diluting the API material with cold WFI to a final concentration of 4.3375 mg/ml. A concentrated formulation buffer stock containing 80 mg/ml glycerol, 15.75 mg/ml meta cresol, and 0.0985 mg/ml zinc chloride at pH 7.5 ± 0.1 is added to the API material in a 1/5 ratio of formulation buffer stock to API. The solution is mixed, followed by sterile filtration into appropriate vials in 10 ml aliquots.

Example 5 – Working Cell Bank

[0122] The preparation of a WCB (working cell bank) for research and development containing cells capable of expressing recombinant Lis-Pro proinsulin is carried out according to the following processes.

[0123] The cloning procedure outlined in Example 1 is utilized to create the initial vector (transfection vector). Purified His Tagged Lis-Pro proinsulin pTrcHis2A(Kan) vector is transformed into competent BL21 cells and plated on sterile LB-Kan plates. From the plates, an isolated colony is used to inoculate sterile LB-Kan media (~100mls). The cells are grown at 37°C to mid log phase (about 4-5 hours) OD_{600nm} of about 1.5-2.0. Culture media containing

cells is then aliquoted into sterile cryovials, combined with glycerol at a 20% final concentration. The vials are then stored at 80°C.

WHAT IS CLAIMED IS:

1. A composition comprising a Lis-Pro proinsulin sequence having a formula I:

$$R_1\text{-}(B_1\text{-}B_{27})\text{-}B_{28}\text{-}B_{29}\text{-}B_{30}\text{-}R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5\text{-}(A_1\text{-}A_{21})\text{-}R_6 \quad \text{Formula I}$$

wherein

R_1 is a tag sequence containing one or more amino acids or R_1 is absent with an Arg or Lys present prior to the start of the B chain;

(B_1-B_{27}) and (A_1-A_{21}) comprise amino acid sequences of native human insulin;

B_{28} is any amino acid other than Pro;

B_{29} is any amino acid other than Lys or Arg;

B_{30} is Thr;

R_2, R_3 and R_5 are Arg;

R_4 is any amino acid other than Gly, Lys or Arg or is absent;

X is a sequence comprises one or more amino acids or is absent, provided that X is not EAEALQVGQVELGGGPGAGSLQPLALEGSLQ (SEQ ID NO: 2) and X does not comprise a C-terminal Gly, Lys, or Arg when R_4 is absent; and

R_6 is a tag sequence containing one or more amino acids or R_6 is absent.

2. The composition of claim 1, wherein R_1 and/or R_6 is present and R_1 is tag sequence of one or more amino acids with a C-terminal Arg or Lys and/or R_6 tag sequence of one or more amino acids with a N-terminal Arg or Lys.

3. The composition of claim 1, wherein R_4 is Ala.

4. The composition of claim 1, wherein the modified Lis-Proproinsulin sequence comprises a connecting peptide sequence of a sequence having the formula

$$R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5 \quad \text{Formula II}$$

wherein

R_2, R_3, R_4, R_5, and X are defined in claim 1.

5. The composition of claim 4, wherein the connecting peptide sequence is RREAEDLQVGQVELGGGPGAGSLQPLALEGSLQAR (SEQ ID NO: 7).

6. The composition of claim 1, wherein the Lis-Pro modified proinsulin sequence is; FVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPLAL EGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 13);

MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGG PGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN (SEQ ID NO: 14);

MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGG PGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNRHHHHHH (SEQ ID NO: 16);

MHHHHHHGGRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGG PGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCNKHHHHHH (SEQ ID NO: 21);

MRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPL ALEGSLQARGIVEQCCTSICSLYQLENYCNRHHHHHH (SEQ ID NO: 17); or

MRFVNQHLCGSHLVEALYLVCGERGFFYTKPTRREAEDLQVGQVELGGGPGAGSLQPL ALEGSLQARGIVEQCCTSICSLYQLENYCNKHHHHHH (SEQ ID NO: 22).

7. An expression vector comprising the nucleic acid sequence of claim 1.

8. The expression vector of claim 17, wherein the expression vector is His Tagged Lis-Pro proinsulin pTrcHis2A(Kan).

9. A microorganism transformed with the vector of claim 17.

10. The microorganism of claim 9, further defined as an E. coli transformed with plasmid His Tagged Lis-Pro proinsulin pTrcHis2A(Kan).

11. A process for producing Lis-Pro insulin analogs comprising the steps of:

(a) culturing transformed *E. coli* comprising a sequence of Formula I under conditions suitable for expression of a modified proinsulin sequence of formula I:

$$R_1\text{-}(B_1\text{-}B_{27})\text{-}B_{28}\text{-}B_{29}\text{-}B_{30}\text{-}R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5\text{-}(A_1\text{-}A_{21})\text{-}R_6 \qquad \text{Formula I}$$

wherein

R_1 is a tag sequence containing one or more amino acids or R_1 is absent with an Arg or Lys present prior to the start of the B chain;

(B_1-B_{27}) and (A_1-A_{21}) comprise amino acid sequences of native human insulin;

B_{28} is any amino acid other than Pro;

B_{29} is any amino acid other than Lys or Arg;

B_{30} is Thr;

R_2, R_3 and R_5 are Arg;

R_4 is any amino acid other than Gly, Lys or Arg or is absent;

X is a sequence comprises one or more amino acids or is absent, provided that X is not EAEALQVGQVELGGGPGAGSLQPLALEGSLQ (SEQ ID NO: 2) and X does not comprise a C-terminal Gly, Lys, or Arg when R_4 is absent; and

R_6 is a tag sequence containing one or more amino acids or R_6 is absent,

(b) disrupting said cultured *E. coli* cells to provide a composition comprising inclusion bodies containing a modified Lis-Pro proinsulin;

(c) solubilizing said composition of inclusion bodies; and

(d) recovering the Lis-Pro insulin analogs from said solubilized composition.

12. The process of claim 11, wherein the solubilization of said composition of inclusion bodies further comprises adjusting the pH to at least 10.5.

13. The process of claim 11, wherein the solubilization of said composition of inclusion bodies by adjusting the pH to 11.8 to 12.

14. The process of claim 11, wherein the solubilization of said composition of inclusion bodies includes one or more reducing agents selected from the group consisting of 2-mercaptoethanol, L-cysteine hydrochloride monohydrate, dithiothreitol, dithierythritol, and mixtures thereof.

15. The process of claim 11, wherein the solubilization of said composition of inclusion bodies includes one or more chaotropic agents selected from the group consisting of urea, thiourea, lithium perchlorate or guanidine hydrochloride and mixtures thereof.

16. The process of claim 11, wherein the step of recovering the Lis-Pro insulin analogs further comprises:

(e) folding said modified Lis-Pro proinsulin to provide a Lis-Pro proinsulin derivative peptide;

(f) purifying said Lis-Pro proinsulin derivative peptide using metal affinity chromatography

(g) enzymatically cleaving the Lis-Pro proinsulin derivative peptide to remove a connecting peptide and provide an intermediate solution comprising Lis-Pro insulin analog; and

(h) purifying said intermediate solution in a chromatography column to yield the Lis-Pro insulin analog.

17. A process for producing Lis-Pro insulin analogs comprising the steps of:

(a) recombinately producing a modified Lis-Pro proinsulin having a formula I:

$$R_1\text{-}(B_1\text{-}B_{27})\text{-}B_{28}\text{-}B_{29}\text{-}B_{30}\text{-}R_2\text{-}R_3\text{-}X\text{-}R_4\text{-}R_5\text{-}(A_1\text{-}A_{21})\text{-}R_6 \qquad \text{Formula I}$$

wherein

R_1 is a tag sequence containing one or more amino acids or R_1 is absent with an Arg or Lys present prior to the start of the B chain;

$(B_1\text{-}B_{27})$ and $(A_1\text{-}A_{21})$ comprise amino acid sequences of native human insulin;

B_{28} is any amino acid other than Pro;

B_{29} is any amino acid other than Lys or Arg;

B_{30} is Thr;

R_2, R_3 and R_5 are Arg;

R_4 is any amino acid other than Gly, Lys or Arg or is absent;

X is a sequence comprises one or more amino acids or is absent, provided that X is not EAEALQVGQVELGGGPGAGSLQPLALEGSLQ (SEQ ID NO: 2) and X does not comprise a C-terminal Gly, Lys, or Arg when R_4 is absent; and

R_6 is a tag sequence containing one or more amino acids or R_6 is absent,

(b) folding said modified Lis-Pro proinsulin to provide a Lis-Pro proinsulin derivative peptide;

(c) purifying said Lis-Pro proinsulin derivative peptide using metal affinity chromatography;

(d) enzymatically cleaving said Lis-Pro proinsulin derivative peptide to remove a connecting peptide and provide an intermediate solution comprising Lis-Pro insulin analog; and

(e) purifying said intermediate solution to yield Lis-Pro insulin analog.

18. The process of claim 16, wherein the step of purifying further comprises eluting the Lis-Pro insulin analog with a buffer containing n-propanol.

pTrcHis2A(Kan) Vector map with Lis-Pro-Proinsulin Insert:

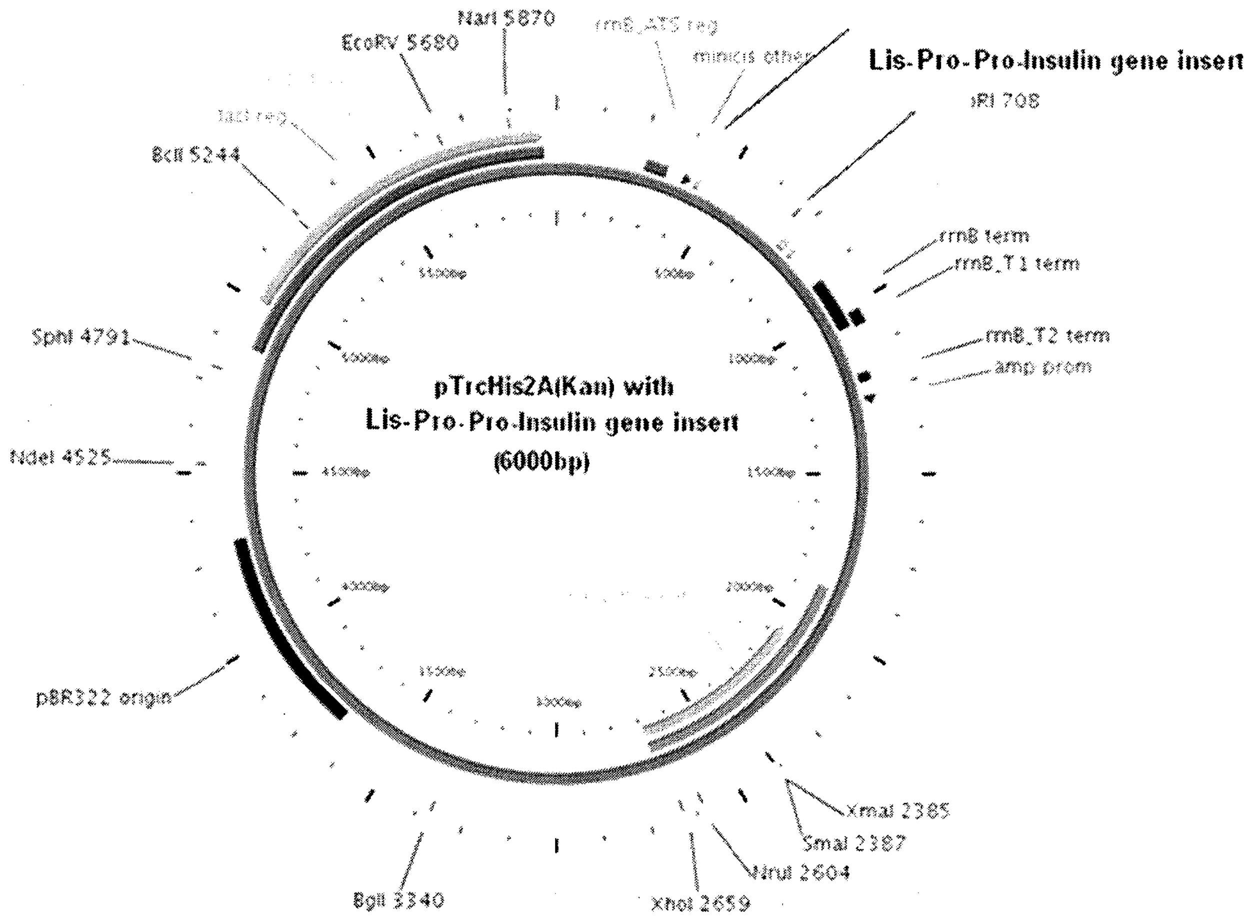


FIG. 1

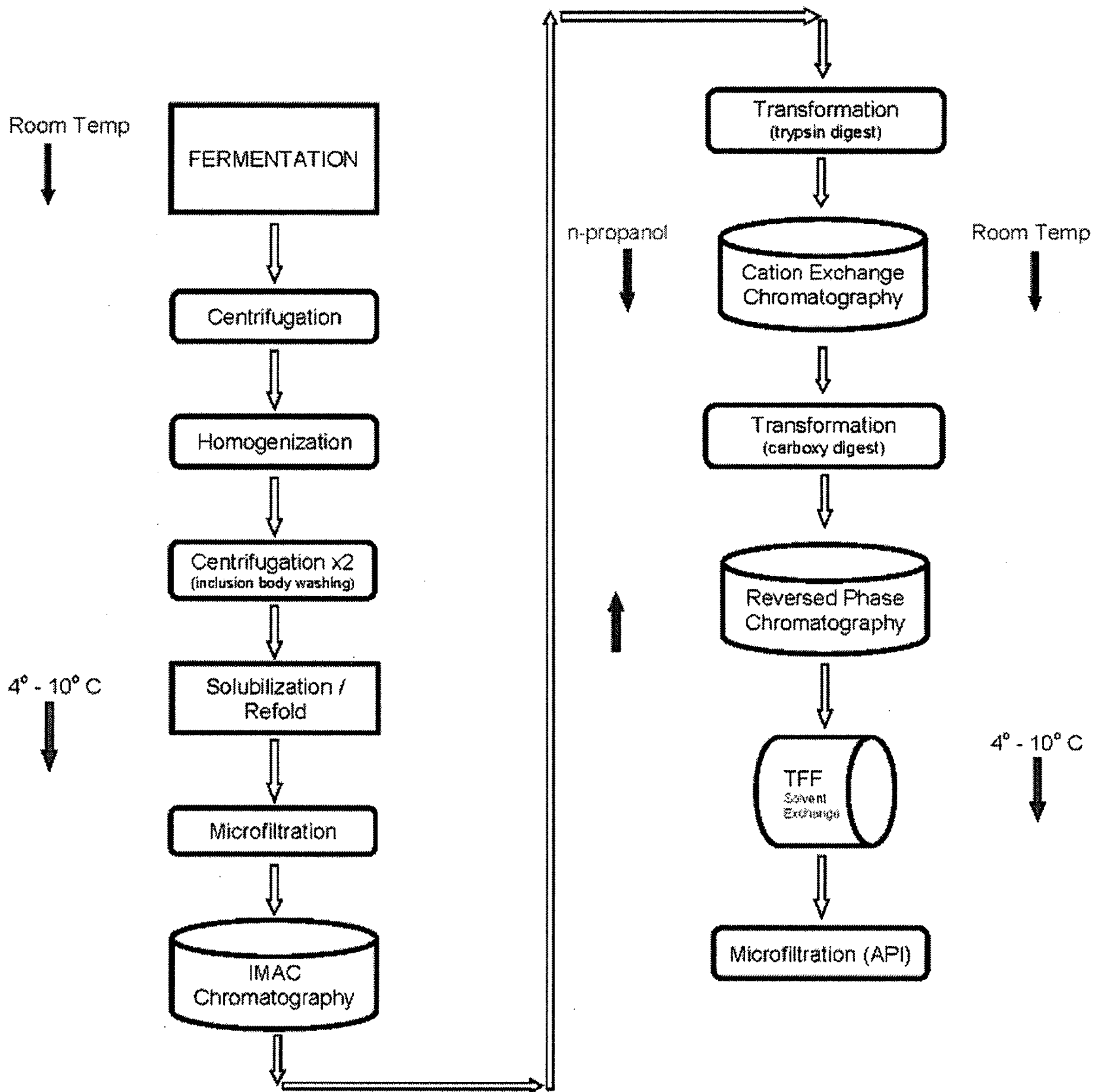


Figure 2a: Production Block Flow Diagram for Insulin Lispro

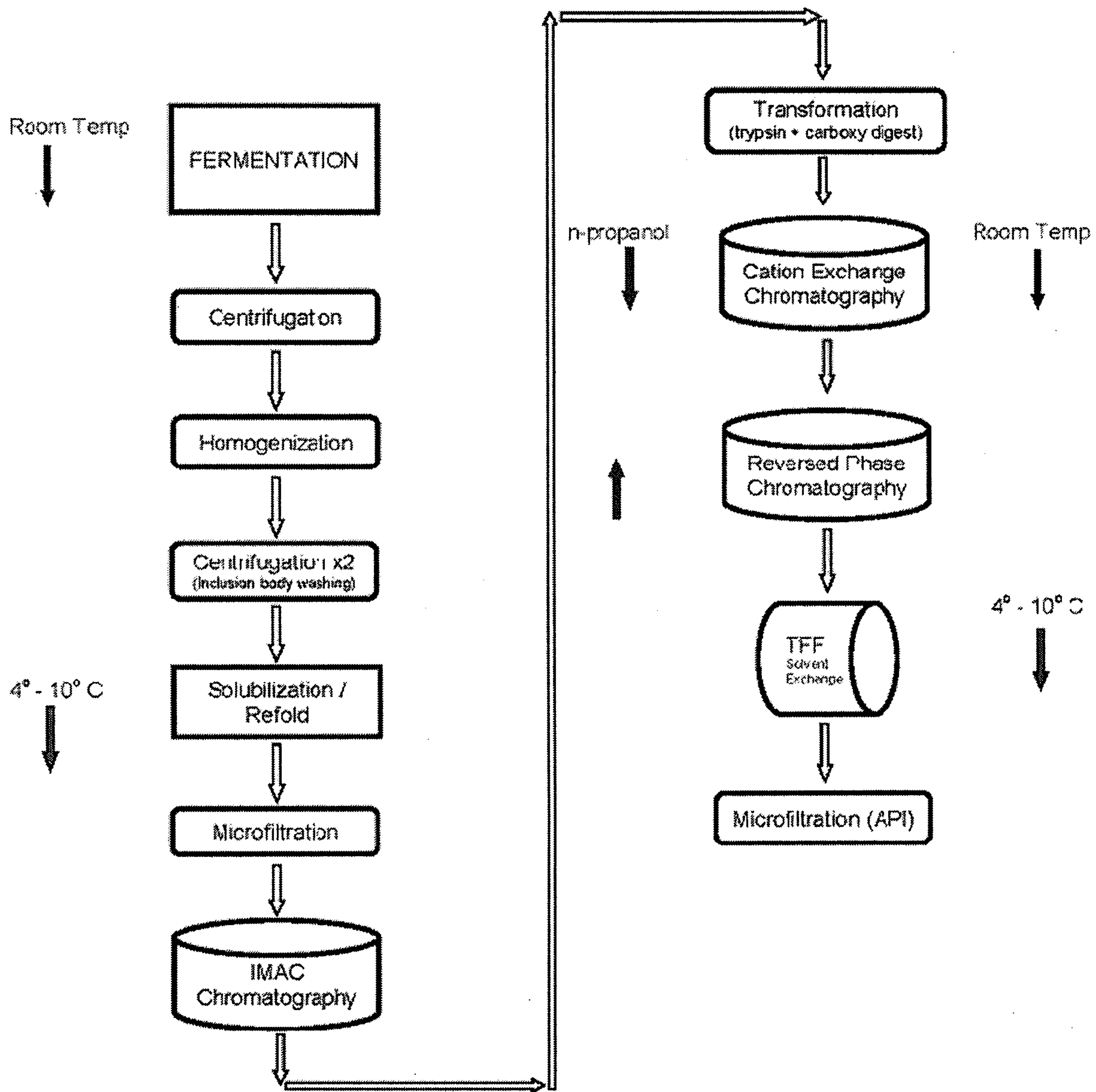


Figure 2b: Alternate Production Block Flow Diagram for Insulin Lispro

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/025934

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/62 C12N15/00 C12P21/02 A61P3/10
ADD.

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)
C07K C12N C12P A61P

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 practical, search terms used)

EPO-Internal, BIOSIS, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2010/210815 A1 (ZIMMERMAN RONALD E [US] ET AL) 19 August 2010 (2010-08-19) paragraphs [0017] - [0032], [0 37], [0 95], [128]; claims 1-19; figure 4; examples 2, 5, 6, 8, 10; sequence 31	1-18
A	CHANG S-G ET AL: "HUMAN INSULIN PRODUCTION FROM A NOVEL MINI-PROINSULIN WHICH HAS HIGH RECEPTOR-BINDING ACTIVITY", BIOCHEMICAL JOURNAL, THE BIOCHEMICAL SOCIETY, LONDON, GB, vol. 329, 1 January 1998 (1998-01-01), pages 631-635, XP001037629, ISSN: 0264-6021	1-18
A	DE 196 05 657 A1 (HANIL SYNTHETIC FIBER CO LTD [KR]) 22 August 1996 (1996-08-22)	1-18

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"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 when the document is taken alone

"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.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 July 2011	04/08/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Habedanck, Robert

2

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/025934

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2001/041787 A1 (KJELDSEN THOMAS BORGLUM [DK] ET AL) 15 November 2001 (2001-11-15) -----	1-18
A	US 5 700 662 A (CHANCE RONALD E [US] ET AL) 23 December 1997 (1997-12-23) -----	1-18

2

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2011/025934

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010210815 A1	19-08-2010	US 2008146492 A1	19-06-2008
DE 19605657 A1	22-08-1996	FR 2730495 A1	14-08-1996
		GB 2298206 A	28-08-1996
		JP 8333395 A	17-12-1996
		US 5962267 A	05-10-1999
US 2001041787 A1	15-11-2001	US 2009017497 A1	15-01-2009
		US 2002137144 A1	26-09-2002
US 5700662 A	23-12-1997	NONE	