



(12) **DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

(13) **A1**

(86) **Date de dépôt PCT/PCT Filing Date:** 2017/09/05
(87) **Date publication PCT/PCT Publication Date:** 2018/03/15
(85) **Entrée phase nationale/National Entry:** 2019/03/05
(86) **N° demande PCT/PCT Application No.:** IB 2017/055336
(87) **N° publication PCT/PCT Publication No.:** 2018/047062
(30) **Priorité/Priority:** 2016/09/06 (GR20160100458)

(51) **Cl.Int./Int.Cl.** **A61K 38/28** (2006.01),
A61P 3/10 (2006.01), **C07K 14/62** (2006.01)
(71) **Demandeur/Applicant:**
CHEMICAL & BIOPHARMACEUTICAL LABORATORIES
OF PATRAS S.A., GR
(72) **Inventeurs/Inventors:**
GATOS, DIMITRIOS, GR;
ANASTASIOU, ALEXANDRA, GR;
BARLOS, KLEOMENIS, GR
(74) **Agent:** BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) **Titre : DERIVES DE PRO-INSULINE**
(54) **Title: PROINSULIN DERIVATIVES**

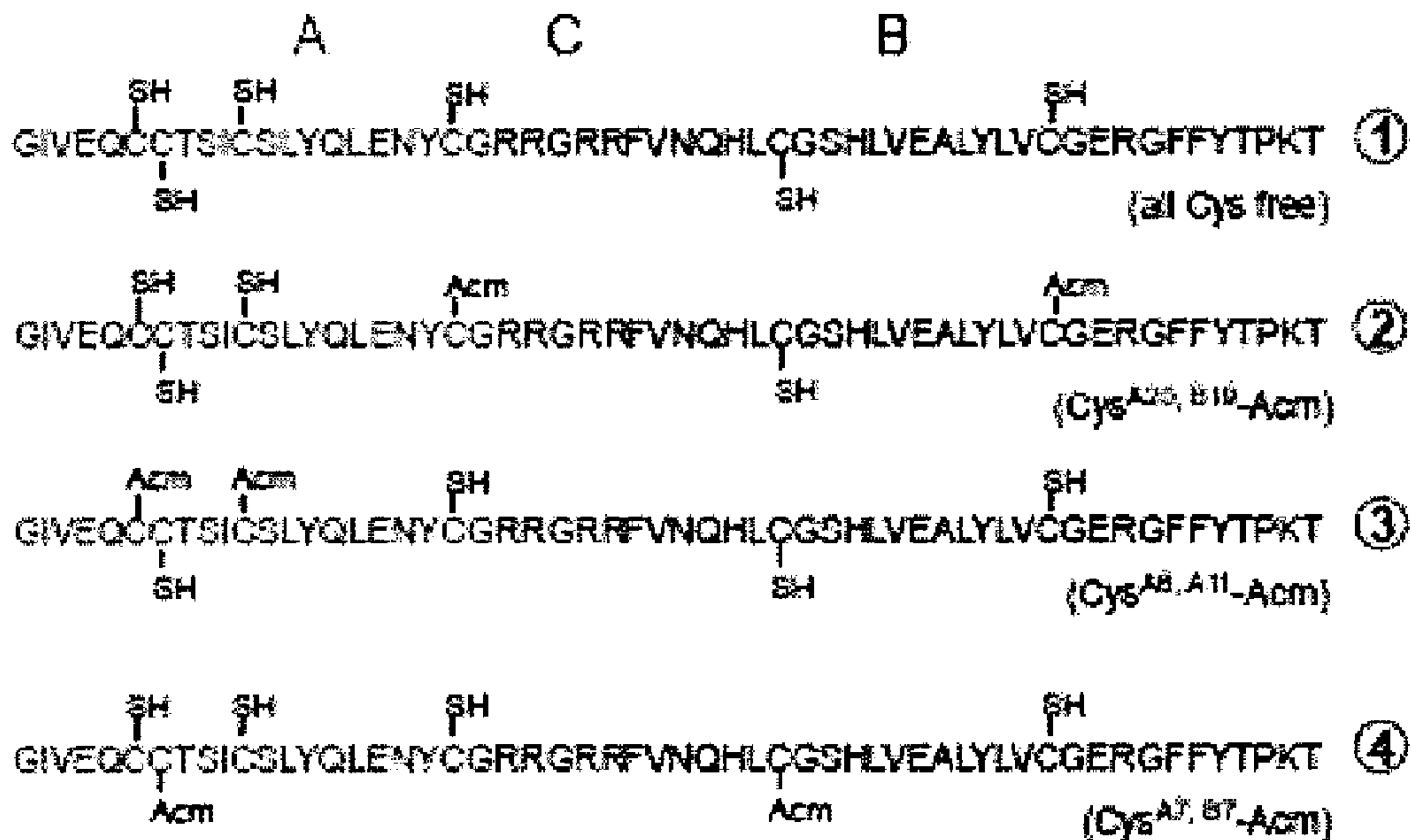


FIGURE 1

(57) **Abrégé/Abstract:**

A first aspect of the invention relates to a polypeptide compound of formula (I): A-C-B (I) wherein: A is the A chain of insulin or a functional derivative or variant thereof; B is the B chain of insulin or a functional derivative or variant thereof; C is a peptide of the formula: (X₁)_p - (X₂)_n - (X₃)_q wherein: each X₁ and X₃ is independently a basic amino acid; each X₂ is independently a natural or unnatural amino acid; p is 1 or 2; q is 0, 1 or 2; n is 0, 1, 2 or 3. Further aspects of the invention relate to pharmaceutical compositions comprising said polypeptide compound, and therapeutic uses thereof. Another aspect relates to the use of said polypeptide compounds in the preparation of insulin and derivatives thereof.

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

(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
15 March 2018 (15.03.2018)(10) International Publication Number
WO 2018/047062 A1

(51) International Patent Classification:

A61K 38/28 (2006.01) *A61P 3/10* (2006.01)
C07K 14/62 (2006.01)

(21) International Application Number:

PCT/IB2017/055336

(22) International Filing Date:

05 September 2017 (05.09.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

20160100458 06 September 2016 (06.09.2016) GR

(71) Applicant: **CHEMICAL & BIOPHARMACEUTICAL LABORATORIES OF PATRAS S.A.** [GR/GR]; Industrial Area of Patras Building, Block 1, GR-26000 Patras (GR).

(72) Inventors: **GATOS, Dimitrios**; c/o Chemical & Biopharmaceutical Laboratories of Patras S.A., Industrial Area of Patras Building, Block 1, GR-26000 Patras (GR). **ANASTASIOU, Alexandra**; c/o Chemical & Biopharmaceutical Laboratories of Patras S.A., Industrial Area of Patras Building, Block 1, GR-26000 Patras (GR). **BARLOS, Kleomenis**; c/o Chemical & Biopharmaceutical Laboratories of

Patras S.A., Industrial Area of Patras Building, Block 1, GR-26000 Patras (GR).

(74) Agent: **CLYDE-WATSON, Zoe**; D Young & Co LLP, 120 Holborn, London Greater London EC1N 2DY (GB).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

(54) Title: PROINSULIN DERIVATIVES

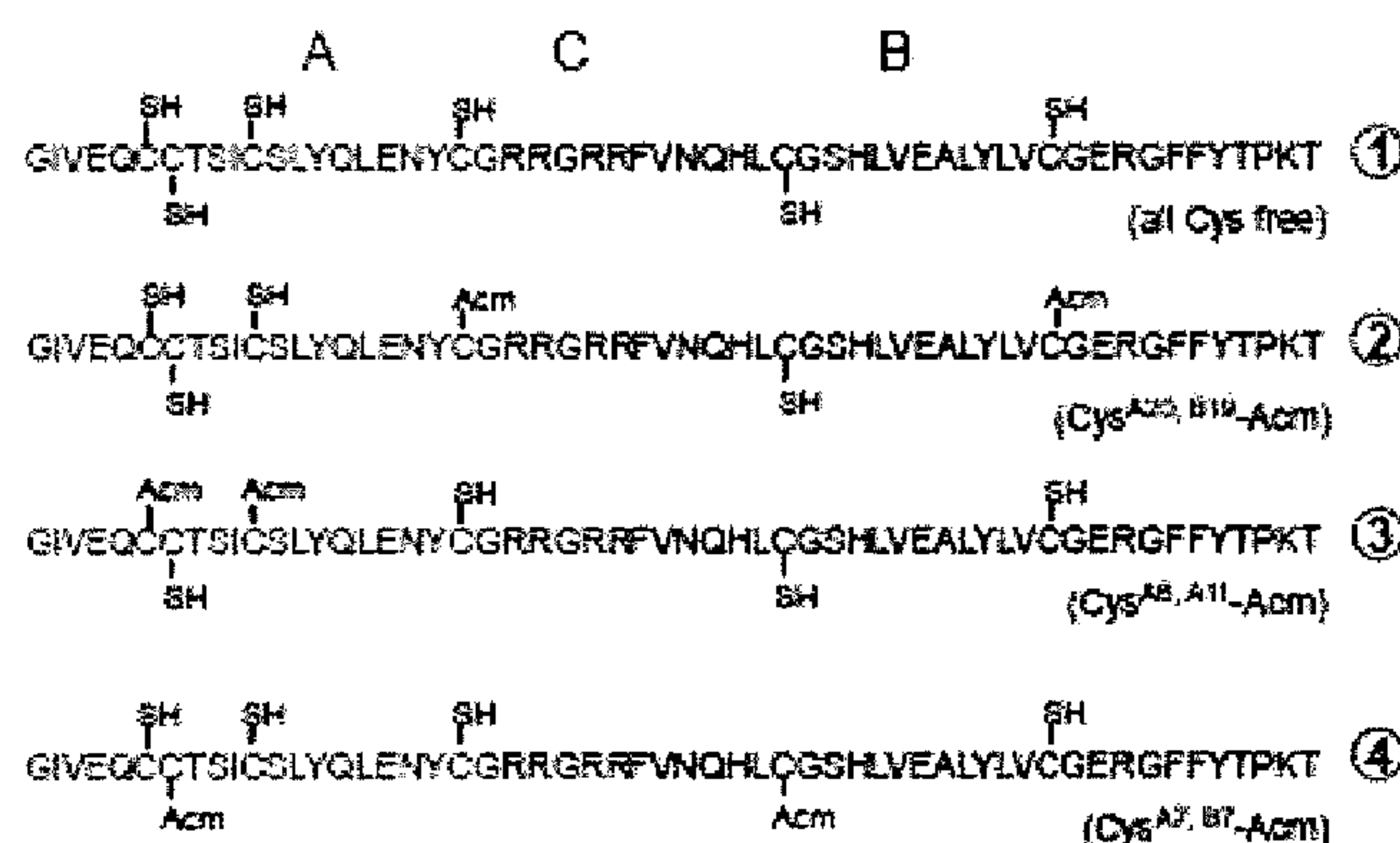


FIGURE 1

(57) Abstract: A first aspect of the invention relates to a polypeptide compound of formula (I): A-C-B (I) wherein: A is the A chain of insulin or a functional derivative or variant thereof; B is the B chain of insulin or a functional derivative or variant thereof; C is a peptide of the formula: $(X_1)_p - (X_2)_n - (X_3)_q$ wherein: each X_1 and X_3 is independently a basic amino acid; each X_2 is independently a natural or unnatural amino acid; p is 1 or 2; q is 0, 1 or 2; n is 0, 1, 2 or 3. Further aspects of the invention relate to pharmaceutical compositions comprising said polypeptide compound, and therapeutic uses thereof. Another aspect relates to the use of said polypeptide compounds in the preparation of insulin and derivatives thereof.

[Continued on next page]

WO 2018/047062 A1

Published:

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

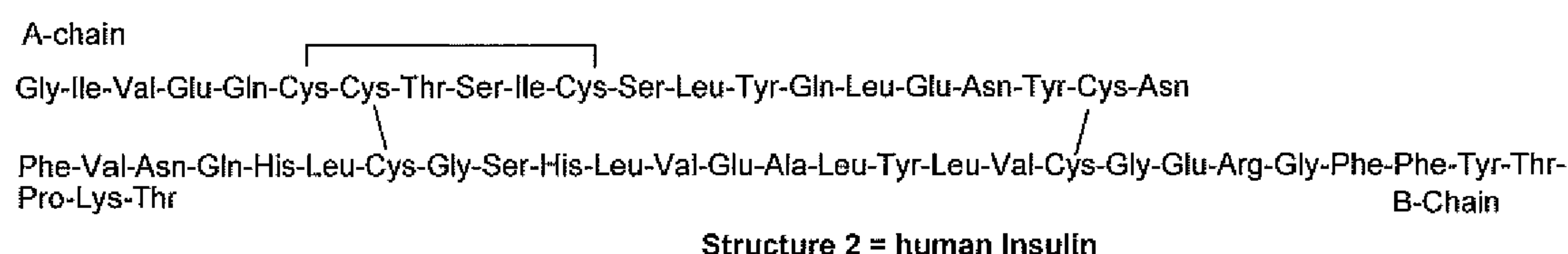
PROINSULIN DERIVATIVES

The present invention relates to proinsulin derivatives, more specifically inverted proinsulin derivatives, that have applications in the preparation of insulin. The proinsulin derivatives of the invention also have potential applications as therapeutic moieties in their own right.

BACKGROUND TO THE INVENTION

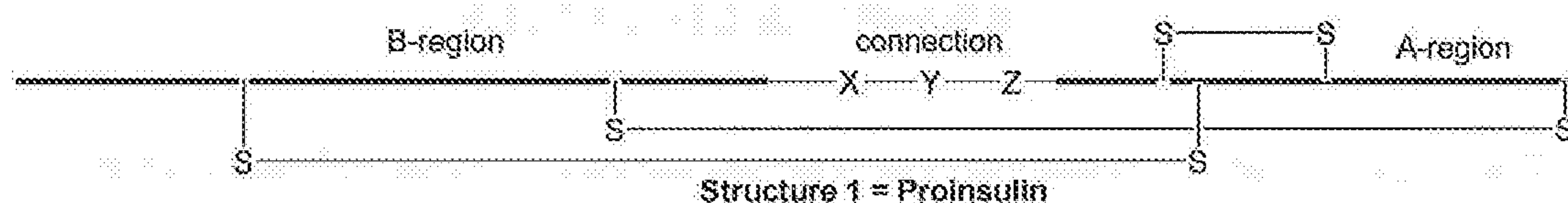
Insulin and its derivatives are the most important drugs for the treatment of diabetes, with annual sales of over 20 billion and with a steadily increasing market.

Insulin is a peptide hormone secreted by the β -cells of the pancreas. It consists of two peptide chains, A and B, which are linked by two intermolecular disulphide bonds. The A-chain also contains an additional intramolecular disulfide bond. Human insulin has the structure 2.



Insulins are typically manufactured on a large scale (hundreds of kg) starting from proinsulins which are recombinantly produced. Subsequently proinsulin is folded and the connecting C-peptide is removed enzymatically using trypsin and carboxypeptidase B to yield the mature insulin peptide.

More specifically, insulin is produced as a single-chain precursor, preproinsulin, which consists of a propeptide of 24 amino acid followed by proinsulin containing 86 amino acids. The sequence of the propeptide is Prepeptide-[B-chain]-Arg-Arg-[connecting peptide]-Lys-Arg-[A-chain], wherein the connecting peptide consists of 31 amino acids. After the enzymatic removal of the prepeptide, the three disulfide bonds are formed and proinsulin is produced.



The mature insulin is then liberated by enzymatic cleavage of the connecting peptide at the Arg-Arg and Lys-Arg sites.

5

Proinsulin has a 100-fold lower affinity for the insulin receptor than native insulin because the essential residues for binding to the receptor, namely the N-terminal amino function of the A-chain and the C-terminal carboxyl function of the B-chain, are blocked. The stability and solubility properties of insulin are important in the context of insulin therapeutics. A number of insulin analogues are known in the art. By way of example, single-chain insulin analogues with insulin activity are disclosed in EP1193272. These single-chain insulins have a modified C-peptide of 5-18 amino acids and are reported to have up to 42% insulin activity.

15 US 5,597,796 discloses insulin analogues in which two or more amino acid residues are substituted by Glu and/or Asp. Similarly, US 20090069216 and WO 2007/096332 disclose fast acting single chain insulins containing a modified B-chain and a connecting peptide. The resulting analogues are particularly well suited for transdermal administration. Fibrillation-resistant insulin and insulin analogues are disclosed in US
20 8,192,957. Pegylated single chain insulins are disclosed in US 2010/0216690, whereas acylated single chain insulins are disclosed in WO 2007/104738.

WO 2005/054291 discloses single chain insulin analogues wherein the A-chain and B-chains are connected by a connecting peptide of 5-11 amino acids. Likewise, WO
25 95/16708 also discloses single chain insulin analogues wherein the A-chain and B-chains are connected by a connecting peptide of 1-15 amino acids, in which the C-terminal amino acid residue is other than Lys or Arg. EP0427296 discloses human insulin precursors of the general formula B(1-29)-X_n-Y-A(1-21), wherein X_n is a peptide chain with n naturally occurring amino acid residues, where n is 0 to 33 and Y is Arg or
30 Lys. Similarly, EP0741188 discloses single chain insulin derivatives of the formula b-BP-a having significant insulin activity, where BP is a bridging peptide of 10 to 14 amino acids; these single chain insulins are reported to have insulin activity but also a high affinity to the IGF-1 receptor. Despite numerous efforts, a chemical and economically

feasible route to insulin has not yet been developed. The methods which have been applied to date include the random mixing of the linear A and B chains and their air oxidation, the mixing of the sulfonated A and B-chains, the site-directed building of the three disulfide bonds and the biomimetic folding of single-chain precursors (Sohma, Y. and Kent, S. B. *J. Am. Chem. Soc.* 2009, 131, 16313–16318; Tofteng, A. P.; Jensen, K. J.; Schaffer, L.; Hoeg-Jensen, T. *Chem-BioChem* 2008, 9, 2989-2996).

A-C-B 'inverted'-proinsulins with a connecting C-peptide are also known to yield insulin derivatives. By way of example, EP0518587 discloses A-C-B proinsulin derivatives wherein the linker peptide C is a fragment of the formula $X_1-X_2-P-X_3-X_4$, wherein X_1 to X_4 are basic amino acids and P is peptide from 4 to 35 amino acids which does not contain a cysteine residue (see also W. F. Heath *et al*, The Journal of Biological Chemistry, Vol. 276, No. 1, p419-425). It is generally understood that the C-peptide of a natural or inverse proinsulin must be of a minimum length in order to have the required flexibility to fold correctly to the natural mature proinsulin or reverse proinsulin.

The present invention seeks to provide alternative proinsulin derivatives that are useful in the preparation of insulin, and also as therapeutic moieties in their own right. In particular, the invention seeks to provide proinsulin derivatives that give rise to one or more advantages in terms of solubility, activity, yield, purity and/or ease of synthesis.

STATEMENT OF INVENTION

The present invention relates to proinsulin derivatives, more specifically inverted proinsulin derivatives of formula A-C-B.

Thus, a first aspect of the invention relates to a polypeptide compound of formula (I):



wherein:

- A is the A chain of insulin or a functional derivative or variant thereof;
- B is the B chain of insulin or a functional derivative or variant thereof;
- C is a connecting peptide of the formula:



wherein:

each X_1 and X_3 is independently a basic amino acid;

5 each X_2 is independently a natural or unnatural amino acid;

p is 1 or 2;

q is 0, 1 or 2;

n is 0, 1, 2 or 3.

10 The presently claimed inverted proinsulin derivatives differ from those known in the art by virtue of the fact that they contain a shorter connecting peptide (they are "super mini" insulins). Surprisingly, the Applicants have found that C-peptides containing 7 amino acids or less (including those with only a single amino acid such as the inverse proinsulin A-Arg-B) and their corresponding protected or partially protected derivatives
15 are useful proinsulin precursors.

DETAILED DESCRIPTION

The presently claimed peptide compounds were developed with a view to improving the solid phase chemical synthesis of proinsulins.

20

Ideally, the peptide must not contain problematic peptide synthesis regions, for example, those that give rise to the formation of β -turns and β -sheets during the on resin peptide chain elongation. The presently claimed peptides are inverted proinsulin derivatives in the order A-C-B, rather than the natural B-C-A order. These inverted
25 proinsulin derivatives have been shown to be synthetically easier to prepare than the corresponding B-C-A proinsulins.

Preferably, the inverted proinsulins of the invention comprise residues in the C-peptide (for example, as amino acid X_2) which disrupt the formation of β -sheets and β -turns.

30

Suitable residues in this regard include Pro, Hyp and pseudoproline. Insertion of these residues into the C-peptide allows the effective synthesis of inverse proinsulins.

35

Preferably, the C peptide contains residues that improve the solubility in solvents used in the purification step. Suitable residues in this regard include Pro, Hyp, basic amino acids such as Arg or Lys, or acidic and hydrophilic amino acids, such as Glu and Ser.

To obtain proinsulins of high purity, smaller protected peptides can be condensed in solution or on solid-phase. Preferably, in order to avoid racemisation in the condensation reactions, the C-terminal amino acids of the fragments are typically selected from Gly, Pro, β -Ala and amino acids which contain an oligo or poly-glycol part in their structure, for example the $-\text{NH}-(\text{CH}_2\text{CH}_2\text{O})_{n'}-\text{CO}-$ structural element, where n' is an integer, for example 1 to 10, more preferably 1 to 5.

Preferably, to facilitate excision of the C-peptide, one or more basic amino acids are positioned at the C-terminus of the A-chain and the amino terminus of the B-chain, i.e. $(\text{X}_1)_p$ where p is 1 or 2, and $(\text{X}_3)_q$, where q is 1 or 2.

As used herein, the term "variant" includes any variation wherein wherein (a) one or more amino acid residues are replaced by a naturally or non-naturally occurring amino acid residue (b) the order of two or more amino acid residues is reversed, (c) one, two or three amino acids are deleted, (d) a spacer group is present between any two amino acid residues, (e) one or more amino acid residues are in peptoid form, (f) the (N-C-C) backbone of one or more amino acid residues of the peptide has been modified, (g) one or more additional amino acids are present at the N-terminus and/or the C-terminus, or any of (a)-(g) in combination. Preferably, the variants arise from one of (a), (b) or (c).

More preferably, one to five, or one to four, or one to three amino acids residues are substituted by one or more other amino acid residues. Even more preferably, two amino acid residues are substituted by another amino acid residue. More preferably still, one amino acid residue is substituted by another amino acid residue. Preferably, the substitution is homologous.

Homologous substitution (substitution and replacement are both used herein to mean the interchange of an existing amino acid residue, with an alternative residue) may occur i.e. like-for-like substitution such as basic for basic, acidic for acidic, polar for polar etc. Non-homologous substitution may also occur i.e. from one class of residue to another or alternatively involving the inclusion of unnatural amino acids such as ornithine, diaminobutyric acid ornithine, norleucine ornithine, pyridylalanine, thienylalanine, naphthylalanine and phenylglycine, a more detailed list of which appears below. More than one amino acid residue may be modified at a time.

As used herein, amino acids are classified according to the following classes;

basic; H, K, R

acidic; D, E

non-polar; A, F, G, I, L, M, P, V, W

5 polar; C, N, Q, S, T, Y,

(using the internationally accepted single letter amino acid notation) and homologous

and non-homologous substitution is defined using these classes. Thus, homologous

substitution is used to refer to substitution from within the same class, whereas non-

homologous substitution refers to substitution from a different class or by an unnatural

10 amino acid.

Suitable spacer groups that may be inserted between any two amino acid residues of the carrier moiety include alkyl groups such as methyl, ethyl or propyl groups in addition to amino acid spacers such as glycine or β -alanine residues. A further form of variation,

15 type (e), involving the presence of one or more amino acid residues in peptoid form, will

be well understood by those skilled in the art. For the avoidance of doubt, "the peptoid

form" is used to refer to variant amino acid residues wherein the α -carbon substituent

group is on the residue's nitrogen atom rather than the α -carbon. Processes for

preparing peptides in the peptoid form are known in the art, for example, Simon RJ *et*

20 *al.*, PNAS (1992) 89(20), 9367-9371 and Horwell DC, Trends Biotechnol. (1995) 13(4),

132-134. Type (f) modification may occur by methods such as those described in

International Application PCT/GB99/01855 (WO 99/64574).

It is preferable for amino acid variation, preferably of type (a) or (b), to occur

25 independently at any position. As mentioned above more than one homologous or non-

homologous substitution may occur simultaneously. Further variation may occur by

virtue of reversing the sequence of a number of amino acid residues within a sequence.

In one embodiment the replacement amino acid residue is selected from the residues of

30 alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine,

histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine,

threonine, tryptophan, tyrosine, and valine.

The replacement amino acid residue may additionally be selected from unnatural amino

35 acids as described below.

As used herein, the term "derivative" refers to insulin that has undergone chemical modification, for example, to the amino acid side chains at the N-terminus and/or the C-terminus. Preferably, the chemical modification serves to alter the absorption, distribution, metabolism, and excretion characteristics of the analogue. Semisynthetic
5 insulins were clinically used for some time based on chemical modification of animal insulins, for example Novo Nordisk enzymatically converted porcine insulin into semisynthetic 'human' insulin by removing the single amino acid that varies from the human variety, and chemically adding the human amino acid.

10 In one preferred embodiment, the insulin is chemically modified to alter its isoelectric point. Normal unmodified insulin is soluble at physiological pH. Modified derivatives of insulin have been created that have a shifted isoelectric point so that they exist in a solubility equilibrium in which most precipitates out but slowly dissolves in the bloodstream and is eventually excreted by the kidneys.

15 As used herein, the term "functional derivative of insulin" refers to any molecule that performs a similar or equivalent biological function to insulin. The functional derivative may or may not be structurally similar to insulin in terms of its chemical structure.

20 In one preferred embodiment, the single chain insulin analogue of the invention is derived from animal insulin.

The amino acid sequence of animal insulins in different mammals may be similar to human insulin (insulin human INN). However, there is considerable variability within vertebrate species. Porcine insulin has only a single amino acid variation from the
25 human variety, and bovine insulin varies by three amino acids. Both are active on the human receptor with approximately the same strength. Bovine insulin and porcine insulin were the first clinically used insulin analogues (naturally occurring, produced by extraction from animal pancreas), at the time when biosynthetic human insulin (insulin human rDNA) was not available. Insulin from sharks and some species of fish may be
30 also effective.

In another preferred embodiment, the single chain insulin analogue of the invention is derived from human insulin, or an analogue thereof. More preferably, the insulin is biosynthetic insulin (insulin human rDNA). In one preferred embodiment, the insulin is a

derivative selected from Insulin Glargin (Lantus), Insulin Lispro (Humalog), Insulin Detemir (Levemir), Insulin Aspart (Novolog), Insulin Degludec and biotinylated insulin.

Other embodiments of this invention include rabbit, monkey, horse, rat I, rat II, porcine,
 5 bovine-lamb, dog, guinea pig, chinchilla, or duck ACB-proinsulin molecules.

In one preferred embodiment, the A chain is the naturally occurring amino acid sequence of the A chain of insulin. In one preferred embodiment, the B chain is the naturally occurring amino acid sequence of the B chain of insulin. It is preferred that the
 10 amino acid sequence of the ACB-proinsulin molecule of these species be the naturally occurring amino acid sequence of the A-chain and the naturally occurring sequence of the B-chain. Other embodiments of the invention may be directed to functional analogs of the proinsulin molecule derived from the aforementioned species.

Thus, in one preferred embodiment, the A chain is a variant or functional derivative of the A chain of insulin. In another preferred embodiment, the B chain is a variant or functional derivative of the B chain of insulin. In another preferred embodiment, the A chain is a variant or functional derivative of the A chain of insulin and the B chain is a variant or functional derivative of the B chain of insulin.

20 Natural amino acids include alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine.

25 As used herein, the term "non-natural amino acid" or "unnatural amino acid" includes alpha and alpha-disubstituted amino acids, N-alkyl amino acids, lactic acid, halide derivatives of natural amino acids such as trifluorotyrosine, p-Cl-phenylalanine, p-F-phenylalanine, p-Br-phenylalanine, p-NO₂-phenylalanine, phenylglycine, sarcosine, penicillamine, D-2-methyltryptophan, phosphoserine, phosphothreonine,
 30 phosphotyrosine, p-I-phenylalanine, L-allyl-glycine, β-alanine, β-aspartic acid, β-cyclohexylalanine, citrulline, homoserine, homocysteine, pyroglutamic acid, L-α-amino butyric acid, L-γ-amino butyric acid, L-α-amino isobutyric acid, α-cyclohexylglycine, diaminobutyric acid, diaminopimelic acid, N-ε-dinitrophenyl-lysine, L-1-naphthylalanine, L-2-naphthylalanine, 3-(2-pyridyl)-L-alanine, 3-(3-pyridyl)-L-alanine, 3-(4-pyridyl)-L-
 35 alanine, N-ε-methyl-lysine, N,N-ε-dimethyl-lysine, N,N,N-ε-trimethyl-lysine, 3-

mercaptopropionic acid, L- ϵ -amino caproic acid, 7-amino heptanoic acid, 6-amino hexanoic acid L-methionine sulfone, ornithine, L-norleucine, L-norvaline, p-nitro-L-phenylalanine, L-hydroxyproline, γ -glutamic acid, γ -amino butyric acid L-thiopropine, methyl derivatives of phenylalanine (Phe) such as 4-methyl-Phe, pentamethyl-Phe, L-Phe (4-amino), L-Tyr (methyl), L-Phe (4-isopropyl), L-Tic (1,2,3,4-tetrahydroisoquinoline-3-carboxyl acid), L-diaminopropionic acid and L-Phe (4-benzyl).

Advantageously, the introduction of one or more unnatural amino acids leads to an increase in the enzymatic stability of the peptides.

10

The insulin analogues of the present invention may comprise amino acids in the L or D form, i.e. one or more residues, preferably all the residues, may be in the L or D form.

In the peptides of the invention, each of X_1 and X_3 is independently a basic amino acid. Having a basic amino acid present in the connecting peptide allows the peptide to be cleaved with trypsin and carboxypeptidase.

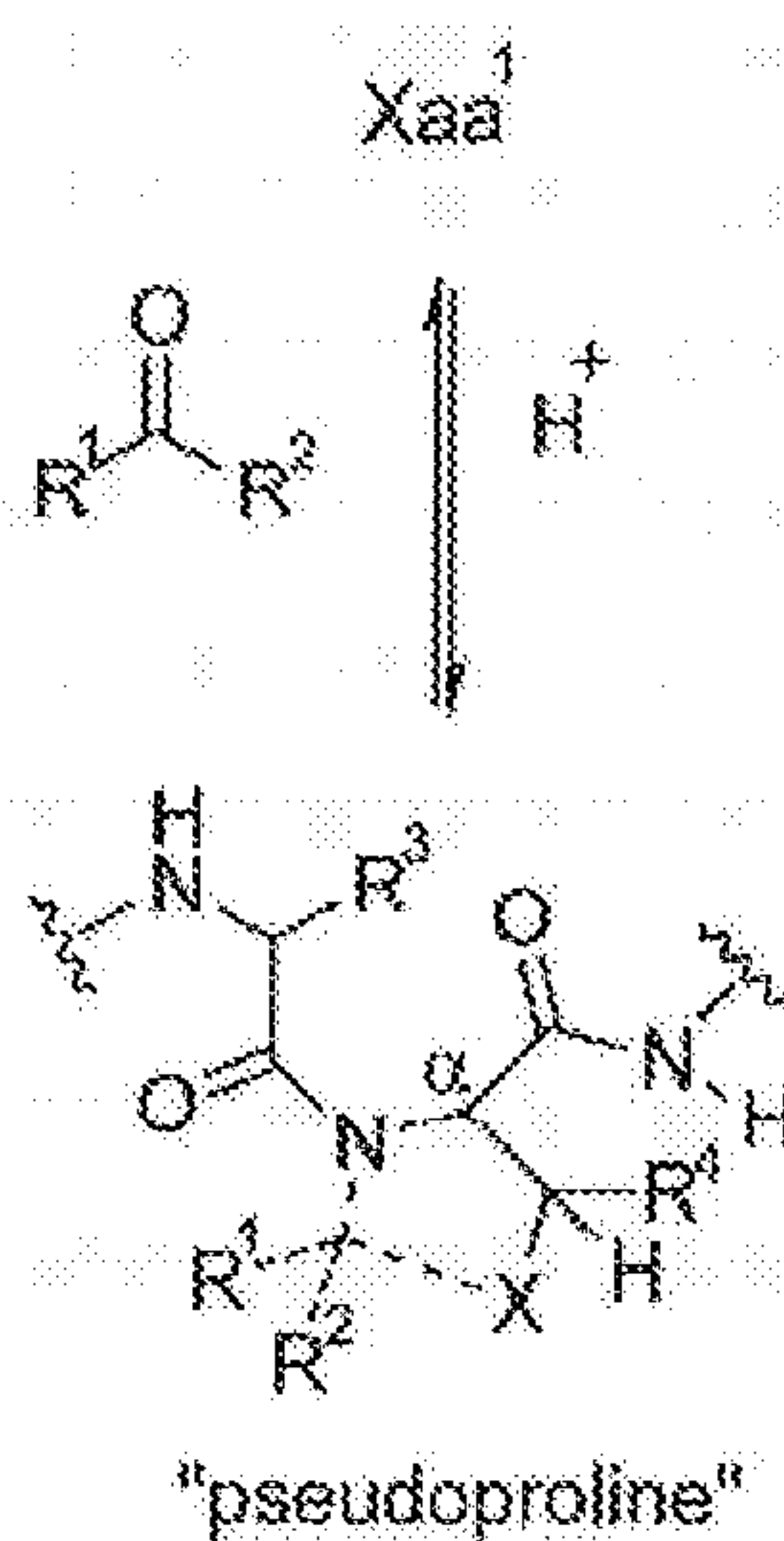
In the peptides of the invention, each X_2 is independently a natural or unnatural amino acid. Where X_2 is a cysteine residue, the skilled person would appreciate the need for protection to avoid unwanted S-S bond formation during subsequent oxidative folding. Suitable cysteine protecting groups will be familiar to the skilled person and include AcM and Trt.

In one preferred embodiment, X_2 is other than cysteine.

25

In one preferred embodiment, each X_2 is independently selected from a basic amino acid, Gly, β -Ala, Pro, Hyp, a pseudoproline, an acidic amino acid and a hydrophilic amino acid.

In one preferred embodiment, X_2 is a pseudoproline. Pseudoprolines are artificially created dipeptides that minimize aggregation during Fmoc solid phase synthesis of peptides. Pseudoprolines consist of serine- (Oxa) or threonine-derived oxazolidines [Oxa(5-Me)] and Cysteine-derived thiazolidines (THz) with Proline-like ring structures (see below).

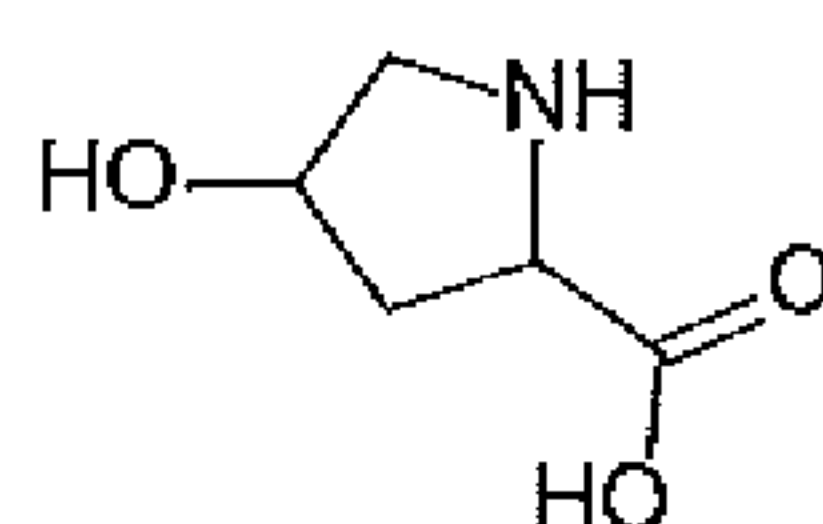


Xaa¹ = Ser, Thr, Cys

Due to the preference for a cis-amide bond with the preceding residue of C2-substituted pseudoprolines, their incorporation results in a kink conformation of the peptide backbone, thereby preventing peptide aggregation, self-association, or β -structure formation. Hence, pseudoprolines fulfil two functions simultaneously: firstly, they serve as temporary side-chain protection for Ser, Thr, and Cys, and secondly they act as solubilizing building blocks to increase solvation and coupling rates during peptide synthesis and in subsequent chain assembly.

Pseudoprolines are obtained by reacting the free amino acids with aldehydes or ketones. Pseudoproline dipeptides can be introduced in the same manner as other amino acid derivatives. Preferably the pseudoproline is derived from a Ser-X, Thr-X or CysX group, where X is a natural or unnatural amino acid. The routine use of pseudoproline (oxazolidine) dipeptides in the Fmoc solid phase peptide synthesis (SPPS) of serine- and threonine-containing peptides leads to significant improvements in quality and yield of crude products. Once the peptide is deprotected, the pseudoproline becomes a conventional dipeptide of the form X-Ser, X-Thr or X-Cys, wherein X is a natural or unnatural amino acid.

In one preferred embodiment, X₂ is Hyp. As used herein, Hyp refers to (2S,4R)-4-hydroxyproline, or L-hydroxyproline, which is a non-proteinogenic amino acid having the following structure:



In one preferred embodiment, X_2 is an acidic or hydrophilic acid. Asp and Glu are examples of acidic amino acids, whereas Ser, Cys, Asn, Gln and Thr are examples of hydrophilic amino acids include. Preferably, the acidic or hydrophilic amino acid is selected from Ser, Asp and Glu.

In one preferred embodiment, X_2 is a basic amino acid. More preferably, the basic amino acid is selected from Lys, Arg, Orn and His. Even more preferably, the basic amino acid is Arg.

In one highly preferred embodiment, each X_2 is independently selected from Gly and Arg. More preferably, X_2 is Gly.

In one preferred embodiment, X_1 and X_3 are both Arg.

In one preferred embodiment, p and q are both 1.

In another preferred embodiment, p and q are both 2.

In one preferred embodiment, n is 1 or 2.

In one particularly preferred embodiment, n is 1.

In one particularly preferred embodiment, p and q are both 2, and n is 1.

In one preferred embodiment, n and q are both 0 and p is 1, i.e. the A chain and the B chain are separated by a single basic amino acid. For this embodiment, preferably X_1 is Arg or Lys, more preferably, Arg.

In another preferred embodiment, X_1 and X_3 are both Arg, and p and q are both 2.

In another preferred embodiment, X_2 is Gly and n is 1.

In one particularly preferred embodiment, the A chain and the B chain are linked by the peptide RRGR, i.e. C is a peptide wherein X_1 is Arg, X_2 is Gly, X_3 is Arg, p and q are both 2, and n is 1.

- 5 In one preferred embodiment, A is the A chain of native insulin, preferably human insulin, or a variant thereof, wherein: (a) one or more amino acid residues are substituted by a naturally or non-naturally occurring amino acid residue, (b) the order of two or more amino acid residues is reversed, (c) one, two or three amino acids are deleted, (d) a spacer group is present between any two amino acid residues, (e) one or
10 more amino acid residues are in peptoid form, (f) the (N-C-C) backbone of one or more amino acid residues of the peptide has been modified, (g) one or more additional amino acids are present at the N-terminus and/or the C-terminus, or any of (a)-(g) in combination.
- 15 In one preferred embodiment, one to five amino acids of the native A chain are substituted, preferably, one to four, more preferably one to three, even more preferably, one or two amino acids. In one preferred embodiment, a single amino acid is substituted.
- 20 In one highly preferred embodiment, the A-chain comprises amino acids 1 to 21 of human insulin counting from the N-terminal end of the A chain.
- In one highly preferred embodiment, the A-chain consists of amino acids 1 to 21 of human insulin counting from the N-terminal end of the A chain.
- 25 In another highly preferred embodiment, the A-chain comprises amino acids 1 to 20 of human insulin counting from the N-terminal end of the A chain.
- In another highly preferred embodiment, the A-chain consists of amino acids 1 to 20 of human insulin counting from the N-terminal end of the A chain.
- 30

In one preferred embodiment, B is the B chain of native insulin, preferably human insulin, or a variant thereof, wherein: (a) one or more amino acid residues are substituted by a naturally or non-naturally occurring amino acid residue, (b) the order of
35 two or more amino acid residues is reversed, (c) one, two or three amino acids are

deleted, (d) a spacer group is present between any two amino acid residues, (e) one or more amino acid residues are in peptoid form, (f) the (N-C-C) backbone of one or more amino acid residues of the peptide has been modified, (g) one or more additional amino acids are present at the N-terminus and/or the C-terminus, or any of (a)-(g) in
 5 combination.

In one preferred embodiment, one to five amino acids of the native B chain are substituted, preferably, one to four, more preferably one to three, even more preferably, one or two amino acids. In one preferred embodiment, a single amino acid is
 10 substituted.

In one preferred embodiment, the B-chain comprises amino acids 1 to 29 of human insulin counting from the N-terminal end of the A chain.

15 In one preferred embodiment, the B-chain consists of amino acids 1 to 29 of human insulin counting from the N-terminal end of the A chain.

In one preferred embodiment, the peptide of the invention is extended by one or more amino acids at the amino and/or the carboxyl terminus. For example, in one highly
 20 preferred embodiment, the peptide comprises two additional amino acids at its C-terminus, for example, Arg-Arg.

In one preferred embodiment, the B-chain of insulin further comprises up to 20 additional natural or unnatural amino acids at the C-terminal end. Preferably, the B-
 25 chain further comprises from 1 to 10, or more preferably from 1 to 5 additional natural or unnatural amino acids at the C-terminal end. In one highly preferred embodiment, the B-chain further comprises 1, 2 or 3 additional natural or unnatural amino acids at the C-terminal end. More preferably still, the B-chain further comprises 1, 2 or 3 additional natural amino acids at the C-terminal end.

30 In one preferred embodiment, the peptide comprises a single amino acid change from the native sequence. Preferably, the single amino acid change is at the position A21, B9, B10, B12, B16, B17, B20, B25, B26, B27, B28 or B30, where A denotes the A chain of insulin, and B denotes the B chain of insulin.

35

Preferably, the single amino acid substitution is selected from the following:

Gly A21, Glu A21, hSer A21, Thr B10, Asp B25, Ser A21, Leu A21, Gly A22, Asp B10, His B25, Ala A21, Met A21, Ala A22, Arg B10, Glu B26, His A21, Tyr A21, Asp B9, Ile B12, Glu B27, Asp A21, Val A21, Asn B9, His B16, Asp B28, Thr A21, Ile A21, His B9, Gln B17, Ala B30, Gln A21, Trp A21, Glu B10, Gln B20, des-B30, Thr30-NH₂ and Ala30-NH₂.

In one preferred embodiment, the peptide comprises two amino acids change from the native sequence. Preferably, the two amino acids are at positions: A21 and B10, A21 and B10, A21 and B27, B27 and B16, B5 and B10, B12 and B13, B14 and B17, B28 and B29, A121 and B27, B10 and B30, B29 and B30, B12 and B30, B10 and B2, B10 and B28, B10 and A13, B27 and A13, B27 and A21, B27 and B1, B27 and B9, or A21 and B30.

Preferably, the two amino acid substitutions are selected from the following:

15 Ser A21 and Asp B10

Asp A21 and Lys B27

Thr A21 and Asp B10

Gly A21 and Arg B27

Ala A21 and Asp B10

20 Asp B5 and Asn B10

Thr A21 and Thr B10

Glu B12 and Gln B13

Ala A21 and Thr B10

Ser B14 and Asp B17

25 His A21 and Thr B10

Ser A21 and Arg B27

His A21 and Asp B10

Thr A21 and Arg B27

Asp A21 and Asp B10

30 Ala A21 and Arg B27

Gly A21 and Thr B10

Gly A21 and Asp B10

His A21 and Lys B27

Glu B27 and Glu B16

35 Ser A21 and Thr B10

- Lys B28 and Pro B29
- Asp A21 and Thr B10
- His A21 and Arg B27
- Gly A21 and Arg B10
- 5 Asp A21 and Arg B27
- Ser A21 and Arg B10
- Glu B12 and des B30
- Thr A21 and Arg B10
- Gly A21 and Lys B27
- 10 Gly A21 and Ala B30
- Ser A21 and Lys B27
- Ser A21 and Ala B30
- Thr A21 and Lys B27
- Thr A21 and Ala B30
- 15 Ala A21 and Lys B27
- Ala A21 and Ala B30
- des B29 and des B30
- hSer A21 and Ala B30
- Asp B10 and Ser B2
- 20 Ala A21 and Arg B10
- Asp B10 and Asp B28
- His A21 and Arg B10
- Glu B10 and Glu A13
- Asp A21 and Arg B10
- 25 Glu B27 and Ser A13
- Asp B10 and des-B30
- Glu B27 and Asp A21
- Thr B10 and des-B30
- Glu B27 and Glu B1
- 30 Arg B10 and des-B30
- Glu B27 and Asp B9

In one preferred embodiment, the peptide comprises three amino acid changes from the native sequence. Preferably, the three amino acid changes are selected from the following:

Gly A21 + Lys B27 + Gln A17
 Ser A21 + Lys B27 + Gln A17
 Thr A21 + Lys B27 + Gln A17
 Ala A21 + Lys B27 + Gln A17
 5 His A21 + Lys B27 + Gln A17
 Asp A21 + Lys B27 + Gln A17
 Gly A21 + Lys B27 + Gln B13
 Ser A21 + Lys B27 + Gln B13
 Thr A21 + Lys B27 + Gln B13
 10 Ala A21 + Lys B27 + Gln B13
 His A21 + Lys B27 + Gln B13
 Asp A21 + Lys B27 + Gln B13
 Gly A21 + Arg B27 + Gln A17
 Ser A21 + Arg B27 + Gln A17
 15 Thr A21 + Arg B27 + Gln A17
 Ala A21 + Arg B27 + Gln A17
 His A21 + Arg B27 + Gln A17
 Asp A21 + Arg B27 + Gln A17
 Gly A21 + Arg B27 + Gln B13
 20 Ser A21 + Arg B27 + Gln B13
 Thr A21 + Arg B27 + Gln B13
 Ala A21 + Arg B27 + Gln B13
 His A21 + Arg B27 + Gln B13
 Asp A21 + Arg B27 + Gln B13
 25 Asp B10 + His A8 + His B25
 Glu B10 + Glu A3 + Glu B22
 Glu B27 + Ser B5 + Asp B5
 Glu B27 + His A5 + Asp B9
 Glu B27 + Asp A21 + Asp B9
 30 des B28 + des B29 + des B30
 Gly A21 + Asp B10 + Ala B30
 Ser A21 + Asp B10 + Ala B30
 Thr A21 + Asp B10 + Ala B30
 Ala A21 + Asp B10 + Ala B30
 35 His A21 + Asp B10 + Ala B30

- Asp A21 + Asp B10 + Ala B30
 Gly A21 + Thr B10 + Ala B30
 Ser A21 + Thr B10 + Ala B30
 Thr A21 + Thr B10 + Ala B30
 5 Ala A21 + Thr B10 + Ala B30
 His A21 + Thr B10 + Ala B30
 Asp A21 + Thr B10 + Ala B30
 Gly A21 + Arg B10 + Ala B30
 Ser A21 + Arg B10 + Ala B30
 10 Thr A21 + Arg B10 + Ala B30
 Ala A21 + Arg B10 + Ala B30
 His A21 + Arg B10 + Ala B30
 Asp A21 + Arg B10 + Ala B30
 Gly A21 + Asp B10 + des B30
 15 Ser A21 + Asp B10 + des B30
 Thr A21 + Asp B10 + des B30
 Ala A21 + Asp B10 + des B30
 His A21 + Asp B10 + des B30
 Asp A21 + Asp B10 + des B30
 20 Gly A21 + Thr B10 + des B30
 Ser A21 + Thr B10 + des B30
 Thr A21 + Thr B10 + des B30
 Ala A21 + Thr B10 + des B30
 His A21 + Thr B10 + des B30
 25 Asp A21 + Thr B10 + des B30
 Gly A21 + Arg B10 + des B30
 Ser A21 + Arg B10 + des B30
 Thr A21 + Arg B10 + des B30
 Ala A21 + Arg B10 + des B30
 30 His A21 + Arg B10 + des B30
 Asp A21 + Arg B10 + des B30
 Thr B10 + Glu B28 + Pro B29
 Arg B10 + Glu B28 + Pro B29
 Asp B10 + Lys B28 + Pro B29
 35 Thr B10 + Lys B28 + Pro B29

Arg B10 + Lys B28 + Pro B29
 Gly A21 + Glu B28 + Pro B29
 Ser A21 + Glu B28 + Pro B29
 Thr A21 + Glu B28 + Pro B29
 5 Ala A21 + Lys B28 + Pro B29
 His A21 + Lys B28 + Pro B29
 Glu B28 + Pro B29 + Ala B30
 Glu B28 + Pro B29 + des B30
 Asp A21 + Lys B28 + Pro B29
 10 Lys B28 + Pro B29 + Ala B30
 Lys B28 + Pro B29 + des B30
 Arg B27 + Gly A21 + Thr B30NH₂

In one preferred embodiment, the peptide comprises four amino acid changes from the
 15 native sequence. Preferably, the four amino acid changes are selected from the
 following:

Ser A21 + Arg B27 + Gln A17 + Gln B13
 Thr A21 + Arg B27 + Gln A17 + Gln B13
 Ala A21 + Arg B27 + Gln A17 + Gln B13
 20 Asp A21 + Arg B27 + Gln A17 + Gln B13
 His A21 + Arg B27 + Gln A17 + Gln B13
 Glu B10 + His A8 + His B4 + His B27
 Gly A21 + Lys B27 + Gln A17 + Gln B13
 Ser A21 + Lys B27 + Gln A17 + Gln B13
 25 Thr A21 + Lys B27 + Gln A17 + Gln B13
 Ala A21 + Lys B27 + Gln A17 + Gln B13
 Asp A21 + Lys B27 + Gln A17 + Gln B13
 His A21 + Lys B27 + Gln A17 + Gln B13
 Gly A21 + Arg B27 + Gln A17 + Gln B13
 30 des B27 + des B28 + des B29 + des B30
 Gly A21 + Asp B10 + Glu B28 + Pro B29
 Ser A21 + Asp B10 + Glu B28 + Pro B29
 Thr A21 + Asp B10 + Glu B28 + Pro B29
 Ala A21 + Asp B10 + Glu B28 + Pro B29
 35 His A21 + Asp B10 + Glu B28 + Pro B29

- Asp A21 + Asp B10 + Glu B28 + Pro B29
 Gly A21 + Thr B10 + Glu B28 + Pro B29
 Ser A21 + Thr B10 + Glu B28 + Pro B29
 Thr A21 + Thr B10 + Glu B28 + Pro B29
 5 Ala A21 + Thr B10 + Glu B28 + Pro B29
 His A21 + Thr B10 + Glu B28 + Pro B29
 Asp A21 + Thr E10 + Glu B28 + Pro B29
 Ala A21 + Arg B10 + Glu B28 + Pro B29
 Ala A21 + Asp B10 + Lys B28 + Pro B29
 10 His A21 + Asp B10 + Lys B28 + Pro B29
 Asp A21 + Asp B10 + Lys B28 + Pro B29
 Gly A21 + Arg B10 + Glu B28 + Pro B29
 Ser A21 + Arg B10 + Glu B28 + Pro B29
 Thr A21 + Arg B10 + Glu B28 + Pro B29
 15 Gly A21 + Thr B10 + Lys B28 + Pro B29
 Ser A21 + Thr B10 + Lys B28 + Pro B29
 His A21 + Arg B10 + Glu B28 + Pro B29
 Asp A21 + Arg B10 + Glu B28 + Pro B29
 Gly A21 + Asp B10 + Lys B28 + Pro B29
 20 Ser A21 + Asp B10 + Lys B29 + Pro B29
 Thr A21 + Asp B10 + Lys B28 + Pro B29
 Ser A21 + Arg B10 + Lys B28 + Pro B29
 Thr A21 + Arg B10 + Lys B28 + Pro B29
 Ala A21 + Arg B10 + Lys B28 + Pro B29
 25 His A21 + Arg B10 + Lys B28 + Pro B29
 Asp A21 + Arg B10 + Lys B28 + Pro B29
 Gly A21 + Glu B28 + Pro B29 + Ala B30
 Ser A21 + Glu B28 + Pro B29 + Ala B30
 Thr A21 + Lys B28 + Pro B29 + Ala B30
 30 Ala A21 + Lys B28 + Pro B29 + Ala B30
 His A21 + Lys B28 + Pro B29 + Ala B30
 Asp A21 + Lys B28 + Pro B29 + Ala B30
 Thr A21 + Thr B10 + Lys B28 + Pro B29
 Ala A21 + Thr B10 + Lys B28 + Pro B29
 35 His A21 + Thr B10 + Lys B28 + Pro B29

Asp A21 + Thr B10 + Lys B28 + Pro B29
 Gly A21 + Arg B10 + Lys B28 + Pro B29
 Gly A21 + Glu B28 + Pro B29 + des B30
 Ser A21 + Glu B28 + Pro B29 + des B30
 5 Thr A21 + Glu B28 + Pro B29 + des B30
 Ala A21 + Glu B28 + Pro B29 + des B30
 His A21 + Glu B28 + Pro B29 + des B30
 Thr B10 + Glu B28 + Pro B29 + des B30
 Arg B10 + Glu B28 + Pro B29 + des B30
 10 Asp B10 + Lys B28 + Pro B29 + des B30
 Thr B10 + Lys B28 + Pro B29 + des B30
 Arg B10 + Lys B28 + Pro B29 + des B30
 des B27 + des B29 + des B29 + des B30
 Asp A21 + Glu B28 + Pro B29 + des B30
 15 Gly A21 + Lys B28 + Pro B29 + des B30
 Ser A21 + Lys B28 + Pro B29 + des B30
 Thr A21 + Lys B28 + Pro B29 + des B30
 Ala A21 + Lys B28 + Pro B29 + des B30
 His A21 + Lys B28 + Pro B29 + des B30
 20 Asp A21 + Lys B28 + Pro B29 + des B30
 Asp B10 + Glu B28 + Pro B29 + Ala B30
 Thr B10 + Glu B28 + Pro B29 + Ala B30
 Arg B10 + Glu B28 + Pro B29 + Ala B30
 Asp B10 + Lys B28 + Pro B29 + Ala B30
 25 Thr A21 + Glu B28 + Pro B29 + Ala B30
 His A21 + Glu B28 + Pro B29 + Ala B30
 Asp A21 + Glu B28 + Pro B29 + Ala B30
 Gly A21 + Lys B28 + Pro B29 + Ala B30
 Ser A21 + Lys B28 + Pro B29 + Ala B30
 30 Thr B10 + Lys B28 + Pro B29 + Ala B30
 Arg B10 + Lys B28 + Pro B29 + Ala B30
 Asp B10 + Glu B28 + Pro B29 + des B30

In one preferred embodiment, the peptide comprises five amino acid changes from the
 35 native sequence. Preferably, the five amino acid changes are at the following positions:

B26, B27, B28, B29 and B30, for example, des B26 + des B27 + des B28 + des B29 + des B30.

In one preferred embodiment, C is a peptide of the sequence shown in SEQ ID NO:1,

5

RRGRR [SEQ ID NO:1]

In one preferred embodiment, A comprises a peptide of the sequence shown in SEQ ID NO:2,

10

¹GIVEQCCTSICSLYQLENYCG²¹ [SEQ ID NO:2]

wherein each amino acid is unprotected, or optionally protected, for example, where the amino acid side chain contains a functional group. Preferred protecting groups include acid cleavable protecting groups such as ^tBu, Acn, O^tBu, Trt, Mmt, Mtt and Pbf.

15

In a more preferred embodiment, A consists of a peptide of SEQ ID NO:2, wherein each amino acid is optionally protected.

In another preferred embodiment, A comprises a peptide of the sequence shown in

20

SEQ ID NO:8,

¹GIVEQCCTSICSLYQLENYCN²¹ [SEQ ID NO:8]

wherein each amino acid is unprotected, or optionally protected, for example, where the amino acid side chain contains a functional group. Preferred protecting groups include acid cleavable protecting groups such as ^tBu, Acn, O^tBu, Trt, Mmt, Mtt and Pbf.

25

In a more preferred embodiment, A consists of a peptide of SEQ ID NO:8, wherein each amino acid is optionally protected.

In one preferred embodiment, B comprises a peptide of sequence shown in SEQ ID NO:3,

30

¹FVNQHLCGSHLVEALYLVCGERGFFYTPKT³⁰ [SEQ ID NO:3]

wherein each amino acid is unprotected, or optionally protected, for example, where the amino acid side chain contains a functional group. Preferred protecting groups include acid cleavable protecting groups such as ^tBu, Acn, O^tBu, Trt, Mmt, Mtt and Pbf.

- 5 In a more preferred embodiment, B consists of a peptide of SEQ ID NO:3, wherein each amino acid is optionally protected.

In one preferred embodiment, the polypeptide compound is of SEQ ID NO:4:

10 SEQ ID NO:4

¹GIVEQCCTSI~~C~~SLYQLENYCGRRGRRFVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

wherein each amino acid is unprotected, or optionally protected, for example, where the amino acid side chain contains a functional group. Preferred protecting groups include acid cleavable protecting groups such as ^tBu, Acn, O^tBu, Trt, Mmt, Mtt and Pbf.

15

In one preferred embodiment, the polypeptide compound is selected from SEQ ID NOS: 4-7,

SEQ ID NO:4

20 ¹GIVEQCCTSI~~C~~SLYQLENYCGRRGRRFVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

SEQ ID NO:5

¹GIVEQCCTSI~~C~~SLYQLENYCGRRGRRFVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

25 SEQ ID NO:6

¹GIVEQCCCTSI~~C~~SLYQLENYCGRRGRRFVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

SEQ ID NO:7

¹GIVEQCCCTSI~~C~~SLYQLENYCGRRGRRFVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

30

wherein in SEQ ID NOS: 4-7, C denotes a cysteine protected with a protecting group and C denotes an unprotected cysteine.

In one preferred embodiment, the cysteine protecting group is Acn or Trt.

35

SH SH SH SH
| | | |
GIVEQCCTSLYQLENYCGRRGRFVNQHLCSHLVEALYLVCGERGFFYTPKT
| |
Acm Acm
(I.1a)

23

In one preferred embodiment, the invention relates to a process for preparing a polypeptide compound as defined above using solid-phase peptide synthesis. The accompanying examples section describes such a synthesis in more detail.

5

Syntheses can be performed using the Fmoc/tBu-protection scheme using the 2-chlorotrityl resin as the solid support. For the protection of the Cys-residues, Trt or the Trt/Acm protecting groups can be used.

10 In alternative embodiment, the peptide compounds of the invention can be prepared recombinantly by methods known in the art. For further details, see for example, the teachings of EP0518587A.

One aspect of the invention therefore relates to a DNA sequence encoding a peptide
15 compound as defined above.

Another aspect of the invention relates to a recombinant DNA vector comprising a DNA sequence as described above.

20 Another aspect of the invention relates to a process for the recombinant preparation of a polypeptide compound as defined above, said method comprising the steps of:

- (i) constructing a DNA sequence encoding a polypeptide compound as defined above,
- (ii) incorporating said DNA sequence into a suitable vector containing a promoter-operator region functional in a host cell,
- 25 (iii) orienting said DNA sequence in said vector so as to achieve transcription and translation of said DNA sequence and further that said DNA sequence is under the transcriptional control of said promoter-operator region,
- (iv) transforming said host cell with said vector,
- 30 (v) culturing said transformed host cell under conditions appropriate so as to induce transcription and translation of said gene, and
- (vi) recovering and purifying the polypeptide product encoded by said DNA sequence.

35 Another aspect of the invention relates to a process for preparing insulin or a derivative

thereof, said method comprising the steps of:

- (i) constructing a DNA sequence encoding a polypeptide compound as defined above;
- (ii) incorporating said DNA sequence into a suitable vector containing a promoter-operator region functional in a host cell,
- (iii) orienting said DNA sequence in said vector so as to achieve transcription and translation of said DNA sequence and further that said DNA sequence is under the transcriptional control of said promoter-operator region,
- (iv) transforming said host cell with said vector,
- (v) culturing said transformed host cell under conditions appropriate so as to induce transcription and translation of said gene, and
- (vi) recovering and purifying the polypeptide product encoded by said DNA sequence.
- (vii) cleaving said polypeptide product encoded by said gene to excise the C-peptide.

PROCESS FOR OXIDATIVE FOLDING OF PROINSULIN DERIVATIVES AND INTERMEDIATES THEREIN

A further aspect of the invention relates to a process for preparing insulin, or a derivative thereof, said process comprising the steps of:

- (i) preparing a polypeptide compound according to the invention, and
- (ii) cleaving said polypeptide to excise the C-peptide.

The folding of the obtained inverse "super mini" proinsulins described herein can be performed either in one step or in two steps. The removal of the C-peptide can be performed in a conventional manner by trypsin and/or carboxypeptidase B. The accompanying examples section describes the synthesis of an ACB-proinsulin derivative with the side-chains of the Cys residues protected either with 6 Trt, or 2 Acn and 4 Trt groups in the three possible combinations. Their folding efficiency was then compared and the nature of the resulting products analysed by HPLC.

The results demonstrated that A-C-B proinsulin peptides can be prepared in high purity and yield. The use of a two-step oxidation procedure was shown to be particularly effective. In this way, folded, single-chain insulin precursors can be transformed into bioactive two-chain insulin by the enzymatic removal of the chain-bridging C-peptide.

A further aspect of the invention relates to a process for preparing insulin, or a derivative thereof, said process comprising the steps of:

- (i) preparing an A-C-B polypeptide, and
- (ii) cleaving said polypeptide to excise the C-peptide.

5

In a preferred embodiment, A and B are defined as for formula (I), and C is the connecting peptide of insulin, or a variant thereof. In another preferred embodiment, C is a peptide as defined in formula (I), i.e. $(X_1)_p-(X_2)_n-(X_3)_q$.

10 In one preferred embodiment, the cysteine residue in the 7-position of the A chain of insulin and the cysteine residue in the 7-position of the B chain of insulin are each protected with a protecting group.

In one preferred embodiment, the cysteine residues in the 6-, 11- and 20-positions of the A chain are unprotected, and the cysteine residue in the 19-position of the B chain is
15 unprotected.

In another preferred embodiment, the cysteine residues in the 6-, 11- and 20-positions of the A chain and the 19-position of the B chain are unprotected, and the cysteine residue in the 7-position of the A chain and the 7-position of the B chain are each
20 protected with a protecting group.

Preferably, the protecting group is Acn or Trt, more preferably Acn.

In one particularly preferred embodiment, the process comprises a stepwise oxidative
25 folding of the peptide compound, said process comprising the steps of:

- (i) oxidising a peptide as defined above to form a first intermediate;
- (ii) oxidising said first intermediate to form a second intermediate;
- (iii) cleaving said second intermediate to excise the C-peptide.

30 Preferably, the peptide is of formula I.1 or I.1a as defined above.

In one preferred embodiment, step (i), the first oxidation step, comprises treating the peptide with DMSO to form a first intermediate. The remaining protecting groups are then removed and the resulting intermediate is then subjected to a further (second)
35 oxidation step to form a second intermediate. The protecting groups of the intermediate

can be removed by any suitable method. Thus, incorporating orthogonal cysteine protecting groups in the starting peptide allows the oxidative folding process to be dictated in a preferential manner.

- 5 More preferably, step (i) comprises treatment with DMSO in Gly buffer, more preferably 20 % DMSO in 0.1M Gly buffer. Preferably, the pH of the mixture is greater than 9, more preferably, greater than 10, more preferably still, about 10.5. Preferably, the oxidation takes place over a period of at least 4 hours, more preferably at least 8 hours, more preferably at least 21 hours, even more preferably, 1 day. Preferably, the reaction
10 is carried out at room temperature.

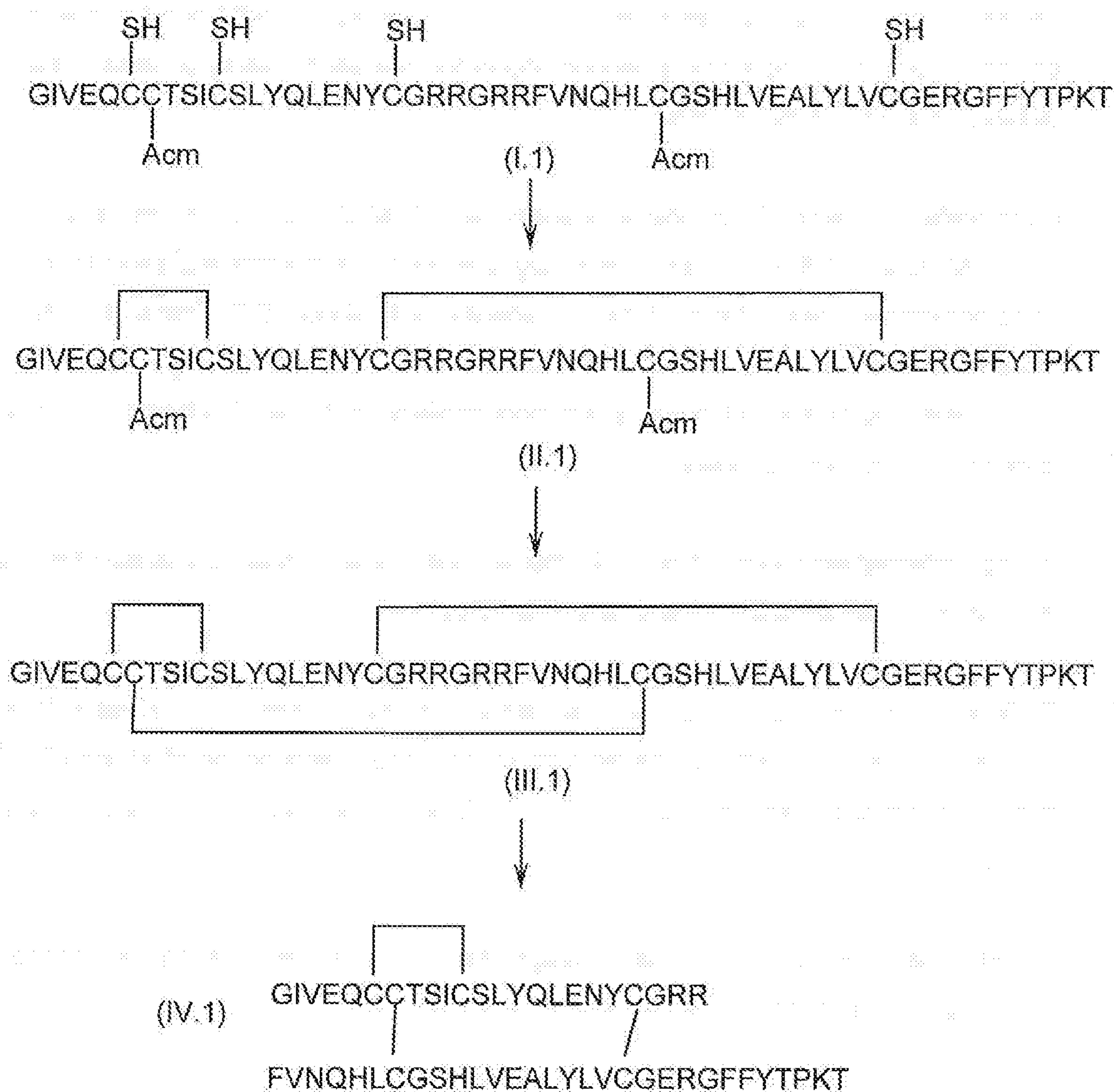
In one preferred embodiment, step (ii), the second oxidation step, comprises treating the first intermediate with iodine in acetic acid/water.

- 15 Preferably, the second oxidation step is carried out using iodine in AcOH/water in a ratio of about 4:1. Preferably, the reaction takes place over a period of at least 30 minutes, more preferably, at least 1 hour. Preferably, the reaction is carried out at room temperature.
- 20 In one preferred embodiment, the cleavage step (iii) comprises treating the second intermediate with trypsin.

Preferably, the cleavage of the C peptide is carried out at room temperature.

- Preferably, the reaction is carried out using trypsin in tris-buffer. Preferably the pH is
25 greater than 7.5, more preferably about 8. Preferably, the reaction is carried out over a period of about 5 minutes.

In one particularly preferred embodiment, the process of the invention is as shown below in Scheme 1:



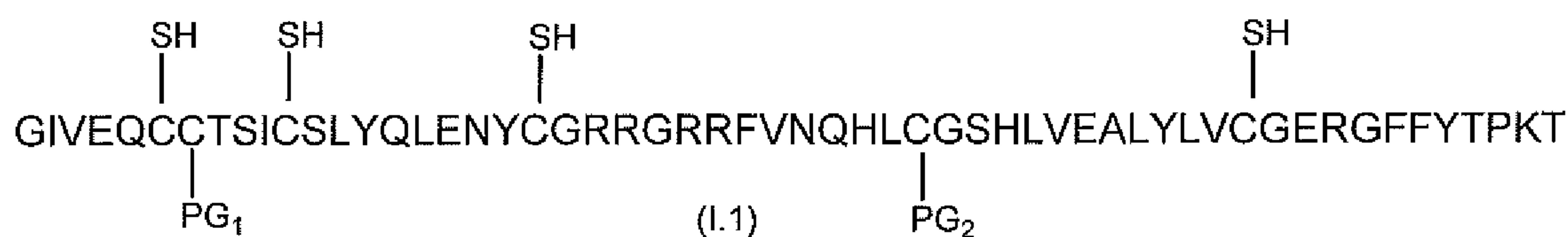
Scheme 1: Two step oxidative folding

- 5 Another aspect of the invention relates to the use of a peptide compound as defined above in the preparation of insulin, or a derivative thereof.

Another aspect of the invention relates to intermediate polypeptide compounds in the oxidative folding process.

10

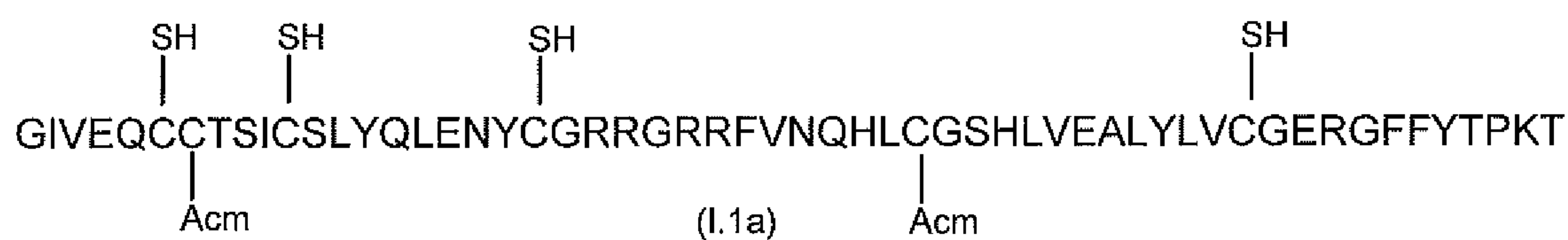
Thus, in one embodiment, the invention relates to a polypeptide of formula (I.1):



wherein PG_1 and PG_2 are each independently a cysteine protecting group, preferably Acn or Trt, more preferably Acn.

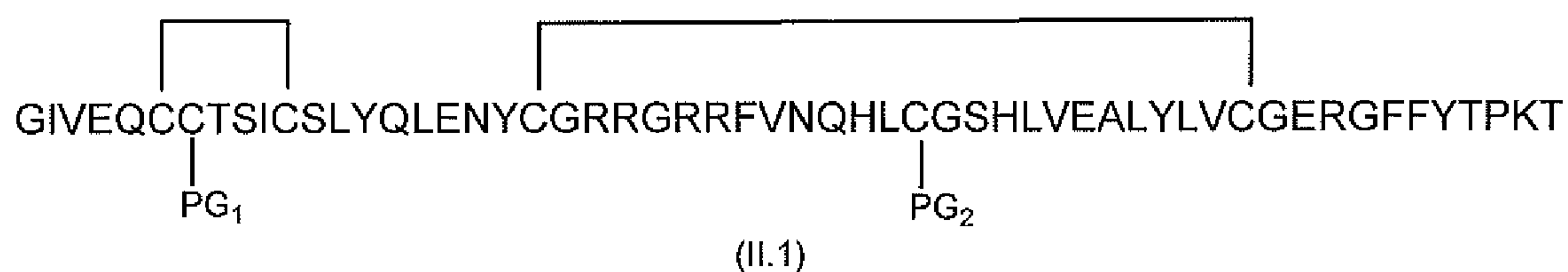
5

Thus, in one preferred embodiment, the invention relates to a polypeptide compound of formula (I.1a):



10

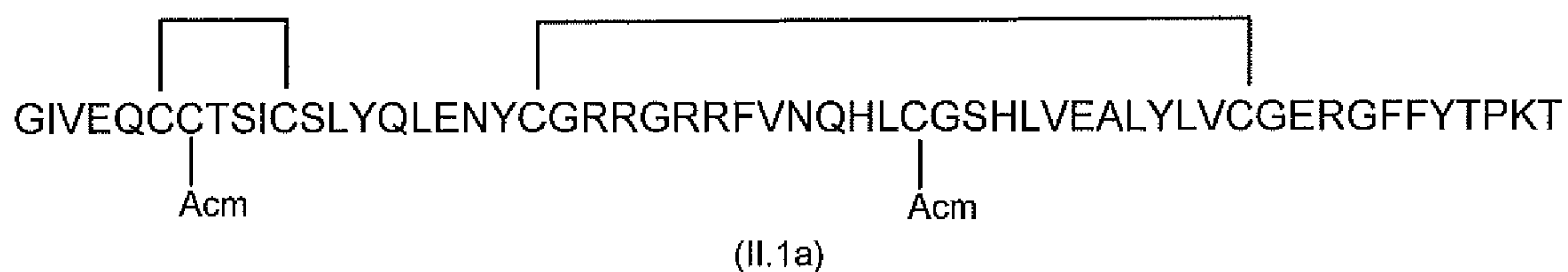
In another preferred embodiment, the invention relates to a polypeptide of formula (II.1):



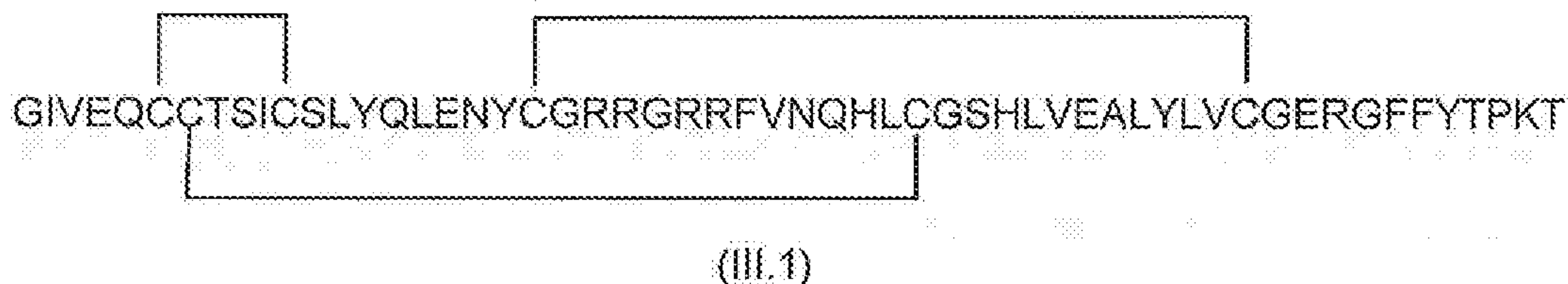
15 wherein PG_1 and PG_2 are each independently a cysteine protecting group, preferably Acn or Trt, more preferably Acn.

Thus, in one preferred embodiment, the invention relates to a peptide compound of formula (II.1a):

20



In another embodiment, the invention relates to a peptide compound of formula of formula (III.1):



- 5 Peptides of formula I.1, I.1a, II.1 and II.1a are useful intermediates in the preparation of insulin.

PHARMACEUTICAL COMPOSITIONS

- One aspect of the invention relates to a pharmaceutical composition comprising a
 10 proinsulin derivative of the invention admixed with a pharmaceutically acceptable diluent, excipient or carrier, or a mixture thereof. Even though the proinsulin derivatives of the present invention (including their pharmaceutically acceptable salts, esters and pharmaceutically acceptable solvates) can be administered alone, they will generally be administered in admixture with a pharmaceutical carrier, excipient or diluent, particularly
 15 for human therapy. The pharmaceutical compositions may be for human or animal usage in human and veterinary medicine.

- For example, compounds of the invention can be admixed with conventional pharmaceutical carriers and excipients and used in the form of tablets, capsules, elixirs,
 20 suspensions, syrups, wafers and the like. The compositions comprising ACB-proinsulin compounds will typically contain from about 0.1 to 90% by weight of the active compound, and more generally from about 10 to 30%. The compositions may contain common carriers and excipients such as corn starch or gelatin, lactose, sucrose, microcrystalline cellulose, kaolin, mannitol, dicalcium phosphate, sodium chloride, and
 25 alginic acid.

- Disintegrators commonly used in the formulations of this invention include croscarmellose, microcrystalline cellulose, corn starch, sodium starch, glycolate and alginic acid.

Examples of suitable excipients for the various different forms of pharmaceutical compositions described herein may be found in the "Handbook of Pharmaceutical Excipients, 2nd Edition, (1994), Edited by A Wade and PJ Weller.

- 5 Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R. Gennaro edit. 1985).

10 Examples of suitable carriers include lactose, starch, glucose, methyl cellulose, magnesium stearate, mannitol, sorbitol and the like. Examples of suitable diluents include ethanol, glycerol and water.

The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The
15 pharmaceutical compositions may comprise as, or in addition to, the carrier, excipient or diluent any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), solubilising agent(s).

20 Examples of suitable binders include starch, gelatin, natural sugars such as glucose, anhydrous lactose, free-flow lactose, beta-lactose, corn sweeteners, natural and synthetic gums, such as acacia, tragacanth or sodium alginate, carboxymethyl cellulose and polyethylene glycol.

25 Tablet binders that can be included are acacia, methyl cellulose, sodium carboxymethyl cellulose, polyvinylpyrrolidone (Povidone), hydroxypropyl methylcellulose, sucrose, starch and ethylcellulose.

30 Examples of suitable lubricants include sodium oleate, sodium stearate, magnesium stearate or other metallic stearates, sodium benzoate, sodium acetate, sodium chloride, stearic acid, silicone fluid, talc, waxes, oils and colloidal silica and the like.

Preservatives, stabilizers, dyes and even flavoring agents may be provided in the pharmaceutical composition. Examples of preservatives include sodium benzoate, sorbic acid and esters of p-hydroxybenzoic acid. Antioxidants and suspending agents
35 may be also used. Suitable flavoring agents include peppermint, oil of wintergreen,

cherry flavoring, or the like. It may be desirable to add a coloring agent to make the dosage form more attractive in appearance or to help identify the product.

For intravenous (IV) use, a water soluble form of compounds of the invention can be dissolved in one of the commonly used intravenous fluids and administered by infusion. Such fluids, for example, physiological saline, Ringer's solution or 5% dextrose solution can be used.

For intramuscular preparations, a sterile formulation of a suitable soluble salt form of the compounds of the invention, for example the hydrochloride salt, can be dissolved and administered in a pharmaceutical diluent such as pyrogen-free water (distilled), physiological saline or 5% glucose solution. A suitable insoluble form of the compound may be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, e.g. an ester of a long chain fatty acid such as ethyl oleate.

For oral use, a sterile formulation of a suitable salt form of ACB-proinsulin, for example, the hydrochloride salt, formulated in a diluent such as distilled or deionized water, is particularly useful.

Alternatively, the unit dosage form of the compound can be solution of the compound, preferably in its salt form, in a suitable diluent in sterile hermetically sealed ampoules. The concentration of the compound in the unit dosage may vary, e.g. from about 1% to about 50% depending on the particular form of the compound and its solubility and the dose desired by the physician.

SALTS/ESTERS

The proinsulin derivatives of the present invention can be present as salts or esters, in particular pharmaceutically acceptable salts or esters.

Pharmaceutically acceptable salts of the proinsulin derivatives of the invention include suitable acid addition or base salts thereof. A review of suitable pharmaceutical salts may be found in Berge et al, J Pharm Sci, 66, 1-19 (1977). Salts are formed, for example with strong inorganic acids such as mineral acids, e.g. sulphuric acid, phosphoric acid or hydrohalic acids; with strong organic carboxylic acids, such as

- alkanecarboxylic acids of 1 to 4 carbon atoms which are unsubstituted or substituted (e.g., by halogen), such as acetic acid; with saturated or unsaturated dicarboxylic acids, for example oxalic, malonic, succinic, maleic, fumaric, phthalic or tetraphthalic; with hydroxycarboxylic acids, for example ascorbic, glycolic, lactic, malic, tartaric or citric acid; with aminoacids, for example aspartic or glutamic acid; with benzoic acid; or with organic sulfonic acids, such as (C₁-C₄)-alkyl- or aryl-sulfonic acids which are unsubstituted or substituted (for example, by a halogen) such as methane- or p-toluene sulfonic acid.
- 10 Esters are formed either using organic acids or alcohols/hydroxides, depending on the functional group being esterified. Organic acids include carboxylic acids, such as alkanecarboxylic acids of 1 to 12 carbon atoms which are unsubstituted or substituted (e.g., by halogen), such as acetic acid; with saturated or unsaturated dicarboxylic acid, for example oxalic, malonic, succinic, maleic, fumaric, phthalic or tetraphthalic; with hydroxycarboxylic acids, for example ascorbic, glycolic, lactic, malic, tartaric or citric acid; with aminoacids, for example aspartic or glutamic acid; with benzoic acid; or with organic sulfonic acids, such as (C₁-C₄)-alkyl- or aryl-sulfonic acids which are unsubstituted or substituted (for example, by a halogen) such as methane- or p-toluene sulfonic acid. Suitable hydroxides include inorganic hydroxides, such as sodium hydroxide, potassium hydroxide, calcium hydroxide, aluminium hydroxide. Alcohols include alkanealcohols of 1-12 carbon atoms which may be unsubstituted or substituted, e.g. by a halogen).

ENANTIOMERS/TAUTOMERS

- 25 In all aspects of the present invention previously discussed, the invention includes, where appropriate all enantiomers and tautomers of the proinsulin derivatives of the invention. The man skilled in the art will recognise compounds that possess an optical properties (one or more chiral carbon atoms) or tautomeric characteristics. The corresponding enantiomers and/or tautomers may be isolated/prepared by methods known in the art.

STEREO AND GEOMETRIC ISOMERS

- Some of the proinsulin derivatives of the invention may exist as stereoisomers and/or geometric isomers – e.g. they may possess one or more asymmetric and/or geometric centres and so may exist in two or more stereoisomeric and/or geometric forms. The

present invention contemplates the use of all the individual stereoisomers and geometric isomers of those analogues, and mixtures thereof. The terms used in the claims encompass these forms, provided said forms retain the appropriate functional activity (though not necessarily to the same degree).

5

The present invention also includes all suitable isotopic variations of the proinsulin derivatives or pharmaceutically acceptable salts thereof. An isotopic variation is defined as one in which at least one atom is replaced by an atom having the same atomic number but an atomic mass different from the atomic mass usually found in nature. Examples of isotopes that can be incorporated into the agent and pharmaceutically acceptable salts thereof include isotopes of hydrogen, carbon, nitrogen, oxygen, phosphorus, sulphur, fluorine and chlorine such as ^2H , ^3H , ^{13}C , ^{14}C , ^{15}N , ^{17}O , ^{18}O , ^{31}P , ^{32}P , ^{35}S , ^{18}F and ^{36}Cl , respectively. Certain isotopic variations, for example, those in which a radioactive isotope such as ^3H or ^{14}C is incorporated, are useful in drug and/or substrate tissue distribution studies. Tritiated, i.e., ^3H , and carbon-14, i.e., ^{14}C , isotopes are particularly preferred for their ease of preparation and detectability. Further, substitution with isotopes such as deuterium, i.e., ^2H , may afford certain therapeutic advantages resulting from greater metabolic stability, for example, increased *in vivo* half-life or reduced dosage requirements and hence may be preferred in some circumstances. Isotopic variations of the analogues of the present invention and pharmaceutically acceptable salts thereof of this invention can generally be prepared by conventional procedures using appropriate isotopic variations of suitable reagents.

25 SOLVATES

The present invention also includes solvate forms of the compounds of the present invention. The terms used in the claims encompass these forms.

POLYMORPHS

30 The invention furthermore relates to compounds of the present invention in their various crystalline forms, polymorphic forms and (an)hydrous forms. It is well established within the pharmaceutical industry that chemical compounds may be isolated in any of such forms by slightly varying the method of purification and or isolation from the solvents used in the synthetic preparation of such compounds.

35

PRODRUGS

The invention further includes proinsulin derivatives of the present invention in prodrug form. Such prodrugs are generally analogues of the invention wherein one or more appropriate groups have been modified such that the modification may be reversed upon administration to a human or mammalian subject. Such reversion is usually performed by an enzyme naturally present in such subject, though it is possible for a second agent to be administered together with such a prodrug in order to perform the reversion in vivo. Examples of such modifications include ester (for example, any of those described above), wherein the reversion may be carried out by an esterase etc. Other such systems will be well known to those skilled in the art.

ADMINISTRATION

The pharmaceutical compositions of the present invention may be adapted for oral, rectal, vaginal, parenteral, intramuscular, intraperitoneal, intraarterial, intrathecal, intrabronchial, subcutaneous, intradermal, intravenous, nasal, buccal or sublingual routes of administration.

For oral administration, particular use is made of compressed tablets, pills, tablets, gellules, drops, and capsules. Preferably, these compositions contain from 1 to 250 mg and more preferably from 10-100 mg, of active ingredient per dose.

Other forms of administration comprise solutions or emulsions which may be injected intravenously, intraarterially, intrathecally, subcutaneously, intradermally, intraperitoneally or intramuscularly, and which are prepared from sterile or sterilisable solutions. The pharmaceutical compositions of the present invention may also be in form of suppositories, pessaries, suspensions, emulsions, lotions, ointments, creams, gels, sprays, solutions or dusting powders.

An alternative means of transdermal administration is by use of a skin patch. For example, the active ingredient can be incorporated into a cream consisting of an aqueous emulsion of polyethylene glycols or liquid paraffin. The active ingredient can also be incorporated, at a concentration of between 1 and 10% by weight, into an ointment consisting of a white wax or white soft paraffin base together with such stabilisers and preservatives as may be required.

Injectable forms may contain between 10 - 1000 mg, preferably between 10 - 250 mg, of active ingredient per dose.

Compositions may be formulated in unit dosage form, i.e., in the form of discrete portions containing a unit dose, or a multiple or sub-unit of a unit dose.

DOSAGE

A person of ordinary skill in the art can easily determine an appropriate dose of one of the instant compositions to administer to a subject without undue experimentation.

Typically, a physician will determine the actual dosage which will be most suitable for an individual patient and it will depend on a variety of factors including the activity of the specific compound employed, the metabolic stability and length of action of that compound, the age, body weight, general health, sex, diet, mode and time of administration, rate of excretion, drug combination, the severity of the particular condition, and the individual undergoing therapy. The dosages disclosed herein are exemplary of the average case. There can of course be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

THERAPEUTIC USE

Another aspect of the invention relates to proinsulin derivatives as described above for use as a medicament.

Another aspect of the invention relates to proinsulin derivatives as described above for use in treating or preventing diabetes, or treating or preventing hyperglycemia.

Preferably, the diabetes is type 2 diabetes mellitus.

Another aspect of the invention relates to proinsulin derivatives as described above in the preparation of a medicament for treating diabetes, or treating or preventing hyperglycemia.

As used herein the phrase "preparation of a medicament" includes the use of an analogue of the invention directly as the medicament in addition to its use in a

screening programme for further therapeutic agents or in any stage of the manufacture of such a medicament.

Another aspect of the invention relates to a method of treating diabetes or treating or
 5 preventing hyperglycemia in a subject in need thereof, said method comprising administering to the subject a therapeutically effective amount of a single chain insulin analogue as described above.

Preferably, the method comprises administering to the organism an amount of ACB-
 10 proinsulin in a dose between about 10 and 1000 µg/kg. A more preferred dose is from about 10 to 100 µg/kg of active compound. A typical daily dose for an adult human is from about 0.5 to 100 mg.

In practicing this method, compounds of the invention can be administered in a single
 15 daily dose or in multiple doses per day. The treatment regime may require administration over extended periods of time. The amount per administered dose or the total amount administered will depend on such factors as the nature and severity of the disease, the age and general health of the patient and the tolerance of the patient to the compound.

20 A convenient method of practicing the treatment method is to administer the compounds of the invention via intravenous infusion. In this procedure a sterile formulation of a suitable soluble salt of the compound is incorporated in a physiological fluid, such as 5% dextrose solution, and the resulting solution is infused slowly IV.
 25 Alternatively, the piggy-back method of IV infusion can also be used.

The present invention is further described with reference to the following figures, wherein:

30 Figure 1 shows ACB-proinsulin peptides 1-4 according to the invention.

Figure 2 shows a schematic representation of the first oxidation step (with DMSO) of peptide 4 (C(Acm)⁷C(Acm)³³).

Figure 3 shows HPLC profiles and ESI-MS of crude reduced C(Acm)⁷C(Acm)³³ reverted proinsulin with C= RRGR (A), the product obtained after its oxidation with DMSO (B), and the HPLC profile of the purified product (bis-oxidized ACB-proinsulin with C = RRGR) (C).

5

Figure 4 shows a schematic representation of the second oxidation step (with I₂) of peptide 4 (C(Acm)⁷C(Acm)³³) reverted proinsulin with C= RRGR and C-peptide excision by trypsinolysis.

10 Figure 5 shows HPLC profiles and corresponding MS-spectra of the crude inverted proinsulin with C-peptide RRGR (A); the same after purification (B); the product mixture obtained after trypsin cleavage (D) and the final purified [Gly^{A21}, Arg^{A22, A23}] insulin derivative (E).

15 The present invention is further described by way of the following non-limiting examples.

EXAMPLES

20 Abbreviations

Acm	5-acetomidomethyl
Boc	t-butyloxycarbonyl
CTC	chlorotriyl chloride
NMP	N-methylpyrrolidone
25 DCM	dichloromethane
TFA	trifluoroacetic acid
RE	rotary evaporator
DEE	diethyl ether
DIC	N,N'-diisopropylcarbodiimide
30 HOBt	hydroxybenzotriazole
HOSu	N-hydroxysuccinimide
Hyp	(2S,4R)-4-hydroxyproline, or L-hydroxyproline
DMF	dimethylformamide
EDAC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
35 RT	room temperature

	DTT	dithiothreitol
	DMSO	dimethylsulfoxide
	MMt	monomethoxytrityl
	Trt	trityl
5	DIPEA/DIEA	N,N-diisopropylethylamine
	Fmoc	fluorenylmethyloxycarbonyl
	MeOH	methanol
	AcOH	acetic acid
	TFE	trifluoroethyl alcohol
10	Dde	N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl)
	IPA	isopropyl alcohol
	TES	triethylsilane

Example 1

15 Solid-phase synthesis of A-Arg-Arg-Gly-Arg-Arg-B proinsulin (ACB-proinsulin)

General procedure:

A1. Preparation of loaded 2-chlorotrityl resins

2-Chlorotrityl chloride resin (CTC-Cl) (2 g; loading 1.6 mmol/g) from CBL-Patras, was placed in a 60 ml peptide synthesis reactor and swelled with 15 ml dichloromethane (DCM) for 15 min at 25 °C. The resin was filtered and a solution of 1 mmol Fmoc-amino acid and 8 mmol diisopropylethylamine (DIEA) in 10 ml DCM was added. The mixture was stirred for 2 hours at 25 °C. The remaining active sites of 2-CTC resin were neutralised by adding 1 ml of methanol (MeOH) and reacting for 1 hour. The resin was filtered and washed three times with 10 ml each of a mixture consisting of DCM/MeOH/DIPEA (85:10:5) and three times with NMP. After washing the resin was treated twice with 10 ml 25% by volume piperidine in NMP for 30 min. The resin was washed five times with 10 ml NMP. The resin was unswelled with 3 washes with 10 ml of isopropanol (IPA) and dried to constant weight. 70-95% mmol of the amino acid used was bound on the resin.

B. Solid-phase synthesis, a general protocol

The solid-phase synthesis was performed at 24 °C, with 1.0 g amino acid esterified to the CTC resin as described in Part A of Example 1. During the whole synthesis the following protocol was used.

B1. Swelling of the resin

The resin was placed in a 20 ml plastic syringe equipped with a porous polypropylene frit and treated twice with 7 ml NMP, followed by filtration.

5 **B2. Activation of the amino acid**

The amino acid (3.0 equiv.) and 1-hydroxybenzotriazol (4.0 equiv.) was dissolved in a 10 ml vial with 2.5 times their volume in NMP and cooled to 0 °C. DIC was then added (3.0 equiv.) and the mixture was stirred for 15 min.

10 **B3. Coupling Reaction**

The solution, which was prepared in B2 was then added to the B 1 reactor. The reactor was washed once with one volume of DCM and was added to the reactor which was stirred for 1-3 h at 25°-30 °C. A Kaiser Test was performed to determine the completion of the reaction. If the coupling reaction was not completed after 3 h (positive Kaiser
15 Test), the reaction mixture was filtered and recoupled with a fresh solution of activated amino acid. After completion of the coupling the reaction mixture was filtered and washed 6 times with NMP (5 volumes per wash).

B4. Removal of the Fmoc-group

20 The resulting resin in B3 was filtered and then treated for 30 min with 5 ml of a solution, which contained 25% by volume of piperidine. The resin is washed 3 x 5 ml NMP.

B5. Elongation of the peptide chain

After the incorporation of each amino acid the steps B1-B5 were repeated until the
25 desired peptide chain was formed.

The following Fmoc-amino acids were used for coupling of the individual amino acid or amino acid fragments: Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Ile-OH, Fmoc-Leu- OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Asp(tBu)-OH, Fmoc-Glu(tBu)-OH,
30 Fmoc-Lys(Boc)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-His(Trt)-OH, Fmoc-Cys(Trt)-OH and Fmoc-Cys(Acm)-OH; and the following Boc-amino acids: Boc-Gly-OH.

C. General method for the cleavage from the CTC-resin of the protected ACB proinsulin peptides and of their protected fragments, which contain Fmoc- or Boc-groups on their N-terminus

The resin-bound peptide or peptide segment which was produced as described above in B1-B5 was washed 4 times with 5 ml NMP, 3 times with 5 ml IPA and finally 5 times with 7 ml DCM to remove completely any residual NMP or other basic components. The resin was then cooled to 0 °C, filtered from DCM and treated six times with a solution of 10 ml 1.0-1.5% TFA in DCM/TES (95:5) at 5 °C.

The mixture was then stirred for 20 min at 0 °C and filtered. The resin was then washed three times with 10 ml DCM. Pyridine was then added to the filtrates (1.3 equiv. relative to TFA) to neutralize the TFA. The cleavage solution in DCM was then mixed with an equal volume of water. The resulting mixture was distilled at reduced pressure to remove DCM (350 Torr at 28 °C). The peptide or peptide fragment precipitated after the removal of DCM. The resulting peptide was washed with water and ether and dried at 30-35 °C under 15 Torr vacuum. Alternatively, DCM was removed in vacuum and the protected peptide was precipitated by the addition of diethyl ether.

Example 2

Deprotection of the protected ACB-proinsulin

General method: The protected ACB-proinsulin, obtained as described above in Example 1 (100 mg, 0.01 mmol) was treated with 10 ml TFA/TES/DTT/DCM (93:3:3:3) for 3 h at 5 °C and for 1 h at 22 °C. The resulting solution was concentrated in vacuum and then the deprotected peptide was precipitated by the addition of diethylether and washed three times with 10 ml diethylether. The resulting solid was dried in vacuum (25 °C, 15 Torr) until constant weight. Yield: 60 mg.

Example 3

One-step (random) oxidative folding of crude ACB-proinsulin (all Cys free)

10 mg of crude deprotected ACB-proinsulin (all Cys free) prepared as described in Example 2 was dissolved in 2 ml DMSO. To this solution a 0.1 M Gly buffer pH 10.5 was added (8 ml) and then solid Gnd.HCl, until a clear solution was obtained (1.9 g, about 2 M). The solution was left overnight at RT. During standing a small amount of precipitate was formed. HPLC-analysis of the supernatant (acidified with CF₃CO₂H to

pH 2) performed on a Waters Alliance 2695 system equipped with Waters 996 PDA detector, using a Purospher RP-8, 125 x 4 mm, 5 μ m column (Merck) and a linear gradient of 10 % B to 60 % B, flow rate 1 ml/min, where A = 1 % TFA in water and B = 1 % TFA in acetonitrile, showed that a complicated mixture of trioxidized isomers was formed.

Example 4

Oxidative folding of [Cys(Acm)7, 33]-ACB-proinsulin

10 mg of crude deprotected [Cys(Acm)7,33]-ACB-proinsulin prepared as described in Example 2 was dissolved in 2 ml DMSO. To this solution a 0.1 M Gly buffer pH 10.5 was added (8 ml) and then solid Gnd.HCl, until a clear solution was obtained (1.9 g, about 2 M). The solution left overnight at RT. During standing a small amount of precipitate was formed. HPLC-analysis of the supernatant (acidified with CF₃CO₂H to pH 2) performed as described in Example 4 above, showed that a a major bis-oxidized product was formed, as confirmed by ESI-MS analysis on a Waters-Micromass ZQ4000 system. M calc. 6558.56 Da, M found 6558.36 Da.

Example 5

Purification of bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin

The solution of bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin, obtained as described in Example 4 above, was loaded after filtration through a 0.2 membrane syringe filter, in two portions on a Purospher RP-18 semi-preparative column, 10 x 250 mm (Merck); phase A = 1% TFA in water, phase B = 1% TFA in acetonitrile; isocratic 30 % B for 5 min and then linear gradient from 30 % B to 40 % B in 30 min, flow rate 4 ml/min. Collected fractions of main product were frozen in liquid nitrogen and lyophilized. The purification yield was 30 %.

Example 6

Determination of disulfide bond-pairing in bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin by Staphylococcus Aureus V8 protease fingerprinting

100 μ g of purified bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin was dissolved in 150 μ L of 0.1 M ammonium acetate, pH 4.0. 50 μ L (5 μ g) of a solution of Staphylococcus Aureus V8 protease (sequencing grade, commercially available from Sigma-Aldrich) in 0.1 M ammonium acetate was added and the digest incubated at 25° C. The progress of digestion was monitored by injecting 20 μ L of the solution on a Purospher RP-8

HPLC column, using a gradient of 10 % B to 60 % B in 30 min. After 54 h the starting peptide was almost disappeared. ESI-MS analysis of the produced peaks confirmed the proper disulfide bond arrangement of the bis-oxidized- [Cys(Acm)7, 33]-ACB-proinsulin.

5 **Example 7**

Oxidation of bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin with iodine

Purified and lyophilized bis-oxidized [Cys(Acm)7, 33]-ACB-proinsulin (2 mg), obtained as described in Example 5 above was dissolved in a mixture of acetic acid/water 4:1 v/v (1 ml). This solution was then added over a period of 10 min in a solution of iodine (4
10 mg) in acetic acid/water 4:1 (1 ml), under shaking. After standing for 1 h at RT with occasional shaking, 2 drops of a 1 M solution of ascorbic acid in water was added to neutralize iodine excess (decoloration). This solution was then loaded in portions on a Purospher RP-8, 125X4 mm HPLC column (Merck) and purified using a gradient of 10 % B to 60 % B in 30 min, flow rate 1 ml/min. Collected fractions of the main product
15 were frozen in liquid nitrogen and lyophilized. Yield: 70 %. ESI-MS: M_{calc}. 6414.39 Da, M_{found} 6413.83 Da.

Example 8

Trypsin cleavage of ACB-proinsulin

20 500 µg ACB-proinsulin was dissolved in 500 µL of 0.1 M Tris.HCl buffer pH 7.9, containing 2 M urea and 5 µg of trypsin (Sigma, slightly cloudy solution) at RT. The reaction was stopped after 5 min by acidification with trifluoroacetic acid (5 µL). The solution, after centrifugation, was loaded in 100 µL portions on a Purospher RP-8 column, as described in Example 7 and the main product was isolated. The product
25 was analyzed by ESI-MS and corresponded to [GlyA21]- insulin plus 2 Arg residues (ESI-MS: M_{calc}. 6062.97 Da, M_{found} 6062.28 Da). After reduction of disulfide bonds with TCEP and LC-MS analysis of the two chains produced the product identified as the [GlyA21, ArgA22, ArgA23]-insulin.

30 Various modifications and variations of the described aspects of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of

44

CLAIMS

1. A polypeptide compound of formula (I):

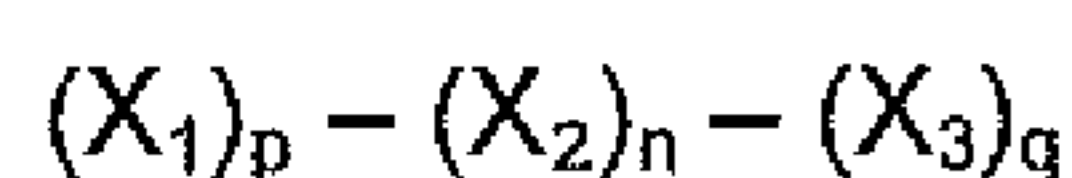


wherein:

A is the A chain of insulin or a functional derivative or variant thereof;

B is the B chain of insulin or a functional derivative or variant thereof;

C is a peptide of the formula:



wherein:

each X_1 and X_3 is independently a basic amino acid;

each X_2 is independently a natural or unnatural amino acid;

p is 1 or 2;

q is 0, 1 or 2;

n is 0, 1, 2 or 3.

2. A polypeptide compound according to claim 1 wherein each X_2 is independently selected from a basic amino acid, Gly, β -Ala, Pro, Hyp, pseudoproline and an acidic or hydrophilic amino acid.
3. A polypeptide compound according to claim 2 wherein the acidic or hydrophilic amino acid is selected from Ser, Asp and Glu.
4. A polypeptide compound according to any preceding claim wherein the basic amino acid is selected from Lys, Arg, Orn and His.
5. A compound according to any preceding claim wherein each X_2 is independently selected from Gly and Arg.

6. A polypeptide compound according to any preceding claim wherein X_1 and X_3 are both Arg.
7. A polypeptide compound according to any preceding claim wherein p and q are both 2.
8. A polypeptide compound according to any preceding claim wherein n is 1 or 2.
9. A polypeptide compound according to any preceding claim wherein n is 1.
10. A polypeptide compound according to any preceding claim wherein A is the A chain of native insulin, preferably human insulin, or a variant thereof, wherein: (a) one or more amino acid residues are substituted by a naturally or non-naturally occurring amino acid residue, (b) the order of two or more amino acid residues is reversed, (c) one, two or three amino acids are deleted, (d) a spacer group is present between any two amino acid residues, (e) one or more amino acid residues are in peptoid form, (f) the (N-C-C) backbone of one or more amino acid residues of the peptide has been modified, (g) one or more additional amino acids are present at the N-terminus and/or the C-terminus, or any of (a)-(g) in combination.
11. A polypeptide compound according to any preceding claim wherein B is the B chain of native insulin, preferably human insulin, or a variant thereof, wherein: (a) one or more amino acid residues are substituted by a naturally or non-naturally occurring amino acid residue, (b) the order of two or more amino acid residues is reversed, (c) one, two or three amino acids are deleted, (d) a spacer group is present between any two amino acid residues, (e) one or more amino acid residues are in peptoid form, (f) the (N-C-C) backbone of one or more amino acid residues of the peptide has been modified, (g) one or more additional amino acids are present at the N-terminus and/or the C-terminus, or any of (a)-(g) in combination.
12. A polypeptide compound according to any preceding claim wherein C is a peptide of the sequence shown in SEQ ID NO:1,

RRGRR [SEQ ID NO:1]

13. A polypeptide compound according to any preceding claim wherein A comprises a peptide of the sequence shown in SEQ ID NO:2,

¹GIVEQCCTSI~~C~~SLYQLENYCG²¹ [SEQ ID NO:2]

wherein each amino acid is unprotected, or optionally protected.

14. A polypeptide compound according to any preceding claim wherein B comprises a peptide of the sequence shown in SEQ ID NO:3,

¹FVNQHLCGSHLVEALYLVCGERGFFYTPKT³⁰ [SEQ ID NO:3]

wherein each amino acid is unprotected, or optionally protected.

15. A polypeptide compound according to any preceding claim which is selected from SEQ ID NOS: 4-7,

SEQ ID NO:4

¹GIVEQCCTSI~~C~~SLYQLENYC~~GRRGRR~~FVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

SEQ ID NO:5

¹GIVEQCCTSI~~C~~SLYQLENYC~~GRRGRR~~FVNQHLCGSHLVEALYLV~~C~~GERGFFYTPKT⁵⁶

SEQ ID NO:6

¹GIVEQCCTSI~~C~~SLYQLENYC~~GRRGRR~~FVNQHLCGSHLVEALYLVCGERGFFYTPKT⁵⁶

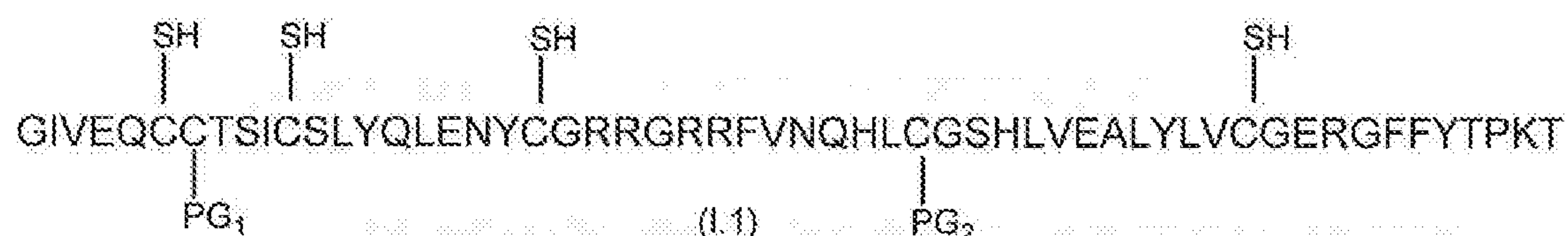
SEQ ID NO:7

¹GIVEQC~~C~~TSI~~C~~SLYQLENYC~~GRRGRR~~FVNQHL~~C~~GSHLVEALYLVCGERGFFYTPKT⁵⁶

wherein in SEQ ID NOS: 4-7, C denotes a cysteine protected with a protecting group and C denotes an unprotected cysteine.

16. A polypeptide compound according to claim 15 wherein the cysteine protecting group is Acn or Trt.

17. A polypeptide according to any preceding claim which is of formula (I.1):



wherein PG₁ and PG₂ are each independently a cysteine protecting group.

18. A process for preparing insulin, or a derivative thereof, said process comprising the steps of:

- (i) preparing a polypeptide compound according to any one of claims 1 to 17, and
- (ii) cleaving said polypeptide to excise the C-peptide.

19. A process according to claim 18 wherein the cysteine residue in the 7-position of the A chain of insulin and the cysteine residue in the 7-position of the B chain of insulin are each protected with a protecting group.

20. A process according to claim 19 wherein the protecting group is Acn or Trt, more preferably Acn.

21. A process according to claim 19 or claim 20 which comprises a stepwise oxidative folding of the polypeptide compound, said process comprising the steps of:

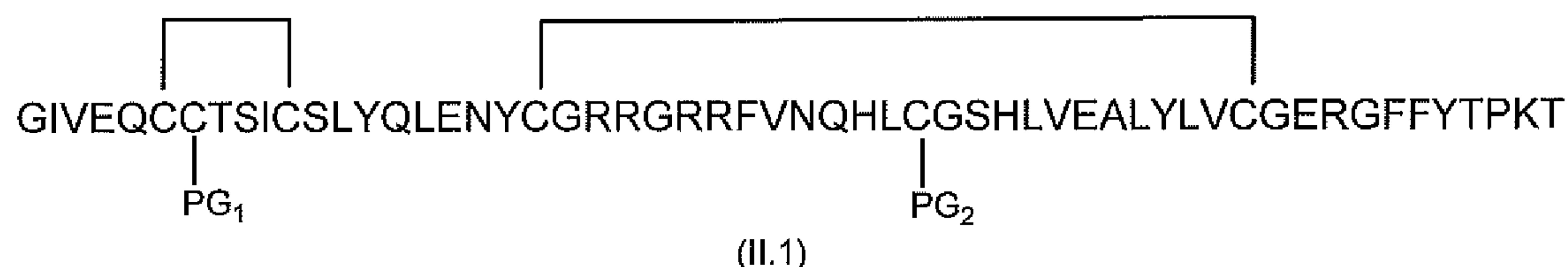
- (i) oxidising a peptide according to claim 19 to form a first intermediate;
- (ii) oxidising said first intermediate to form a second intermediate;
- (iii) cleaving said second intermediate to excise the C-peptide.

22. A process according to claim 21 wherein step (i) comprises treating the peptide according to claim 19 with DMSO.

23. A process according to claim 21 or claim 22 wherein step (ii) comprises treating the first intermediate with iodine in acetic acid/water.

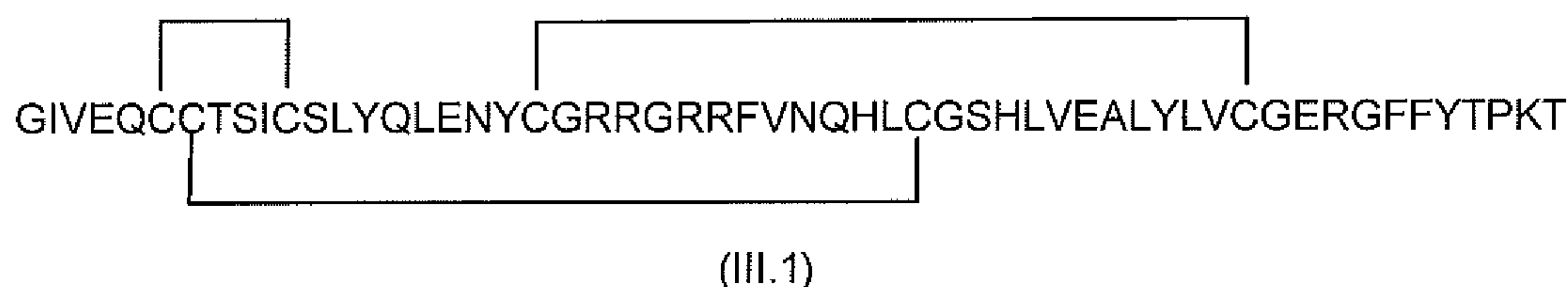
24. A process according to any one of claims 21 to 23 wherein step (iii) comprises treating the second intermediate with trypsin.

25. A polypeptide compound of formula (II.1):



wherein PG₁ and PG₂ are each independently a cysteine protecting group.

26. A polypeptide compound of formula (III.1):



27. Use of a peptide compound according to any one of claims 1 to 17 in the preparation of insulin, or a derivative thereof.

28. A pharmaceutical formulation comprising as an active ingredient a polypeptide compound as claimed in any one of claims 1 to 17, 25 or 26, associated with one or more pharmaceutically acceptable carriers, excipients or diluents therefor.

29. A DNA sequence encoding any of the peptide compounds of claims 1 to 17.

30. A recombinant DNA vector comprising a DNA compound of claim 29.

31. A process for the recombinant preparation of a polypeptide compound according to any one of claims 1 to 17, said method comprising the steps of:

- (i) constructing a DNA sequence encoding a polypeptide compound according to any one of claims 1 to 17,
- (ii) incorporating said DNA sequence into a suitable vector containing a promoter-operator region functional in a host cell,
- (iii) orienting said DNA sequence in said vector so as to achieve transcription and translation of said DNA sequence and further that said DNA sequence is under the

transcriptional control of said promoter-operator region,

- (iv) transforming said host cell with said vector,
- (v) culturing said transformed host cell under conditions appropriate so as to induce transcription and translation of said gene, and
- (vi) recovering and purifying the polypeptide product encoded by said DNA sequence.

32. A process for preparing a polypeptide compound according to any one of claims 1 to 17 using solid-phase peptide synthesis.

33. A process for preparing insulin or a derivative thereof, said method comprising the steps of:

- (i) constructing a DNA sequence encoding a polypeptide compound according to any one of claims 1 to 17;
- (ii) incorporating said DNA sequence into a suitable vector containing a promoter-operator region functional in a host cell,
- (iii) orienting said DNA sequence in said vector so as to achieve transcription and translation of said DNA sequence and further that said DNA sequence is under the transcriptional control of said promoter-operator region,
- (iv) transforming said host cell with said vector,
- (v) culturing said transformed host cell under conditions appropriate so as to induce transcription and translation of said gene, and
- (vi) recovering and purifying the polypeptide product encoded by said DNA sequence.
- (vii) cleaving said polypeptide product encoded by said gene to excise the C-peptide.

34. A polypeptide compound according to any one of claims 1 to 17, 25 or 26 for use in medicine.

35. A polypeptide compound according to any one of claims 1 to 17, 25 or 26 for use in treating diabetes.

36. A polypeptide compound for use according to claim 35, for use in treating type 2 diabetes mellitus.

37. A method of treating diabetes or treating or preventing hyperglycemia in a subject in need thereof, said method comprising administering to the subject a therapeutically effective amount of polypeptide compound according to any one of claims 1 to 17.

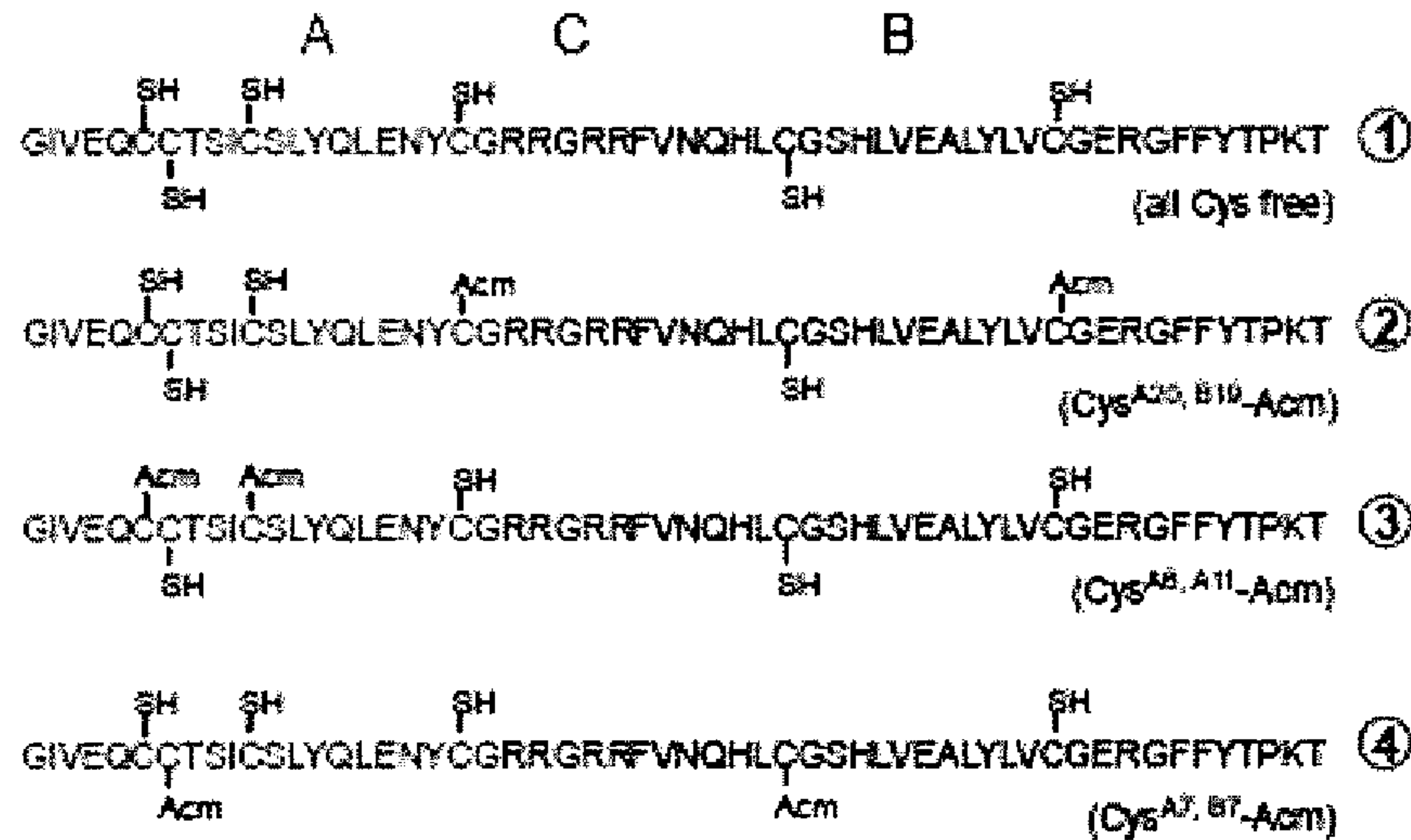


FIGURE 1

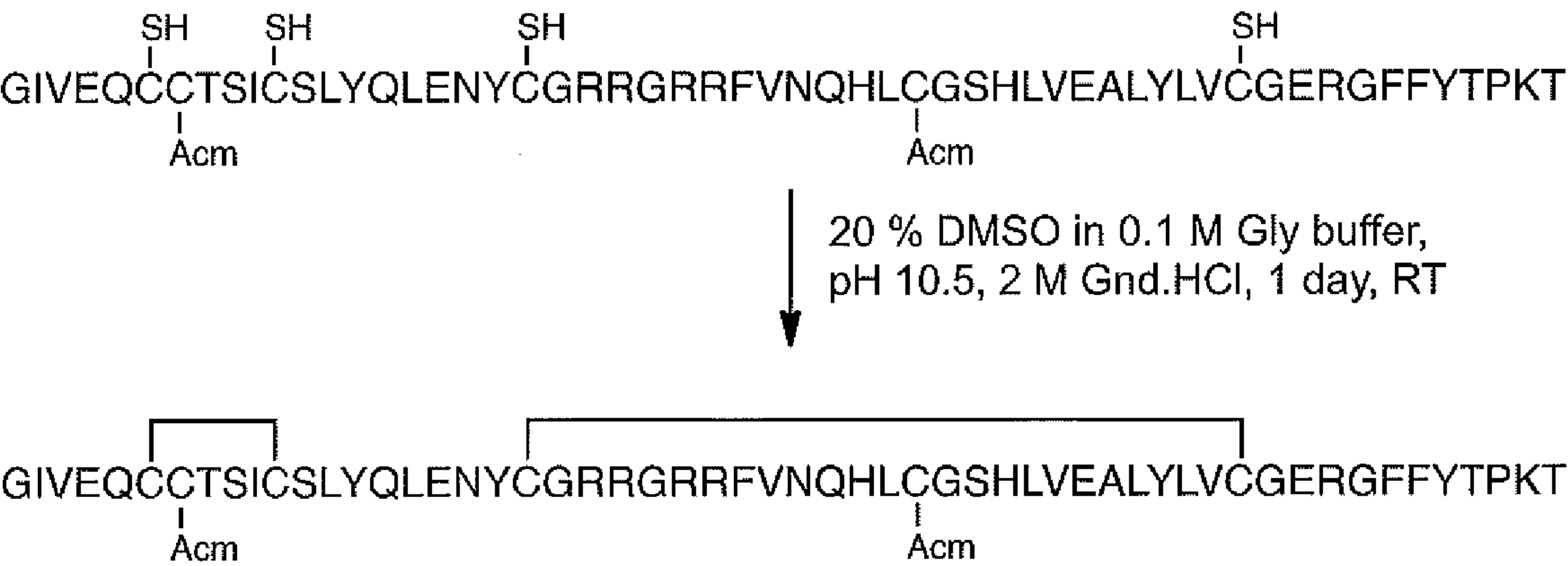


FIGURE 2

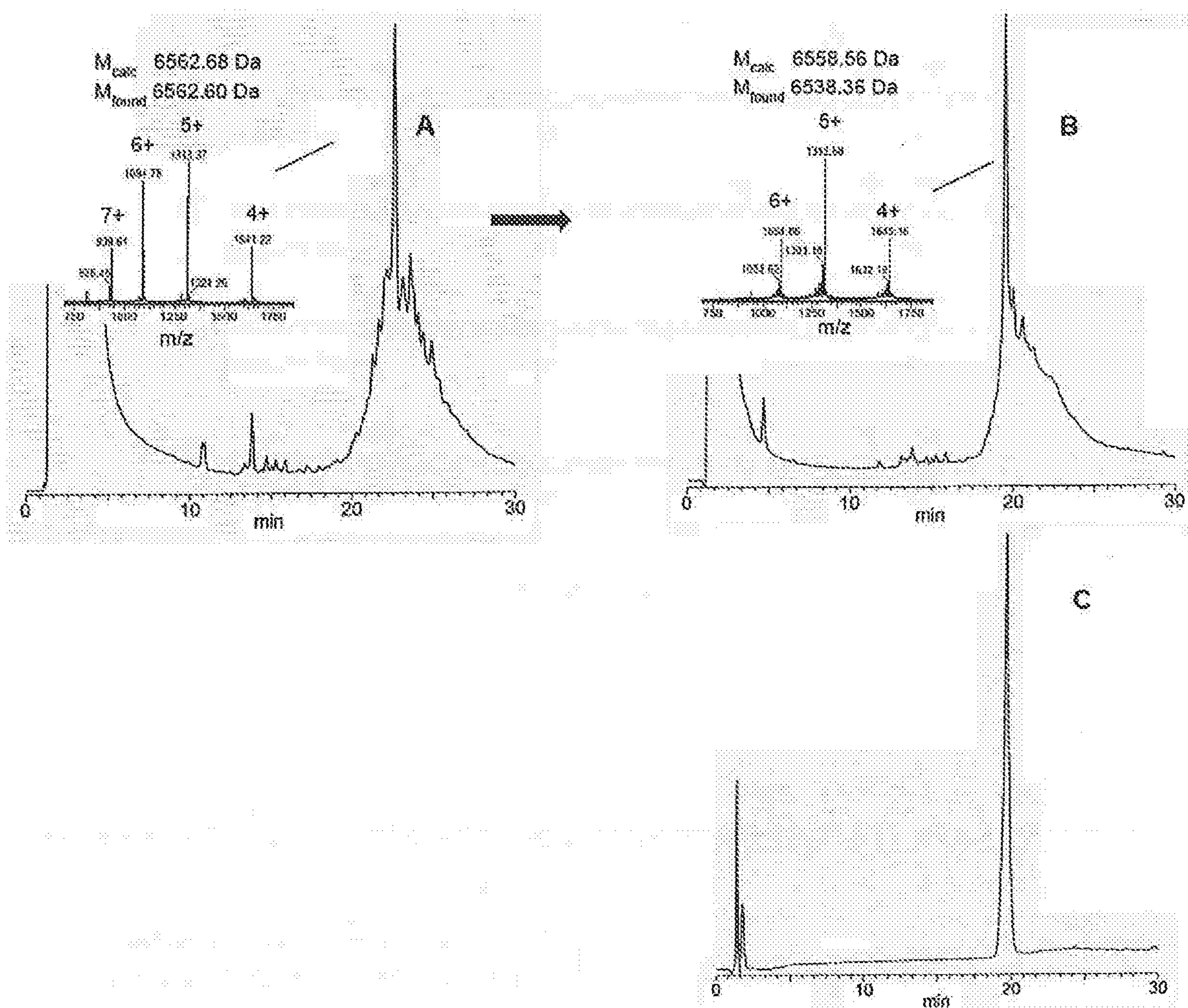


FIGURE 3

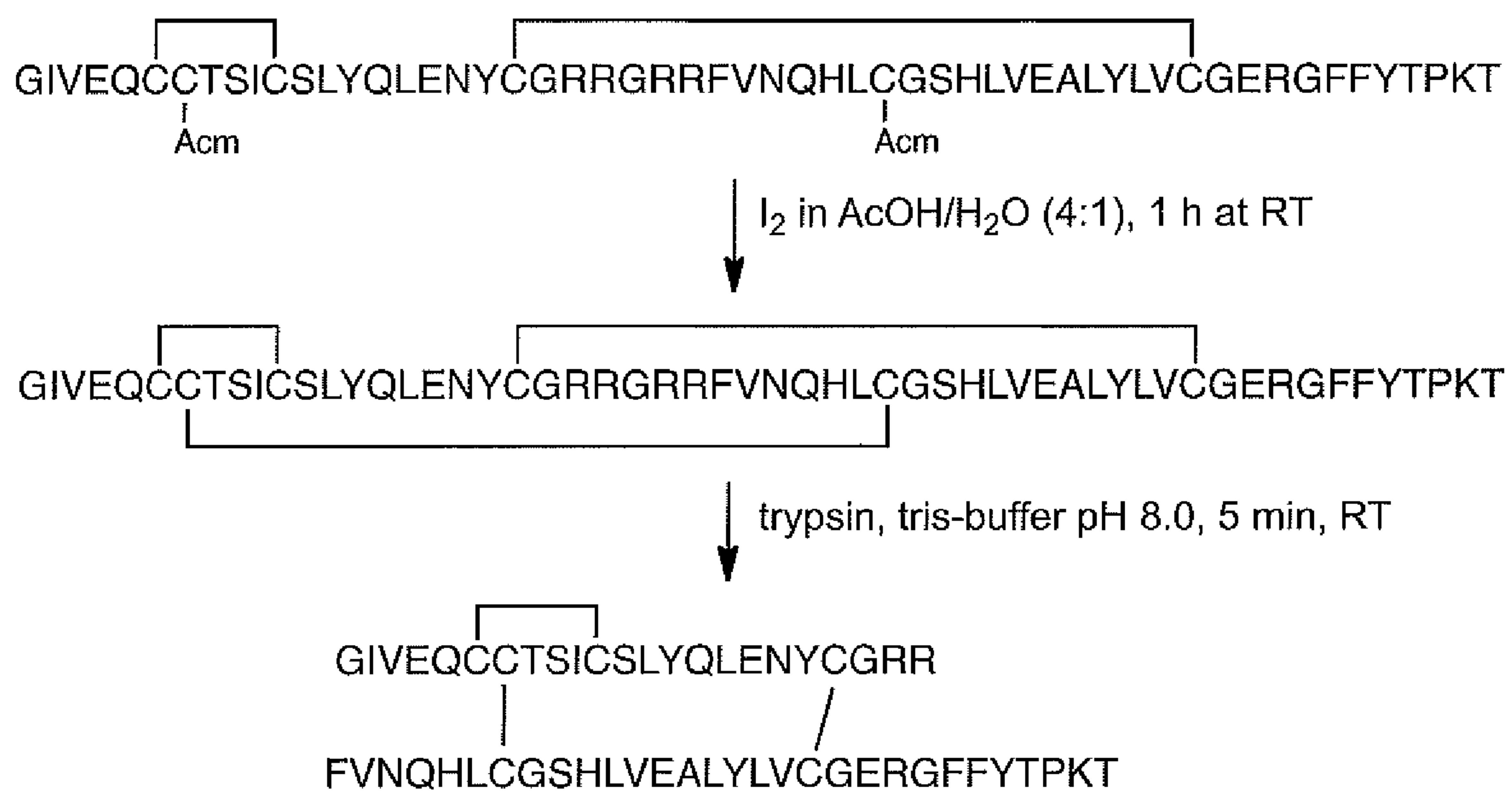


FIGURE 4

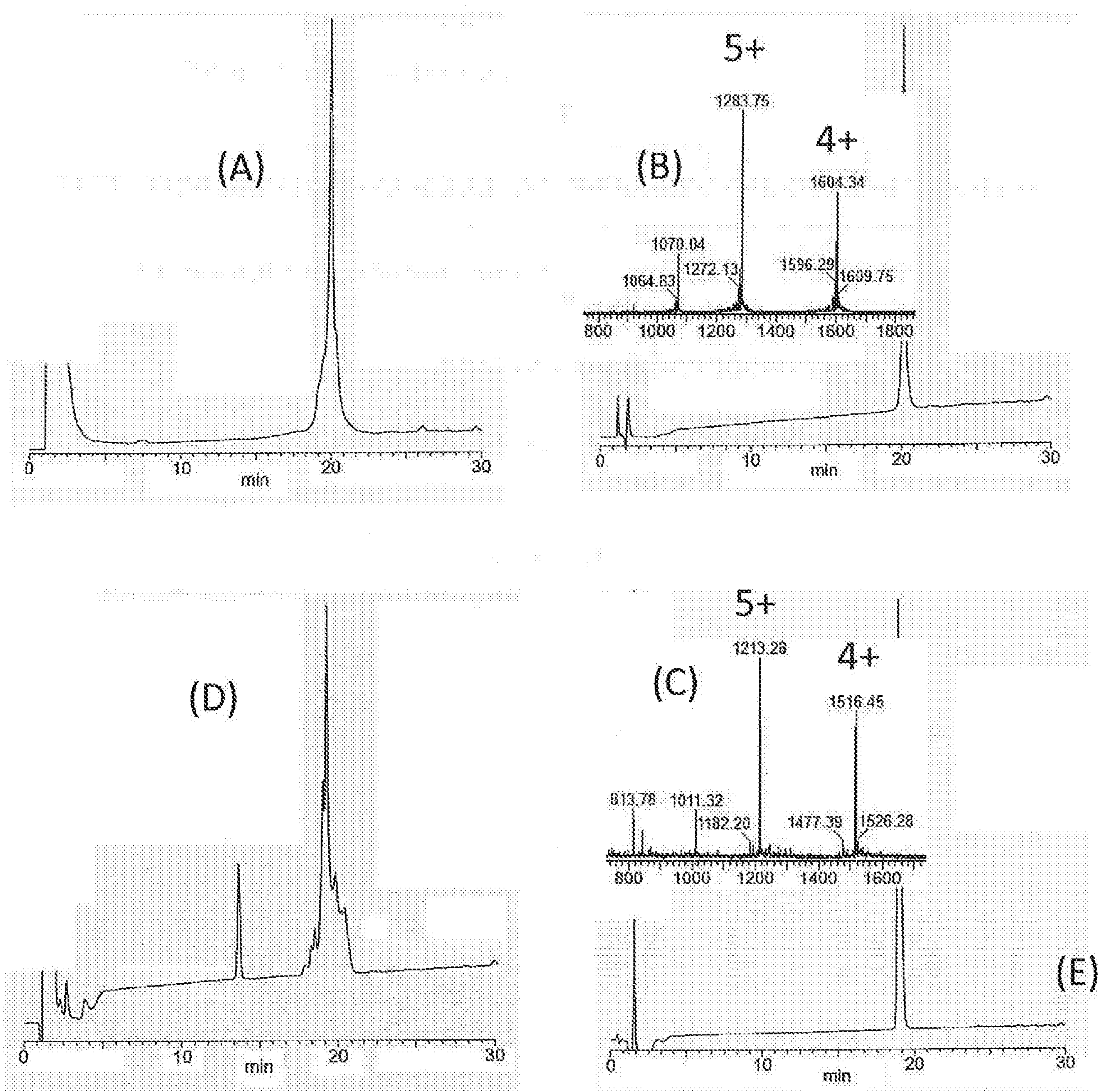


FIGURE 5

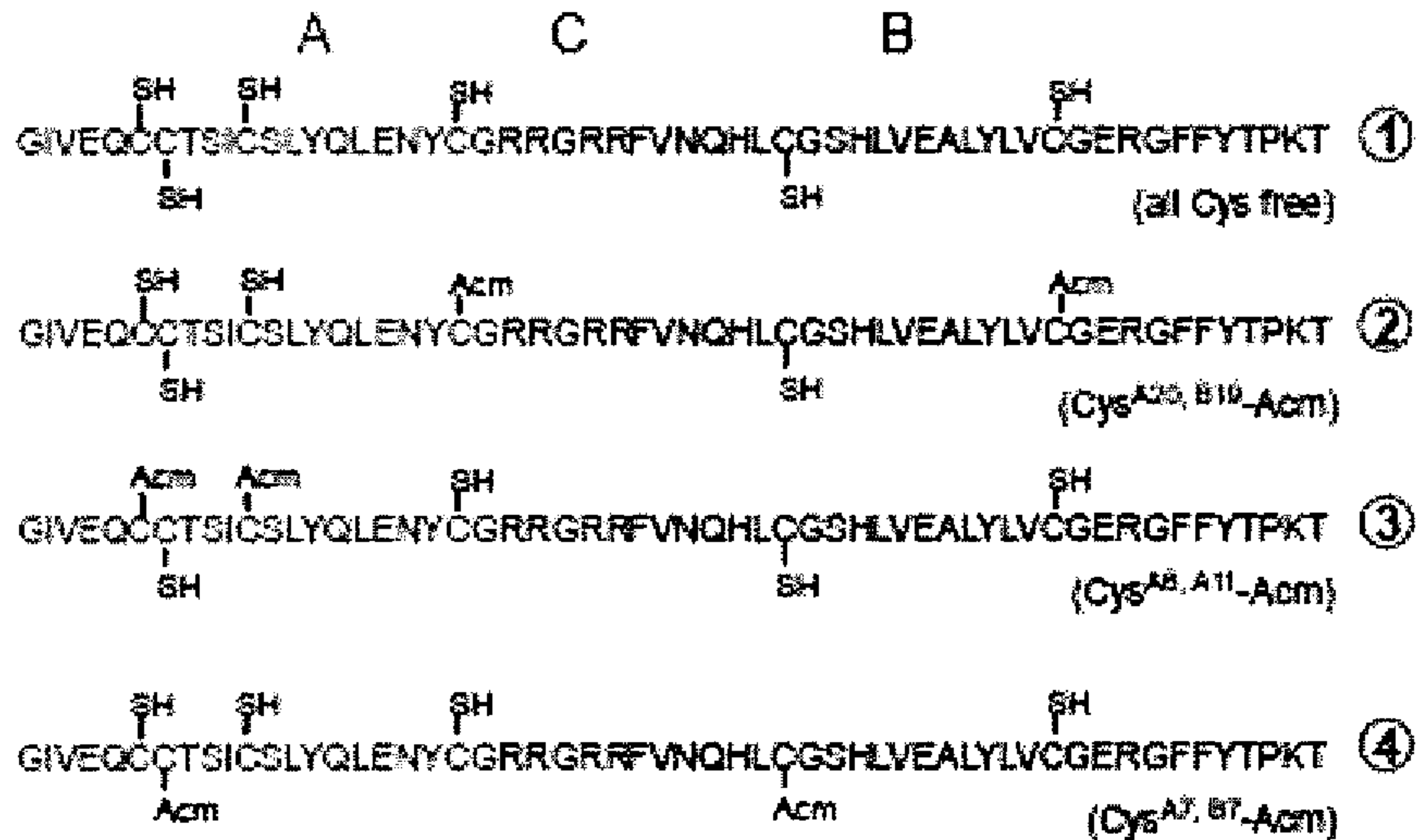


FIGURE 1