

**(12) STANDARD PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

(11) Application No. **AU 2009335715 B2**

(54) Title  
**Amide-based insulin prodrugs**

(51) International Patent Classification(s)  
**C07K 1/00 (2006.01)**

(21) Application No: **2009335715**

(22) Date of Filing: **2009.12.18**

(87) WIPO No: **WO10/080609**

(30) Priority Data

(31) Number  
**61/139,218**

(32) Date  
**2008.12.19**

(33) Country  
**US**

(43) Publication Date: **2010.07.15**

(44) Accepted Journal Date: **2016.09.15**

(71) Applicant(s)  
**Indiana University Research and Technology Corporation**

(72) Inventor(s)  
**Dimarchi, Richard D.;Kou, Binbin;Cheng, Shujiang**

(74) Agent / Attorney  
**Spruson & Ferguson, L 35 St Martins Tower 31 Market St, Sydney, NSW, 2000**

(56) Related Art  
**DE, Arnab et al., Int J Pept Res Ther, 2008, vol 14, pages 255-262**  
**US 20060024053 A1**  
**WO 1990012814 A1**

(51) International Patent Classification:  
C07K 1/00 (2006.01)(21) International Application Number:  
PCT/US2009/068716(22) International Filing Date:  
18 December 2009 (18.12.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
61/139,218 19 December 2008 (19.12.2008) US(71) Applicant (for all designated States except US): **INDIANA UNIVERSITY RESEARCH AND TECHNOLOGY CORPORATION** [US/US]; 351 West 10th Street, Indianapolis, IN 46202 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DIMARCHI, Richard, D.** [US/US]; 10890 Wilmington Drive, Carmel, IN 46033 (US). **KOU, Binbin** [CN/US]; 5684 W. Bedrock Road, Bloomington, IN 47403 (US). **CHENG, Shujiang** [CN/US]; 1072 Sugar Maple Place, Bloomington, IN 47403 (US).(74) Agent: **BREEN, John, P.**; Barnes & Thornburg LLP, 11 South Meridian Street, Indianapolis, IN 46204 (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, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Declarations under Rule 4.17:

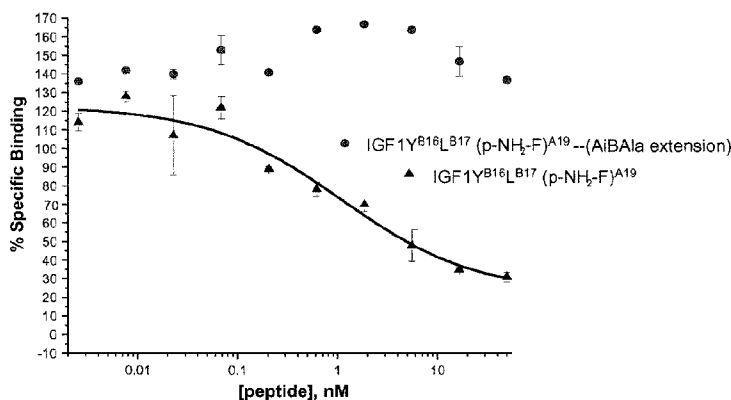
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))
- of inventorship (Rule 4.17(iv))

[Continued on next page]

(54) Title: AMIDE-BASED INSULIN PRODRUGS

Fig. 6

Insulin Receptor, 12 hr



(57) Abstract: Prodrug formulations of insulin and insulin analogs are provided wherein the insulin peptide has been modified by an amide bond linkage of a dipeptide prodrug element. The prodrugs disclosed herein have extended half lives of at least 10 hours, and more typically greater than 2 hours, 20 hours and less than 70 hours, and are converted to the active form at physiological conditions through a non-enzymatic reaction driven by chemical instability.



---

**Published:**

- with international search report (*Art. 21(3)*)
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (*Rule 48.2(h)*)
- with sequence listing part of description (*Rule 5.2(a)*)

## AMIDE-BASED INSULIN PRODRUGS

## CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Patent Application No.  
5 61/139,218 filed on December 19, 2008, the disclosure of which is hereby expressly  
incorporated by reference in its entirety.

## BACKGROUND

Insulin is a peptide hormone comprised of a two chain heterodimer that is  
10 biosynthetically derived from a low potency single chain proinsulin precursor through  
enzymatic processing. Human insulin is comprised of two peptide chains (an "A  
chain" (SEQ ID NO: 1) and "B chain" (SEQ ID NO: 2)) bound together by disulfide  
bonds and having a total of 51 amino acids. The C-terminal region of the B-chain and  
the two terminal regions of the A-chain associate in a three-dimensional structure to  
15 assemble a site for high affinity binding to the insulin receptor.

Insulin demonstrates unparalleled ability to lower glucose in virtually all  
forms of diabetes. Unfortunately, its pharmacology is not glucose sensitive and as  
such it is capable of excessive action that can lead to life-threatening hypoglycemia.  
Inconsistent pharmacology is a hallmark of insulin therapy such that it is extremely  
20 difficult to normalize blood glucose without occurrence of hypoglycemia.  
Furthermore, native insulin is of short duration of action and requires modification to  
render it suitable for use in control of basal glucose. Established approaches to delay  
the onset of insulin action include reduction in solubility, and albumin binding.

For example, two commercially available insulin derivatives have been  
25 prepared to provide a longer action profile. More particularly, the insulin derivative  
[GlyA21, ArgB31, ArgB32]insulin was prepared to shift insulin's pI from 5.4 to 6.7  
resulting in the peptide being precipitated at physiological pH and thus delaying  
adoption and time of action (see Bolli et al., *Diabetologia* 1999, 42, 1151-1167).  
However, this insulin derivative has enhanced IGF-1 affinity, leading to increased  
30 proliferative actions and the possibility of tumorigenesis. Another commercially  
available insulin derivative is [LysB29-tetradecanoyl, des(B30)]insulin, wherein  
LysB29 has been acylated with a C<sub>14</sub> fatty acid (Mayer et al., *Peptide Science*, 88, 5,  
687-713). The presence of the fatty acid chain enhances binding of the peptide to

serum albumin, resulting in increased plasma half life. However, this derivative suffers the disadvantage of having reduced potency *in vivo*. In addition, both insulin derivatives exhibit variability in biological action from one patient to the next.

Prodrug chemistry offers the opportunity to precisely control the onset and  
5 duration of insulin action after clearance from the site of administration and equilibration in the plasma at a highly defined concentration. The central virtue of such an approach, relative to current long-acting insulin analogs and formulations, is that the insulin reservoir is not the subcutaneous fatty tissue where injection occurs, but rather the blood compartment. This removes the variability in absorption  
10 encountered with prior art delayed onset insulin derivatives. It also enables administration of the peptide hormone by routes other than a subcutaneous injection.

Binding of insulin to its receptor will result in biological stimulation, but will also initiate the subsequent deactivation of insulin induced pharmacology through the enzymatic degradation of the insulin peptide. An added advantage of using a prodrug  
15 derivative of insulin is that such an approach also extends insulin's biological half life based on a strategy of inhibiting recognition of the prodrug by the corresponding receptor. In spite of these advantages associated with prodrug derivatives, the complex nature of preparing such prodrugs has, until now, prevented the preparation of an efficacious prodrug derivative of insulin. To build a successful prodrug-  
20 hormone, an active site structural address is needed that can form the basis for the reversible attachment of a prodrug structural element. The structural address needs to offer two key features; (1) the potential for selective chemical modification and (2) the ability to provide a high degree of activity in the native form upon removal of the prodrug structural element. The insulin prodrugs disclosed herein are chemically  
25 converted to structures that can be recognized by the receptor, wherein the speed of this chemical conversion will determine the time of onset and duration of *in vivo* biological action. The prodrug chemistry disclosed in this application relies upon an intramolecular chemical reaction that is not dependent upon additional chemical additives, or enzymes or enzyme inhibitors.

30 The ideal prodrug should be soluble in water at physiological conditions (for example, a pH of 7.2 and 37 °C), and it should be stable in the powder form for long term storage. It should also be immunologically silent and exhibit a low activity relative to the parent drug. Typically the prodrug will exhibit no more than 10% of

the activity of the parent drug, in one embodiment the prodrug exhibits less than 10%, less than 5%, about 1%, or less than 1% activity relative to the parent drug. Furthermore, the prodrug, when injected in the body, should be quantitatively converted to the active drug within a defined period of time. Applicants are the first to disclose insulin prodrug analogs that meet each of these objectives.

## SUMMARY

Peptide-based drugs are highly effective medicines with relatively short duration of action and variable therapeutic index. The present disclosure is directed to insulin prodrugs wherein the prodrug derivative is designed to delay onset of action and extend the half life of the drug. The delayed onset of action is advantageous in that it allows systemic distribution of the prodrug prior to its activation. Accordingly, the administration of prodrugs eliminates complications caused by peak activities upon administration and increases the therapeutic index of the parent drug.

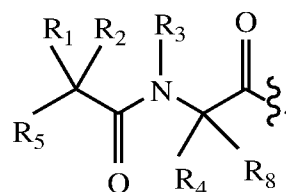
As disclosed herein, a prodrug derivative of insulin may be prepared by covalently linking a dipeptide to an active site of the insulin peptide via an amide linkage. In one embodiment the dipeptide is covalently bound to the insulin peptide at a position that interferes with insulin's ability to interact with the insulin and IGF-1 receptors. Subsequent removal of the dipeptide via an intramolecular reaction, resulting in diketopiperazine or diketomorpholine formation, under physiological conditions and in the absence of enzymatic activity, restores full activity to the polypeptide.

Also herein disclosed is an insulin prodrug having the general structure of U-O-insulin, wherein U is an amino acid or a hydroxyl acid and O is an N-alkylated amino acid linked to the insulin peptide through formation of an amide bond between U-O and an amine of an insulin peptide. The U-O dipeptide may be bound at the N-terminus, or at the side chain of an amino acid corresponding to positions A19, B16 or B25, of the respective A chain or B chain via an amide bond. The structure of U-O may be selected wherein chemical cleavage of U-O from the insulin peptide is at least about 90% complete within about 1 to about 720 hours in PBS under physiological conditions. The chemical cleavage half-life ( $t_{1/2}$ ) of U-O from the insulin peptide may be at least about 1 hour to about 1 week in PBS under physiological conditions.

U and O may be selected to inhibit enzymatic cleavage of the U-O dipeptide from an insulin peptide by enzymes found in mammalian serum. U and/or O may be selected such that the cleavage half-life of U-O from the insulin peptide, in PBS under physiological conditions, is not more than two fold the cleavage half-life of U-O from the insulin peptide in a solution comprising a DPP-IV protease (i.e., cleavage of U-O from the insulin prodrug does not occur at a rate more than 2x faster in the presence of DPP-IV protease and physiological conditions relative to identical conditions in the absence of the enzyme). U, O, or the amino acid of the insulin peptide to which U-O is linked may be a non-coded amino acid. U and/or O may be an amino acid in the D stereoisomer configuration. According to one disclosure, U may be an amino acid in the D stereoisomer configuration and O may be an amino acid in the

L stereoisomer configuration. Alternatively, U may be an amino acid in the L stereoisomer configuration and O is an amino acid in the D stereoisomer configuration. Alternatively, U may be an amino acid in the D stereoisomer configuration and O may be an amino acid in the D stereoisomer configuration. Alternatively, O may be an N-alkylated amino acid but not proline.

The dipeptide prodrug element may comprise a compound having the general structure of Formula I:



wherein

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl or aryl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

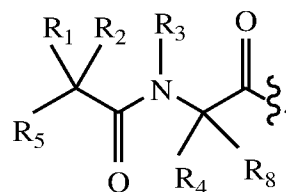
R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH.

In one embodiment the dipeptide prodrug element comprises a compound having the general structure of Formula I:



wherein

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to

which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

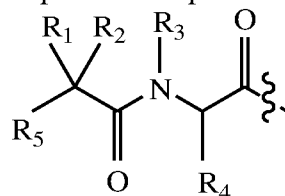
R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo.

The dipeptide extension may comprise a compound of the general structure:



wherein R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>8</sub> cycloalkyl ring;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

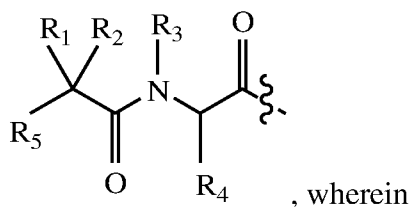
R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH.

R<sub>3</sub> may be C<sub>1</sub>-C<sub>8</sub> alkyl and R<sub>4</sub> may be selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached may form a 5 or 6 member heterocyclic ring.

Also herein disclosed is an insulin prodrug analog comprising an A chain and a B chain, wherein the A chain comprises the sequence Z-GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and the B chain comprises the sequence of J-X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14). The Z and J designations of the A and B chain formulas are independently H (forming an N-terminal amine) or a dipeptide comprising the general structure:

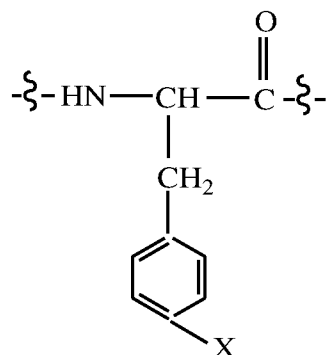




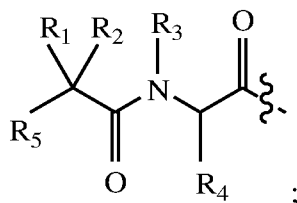
X<sub>14</sub> is either a bond joining the "J" element to the N-terminus of the X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCG ERGFX<sub>8</sub> (SEQ ID NO: 14) sequence or X<sub>14</sub> represents a 1 to 4 amino acid sequence selected from the group consisting of a FVNQ (SEQ ID NO: 11), VNQ, NQ and Q that joins the "J" element to the N-terminus of the X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCG ERGFX<sub>8</sub> (SEQ ID NO: 14) sequence;

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>, wherein R<sub>10</sub> is a dipeptide comprising the general structure:



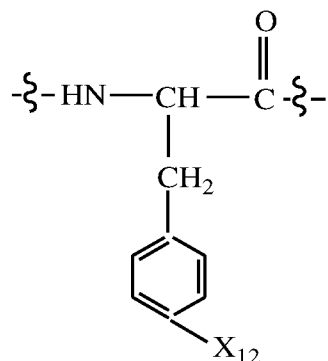
X<sub>3</sub> is selected from the group consisting of asparagine, ornithine, glycine, alanine, threonine, and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

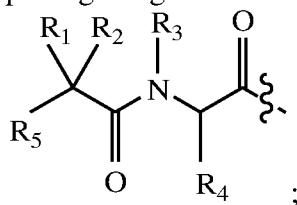
X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

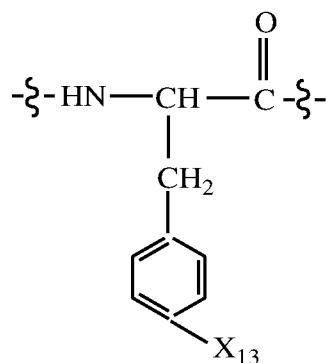
X<sub>7</sub> is an amino acid of the general structure



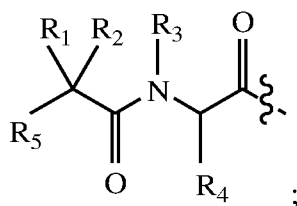
wherein  $X_{12}$  is selected from the group consisting of OH,  $NH_2$ ,  $NHR_{11}$  and  $OCH_3$ , wherein  $R_{11}$  is a dipeptide comprising the general structure:



$X_8$  is an amino acid of the general structure



wherein  $X_{13}$  is selected from the group consisting of H, OH,  $NH_2$ ,  $NHR_{12}$  and  $OCH_3$ , wherein  $R_{12}$  is a dipeptide comprising the general structure:



wherein  $R_1$  is selected from the group consisting of H and  $C_1$ - $C_8$  alkyl;

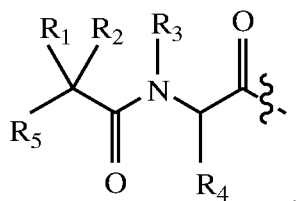
$R_2$  and  $R_4$  are independently selected from the group consisting of H,  $C_1$ - $C_8$  alkyl,  $C_2$ - $C_8$  alkenyl,  $(C_1$ - $C_4$  alkyl)OH,  $(C_1$ - $C_4$  alkyl)SH,  $(C_2$ - $C_3$  alkyl)SCH<sub>3</sub>,  $(C_1$ - $C_4$  alkyl)CONH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)COOH,  $(C_1$ - $C_4$  alkyl)NH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)( $C_3$ - $C_6$  cycloalkyl),  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl)R<sub>7</sub>, and  $CH_2$ ( $C_5$ - $C_9$  heteroaryl);

$R_3$  is selected from the group consisting of  $C_1$ - $C_8$  alkyl,  $(C_1$ - $C_4$  alkyl)OH,  $(C_1$ - $C_4$  alkyl)SH,  $(C_3$ - $C_6$ )cycloalkyl or  $R_4$  and  $R_3$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

$R_5$  is  $NHR_6$  or OH;

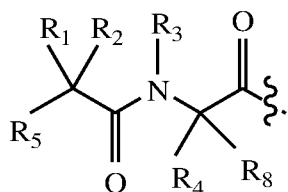
$R_6$  is H, or  $R_6$  and  $R_2$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

$R_7$  is selected from the group consisting of H and OH, with the proviso that one and only one of X,  $X_{12}$ ,  $X_{13}$ , J and Z comprises a dipeptide of the general structure:



In one disclosure, when J or Z comprise the dipeptide of Formula I, and  $R_4$  and  $R_3$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring, then both  $R_1$  and  $R_2$  are not hydrogen.

The dipeptide present at Z, J,  $R_{10}$ ,  $R_{11}$  or  $R_{12}$  may comprise a compound having the general structure of Formula I:



wherein

$R_1$ ,  $R_2$ ,  $R_4$  and  $R_8$  are independently selected from the group consisting of H,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_1$ - $C_{18}$  alkyl)OH,  $(C_1$ - $C_{18}$  alkyl)SH,  $(C_2$ - $C_3$  alkyl)SCH<sub>3</sub>,  $(C_1$ - $C_4$  alkyl)CONH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)COOH,  $(C_1$ - $C_4$  alkyl)NH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)( $C_3$ - $C_6$  cycloalkyl),  $(C_0$ - $C_4$  alkyl)( $C_2$ - $C_5$  heterocyclic),  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ ,  $(C_1$ - $C_4$  alkyl)( $C_3$ - $C_9$  heteroaryl), and  $C_1$ - $C_{12}$  alkyl( $W_1$ ) $C_1$ - $C_{12}$  alkyl, wherein  $W_1$  is a heteroatom selected from the group consisting of N, S and O, or  $R_1$  and  $R_2$  together with the atoms to which they are attached form a  $C_3$ - $C_{12}$  cycloalkyl; or  $R_4$  and  $R_8$  together with the atoms to which they are attached form a  $C_3$ - $C_6$  cycloalkyl;

$R_3$  is selected from the group consisting of  $C_1$ - $C_{18}$  alkyl,  $(C_1$ - $C_{18}$  alkyl)OH,

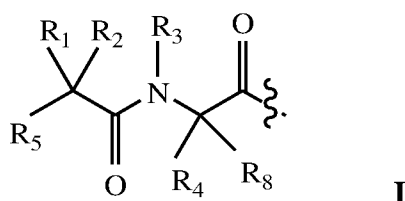
$(C_1$ - $C_{18}$  alkyl)NH<sub>2</sub>,  $(C_1$ - $C_{18}$  alkyl)SH,  $(C_0$ - $C_4$  alkyl)( $C_3$ - $C_6$ )cycloalkyl,  $(C_0$ - $C_4$  alkyl)( $C_2$ - $C_5$  heterocyclic),  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ , and  $(C_1$ - $C_4$  alkyl)( $C_3$ - $C_9$  heteroaryl) or  $R_4$  and  $R_3$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

$R_5$  is NHR<sub>6</sub> or OH;

$R_6$  is H,  $C_1$ - $C_8$  alkyl or  $R_6$  and  $R_1$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl)CONH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)COOH,  $(C_0$ - $C_4$  alkyl)NH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)OH, and halo.

The dipeptide present at Z, J,  $R_{10}$ ,  $R_{11}$  or  $R_{12}$  may comprise a compound having the general structure of Formula I:



wherein

$R_1$  and  $R_8$  are independently H or  $C_1$ - $C_8$  alkyl;

$R_2$  and  $R_4$  are independently selected from the group consisting of H,  $C_1$ - $C_8$  alkyl,  $C_2$ - $C_8$  alkenyl,  $(C_1$ - $C_4$  alkyl)OH,  $(C_1$ - $C_4$  alkyl)SH,  $(C_2$ - $C_3$  alkyl)SCH<sub>3</sub>,  $(C_1$ - $C_4$  alkyl)CONH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)COOH,  $(C_1$ - $C_4$  alkyl)NH<sub>2</sub>,  $(C_1$ - $C_4$  alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)( $C_3$ - $C_6$  cycloalkyl),  $(C_0$ - $C_4$  alkyl)( $C_2$ - $C_5$  heterocyclic),  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ , and CH<sub>2</sub>( $C_3$ - $C_9$

heteroaryl), or  $R_1$  and  $R_2$  together with the atoms to which they are attached form a  $C_3$ - $C_{12}$  cycloalkyl;

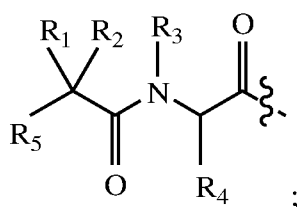
$R_3$  is  $C_1$ - $C_{18}$  alkyl;

$R_5$  is  $NHR_6$ ;

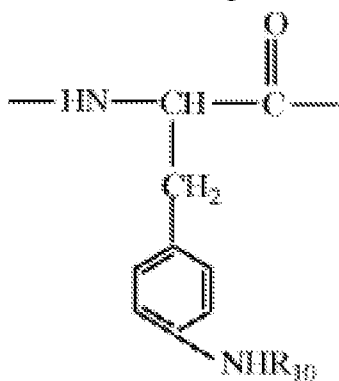
$R_6$  is H or  $C_1$ - $C_8$  alkyl; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl, ( $C_0$ - $C_4$  alkyl) $CONH_2$ , ( $C_0$ - $C_4$  alkyl) $COOH$ , ( $C_0$ - $C_4$  alkyl) $NH_2$ , ( $C_0$ - $C_4$  alkyl) $OH$ , and halo.

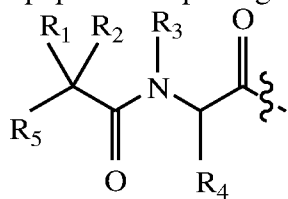
Also herein disclosed is an insulin analog wherein the A chain of the insulin peptide comprises the sequence Z-GIVEQCCTSICSLYQLENX<sub>2</sub>CN (SEQ ID NO: 6) and the B chain comprising a sequence selected from the group consisting of HLCGSHLVEALYLVCGERGFF (SEQ ID NO: 7), FVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO: 8) and FVNQHLCGSHLVEALYLVCGERGFFYTKPT (SEQ ID NO: 9) wherein Z is H or a dipeptide comprising the general structure:



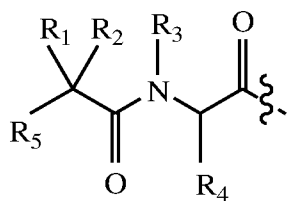
$X_2$  is an amino acid of the general structure



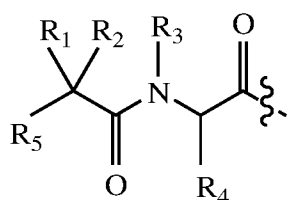
wherein  $R_{10}$  is H or a dipeptide comprising the general structure:



, with the proviso that Z and  $R_{10}$  are not both H and are not both a dipeptide comprising the general structure:



Also herein disclosed are single-chain insulin prodrug analogs. The carboxy terminus of the human insulin B chain, or a functional analog thereof, may be covalently linked to the N-terminus of the human insulin A chain, or functional analog thereof, wherein a dipeptide prodrug moiety having the general structure:



is covalently bound at the N-terminus of the peptide or at the side chain of an amino acid corresponding to positions A19, B 16 or B25 of the respective A chain or B chain via an amide bond. In one embodiment the B chain is linked to the A chain via peptide linker of 4-12 or 4-8 amino acids.

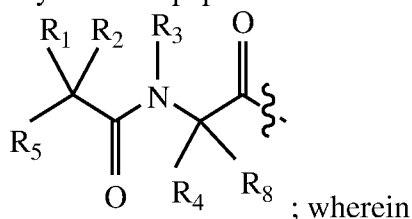
According to an embodiment of the invention, there is provided an insulin analog comprising an A chain, a B chain and a dipeptide element, wherein

said A chain has the general formula Z-peptide A, wherein peptide A comprises the sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3), or an analog thereof comprising a sequence that differs from SEQ ID NO: 3 by 1 to 3 amino acid modifications, selected from positions A5, A8, A9, A10, A14, A15, A17, and A18; and

said B chain has the general formula J-X<sub>14</sub>-peptide B, wherein peptide B comprises the sequence of X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14), or an analog thereof comprising a sequence that differs from SEQ ID NO: 14 by 1 to 3 amino acid modifications, selected from positions B1, B2, B3, B4, B5, B13, B14, B17, B20, B22, and B23;

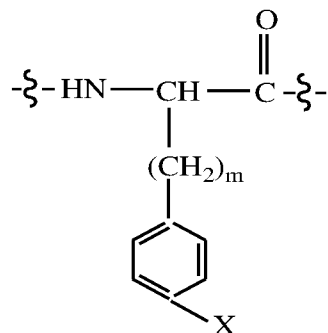
said B chain being linked to said A chain through intermolecular disulfide linkages, wherein

Z and J are independently H or a dipeptide element of the structure of Formula I:



X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the structure of:



wherein m is an integer selected from the group consisting of 0-3, and X is selected from the group consisting of OH, OCH<sub>3</sub>, NH<sub>2</sub> and NHR<sub>10</sub>, wherein R<sub>10</sub> is a dipeptide element comprising the general structure of Formula I;

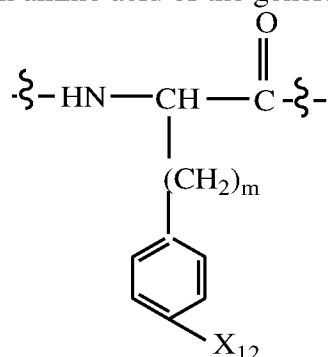
X<sub>3</sub> is selected from the group consisting of asparagine, glycine and alanine, threonine and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

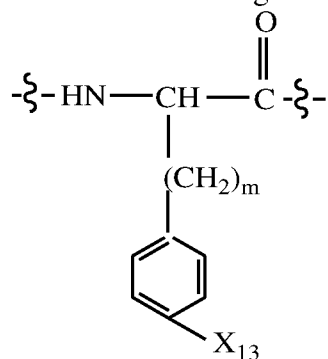
X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

X<sub>7</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X<sub>12</sub> is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>11</sub> and OCH<sub>3</sub>, wherein R<sub>11</sub> is a dipeptide element comprising the structure of Formula I;

X<sub>8</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X<sub>13</sub> is selected from the group consisting of H, OH, NH<sub>2</sub>, NHR<sub>12</sub> and OCH<sub>3</sub>, wherein R<sub>12</sub> is a dipeptide element of the structure of Formula I;

X<sub>14</sub> is selected from the group consisting of a bond, X<sub>9</sub>VNQ (SEQ ID NO: 21), VNQ, NQ and Q, wherein X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine;

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-

C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH; and

R<sub>13</sub> is COOH or CONH<sub>2</sub>, with the proviso that one or more of X, X<sub>12</sub>, X<sub>13</sub>, J and Z is a dipeptide element of the structure of Formula I and at least one of R<sub>1</sub> and R<sub>2</sub> is other than H;

said dipeptide element being linked to said A or B chain through an amide bond wherein the chemical cleavage half-life (t<sub>1/2</sub>) of said dipeptide from said A or B chain is at least about 1 hour to about 1 week in PBS under physiological conditions.

According to another embodiment of the invention, there is provided a dimer or multimer comprising an insulin analog of the invention as described above.

According to another embodiment of the invention, the first amino acid of the dipeptide element is an amino acid in the D-stereochemical configuration.

The solubility of the insulin prodrug analogs may be enhanced by the covalent linkage of a hydrophilic moiety to the peptide. The hydrophilic moiety may be linked to either the N-terminal amino acid of the B chain or to the amino acid at position 28 of SEQ ID NO: 9 or the amino acid at position 29 of SEQ ID NO: 8. The hydrophilic moiety may be a polyethylene glycol (PEG) chain, having a molecular weight selected from the range of about 500 to about 40,000 Daltons. The polyethylene glycol chain may have a molecular weight selected from the range of about 500 to about 5,000 Daltons. The polyethylene glycol chain may have a molecular weight of about 10,000 to about 20,000 Daltons.

Acylation or alkylation can increase the half-life of the insulin peptides in circulation. Acylation or alkylation can advantageously delay the onset of action and/or extend the duration of action at the insulin receptors upon activation of the prodrug. The insulin analogs may be acylated or alkylated at the same amino acid position where a hydrophilic moiety is linked, or at a different amino acid position. Thus, according to an embodiment, the analog is acylated at one or more positions selected from A9, A14, A15, B22, B28 or B29.

In accordance with one embodiment a pharmaceutical composition is provided comprising any of the novel insulin prodrug analogs disclosed herein, preferably at a purity level of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%, and a pharmaceutically acceptable diluent, carrier or excipient. Similarly, in accordance with

another embodiment, there is provided the use of a compound of the invention, in the manufacture of a medicament for the treatment of hyperglycemia. Such compositions may contain an A19 insulin analog as disclosed herein at a concentration of at least 0.5 mg/ml, 1 mg/ml, 2 mg/ml, 3 mg/ml, 4 mg/ml, 5 mg/ml, 6 mg/ml, 7 mg/ml, 8 mg/ml, 9 mg/ml, 10 mg/ml, 11 mg/ml, 12 mg/ml, 13 mg/ml, 14 mg/ml, 15 mg/ml, 16 mg/ml, 17 mg/ml, 18 mg/ml, 19 mg/ml, 20 mg/ml, 21 mg/ml, 22 mg/ml, 23 mg/ml, 24 mg/ml, 25 mg/ml or higher. In one embodiment the pharmaceutical compositions comprise aqueous solutions that are sterilized and optionally stored contained within various package containers. In other embodiments the pharmaceutical compositions comprise a lyophilized powder. The pharmaceutical compositions can be further packaged as part of a kit that includes a disposable device for administering the composition to a patient. The containers or kits may be labeled for storage at ambient room temperature or at refrigerated temperature.

In accordance with one embodiment an improved method of regulating blood glucose levels in insulin dependent patients is provided. The method comprises the steps of administering an insulin prodrug analog of the present disclosure in an amount therapeutically effective for the control of diabetes. In one embodiment the insulin prodrug analog is pegylated with a PEG chain having a molecular weight selected from the range of about 5,000 to about 40,000 Daltons. Thus, according to another embodiment of the invention, there is provided the use of a compound of the invention to treat diabetes.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. is a schematic overview of the two step synthetic strategy for preparing human insulin. Details of the procedure are provided in Example 1.

Fig. 2 is a graph comparing insulin receptor specific binding of synthetic human insulin relative to purified native insulin. As indicated by the data presented in the graph, the two molecules have similar binding activities.

Fig. 3 is a graph comparing relative insulin receptor binding of native insulin and the A19 insulin analog (Insulin(p-NH<sub>2</sub>-F)<sup>A19</sup>). As indicated by the data presented in the graph, the two molecules have similar binding activities.

Fig. 4 is a graph comparing relative insulin receptor binding of native insulin and the IGF1(YL)<sup>B16B17</sup> analog. As indicated by the data presented in the graph, the two molecules have similar binding activities.

Fig. 5 is a schematic drawing of the synthetic scheme used to prepare the IGF1(Y<sup>B16</sup>L<sup>B17</sup>)(p-NH<sub>2</sub>-F)<sup>A19</sup> analog (will not be activated).

Fig. 6 is a graph comparing relative insulin receptor binding of IGF1(Y<sup>B16</sup>L<sup>B17</sup>)(p-NH<sub>2</sub>-F)<sup>A19</sup> and the dipeptide extended form of IGF1(Y<sup>B16</sup>L<sup>B17</sup>)(p-



$\text{NH}_2\text{-F}^{\text{A19}}$ , wherein the dipeptide AiBAla is bound at position A19 (i.e.  $\text{IGF1}(\text{Y}^{\text{B16}}\text{L}^{\text{B17}})^{\text{A19}}\text{-AiBAla}$ ).

Fig. 7A-7C provides the activity of a dimer prepared in accordance with the present disclosure. Fig 7A shows the structure of an IGF-1 single chain dimer that  
 5 comprises two single chain  $\text{IGF}^{\text{B16B17}}$  derivative peptides (IGF-1B chain $[\text{C}^0\text{H}^5\text{Y}^{16}\text{L}^{17}\text{O}^{22}]$ -A chain $[\text{O}^{9,14,15}\text{N}^{18,21}]$ ; SEQ ID NO: 68) linked together by a disulfide bond between the side chains of the amino terminus of the B chains. Fig 7B is a graph demonstrating the relative insulin receptor binding of insulin, IGF-1, a single chain  $\text{IGF}^{\text{B16B17}}$  derivative peptide dimer and a two chain  $\text{IGF}^{\text{B16B17}}$  derivative  
 10 peptide dimer. Fig 7C is a graph demonstrating the relative activity of insulin, IGF-1, and a two chain  $\text{IGF}^{\text{B16B17}}$  derivative peptide dimer to induce insulin receptor phosphorylation.

Fig 8A-8C shows the degradation of a prodrug form of an  $\text{IGF}^{\text{B16B17}}$  derivative peptide: (Aib-Pro on  $(\text{pNH}_2\text{-F})^{19}$  of  $\text{IGF1A}(\text{Ala})^{6,7,11,20}$  amide. The dipeptide was  
 15 incubated in PBS, pH 7.4 at  $37^\circ\text{C}$  for predetermined lengths of time. Aliquots were taken at 20 minutes (Fig. 8A), 81 minutes (Fig 8B) and 120 minutes (Fig. 8C) after beginning the incubation, were quenched with 0.1% TFA and tested by analytical HPLC. Peak a ( $\text{IGF1A}(\text{Ala})^{6,7,11,20}(\text{pNH}_2\text{-F})^1$  amide) and b ( $\text{IGF1A}(\text{Ala})^{6,7,11,20}(\text{Aib-Pro-pNH-F})^{19}$  amide) were identified with LC-MS and quantified by integration of  
 20 peak area. The data indicate the spontaneous, non-enzymatic conversion of  $\text{IGF1A}(\text{Ala})^{6,7,11,20}(\text{Aib-Pro-pNH-F})^{19}$  amide to  $\text{IGF1A}(\text{Ala})^{6,7,11,20}(\text{pNH}_2\text{-F})^1$  amide over time.

Fig. 9A & 9B are graphs depicting the *in vitro* activity of the prodrug Aib,dPro-IGF1YL (dipeptide linked through the A19 4-aminoPhe). Fig 9A is a  
 25 graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at  $4^\circ\text{C}$ ) and the A19 IGF prodrug analog (Aib,dPro-IGF1YL) over time (0 hours, 2.5 hours and 10.6 hours) incubated in PBS. Fig 9B is a graph comparing relative insulin receptor binding of native insulin (measured at 1.5 hour at  $4^\circ\text{C}$ ) and the A19 IGF prodrug analog (Aib,dPro-IGF1YL) over time (0 hours, 1.5 hours and 24.8 hours)  
 30 incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered from the A19 IGF prodrug analog sample as the prodrug form is converted to the active IGF1YL peptide.

Fig. 10A & 10B are graphs depicting the *in vitro* activity of the prodrug dK,(N-isobutylG)-IGF1YL (di-peptide linked through the A19 4-aminoPhe). Fig 10A is a graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK,(N-isobutylG) over time (0 hours, 5 hours and 52 hours) incubated in PBS. Fig 10B is a graph comparing relative insulin receptor binding of native insulin (measured at 1.5 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK,(N-isobutylG) over time (0 hours, 3.6 hours and 24.8 hours) incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered from the A19 IGF prodrug analog sample as the prodrug form is converted to the active IGF1YL peptide.

Fig. 11A & 11B are graphs depicting the *in vitro* activity of the prodrug dK(e-acetyl),Sar)-IGF1YL (di-peptide linked through the A19 4-aminoPhe). Fig 11A is a graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK(e-acetyl),Sar) over time (0 hours, 7.2 hours and 91.6 hours) incubated in PBS. Fig 11B is a graph comparing relative insulin receptor binding of native insulin (measured at 1.5 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK(e-acetyl),Sar) over time (0 hours, 9 hours and 95 hours) incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered from the A19 IGF prodrug analog sample as the prodrug form is converted to the active IGF1YL peptide.

## DETAILED DESCRIPTION

### DEFINITIONS

In describing and claiming the invention, the following terminology will be used in accordance with the definitions set forth below.

As used herein, the term “prodrug” is defined as any compound that undergoes chemical modification before exhibiting its pharmacological effects.

A “bioactive polypeptide” refers to polypeptides which are capable of exerting a biological effect *in vitro* and/or *in vivo*.

As used herein the term "amino acid" encompasses any molecule containing both amino and carboxyl functional groups, wherein the amino and carboxylate groups are attached to the same carbon (the alpha carbon). The alpha carbon optionally may have one or two further organic substituents. Designation of an amino

acid without specifying its stereochemistry is intended to encompass either the L or D form of the amino acid or a racemic mixture. However, in the instance where an amino acid is designated by its three letter code and includes a superscript number, the D form of the amino acid is specified by inclusion of a lower case d before the three letter code and superscript number (e.g., dLys<sup>-1</sup>), wherein the designation lacking the lower case d (e.g., Lys<sup>-1</sup>) is intended to specify the native L form of the amino acid. In this nomenclature, the inclusion of the superscript number designates the position of the amino acid in the IGF peptide sequence, wherein amino acids that are located within the IGF sequence are designated by positive superscript numbers numbered consecutively from the N-terminus. Additional amino acids linked to the IGF peptide either at the N-terminus or through a side chain are numbered starting with 0 and increasing in negative integer value as they are further removed from the IGF sequence. For example, the position of an amino acid within a dipeptide prodrug linked to the N-terminus of IGF is designated aa<sup>-1</sup>-aa<sup>0</sup>-IGF wherein aa<sup>0</sup> represents the carboxy terminal amino acid of the dipeptide and aa<sup>-1</sup> designates the amino terminal amino acid of the dipeptide.

As used herein the term "hydroxyl acid" refers to amino acids that have been modified to replace the alpha carbon amino group with a hydroxyl group.

As used herein the term "non-coded amino acid" encompasses any amino acid that is not an L-isomer of any of the following 20 amino acids: Ala, Cys, Asp, Glu, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, Tyr.

A "dipeptide" is a compound formed by linkage of an alpha amino acid or an alpha hydroxyl acid to another amino acid, through a peptide bond.

As used herein the term "chemical cleavage" absent any further designation encompasses a non-enzymatic reaction that results in the breakage of a covalent chemical bond.

A "bioactive polypeptide" refers to polypeptides which are capable of exerting a biological effect *in vitro* and/or *in vivo*.

As used herein a general reference to a peptide is intended to encompass peptides that have modified amino and carboxy termini. For example, an amino acid sequence designating the standard amino acids is intended to encompass standard amino acids at the N- and C- terminus as well as a corresponding hydroxyl acid at the

N-terminus and/or a corresponding C-terminal amino acid modified to comprise an amide group in place of the terminal carboxylic acid.

As used herein an "acylated" amino acid is an amino acid comprising an acyl group which is non-native to a naturally-occurring amino acid, regardless by the means by which it is produced. Exemplary methods of producing acylated amino acids and acylated peptides are known in the art and include acylating an amino acid before inclusion in the peptide or peptide synthesis followed by chemical acylation of the peptide. In one embodiment, the acyl group causes the peptide to have one or more of (i) a prolonged half-life in circulation, (ii) a delayed onset of action, (iii) an extended duration of action, (iv) an improved resistance to proteases, such as DPP-IV, and (v) increased potency at the insulin peptide receptor.

As used herein, an "alkylated" amino acid is an amino acid comprising an alkyl group which is non-native to a naturally-occurring amino acid, regardless of the means by which it is produced. Exemplary methods of producing alkylated amino acids and alkylated peptides are known in the art and including alkylating an amino acid before inclusion in the peptide or peptide synthesis followed by chemical alkylation of the peptide. Without being held to any particular theory, it is believed that alkylation of peptides will achieve similar, if not the same, effects as acylation of the peptides, e.g., a prolonged half-life in circulation, a delayed onset of action, an extended duration of action, an improved resistance to proteases, such as DPP-IV, and increased potency at the insulin peptide receptor.

As used herein, the term "pharmaceutically acceptable carrier" includes any of the standard pharmaceutical carriers, such as a phosphate buffered saline solution, water, emulsions such as an oil/water or water/oil emulsion, and various types of wetting agents. The term also encompasses any of the agents approved by a regulatory agency of the US Federal government or listed in the US Pharmacopeia for use in animals, including humans.

As used herein the term "pharmaceutically acceptable salt" refers to salts of compounds that retain the biological activity of the parent compound, and which are not biologically or otherwise undesirable. Many of the compounds disclosed herein are capable of forming acid and/or base salts by virtue of the presence of amino and/or carboxyl groups or groups similar thereto.

Pharmaceutically acceptable base addition salts can be prepared from

inorganic and organic bases. Salts derived from inorganic bases, include by way of example only, sodium, potassium, lithium, ammonium, calcium and magnesium salts. Salts derived from organic bases include, but are not limited to, salts of primary, secondary and tertiary amines.

5           Pharmaceutically acceptable acid addition salts may be prepared from inorganic and organic acids. Salts derived from inorganic acids include hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid, and the like. Salts derived from organic acids include acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, malic acid, malonic acid, succinic acid, maleic acid, fumaric acid,  
10   tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluene-sulfonic acid, salicylic acid, and the like.

As used herein, the term "treating" includes prophylaxis of the specific disorder or condition, or alleviation of the symptoms associated with a specific disorder or condition and/or preventing or eliminating said symptoms. For example,  
15   as used herein the term "treating diabetes" will refer in general to maintaining glucose blood levels near normal levels and may include increasing or decreasing blood glucose levels depending on a given situation.

As used herein an "effective" amount or a "therapeutically effective amount" of a prodrug refers to a nontoxic but sufficient amount of the prodrug to provide the  
20   desired effect. For example one desired effect would be the prevention or treatment of hyperglycemia. The amount that is "effective" will vary from subject to subject, depending on the age and general condition of the individual, mode of administration, and the like. Thus, it is not always possible to specify an exact "effective amount." However, an appropriate "effective" amount in any individual case may be  
25   determined by one of ordinary skill in the art using routine experimentation.

The term, "parenteral" means not through the alimentary canal but by some other route such as intranasal, inhalation, subcutaneous, intramuscular, intraspinal, or intravenous.

As used herein the term "native insulin peptide" is intended to designate the 51  
30   amino acid heterodimer comprising the A chain of SEQ ID NO: 1 and the B chain of SEQ ID NO: 2, as well as single-chain insulin analogs that comprise SEQ ID NOS: 1 and 2. The term "insulin peptide" as used herein, absent further descriptive language is intended to encompass the 51 amino acid heterodimer comprising the A chain of

SEQ ID NO: 1 and the B chain of SEQ ID NO: 2, as well as single-chain insulin analogs thereof (including for example those disclosed in published international application WO96/34882 and US Patent No. 6,630,348, the disclosures of which are incorporated herein by reference), and includes heterodimers and single-chain analogs  
5 that comprise modified derivatives of the native A chain and/or B chain, including modification of the amino acid at position A19, B16 or B25 to a 4-amino phenylalanine or one or more amino acid substitutions at positions selected from A5, A8, A9, A10, A12, A14, A15, A17, A18, A21, B1, B2, B3, B4, B5, B9, B10, B13, B14, B17, B20, B21, B22, B23, B26, B27, B28, B29 and B30 or deletions of any or  
10 all of positions B1-4 and B26-30. An "insulin prodrug analog" as used herein refers to an insulin peptide (or an IGF1-based insulin analog as disclosed in Example 9) that has been modified by the covalent attachment of a dipeptide, via an amide linkage, at a location that interferes with insulin's or IGF1-based insulin analog's activity (e.g., the ability to interact with the insulin and IGF-1 receptors).

15 As used herein, the term "single-chain insulin analog" encompasses a group of structurally-related proteins wherein the insulin A and B chains are covalently linked.

As used herein an amino acid "modification" refers to a substitution, addition or deletion of an amino acid, or the derivation of an amino acid by the addition and/or removal of chemical groups to/from the amino acid, and includes substitution with or  
20 addition of any of the 20 amino acids commonly found in human proteins, as well as atypical or non-naturally occurring amino acids. Commercial sources of atypical amino acids include Sigma-Aldrich (Milwaukee, WI), ChemPep Inc. (Miami, FL), and Genzyme Pharmaceuticals (Cambridge, MA). Atypical amino acids may be purchased from commercial suppliers, synthesized de novo, or chemically modified or  
25 derivatized from naturally occurring amino acids.

As used herein an amino acid "substitution" refers to the replacement of one amino acid residue by a different amino acid residue. Throughout the application, all references to a particular amino acid position by letter and number (e.g. position A5) refer to the amino acid at that position of either the A chain (e.g. position A5) or  
30 the B chain (e.g. position B5) in the respective native human insulin A chain (SEQ ID NO: 1) or B chain (SEQ ID NO: 2), or the corresponding amino acid position in any analogs thereof. For example, a reference herein to "position B28" absent any further

elaboration would mean the corresponding position B27 of the B chain of an insulin analog in which the first amino acid of SEQ ID NO: 2 has been deleted.

As used herein, the term "conservative amino acid substitution" is defined herein as exchanges within one of the following five groups:

- 5 I. Small aliphatic, nonpolar or slightly polar residues:  
Ala, Ser, Thr, Pro, Gly;
- II. Polar, negatively charged residues and their amides:  
Asp, Asn, Glu, Gln;
- III. Polar, positively charged residues:  
10 His, Arg, Lys; Ornithine (Orn)
- IV. Large, aliphatic, nonpolar residues:  
Met, Leu, Ile, Val, Cys, Norleucine (Nle), homocysteine
- V. Large, aromatic residues:  
Phe, Tyr, Trp, acetyl phenylalanine

15

As used herein the general term "polyethylene glycol chain" or "PEG chain", refers to mixtures of condensation polymers of ethylene oxide and water, in a branched or straight chain, represented by the general formula  $H(OCH_2CH_2)_nOH$ , wherein n is at least 9. Absent any further characterization, the term is intended to  
20 include polymers of ethylene glycol with an average total molecular weight selected from the range of 500 to 80,000 Daltons. "Polyethylene glycol chain" or "PEG chain" is used in combination with a numeric suffix to indicate the approximate average molecular weight thereof. For example, PEG-5,000 refers to polyethylene glycol chain having a total molecular weight average of about 5,000 Daltons.

25

As used herein the term "pegylated" and like terms refers to a compound that has been modified from its native state by linking a polyethylene glycol chain to the compound. A "pegylated polypeptide" is a polypeptide that has a PEG chain covalently bound to the polypeptide.

30

As used herein a "linker" is a bond, molecule or group of molecules that binds two separate entities to one another. Linkers may provide for optimal spacing of the two entities or may further supply a labile linkage that allows the two entities to be separated from each other. Labile linkages include photocleavable groups, acid-labile

moieties, base-labile moieties and enzyme-cleavable groups.

As used herein an "insulin dimer" is a complex comprising two insulin peptides covalently bound to one another via a linker. The term insulin dimer, when used absent any qualifying language, encompasses both insulin homodimers and  
5 insulin heterodimers. An insulin homodimer comprises two identical subunits (each comprising an A and B chain), whereas an insulin heterodimer comprises two subunits that differ, although the two subunits are substantially similar to one another.

The term " $C_1$ - $C_n$  alkyl" wherein n can be from 1 through 6, as used herein, represents a branched or linear alkyl group having from one to the specified number  
10 of carbon atoms. Typical  $C_1$ - $C_6$  alkyl groups include, but are not limited to, methyl, ethyl, n-propyl, iso-propyl, butyl, iso-butyl, sec-butyl, tert-butyl, pentyl, hexyl and the like.

The terms " $C_2$ - $C_n$  alkenyl" wherein n can be from 2 through 6, as used herein, represents an olefinically unsaturated branched or linear group having from 2 to the  
15 specified number of carbon atoms and at least one double bond. Examples of such groups include, but are not limited to, 1-propenyl, 2-propenyl ( $-CH_2-CH=CH_2$ ), 1,3-butadienyl, ( $-CH=CHCH=CH_2$ ), 1-butenyl ( $-CH=CHCH_2CH_3$ ), hexenyl, pentenyl, and the like.

The term " $C_2$ - $C_n$  alkynyl" wherein n can be from 2 to 6, refers to an  
20 unsaturated branched or linear group having from 2 to n carbon atoms and at least one triple bond. Examples of such groups include, but are not limited to, 1-propynyl, 2-propynyl, 1-butylnyl, 2-butylnyl, 1-pentylnyl, and the like.

As used herein the term "aryl" refers to a mono- or bicyclic carbocyclic ring system having one or two aromatic rings including, but not limited to, phenyl,  
25 naphthyl, tetrahydronaphthyl, indanyl, indenyl, and the like. The size of the aryl ring and the presence of substituents or linking groups are indicated by designating the number of carbons present. For example, the term " $(C_1$ - $C_3$  alkyl)( $C_6$ - $C_{10}$  aryl)" refers to a 5 to 10 membered aryl that is attached to a parent moiety via a one to three membered alkyl chain.

30 The term "heteroaryl" as used herein refers to a mono- or bi- cyclic ring system containing one or two aromatic rings and containing at least one nitrogen, oxygen, or sulfur atom in an aromatic ring. The size of the heteroaryl ring and the presence of substituents or linking groups are indicated by designating the number of



carbons present. For example, the term "(C<sub>1</sub>-C<sub>n</sub> alkyl)(C<sub>5</sub>-C<sub>6</sub> heteroaryl)" refers to a 5 or 6 membered heteroaryl that is attached to a parent moiety via a one to "n" membered alkyl chain.

The term "C<sub>3</sub>-C<sub>n</sub> cycloalkyl" refers to a non-aromatic, monocyclic or polycyclic ring comprising carbon and hydrogen atoms with the subscript number indicating the number of carbon atoms present. For example the term C<sub>3</sub>-C<sub>8</sub> cycloalkyl represents the compounds cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, and cyclooctyl.

The term "C<sub>3</sub>-C<sub>n</sub> heterocyclic" refers to a cycloalkyl ring system containing from one to "n - 1" heteroatoms wherein the heteroatoms are selected from the group consisting of oxygen, sulfur, and nitrogen. For example the phrase "5-membered heterocycle" or "C<sub>5</sub> heterocycle" includes, but is not limited to, 5-membered heterocycles having one hetero atom (e.g. thiophenes, pyrroles, furans); 5-membered heterocycles having two heteroatoms in 1,2 or 1,3 positions (e.g. oxazoles, pyrazoles, imidazoles, thiazoles, purines); 5-membered heterocycles having three heteroatoms (e.g. triazoles, thiadiazoles).

The term "C<sub>3</sub>-C<sub>n</sub> membered ring" as used herein refers to a saturated or unsaturated hydrocarbon ring structure comprising a total of three to "n" number of elements linked to one another to form a ring, wherein the ring elements are selected from the group consisting of C, O, S and N. The term is intended to encompass cycloalkyls, heterocycles, aryls and heteroaryls.

As used herein, the term "halo" refers to one or more members of the group consisting of fluorine, chlorine, bromine, and iodine.

As used herein the term "patient" without further designation is intended to encompass any warm blooded vertebrate domesticated animal (including for example, but not limited to livestock, horses, cats, dogs and other pets) and humans.

## EMBODIMENTS

The present disclosure provides insulin prodrug derivatives that are formulated to delay onset of action and enhance the half life of the insulin peptide, thus improving the therapeutic index of the underlying insulin peptide. The insulin prodrug chemistry disclosed herein allows for activation of the prodrug via a non-enzymatic degradation mechanism. The disclosed prodrug chemistry can be

chemically conjugated to active site amines to form amides that revert to the parent amine upon diketopiperazine formation and release of the prodrug element. This novel biologically friendly prodrug chemistry spontaneously degrades under physiological conditions (e.g. pH of about 7, at 37 °C in an aqueous environment) and is not reliant on enzymatic degradation. The duration of the prodrug derivative is determined by the selection of the dipeptide prodrug sequence, and thus allows for flexibility in prodrug formulation.

In one embodiment a prodrug is provided having a non-enzymatic activation half time ( $t_{1/2}$ ) of between 1-100 hrs under physiological conditions. Physiological conditions as disclosed herein are intended to include a temperature of about 35 to 40 °C and a pH of about 7.0 to about 7.4 and more typically include a pH of 7.2 to 7.4 and a temperature of 36 to 38 °C in an aqueous environment. In one embodiment a dipeptide, capable of undergoing diketopiperazine formation under physiological conditions, is covalently linked through an amide linkage to the insulin peptide.

Advantageously, the rate of cleavage, and thus activation of the prodrug, depends on the structure and stereochemistry of the dipeptide pro-moiety and also on the strength of the nucleophile. The prodrugs disclosed herein will ultimately be chemically converted to structures that can be recognized by the native receptor of the drug, wherein the speed of this chemical conversion will determine the time of onset and duration of *in vivo* biological action. The prodrug chemistry disclosed in this application relies upon an intramolecular chemical reaction that is not dependent upon additional chemical additives, or enzymes. The speed of conversion is controlled by the chemical nature of the dipeptide substituent and its cleavage under physiological conditions. Since physiological pH and temperature are tightly regulated within a highly defined range, the speed of conversion from prodrug to drug will exhibit high intra and interpatient reproducibility.

As disclosed herein prodrugs are provided wherein the bioactive polypeptides have extended half lives of at least 1 hour, and more typically greater than 20 hours but less than 100 hours, and are converted to the active form at physiological conditions through a non-enzymatic reaction driven by inherent chemical instability. In one embodiment the a non-enzymatic activation  $t_{1/2}$  time of the prodrug is between 1-100 hrs, and more typically between 12 and 72 hours, and in one embodiment the  $t_{1/2}$  is between 24-48 hrs as measured by incubating the prodrug in a

phosphate buffer solution (e.g., PBS) at 37 °C and pH of 7.2. The half lives of the various prodrugs are calculated by using the formula  $t_{1/2} = .693/k$ , where 'k' is the first order rate constant for the degradation of the prodrug. In one embodiment, activation of the prodrug occurs after cleavage of an amide bond linked dipeptide, and  
5 formation of a diketopiperazine or diketomorpholine, and the active insulin peptide.

Specific dipeptides composed of natural or synthetic amino acids have been identified that facilitate intramolecular decomposition under physiological conditions to release active insulin peptides. The dipeptide can be linked (via an amide bond) to an amino group present on native insulin, or an amino group introduced into the  
10 insulin peptide by modification of the native insulin peptide. In one embodiment the dipeptide structure is selected to resist cleavage by peptidases present in mammalian sera, including for example dipeptidyl peptidase IV (DPP-IV). Accordingly, in one embodiment the rate of cleavage of the dipeptide prodrug element from the bioactive peptide is not substantially enhanced (e.g., greater than 2X) when the reaction is  
15 conducted using physiological conditions in the presence of serum proteases, relative to conducting the reaction in the absence of the proteases. Thus the cleavage half-life of the dipeptide prodrug element from the insulin peptide (in PBS under physiological conditions) is not more than two, three, four or five fold the cleavage half-life of the dipeptide prodrug element from the insulin peptide in a solution comprising a DPP-IV  
20 protease. In one embodiment the solution comprising a DPP-IV protease is serum, more particularly mammalian serum, including human serum.

In accordance with one embodiment the dipeptide prodrug element comprises the structure U-O, wherein U is an amino acid or a hydroxyl acid and O is an N-alkylated amino acid. In one embodiment U, O, or the amino acid of the insulin  
25 peptide to which U-O is linked is a non-coded amino acid. In one embodiment U and/or O is an amino acid in the D stereoisomer configuration. In some exemplary embodiments, U is an amino acid in the D stereoisomer configuration and O is an amino acid in the L stereoisomer configuration. In some exemplary embodiments, U is an amino acid in the L stereoisomer configuration and O is an amino acid in the D  
30 stereoisomer configuration. In some exemplary embodiments, U is an amino acid in the D stereoisomer configuration and O is an amino acid in the D stereoisomer configuration. In one embodiment O is an N-alkylated amino acid but is not proline. In one embodiment the N-alkylated group of amino acid O is a C<sub>1</sub>-C<sub>18</sub> alkyl, and in

one embodiment N-alkylated group is C<sub>1</sub>-C<sub>6</sub> alkyl. In one embodiment U-O is a dipeptide comprising the structure of Formula I as defined herein.

In one embodiment the dipeptide is linked to the insulin peptide at an amino group selected from the N-terminal amino group of the A or B chain, or the side chain amino group of an amino acid present at an active site of the insulin peptide. In accordance with one embodiment the dipeptide extension is covalently linked to an insulin peptide through the side chain amine of a lysine residue that resides at or near the active site. In one embodiment the dipeptide extension is attached through a synthetic amino acid or a modified amino acid, wherein the synthetic amino acid or modified amino acid exhibits a functional group suitable for covalent attachment of the dipeptide extension (e.g., the aromatic amine of amino-phenylalanine). In accordance with one embodiment the dipeptide is linked to the insulin peptide at an amino group selected from the N-terminal amino group of the A or B chain, or the side chain amino group of an aromatic amine (e.g., a 4-amino-phenylalanine residue) present at position A19, B16 or B25. In one embodiment the U-O dipeptide is bound at position A19 through a 4-amino phenylalanine present at position A19.

The dipeptide prodrug element is designed to spontaneously cleave its amide linkage to the insulin analog under physiological conditions and in the absence of enzymatic activity. In one embodiment the N-terminal amino acid of the dipeptide extension comprises a C-alkylated amino acid (e.g. amino isobutyric acid). In one embodiment the C-terminal amino acid of the dipeptide comprises an N-alkylated amino acid (e.g., proline or N-methyl glycine). In one embodiment the dipeptide comprises the sequence of an N-terminal C-alkylated amino acid followed by an N-alkylated amino acid.

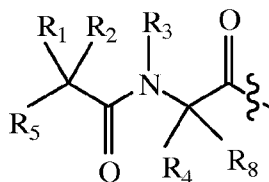
Applicants have discovered that the selective insertion of a 4-amino phenyl amino acid moiety for the native tyrosine at position 19 of the A chain can be accommodated without loss in potency of the insulin peptide (see Fig. 3). Subsequent chemical amidation of this active site amino group with the dipeptide prodrug moiety disclosed herein dramatically lessens insulin receptor binding activity and thus provides a suitable prodrug of insulin (see Fig. 6, data provided for the IGF1Y<sup>16</sup>L<sup>17</sup> (p-NH<sub>2</sub>-F)<sup>A19</sup> analog which has been demonstrated to have comparable activity as insulin (p-NH<sub>2</sub>-F)<sup>A19</sup>, see Fig. 4). Accordingly, in one embodiment the dipeptide prodrug element is linked to the aromatic ring of an A19 4-aminophenylalanine via an

amide bond, wherein the C-terminal amino acid of the dipeptide comprises an N-alkylated amino acid and the N-terminal amino acid of the dipeptide is any amino acid.

The dipeptide prodrug moiety can also be attached to additional sites of an insulin peptide to prepare a prodrug or depot analog of insulin. In accordance with one embodiment an insulin prodrug/depot analog is provided comprising an A chain, a B chain, and a dipeptide linked via an amide bond to one or more sites selected from the group consisting of the N-terminal amino group of the A chain or B chain, or the side chain amino group of an internal amino acid including for example, linked to the aromatic amine of a 4-amino-phenylalanine residue present at position A19, B16 or B25. In one embodiment the insulin peptide comprises two dipeptide elements, wherein the dipeptide elements are optionally pegylated, alkylated, acylated or linked to a depot polymer. In one embodiment the dipeptide comprises an N-terminal C-alkylated amino acid followed by an N-alkylated amino acid.

The A chain and B chain comprising the insulin prodrug analog may comprise the native sequence of the respective peptides (i.e., SEQ ID NO: 1 and SEQ ID NO: 2) or may comprise a derivative of SEQ ID NO: 1 and/or SEQ ID NO: 2 wherein the derivative includes modification of the amino acid at position A19, B16 or B25 to a 4-amino phenylalanine and/or one or more amino acid substitutions at positions selected from A5, A8, A9, A10, A14, A15, A17, A18, A19 and A21, B1, B2, B3, B4, B5, B9, B10, B13, B14, B17, B20, B22, B23, B26, B27, B28, B29 and B30 or deletions of any or all of positions B1-4 and B26-30. In one embodiment the dipeptide prodrug element is linked to an N-terminal amino group of the A or B chain, wherein the C-terminal amino acid of the dipeptide prodrug element comprises an N-alkylated amino acid and the N-terminal amino acid of the dipeptide prodrug element is any amino acid, with the proviso that when the C-terminal amino acid of the dipeptide is proline, the N-terminal amino acid of the dipeptide comprises a C-alkylated amino acid.

In one embodiment the dipeptide prodrug element comprises the general structure of Formula I:



wherein

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl or aryl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

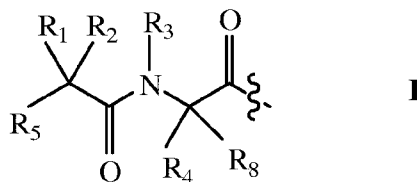
R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH, with the proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are other than H.

In another embodiment the dipeptide prodrug element comprises the general structure:



wherein

R<sub>1</sub> and R<sub>8</sub> are independently H or C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub>

alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>3</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the

5 atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH and halo, provided that when the dipeptide of Formula I is linked through the N-terminal amine of a peptide and R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are not H. In one embodiment either the first amino acid and/or the second amino acid of the dipeptide prodrug element is an amino acid in the D stereoisomer configuration.

In one embodiment the prodrug element of Formula I is provided wherein R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>8</sub> cycloalkyl ring;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>7</sub> is selected from the group consisting of H and OH and R<sub>8</sub> is H. In one embodiment R<sub>3</sub> is C<sub>1</sub>-C<sub>8</sub> alkyl and R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, CH<sub>2</sub>OH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub>

and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring. In a further embodiment R<sub>5</sub> is NHR<sub>6</sub> and R<sub>8</sub> is H.

In a further embodiment the prodrug element of Formula I is provided wherein

R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

5 R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are  
10 attached form a C<sub>3</sub>-C<sub>8</sub> cycloalkyl ring;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

15 R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH and halo; and R<sub>8</sub> is H, provided that when R<sub>4</sub> and R<sub>3</sub> together with the  
20 atoms to which they are attached form a 5 or 6 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are not H. In one embodiment either the first amino acid and/or the second amino acid of the dipeptide prodrug element is a non-coded amino acid and in one embodiment is an amino acid in the D stereoisomer configuration.

25 In other embodiments the dipeptide prodrug element has the structure of Formula I, wherein

R<sub>1</sub> and R<sub>8</sub> are independently H or C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>3</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the  
30 atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl;



R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

R<sub>5</sub> is NHR<sub>6</sub>;

R<sub>6</sub> is H or C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo and R<sub>8</sub> is H.

In a further embodiment the dipeptide prodrug element has the structure of Formula I, wherein

R<sub>1</sub> and R<sub>2</sub> are independently C<sub>1</sub>-C<sub>18</sub> alkyl or (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>; or  
 10 R<sub>1</sub> and R<sub>2</sub> are linked through -(CH<sub>2</sub>)<sub>p</sub>, wherein p is 2-9;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

R<sub>4</sub> and R<sub>8</sub> are each hydrogen;

R<sub>5</sub> is NH<sub>2</sub>; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo.

In a further embodiment the dipeptide prodrug element has the structure of Formula I, wherein

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>,  
 20 or R<sub>1</sub> and R<sub>2</sub> are linked through (CH<sub>2</sub>)<sub>p</sub>, wherein p is 2-9;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-12 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NH<sub>2</sub>; and

R<sub>7</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo, with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen and provided that at least  
 30 one of R<sub>4</sub> or R<sub>8</sub> is hydrogen.

In another embodiment the dipeptide prodrug element has the structure of Formula I, wherein

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, or R<sub>1</sub> and R<sub>2</sub> are linked through (CH<sub>2</sub>)<sub>p</sub>, wherein p is 2-9;

5 R<sub>3</sub> is C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-6 heterocyclic ring;

R<sub>4</sub> is selected from the group consisting of hydrogen and C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>8</sub> is hydrogen; and

R<sub>5</sub> is NH<sub>2</sub>, with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen.

10 In a further embodiment the dipeptide prodrug element has the structure of Formula I, wherein

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>;

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl;

R<sub>4</sub> and R<sub>8</sub> are each hydrogen; and

15 R<sub>5</sub> is NH<sub>2</sub>, with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen.

In another embodiment the dipeptide prodrug element has the structure of Formula I, wherein

20 R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen and C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, or R<sub>1</sub> and R<sub>2</sub> are linked through (CH<sub>2</sub>)<sub>p</sub>, wherein p is 2-9;

R<sub>3</sub> is C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>4</sub> is (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NH<sub>2</sub>;

25 R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)OH; and

R<sub>8</sub> is hydrogen, with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen.

In another embodiment the dipeptide prodrug element has the structure of Formula I, wherein

30 R<sub>1</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>2</sub> is hydrogen;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

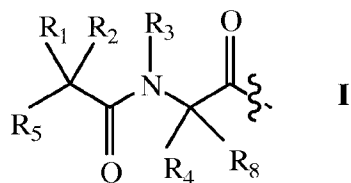
R<sub>4</sub> and R<sub>8</sub> are each hydrogen;

$R_5$  is  $NHR_6$  or  $OH$ ;

$R_6$  is  $H$ ,  $C_1$ - $C_8$  alkyl, or  $R_6$  and  $R_1$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl) $CONH_2$ ,  $(C_0$ - $C_4$  alkyl) $COOH$ ,  $(C_0$ - $C_4$  alkyl) $NH_2$ ,  $(C_0$ - $C_4$  alkyl) $OH$ , and halo, with the proviso that, if  $R_1$  is alkyl or  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ , then  $R_1$  and  $R_5$  together with the atoms to which they are attached form a 4-11 heterocyclic ring.

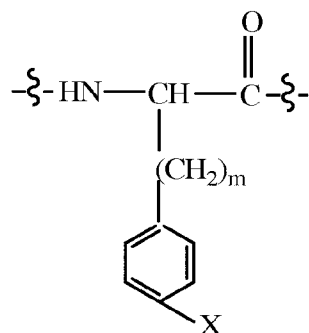
In accordance with one embodiment an insulin prodrug analog is provided comprising an insulin peptide and an amide linked dipeptide. More particularly, the insulin prodrug analog comprises an A chain sequence and a B chain sequence wherein the A chain comprises the sequence  $Z$ -GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub>- $R_{13}$  (SEQ ID NO: 3), or an analog thereof comprising a sequence that differs from SEQ ID NO: 3 by 1 to 9, 1 to 5 or 1 to 3 amino acid modifications, selected from positions A5, A8, A9, A10, A14, A15, A17, A18 (relative to the native insulin A chain), and the B chain sequence comprises the sequence of  $J$ -X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14), or an analog thereof comprising a sequence that differs from SEQ ID NO: 14 sequence by 1 to 10, 1 to 5 or 1 to 3 amino acid modifications, selected from positions B1, B2, B3, B4, B5, B13, B14, B17, B20, B22, B23, B26, B27, B28, B29 and B30 (relative to the native insulin B chain; i.e., amino acid X<sub>4</sub> of SEQ ID NO: 14 corresponds to position B5 in native insulin).  $Z$  and  $J$  are independently  $H$  or a dipeptide comprising the general structure of Formula I:



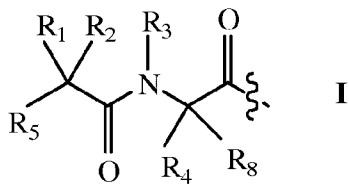
$X_{14}$  is either a bond joining the "J" element to the X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) sequence or  $X_{14}$  represents a 1 to 4 amino acid sequence selected from the group consisting of a X<sub>9</sub>VNQ (SEQ ID NO: 21), VNQ, NQ and Q that joins the "J" element to the X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) sequence.

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>, wherein R<sub>10</sub> is H or a dipeptide comprising the general structure of Formula I:



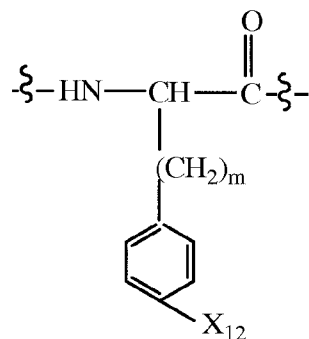
X<sub>3</sub> is selected from the group consisting of asparagine, glycine, alanine, threonine and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

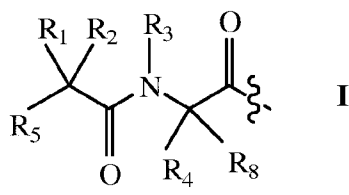
X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

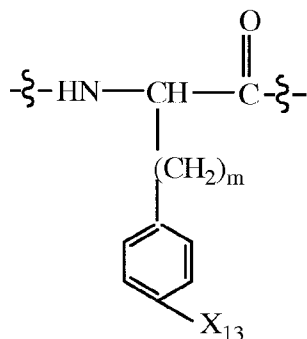
X<sub>7</sub> is an amino acid of the general structure



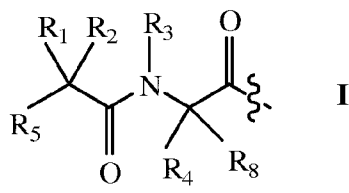
wherein m is an integer selected from 0-3 and X<sub>12</sub> is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>11</sub> and OCH<sub>3</sub>, wherein R<sub>11</sub> is H or a dipeptide comprising the general structure of Formula I:



X<sub>8</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X<sub>13</sub> is selected from the group consisting of H, OH, NH<sub>2</sub>, NHR<sub>12</sub> and OCH<sub>3</sub>, wherein R<sub>12</sub> is H or a dipeptide comprising the general structure of Formula I:



X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine; wherein

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl or aryl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub>

heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

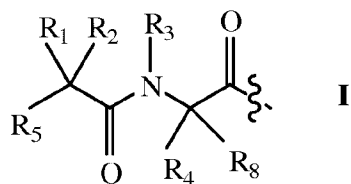
R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are  
5 attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH; and

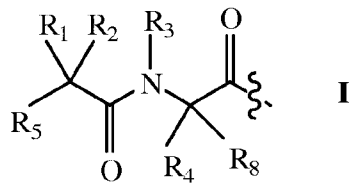
R<sub>13</sub> is COOH or CONH<sub>2</sub>. In one embodiment R<sub>13</sub> is COOH and the carboxy terminal amino acid of the B chain has an amide (CONH<sub>2</sub>) in place of the natural alpha carbon carboxy group. In one embodiment one or more of X, X<sub>12</sub>, X<sub>13</sub>, J and Z

10 is a dipeptide comprising the general structure of Formula I:



, and in one embodiment two of X, X<sub>12</sub>,

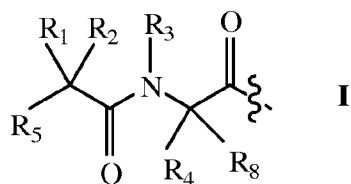
X<sub>13</sub>, J and Z comprise a dipeptide of the general structure of Formula I:



. In accordance with one embodiment at

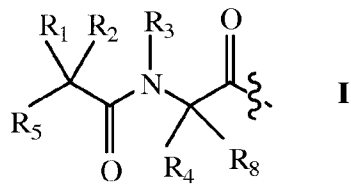
least one of X, X<sub>12</sub>, X<sub>13</sub>, J and Z is a dipeptide comprising the general structure of

15 Formula I:



, and in one embodiment one and only one

of X, X<sub>12</sub>, X<sub>13</sub>, J and Z comprises a dipeptide of the general structure of Formula I:



(i.e., only one dipeptide prodrug element is

attached to the insulin peptide). Furthermore, when the dipeptide prodrug element is

20 linked to the N-terminus of the A or B chain (i.e., either J or Z comprise the dipeptide)

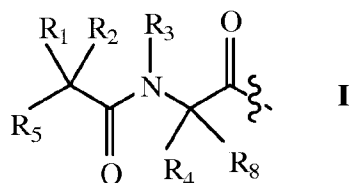
and R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6

member heterocyclic ring, then at least one of  $R_1$  and  $R_2$  are other than H, and in one embodiment both  $R_1$  and  $R_2$  are other than H. In one embodiment J and Z are both H,  $X_{12}$  is OH,  $X_{13}$  is H or OH, and X is  $NHR_{10}$  wherein  $R_{10}$  a dipeptide comprising the general structure of Formula I. In a further embodiment, the A chain comprises the

5 sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub>-R<sub>13</sub> (SEQ ID NO: 3), the B chain sequence comprises the sequence of X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14), m is 1,  $X_{12}$  is OH,  $X_{13}$  is H or OH, X is  $NHR_{10}$ , wherein  $R_{10}$  a dipeptide comprising the general structure of Formula I,  $R_{13}$  is COOH and the

10 carboxy terminal amino acid of the B chain has an amide (CONH<sub>2</sub>) in place of the natural alpha carbon carboxy group, with the remaining designations defined as immediately above.

In one embodiment the dipeptide present at X,  $X_{12}$ ,  $X_{13}$ , J and Z is a dipeptide comprising the general structure of Formula I:



15 wherein

$R_1$ ,  $R_2$ ,  $R_4$  and  $R_8$  are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group consisting of N, S and O, or  $R_1$  and  $R_2$  together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl; or  $R_4$  and  $R_8$  together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

20

25  $R_3$  is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or  $R_4$  and  $R_3$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

30  $R_5$  is  $NHR_6$  or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo provided that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are not H. In one embodiment when J or Z comprise the dipeptide of Formula I, and R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring, then both R<sub>1</sub> and R<sub>2</sub> are not hydrogen.

10 In one embodiment R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl);

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

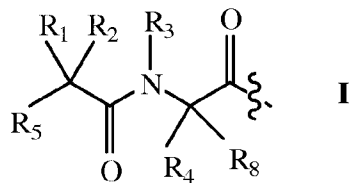
R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>7</sub> is selected from the group consisting of H and OH and R<sub>8</sub> is H. In accordance with another embodiment, m is 1, R<sub>8</sub> is H, R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl and R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub> aryl)R<sub>7</sub>, or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring. In one embodiment, m is 1, R<sub>8</sub> is H, R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl and R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring. In a further embodiment X<sub>9</sub> is phenylalanine and a single dipeptide extension is linked to the insulin peptide through an amide bond to the N-terminus of the A chain or the B chain. In an alternative



embodiment the insulin peptide comprises a 4 amino phenylalanine substitution at position 19 of the A chain and a single dipeptide extension is linked to the insulin peptide through an amide bond formed at the aromatic amine of the 4 amino phenylalanine. In one embodiment insulin analogs disclosed herein comprise a C-terminal amide or ester in place of a C-terminal carboxylate on the A chain and/or B chain.

In accordance with one embodiment the dipeptide of Formula I is further modified to comprise a large polymer that interferes with the insulin analog's ability to interact with the insulin or IGF-1 receptor. Subsequent cleavage of the dipeptide releases the insulin analog from the dipeptide complex wherein the released insulin analog is fully active. In accordance with one embodiment J comprises the dipeptide of Formula I, wherein the dipeptide of Formula I is further modified to comprises a large polymer that interferes with the bound insulin analog's ability to interact with the insulin or IGF-1 receptor. In accordance with one embodiment one of X, X<sub>12</sub>, X<sub>13</sub>, J and Z comprises a dipeptide of the general structure of Formula I:



, wherein the dipeptide of Formula I is pegylated, alkylated or acylated. In one embodiment either J, Z or X comprises an acylated or pegylated dipeptide of Formula I, and in one embodiment J comprises an acylated or pegylated dipeptide of Formula I.

In one embodiment, the dipeptide prodrug element is covalently bound to the insulin peptide via an amide linkage at the N-terminus of the A chain or the B chain, or attached to an amine bearing side chain of an internal amino acid, wherein the dipeptide further comprises a depot polymer linked to dipeptide. In one embodiment a native amino acid of the insulin peptide is substituted with an amino acid suitable for forming an amide bond with the dipeptide of Formula I. In one embodiment a depot bearing dipeptide is linked at a position selected from A14, A19, B16, B28 and B29. In one embodiment two or more depot polymers are linked to a single dipeptide element. The depot polymer is selected to be biocompatible and of sufficient size that the insulin peptide modified by covalent attachment of the dipeptide remains sequestered at an injection site and/or incapable of interacting with its corresponding

receptor upon administration to a patient. Subsequent cleavage of the dipeptide releases the insulin peptide to interact with its intended target.

In accordance with one embodiment the depot polymer is selected from biocompatible polymers known to those skilled in the art. The depot polymers typically have a size selected from a range of about 20,000 to 120,000 Daltons. In one embodiment the depot polymer has a size selected from a range of about 40,000 to 100,000 or about 40,000 to 80,000 Daltons. In one embodiment the depot polymer has a size of about 40,000, 50,000, 60,000, 70,000 or 80,000 Daltons. Suitable depot polymers include but are not limited to dextrans, polylactides, polyglycolides, caprolactone-based polymers, poly(caprolactone), polyanhydrides, polyamines, polyesteramides, polyorthoesters, polydioxanones, polyacetals, polyketals, polycarbonates, polyphosphoesters, polyesters, polybutylene terephthalate, polyorthocarbonates, polyphosphazenes, succinates, poly(malic acid), poly(amino acids), polyvinylpyrrolidone, polyethylene glycol, polyhydroxycellulose, polysaccharides, chitin, chitosan, hyaluronic acid, and copolymers, terpolymers and mixtures thereof, and biodegradable polymers and their copolymers including caprolactone-based polymers, polycaprolactones and copolymers which include polybutylene terephthalate. In one embodiment the depot polymer is selected from the group consisting of polyethylene glycol, dextran, polylactic acid, polyglycolic acid and a copolymer of lactic acid and glycolic acid, and in one specific embodiment the depot polymer is polyethylene glycol. In one embodiment the depot polymer is polyethylene glycol and the combined molecular weight of depot polymer(s) linked to the dipeptide element is about 40,000 to 80,000 Daltons.

In accordance with one embodiment the dipeptide of Formula I further comprises a polyethylene oxide, alkyl or acyl group. In one embodiment one or more polyethylene oxide chains are linked to the dipeptide of Formula I wherein the combined molecular weight of the polyethylene oxide chains ranges from about 20,000 to about 80,000 Daltons, or 40,000 to 80,000 Daltons or 40,000 to 60,000 Daltons. In one embodiment the polyethylene oxide is polyethylene glycol. In one embodiment at least one polyethylene glycol chain having a molecular weight of about 40,000 Daltons is linked to the dipeptide of Formula I. In another embodiment the dipeptide of Formula I is acylated with an acyl group of sufficient size to bind serum albumin, and thus inactivate the insulin analog upon administration. The acyl

group can be linear or branched, and in one embodiment is a C16 to C30 fatty acid. For example, the acyl group can be any of a C16 fatty acid, C18 fatty acid, C20 fatty acid, C22 fatty acid, C24 fatty acid, C26 fatty acid, C28 fatty acid, or a C30 fatty acid. In one embodiment, the acyl group is a C16 to C20 fatty acid, e.g., a C18 fatty acid or  
 5 a C20 fatty acid.

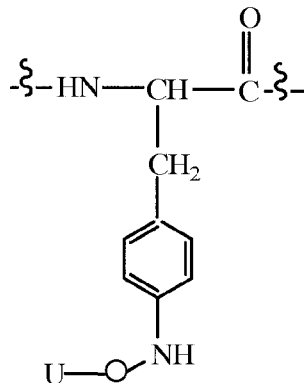
In accordance with one embodiment an insulin prodrug analog is provided comprising an A chain sequence and a B chain sequence wherein the A chain comprises the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub>-R<sub>13</sub> (SEQ ID NO: 3), and the B chain sequence comprises the sequence of X<sub>14</sub>-

10 X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14), wherein

X<sub>14</sub> is selected from the group consisting of an N-terminal amine, X<sub>9</sub>VNQ (SEQ ID NO: 21), VNQ, NQ and Q that joins the "J" element to the X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>-R<sub>14</sub> (SEQ ID NO: 14) sequence.

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

15 X<sub>2</sub> is an amino acid of the general structure



wherein U is an amino acid or a hydroxyl acid and O is an N-alkylated amino acid linked through an amide bond;

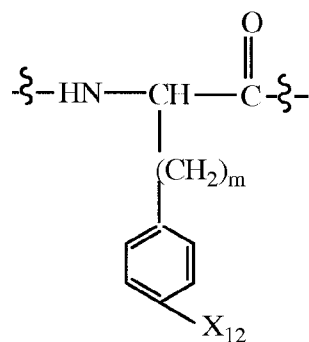
X<sub>3</sub> is selected from the group consisting of asparagine, ornithine, glycine,  
 20 alanine, threonine and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

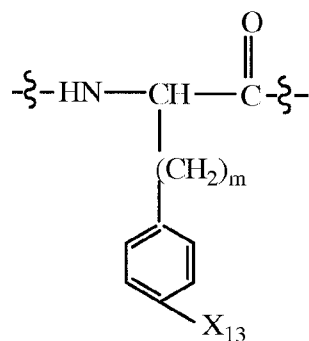
X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

25 X<sub>7</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and  $\text{X}_{12}$  is selected from the group consisting of OH,  $\text{NH}_2$  and  $\text{OCH}_3$ ;

$\text{X}_8$  is an amino acid of the general structure



5

wherein m is an integer selected from 0-3 and  $\text{X}_{13}$  is selected from the group consisting of H, OH,  $\text{NH}_2$  and  $\text{OCH}_3$ ;

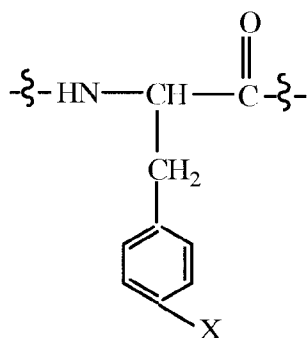
$\text{X}_9$  is selected from the group consisting of phenylalanine and desamino-phenylalanine; and  $\text{R}_{13}$  and  $\text{R}_{14}$  are independently  $\text{COOH}$  or  $\text{CONH}_2$ . In one

10 embodiment,  $\text{X}_7$  and  $\text{X}_8$  are both tyrosine,  $\text{R}_{13}$  is  $\text{COOH}$  and the carboxy terminal amino acid of the B chain has an amide ( $\text{CONH}_2$ ) in place of the natural alpha carbon carboxy group.

In accordance with one embodiment a compound comprising the sequence Z-GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3), or an analog thereof comprising

15 a sequence that differs from SEQ ID NO: 3 by 1 to 3 amino acid modifications, selected from positions A5, A8, A9, A10, A14, A15, A17, A18 (relative to the native insulin A chain sequence) is provided. In this embodiment  $\text{X}_1$  is selected from the group consisting of threonine and histidine;

$\text{X}_2$  is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>,

further wherein R<sub>10</sub> and Z are independently H or a dipeptide of the structure

5 U-O, wherein U is an amino acid or a hydroxyl acid and O is an N-alkylated amino acid, wherein O is linked to the peptide of SEQ ID NO: 3 through formation of an amide bond, with the proviso that R<sub>10</sub> and Z are not the same and that U, O or the amino acid of SEQ ID NO: 3 to which U-O is linked is a non-coded amino acid. In one embodiment the chemical cleavage of the dipeptide from SEQ ID NO: 3 is at

10 least about 90% complete within about 1 to about 720 hours in PBS under physiological conditions. In another embodiment the chemical cleavage half-life (t<sub>1/2</sub>) of U-O from SEQ ID NO: 3 is at least about 1 hour to about 1 week in PBS under physiological conditions. The compound in one embodiment further comprises an insulin B chain linked to the A chain of SEQ ID NO: 3 either through intermolecular

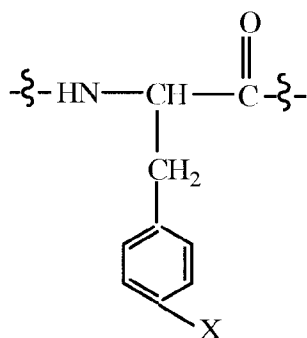
15 disulfide linkages or as a recombinant single chain polypeptide.

Selection of the substituents on the dipeptide element and the attachment site of the dipeptide prodrug element can impact the rate of chemical cleavage of the dipeptide prodrug element from the insulin peptide. In one embodiment an insulin prodrug is provided comprising an A chain sequence and a B chain sequence, wherein

20 the A chain comprises the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and the B chain sequence comprises the sequence of X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFF (SEQ ID NO: 4) wherein

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, and OCH<sub>3</sub>;

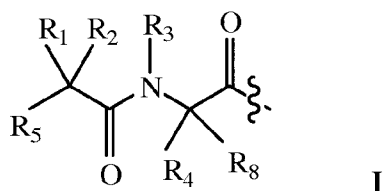
X<sub>3</sub> is selected from the group consisting of asparagine, glycine, alanine, threonine and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid; and

X<sub>14</sub> is selected from the group consisting of a bond, X<sub>9</sub>VNQ (SEQ ID NO: 21), VNQ, NQ and Q, and X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine, further wherein a dipeptide prodrug element comprises the structure:



is linked to the alpha amino group of the N-terminal amino acid of the peptide of SEQ ID NO: 3 or SEQ ID NO: 4 with the proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring, then both R<sub>1</sub> and R<sub>2</sub> are not hydrogen. In this embodiment, compounds having a t<sub>1/2</sub> of about 1 hour in PBS under physiological conditions are provided wherein

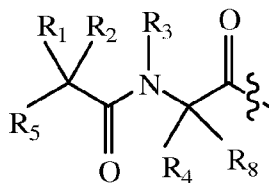
R<sub>1</sub> and R<sub>2</sub> are independently C<sub>1</sub>-C<sub>18</sub> alkyl or aryl; or R<sub>1</sub> and R<sub>2</sub> are linked through -(CH<sub>2</sub>)<sub>p</sub>-, wherein p is 2-9;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

R<sub>4</sub> and R<sub>8</sub> are each hydrogen; and

R<sub>5</sub> is an amine.

In other embodiments, prodrugs having the prodrug element linked at the N-terminus and having a  $t_{1/2}$  of, e.g., about 1 hour comprise a dipeptide prodrug element with the structure:



5 wherein

$R_1$  and  $R_2$  are independently  $C_1$ - $C_{18}$  alkyl or  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ ; or

$R_1$  and  $R_2$  are linked through  $-(CH_2)_p$ , wherein  $p$  is 2-9;

$R_3$  is  $C_1$ - $C_{18}$  alkyl;

$R_4$  and  $R_8$  are each hydrogen;

10  $R_5$  is  $NH_2$ ; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl) $CONH_2$ ,  $(C_0$ - $C_4$  alkyl) $COOH$ ,  $(C_0$ - $C_4$  alkyl) $NH_2$ ,  $(C_0$ - $C_4$  alkyl) $OH$ , and halo.

Alternatively, an insulin prodrug is provided comprising an A chain sequence  
 15 of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and a B chain sequence of  
 X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein the dipeptide  
 prodrug element is linked to the alpha amino group of the N-terminal amino acid of  
 the peptide of SEQ ID NO: 3 or SEQ ID NO: 14 and exhibits a  $t_{1/2}$  between about 6 to  
 about 24 hours in PBS under physiological conditions. In one embodiment such  
 20 compounds comprise a prodrug element of Formula I, wherein

$R_1$  and  $R_2$  are independently selected from the group consisting of hydrogen,  
 $C_1$ - $C_{18}$  alkyl and aryl, or  $R_1$  and  $R_2$  are linked through  $-(CH_2)_p$ , wherein  $p$  is 2-9;

$R_3$  is  $C_1$ - $C_{18}$  alkyl or  $R_3$  and  $R_4$  together with the atoms to which they are  
 attached form a 4-12 heterocyclic ring;

25  $R_4$  and  $R_8$  are independently selected from the group consisting of hydrogen,  
 $C_1$ - $C_8$  alkyl and aryl; and

$R_5$  is an amine, with the proviso that both  $R_1$  and  $R_2$  are not hydrogen and  
 provided that one of  $R_4$  or  $R_8$  is hydrogen.

In a further embodiment an insulin prodrug is provided comprising an A chain  
 30 sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and a B chain

sequence of X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein the dipeptide prodrug element is linked to the alpha amino group of the N-terminal amino acid of the peptide of SEQ ID NO: 3 or SEQ ID NO: 14 and exhibits a t<sub>1/2</sub> of about 72 to about 168 hours in PBS under physiological conditions. In one

5 embodiment such compounds comprise a prodrug element of Formula I, wherein

R<sub>1</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and aryl;

R<sub>2</sub> is H;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

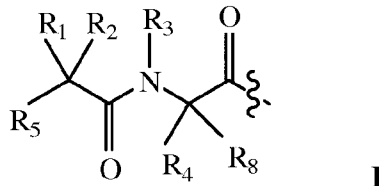
R<sub>4</sub> and R<sub>8</sub> are each hydrogen; and

10 R<sub>5</sub> is an amine or N-substituted amine or a hydroxyl;

with the proviso that, if R<sub>1</sub> is alkyl or aryl, then R<sub>1</sub> and R<sub>5</sub> together with the atoms to which they are attached form a 4-11 heterocyclic ring.

In one embodiment, prodrugs having the dipeptide prodrug element linked to the N-terminal alpha amino acid of the insulin A or B chain peptide and having a t<sub>1/2</sub>,

15 e.g., between about 12 to about 72 hours, or in one embodiment between about 12 to about 48 hours, comprise a dipeptide prodrug element with the structure:



wherein R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>,  
20 or R<sub>1</sub> and R<sub>2</sub> are linked through (CH<sub>2</sub>)<sub>p</sub>, wherein p is 2-9;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-12 heterocyclic ring;

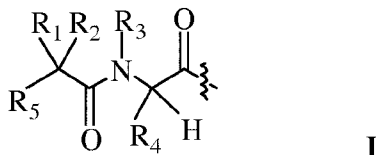
R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

25 R<sub>5</sub> is NH<sub>2</sub>; and

R<sub>7</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo; with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen and provided that at least one of R<sub>4</sub> or R<sub>8</sub> is hydrogen.



In one embodiment, prodrugs having the dipeptide prodrug element linked to the N-terminal amino acid of the insulin A or B chain peptide and having a  $t_{1/2}$ , e.g., between about 12 to about 72 hours, or in one embodiment between about 12 to about 48 hours, comprise a dipeptide prodrug element with the structure:



5

wherein  $R_1$  and  $R_2$  are independently selected from the group consisting of hydrogen,  $C_1$ - $C_8$  alkyl and  $(C_1$ - $C_4$  alkyl) $NH_2$ , or  $R_1$  and  $R_2$  are linked through  $(CH_2)_p$ , wherein  $p$  is 2-9;

$R_3$  is  $C_1$ - $C_8$  alkyl or  $R_3$  and  $R_4$  together with the atoms to which they are attached form a 4-6 heterocyclic ring;

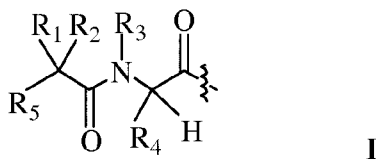
10

$R_4$  is selected from the group consisting of hydrogen and  $C_1$ - $C_8$  alkyl; and

$R_5$  is  $NH_2$ , with the proviso that both  $R_1$  and  $R_2$  are not hydrogen.

In other embodiments, prodrugs having the dipeptide prodrug element linked to the N-terminal amino acid of the insulin A or B chain peptide and having a  $t_{1/2}$ , e.g., between about 12 to about 72 hours, or in one embodiment between about 12 to about 48 hours, comprise a dipeptide prodrug element with the structure:

15



wherein

$R_1$  and  $R_2$  are independently selected from the group consisting of hydrogen,  $C_1$ - $C_8$  alkyl and  $(C_1$ - $C_4$  alkyl) $NH_2$ ;

20

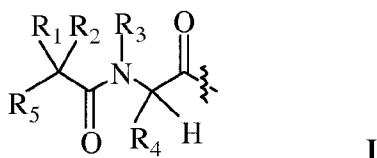
$R_3$  is  $C_1$ - $C_6$  alkyl;

$R_4$  is hydrogen; and

$R_5$  is  $NH_2$ , with the proviso that both  $R_1$  and  $R_2$  are not hydrogen.

In one embodiment, prodrugs having the dipeptide prodrug element linked to the N-terminal amino acid of the insulin A or B chain peptide and having a  $t_{1/2}$ , e.g., between about 12 to about 72 hours, or in one embodiment between about 12 to about 48 hours, comprise a dipeptide prodrug element with the structure:

25



wherein

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen and C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, or R<sub>1</sub> and R<sub>2</sub> are linked through (CH<sub>2</sub>)<sub>p</sub>, wherein p

5 is 2-9;

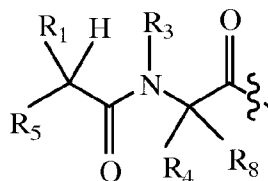
R<sub>3</sub> is C<sub>1</sub>-C<sub>8</sub> alkyl;

R<sub>4</sub> is (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NH<sub>2</sub>; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)OH. with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen.

In addition a prodrug having the dipeptide prodrug element linked to the N-terminal alpha amino acid of the insulin A or B chain peptide and having a t<sub>1/2</sub>, e.g., of about 72 to about 168 hours is provided wherein the dipeptide prodrug element has the structure:



15

wherein R<sub>1</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;

R<sub>4</sub> and R<sub>8</sub> are each hydrogen;

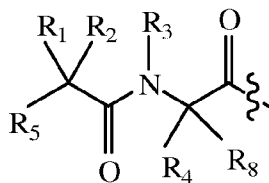
20 R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl, or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub>

25 alkyl)OH, and halo. with the proviso that, if R<sub>1</sub> is alkyl or (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, then R<sub>1</sub> and R<sub>5</sub> together with the atoms to which they are attached form a 4-11 heterocyclic ring.

In one embodiment the dipeptide prodrug element is linked to a side chain amine of an internal amino acid of the insulin peptide. In this embodiment prodrugs having a  $t_{1/2}$ , e.g., of about 1 hour have the structure:



5 wherein

$R_1$  and  $R_2$  are independently  $C_1$ - $C_8$  alkyl or  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ ; or  $R_1$  and  $R_2$  are linked through  $-(CH_2)_p-$ , wherein  $p$  is 2-9;

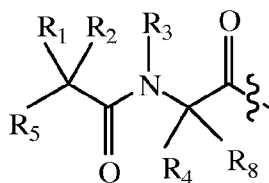
$R_3$  is  $C_1$ - $C_{18}$  alkyl;

$R_4$  and  $R_8$  are each hydrogen;

10  $R_5$  is  $NH_2$ ; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl) $CONH_2$ ,  $(C_0$ - $C_4$  alkyl) $COOH$ ,  $(C_0$ - $C_4$  alkyl) $NH_2$ ,  $(C_0$ - $C_4$  alkyl) $OH$ , and halo.

Furthermore, prodrugs having a  $t_{1/2}$ , e.g., between about 6 to about 24 hours and having the dipeptide prodrug element linked to an internal amino acid side chain comprise a dipeptide prodrug element with the structure:



wherein  $R_1$  and  $R_2$  are independently selected from the group consisting of hydrogen,  $C_1$ - $C_8$  alkyl, and  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ , or  $R_1$  and  $R_2$  are linked through  $-(CH_2)_p-$ , wherein  $p$  is 2-9;

$R_3$  is  $C_1$ - $C_{18}$  alkyl or  $R_3$  and  $R_4$  together with the atoms to which they are attached form a 4-12 heterocyclic ring;

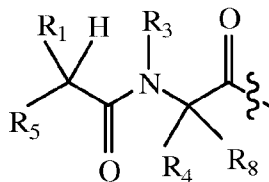
$R_4$  and  $R_8$  are independently hydrogen,  $C_1$ - $C_{18}$  alkyl or  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ ;

25  $R_5$  is  $NHR_6$ ;

$R_6$  is H or  $C_1$ - $C_8$  alkyl, or  $R_6$  and  $R_2$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

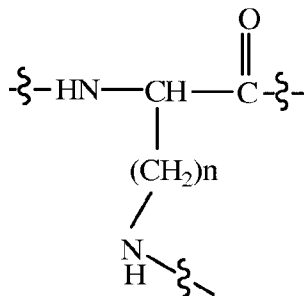
R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo, with the proviso that both R<sub>1</sub> and R<sub>2</sub> are not hydrogen and provided that at least one of R<sub>4</sub> or R<sub>8</sub> is hydrogen.

- 5 In addition a prodrug having a t<sub>1/2</sub>, e.g., of about 72 to about 168 hours and having the dipeptide prodrug element linked to an internal amino acid side chain of the insulin peptide is provided wherein the dipeptide prodrug element has the structure:



- 10 wherein R<sub>1</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;  
 R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl;  
 R<sub>4</sub> and R<sub>8</sub> are each hydrogen;  
 R<sub>5</sub> is NHR<sub>6</sub> or OH;  
 15 R<sub>6</sub> is H or C<sub>1</sub>-C<sub>8</sub> alkyl, or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and  
 R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo; with the proviso that, if R<sub>1</sub> and R<sub>2</sub> are both independently an alkyl or (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, either R<sub>1</sub> or R<sub>2</sub> is linked through (CH<sub>2</sub>)<sub>p</sub> to R<sub>5</sub>, wherein p is 2-9.  
 20

In one embodiment the dipeptide prodrug element is linked to a side chain amine of an internal amino acid of the insulin peptide wherein the internal amino acid comprises the structure of Formula IV:



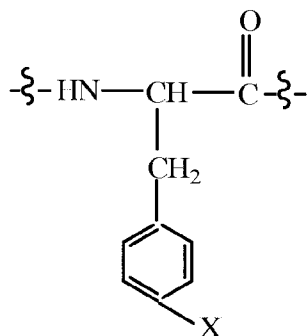
- 25 wherein

n is an integer selected from 1 to 4. In one embodiment n is 3 or 4 and in one embodiment the internal amino acid is lysine. In one embodiment the dipeptide prodrug element is linked to a primary amine on a side chain of an amino acid located at position 28, or 29 of the B-chain of the insulin peptide.

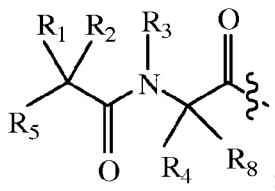
- 5 In one embodiment an insulin prodrug is provided comprising an A chain sequence and a B chain sequence, wherein the A chain comprises the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and the B chain sequence comprises the sequence of X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein

- 10 X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>, wherein R<sub>10</sub> is H or a dipeptide comprising the general structure:



- 15

wherein

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl, and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, or R<sub>1</sub> and R<sub>2</sub> are linked through -(CH<sub>2</sub>)<sub>p</sub>-, wherein p is 2-9;

- 20 R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-12 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are independently hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl or (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NHR<sub>6</sub>;

$R_6$  is H or  $C_1$ - $C_8$  alkyl, or  $R_6$  and  $R_2$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

$R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl) $CONH_2$ ,  $(C_0$ - $C_4$  alkyl) $COOH$ ,  $(C_0$ - $C_4$  alkyl) $NH_2$ ,  $(C_0$ - $C_4$  alkyl) $OH$ , and halo;

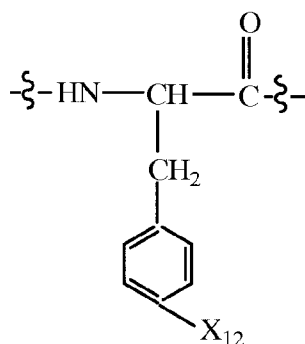
$X_3$  is selected from the group consisting of asparagine, glycine, alanine, threonine and serine;

$X_4$  is selected from the group consisting of histidine and threonine;

$X_5$  is selected from the group consisting of alanine, glycine and serine;

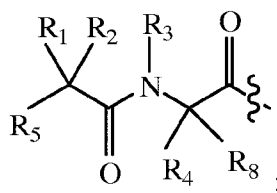
$X_6$  is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

$X_7$  is an amino acid of the general structure

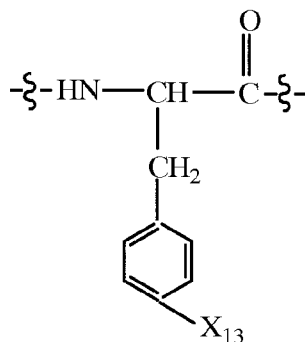


wherein  $X_{12}$  is selected from the group consisting of OH,  $NHR_{11}$

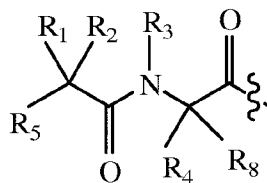
and  $OCH_3$ , wherein  $R_{11}$  is H or a dipeptide comprising the general structure:



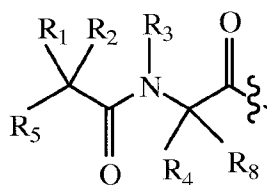
$X_8$  is an amino acid of the general structure



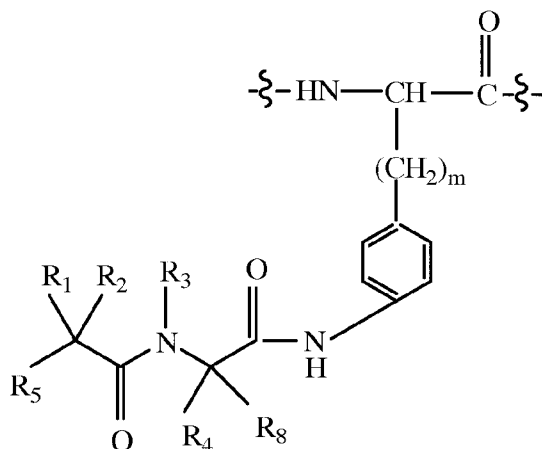
wherein  $X_{13}$  is selected from the group consisting of H, OH,  $NH_2$ ,  $NHR_{12}$  and  $OCH_3$ , wherein  $R_{12}$  is H or a dipeptide comprising the general structure:



$X_{14}$  is selected from the group consisting of a Hydrogen (forming an N-terminal amine),  $X_9VNQ$  (SEQ ID NO: 21),  $VNQ$ ,  $NQ$  and  $Q$ , and  $X_9$  is selected from the group consisting of phenylalanine and desamino-phenylalanine, with the proviso that one and only one of  $R_{10}$ ,  $R_{11}$  and  $R_{12}$  is a dipeptide comprising the general structure:



10 In accordance with one embodiment wherein the dipeptide prodrug element is linked to an amino substituent of an aryl group of an aromatic amino acid, prodrug formulations can be provided having the desired time of activation. For example, an insulin prodrug comprising the structure of Formula III:



15 wherein  $m$  is an integer from 0 to 3 and having a  $t_{1/2}$  of about 1 hour in PBS under physiological conditions is provided. In one embodiment where an insulin prodrug comprises the structure of formula III and exhibits such a half life,

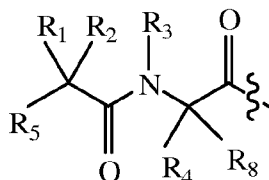
$R_1$  and  $R_2$  are independently  $C_1$ - $C_{18}$  alkyl or aryl;

$R_3$  is  $C_1$ - $C_{18}$  alkyl or  $R_3$  and  $R_4$  together with the atoms to which they are  
20 attached form a 4-12 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and aryl; and

R<sub>5</sub> is an amine or a hydroxyl. In one embodiment m is 1.

In one embodiment, the dipeptide prodrug element is linked to the insulin peptide via an amine present on an aryl group of an aromatic amino acid of the insulin peptide, wherein prodrugs having a t<sub>1/2</sub>, e.g., of about 1 hour have a dipeptide structure of:



wherein R<sub>1</sub> and R<sub>2</sub> are independently C<sub>1</sub>-C<sub>18</sub> alkyl or (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-12 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NH<sub>2</sub> or OH; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo.

In another embodiment an insulin prodrug comprising the structure of Formula III, wherein m is an integer from 0 to 3 and has a t<sub>1/2</sub> of about 6 to about 24 hours in PBS under physiological conditions, is provided. In one embodiment where an insulin prodrug comprises the structure of Formula III and exhibits such a half life,

R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and aryl, or R<sub>1</sub> and R<sub>2</sub> are linked through -(CH<sub>2</sub>)<sub>p</sub>-, wherein p is 2-9;

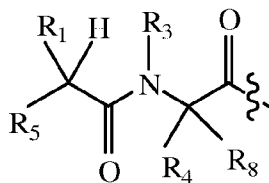
R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-6 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and aryl; and

R<sub>5</sub> is an amine or N-substituted amine. In one embodiment m is 1.



In one embodiment, prodrugs having the dipeptide prodrug element linked via an amine present on an aryl group of an aromatic amino acid and having a  $t_{1/2}$ , e.g., of about 6 to about 24 hours are provided wherein the dipeptide comprises a structure of:



5 wherein

R<sub>1</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-6 heterocyclic ring;

10 R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>;

R<sub>5</sub> is NHR<sub>6</sub>;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl, or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

15 R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo.

In another embodiment an insulin prodrug comprising the structure of Formula III, wherein m is an integer from 0 to 3 and has a  $t_{1/2}$  of about 72 to about 168 hours  
20 in PBS under physiological conditions, is provided. In one embodiment where an insulin prodrug comprises the structure of formula III and exhibits such a half life,

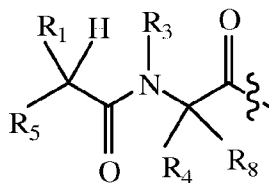
R<sub>1</sub> and R<sub>2</sub> are independently selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>8</sub> alkyl and aryl;

25 R<sub>3</sub> is C<sub>1</sub>-C<sub>18</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 4-6 heterocyclic ring;

R<sub>4</sub> and R<sub>8</sub> are each hydrogen; and

R<sub>5</sub> is selected from the group consisting of amine, N-substituted amine and hydroxyl. In one embodiment m is 1.

In one embodiment, prodrugs having the dipeptide prodrug element linked via an aromatic amino acid and having a  $t_{1/2}$ , e.g., of about 72 to about 168 hours are provided wherein the dipeptide comprises a structure of:



- 5 wherein  $R_1$  is selected from the group consisting of hydrogen,  $C_1$ - $C_8$  alkyl,  $(C_1$ - $C_4$  alkyl)COOH, and  $(C_0$ - $C_4$  alkyl)( $C_6$ - $C_{10}$  aryl) $R_7$ , or  $R_1$  and  $R_5$  together with the atoms to which they are attached form a 4-11 heterocyclic ring;

$R_3$  is  $C_1$ - $C_{18}$  alkyl or  $R_3$  and  $R_4$  together with the atoms to which they are attached form a 4-6 heterocyclic ring;

- 10  $R_4$  is hydrogen or forms a 4-6 heterocyclic ring with  $R_3$ ;

$R_8$  is hydrogen;

$R_5$  is  $NHR_6$  or OH;

$R_6$  is H or  $C_1$ - $C_8$  alkyl, or  $R_6$  and  $R_1$  together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

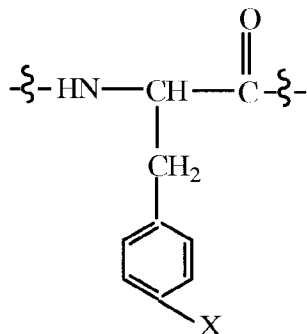
- 15  $R_7$  is selected from the group consisting of hydrogen,  $C_1$ - $C_{18}$  alkyl,  $C_2$ - $C_{18}$  alkenyl,  $(C_0$ - $C_4$  alkyl)CONH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)COOH,  $(C_0$ - $C_4$  alkyl)NH<sub>2</sub>,  $(C_0$ - $C_4$  alkyl)OH, and halo.

In one embodiment the insulin prodrug analog comprises an A chain sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub>- $R_{13}$  (SEQ ID NO: 3) and a B chain sequence

- 20 of X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein

X<sub>1</sub> is selected from the group consisting of threonine, histidine, arginine and lysine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, and OCH<sub>3</sub>;

X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine.

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

5 X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

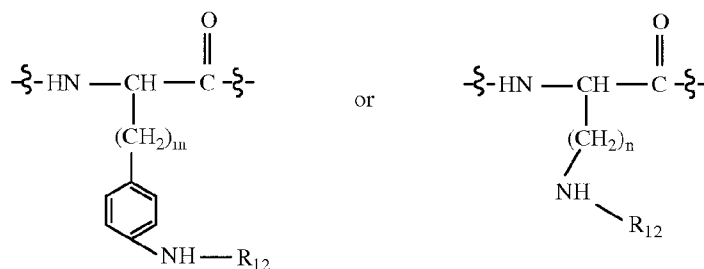
X<sub>7</sub> is tyrosine;

X<sub>8</sub> is tyrosine or phenylalanine;

10 X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine;

X<sub>10</sub> is aspartate-lysine dipeptide, a lysine-proline dipeptide, or a proline-lysine dipeptide;

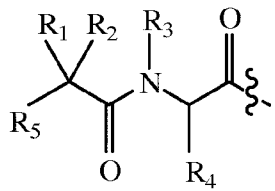
15 X<sub>11</sub> is threonine, alanine, or a threonine-arginine-arginine tripeptide; further wherein the B chain comprises a carboxy terminus extension of 1 to 4 amino acids wherein said carboxy terminal extension comprises an amino acid having the structure of



wherein m is an integer from 0-3;

20 n is an integer from 1-4;

R<sub>12</sub> is a dipeptide comprising the general structure:



R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

25 R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>,

(C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl);

- 5 R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

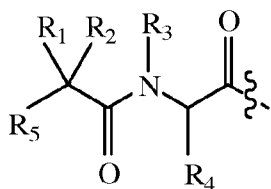
R<sub>5</sub> is NHR<sub>6</sub> or OH;

- 10 R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo; and

R<sub>13</sub> is COOH or CONH<sub>2</sub>,

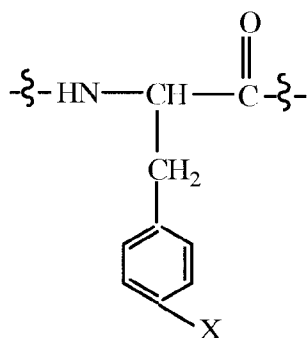
- 15 In one embodiment the insulin prodrug analog comprises an A chain sequence that includes the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub>-R<sub>13</sub> (SEQ ID NO: 3) and a B chain sequence that includes the sequence of J-X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) or J-X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTX<sub>10</sub>X<sub>11</sub>-R<sub>14</sub> (SEQ ID NO: 5)  
20 wherein J is H (forming an N-terminal amine) or a dipeptide comprising the general structure:



- wherein X<sub>14</sub> is either a bond joining the "J" element to SEQ ID NO: 14 or X<sub>14</sub> represents a 1 to 4 amino acid sequence selected from the group consisting of a  
25 FVNQ (SEQ ID NO: 11), VNQ, NQ and Q that joins the "J" element to SEQ ID NO: 14;

X<sub>1</sub> is selected from the group consisting of threonine, histidine, arginine and lysine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, and OCH<sub>3</sub>;

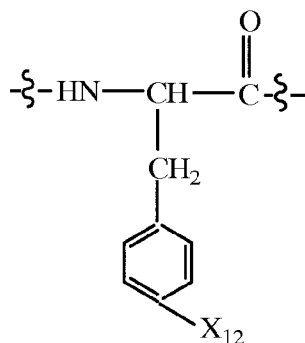
X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine.

5 X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

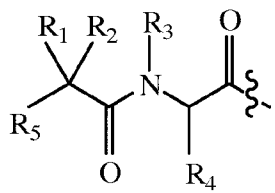
X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

X<sub>7</sub> is an amino acid of the general structure

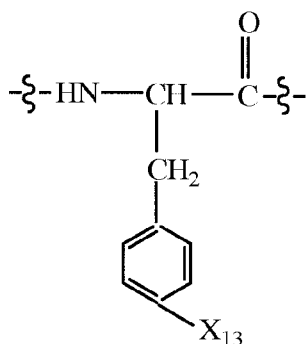


10

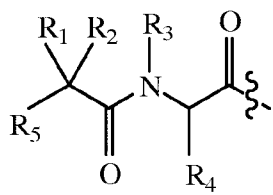
wherein X<sub>12</sub> is selected from the group consisting of OH, OCH<sub>3</sub>, NH<sub>2</sub> and NHR<sub>11</sub>, wherein R<sub>11</sub> is a dipeptide comprising the general structure:



X<sub>8</sub> is an amino acid of the general structure



wherein  $\text{X}_{13}$  is selected from the group consisting of H, OH,  $\text{OCH}_3$ ,  $\text{NH}_2$ , and  $\text{NHR}_{12}$ , wherein  $\text{R}_{12}$  is a dipeptide comprising the general structure:



5  $\text{X}_9$  is selected from the group consisting of phenylalanine and desamino-phenylalanine;

$\text{X}_{10}$  is aspartate-lysine dipeptide, a lysine-proline dipeptide, or a proline-lysine dipeptide;

$\text{X}_{11}$  is threonine, alanine, or a threonine-arginine-arginine tripeptide;

10 wherein  $\text{R}_1$  is selected from the group consisting of H and  $\text{C}_1$ - $\text{C}_8$  alkyl; and

$\text{R}_2$  and  $\text{R}_4$  are independently selected from the group consisting of H,  $\text{C}_1$ - $\text{C}_8$  alkyl,  $\text{C}_2$ - $\text{C}_8$  alkenyl,  $(\text{C}_1$ - $\text{C}_4$  alkyl)OH,  $(\text{C}_1$ - $\text{C}_4$  alkyl)SH,  $(\text{C}_2$ - $\text{C}_3$  alkyl) $\text{SCH}_3$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{CONH}_2$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{COOH}$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NH}_2$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NHC}(\text{NH}_2^+) \text{NH}_2$ ,  $(\text{C}_0$ - $\text{C}_4$  alkyl)( $\text{C}_3$ - $\text{C}_6$  cycloalkyl),  $(\text{C}_0$ - $\text{C}_4$  alkyl)( $\text{C}_6$ - $\text{C}_{10}$  aryl) $\text{R}_7$ , and  $\text{CH}_2(\text{C}_5$ - $\text{C}_9$  heteroaryl);

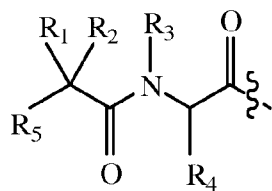
$\text{R}_3$  is selected from the group consisting of  $\text{C}_1$ - $\text{C}_8$  alkyl,  $(\text{C}_1$ - $\text{C}_4$  alkyl)OH,  $(\text{C}_1$ - $\text{C}_4$  alkyl)SH,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NH}_2$ ,  $(\text{C}_3$ - $\text{C}_6$ )cycloalkyl or  $\text{R}_4$  and  $\text{R}_3$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

$\text{R}_5$  is  $\text{NHR}_6$  or OH;

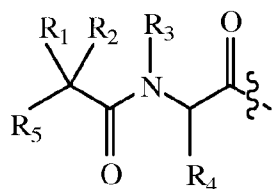
$\text{R}_6$  is H, or  $\text{R}_6$  and  $\text{R}_2$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

$\text{R}_7$  is selected from the group consisting of H and OH; and

$R_{13}$  and  $R_{14}$  are independently COOH or CONH<sub>2</sub>, with the proviso that one and only one of  $X_{12}$ ,  $X_{13}$ , or J is a dipeptide comprising the general structure:



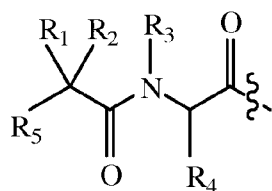
- (i.e., only one dipeptide prodrug element is attached to the insulin peptide). In one embodiment J is a dipeptide comprising the general structure:



- , and X and  $X_{12}$  are each OH and  $X_{13}$  is H, with the further proviso that when  $R_4$  and  $R_3$  together with the atoms to which they are attached form a 5 member heterocyclic ring, both  $R_1$  and  $R_2$  are other than H. In an alternative embodiment  $X_{12}$  is NHR<sub>11</sub>, J and  $X_{13}$  are each H and X is OH. In another alternative embodiment  $X_{13}$  is NHR<sub>12</sub>, X and  $X_{12}$  are each OH and J is H. In one embodiment the B chain comprises the sequence J-  
 $X_9$ VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTPKT (SEQ ID NO: 15) or J-  
 $X_9$ VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTKPT (SEQ ID NO: 16), wherein J,  
 $X_4$ ,  $X_5$ ,  $X_6$ ,  $X_7$ ,  $X_8$  and  $X_9$  are defined as immediately above. In a further embodiment  $R_3$  is C<sub>1</sub>-C<sub>6</sub> alkyl and  $R_4$  is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl, or  $R_3$  and  $R_4$  together with the atoms to which they are attached form a 5 member heterocyclic ring. In another further embodiment  $X_4$  is histidine,  $X_5$  is serine and  $X_6$  is histidine.

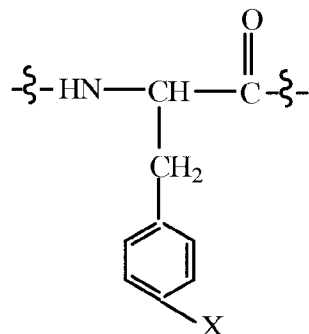
- In another embodiment an insulin prodrug analog is provided comprising an A chain sequence of Z-GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and a B chain sequence comprising a sequence of  $X_4$ LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFF (SEQ ID NO: 4) wherein

Z is H or a dipeptide comprising the general structure:

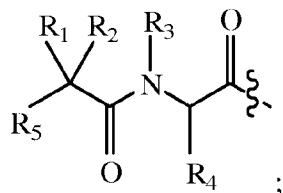


X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



- 5                    wherein X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>, wherein R<sub>10</sub> is a dipeptide comprising the general structure:



X<sub>3</sub> is selected from the group consisting of asparagine, glycine, alanine, threonine, or serine;

- 10                    X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

wherein R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl;

- 15                    R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are  
20                    attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

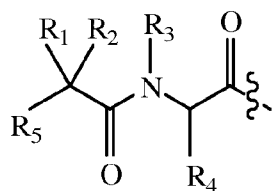


R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

5 R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH, with the proviso that X and Z are not both dipeptides and Z is not H when X is OH. In one embodiment when Z is a dipeptide comprising the general structure:

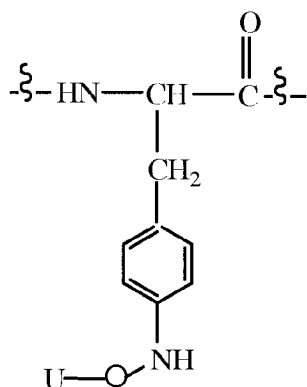


10 and R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring, then at least one of R<sub>1</sub> and R<sub>2</sub> are other than H. In one embodiment the A chain comprises a sequence of Z-GIVEQCCX<sub>1</sub>SICSLYQLENYCX<sub>3</sub> (SEQ ID NO: 17) and the B chain sequence comprises the sequence X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTPKT (SEQ ID NO: 12) or X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTKPT (SEQ ID NO: 13).

In an alternative embodiment an insulin prodrug analog is provided comprising an A chain sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and a B chain sequence comprising the sequence

20 X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTPKT (SEQ ID NO: 12) or X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTKPT (SEQ ID NO: 13), wherein X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein U is an amino acid or a hydroxyl acid and O is an N-alkylated amino acid;

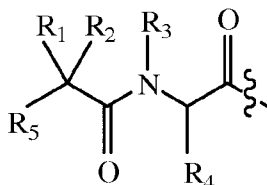
X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine;

5 X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

10 X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine. In one embodiment U-O represent a dipeptide of the general structure;



wherein R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl;

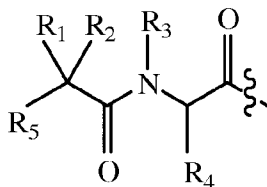
R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

20 R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH. In a further embodiment X<sub>7</sub> is tyrosine, X<sub>8</sub> is phenylalanine and X<sub>9</sub> is phenylalanine, and in an additional further embodiment X<sub>4</sub> is histidine, X<sub>5</sub> is serine and X<sub>6</sub> is histidine. In a further embodiment U-O represent a dipeptide of the general structure;



wherein

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

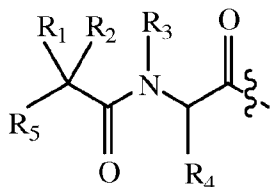
R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo.

In another embodiment an insulin prodrug analog is provided comprising an A chain sequence of Z-GIVEQCCTSICSLYQLENX<sub>2</sub>CX<sub>3</sub>-R<sub>13</sub> (SEQ ID NO: 18) and a B

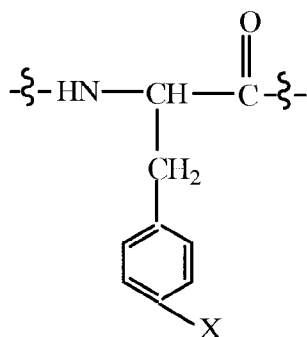
chain sequence comprising a sequence of X<sub>4</sub>LCGSHLVEALYLVCGERGFF-R<sub>14</sub>  
(SEQ ID NO: 19) wherein

Z is H or an amide linked dipeptide comprising the general structure:



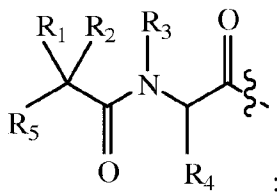
5

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH and NHR<sub>10</sub>;

wherein R<sub>10</sub> is a dipeptide comprising the general structure:



10

X<sub>3</sub> is serine, asparagine or glycine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl;

15 R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

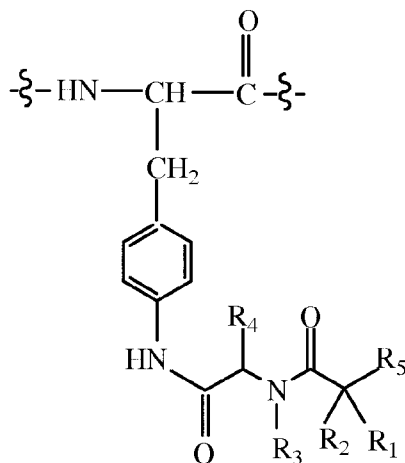
R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

5 R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>7</sub> is selected from the group consisting of H and OH; and

R<sub>13</sub> and R<sub>14</sub> are independently COOH or CONH<sub>2</sub>; with the proviso that when Z is H, X is not OH and when X is OH, Z is not H. In one embodiment R<sub>13</sub> is COOH and R<sub>14</sub> is CONH<sub>2</sub>. In a further embodiment X<sub>2</sub> is an amino acid of the general structure of Formula III:



Z is H, X<sub>3</sub> is serine, X<sub>4</sub> is histidine, R<sub>13</sub> is COOH and R<sub>14</sub> is CONH<sub>2</sub>. In an additional further embodiment

15 R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl;

R<sub>2</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>2</sub> and R<sub>6</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring;

20                    R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl;

R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring. In a further embodiment R<sub>3</sub> is CH<sub>3</sub>, R<sub>4</sub> is H and R<sub>5</sub> is NH<sub>2</sub>, or alternatively, R<sub>5</sub> is NH<sub>2</sub> and R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member

heterocyclic ring. In accordance with one embodiment the B chain of the insulin prodrug analog comprises the sequence

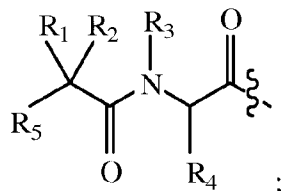
FVNQHLCGSHLVEALYLVCGERGFFYTPKT-R<sub>14</sub> (SEQ ID NO: 8),

FVNQHLCGSHLVEALYLVCGERGFFYTKPT-R<sub>14</sub> (SEQ ID NO: 9) or

- 5 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRR-R<sub>14</sub> (SEQ ID NO: 10), wherein R<sub>14</sub> is COOH or CONH<sub>2</sub>, and in one embodiment R<sub>14</sub> is CONH<sub>2</sub>.

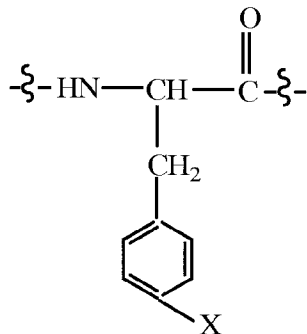
In one embodiment an insulin prodrug analog is provided comprising a polypeptide of the sequence Z-GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3), wherein

- 10 Z is a dipeptide comprising the general structure:



X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



- 15 wherein X is selected from the group consisting of OH, NH<sub>2</sub>, and OCH<sub>3</sub>;

X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine;

R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

- 20 R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

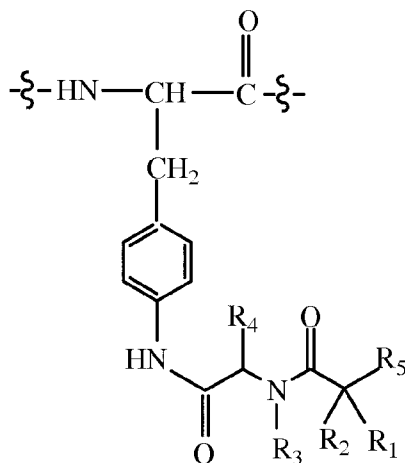
5 R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH. In one embodiment X<sub>1</sub> is threonine and X<sub>3</sub> is asparagine or glycine and in a further embodiment R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl and R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub> aryl)R<sub>7</sub>, or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring, with the proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are both other than H.

15 In one embodiment an insulin prodrug analog is provided comprising a polypeptide of the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3), wherein

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

$X_2$  is an amino acid of the general structure



X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine;

R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub>

alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl), or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

- 5 R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

- 10 R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH. In one embodiment the A chain of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) as defined immediately above is linked, either by disulfide bonds or as a single chain polypeptide, to a B chain comprising the sequence X<sub>14</sub>-

- 15 X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

- 20 X<sub>7</sub> is tyrosine; and

X<sub>8</sub> is phenylalanine.

In a further embodiment the dipeptide prodrug element has the structure of Formula I wherein

- 25 R<sub>1</sub>, R<sub>2</sub>, and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group  
30 consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub>



alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

5 R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>0</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)OH, and halo. In a further embodiment R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl and R<sub>4</sub> is selected  
10 from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl, or alternatively, R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring. In one embodiment X<sub>1</sub> is threonine and X<sub>3</sub> is asparagine or glycine.

In one embodiment R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>6</sub> alkyl, R<sub>2</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub>  
15 alkyl)C(O)NH<sub>2</sub>, CH<sub>2</sub>OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub> cycloalkyl) and CH<sub>2</sub>(C<sub>6</sub> aryl)R<sub>7</sub> or R<sub>2</sub> and R<sub>6</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring;

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl;

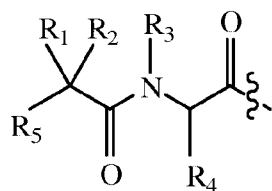
R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl, (C<sub>3</sub>-  
20 C<sub>6</sub>)cycloalkyl, and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

25 R<sub>7</sub> is selected from the group consisting of H and OH.

In accordance with one embodiment a single-chain insulin prodrug analog is provided wherein the carboxy terminus of the human insulin B chain, or a functional analog thereof, is covalently linked to the N-terminus of the human insulin A chain, or functional analog thereof, and further wherein a dipeptide prodrug moiety having the  
30 general structure:



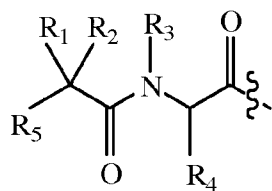
is covalently bound at the N-terminus of the peptide, or at the side chain of an amino acid via an amide bond, including for example at positions corresponding to A19, B16 or B25 of the respective native insulin A chain or B chain,. In accordance with one embodiment the single-chain

5 insulin analog comprises a compound of the formula: B-P-A, wherein: B represents the B-chain of human insulin or one of the functional or prodrug analogs of a B chain as disclosed herein, A represents the A chain of human insulin or one of the functional or prodrug analogs of an A chain as disclosed herein, and P represents a linker, including a peptide linker, that covalently joins the A chain to the B chain. In one

10 embodiment the linker is a peptide linker of about 5 to about 18, or about 10 to about 14, or about 4 to about 8, or about 6 amino acids. In one embodiment the B chain is linked to the A chain via peptide linker of 4-12 or 4-8 amino acids.

In one embodiment the single-chain insulin prodrug analog of the formula B-P-A comprises an A chain having the sequence GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and a B chain sequence that includes the sequence of J-X<sub>14</sub>-X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14) wherein J is H or a dipeptide comprising the general structure:

15

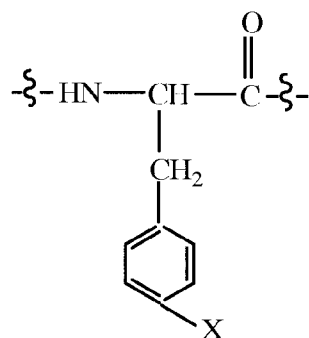


wherein X<sub>14</sub> is either a bond or a 1 to 4 amino acid sequence selected from the

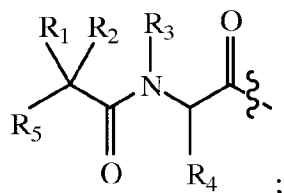
20 group consisting of a FVNQ (SEQ ID NO: 11), VNQ, NQ and Q .

X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NHR<sub>10</sub> and OCH<sub>3</sub>; wherein R<sub>10</sub> is H or a dipeptide comprising the general structure:



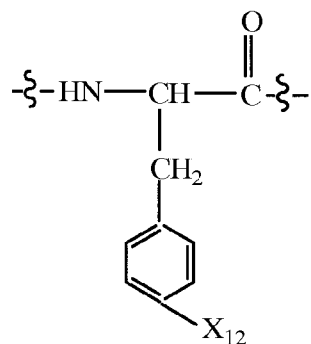
5 X<sub>3</sub> is asparagine, glycine, alanine, threonine, or serine.

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

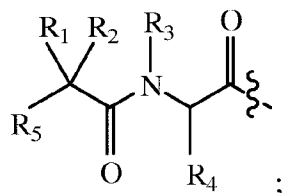
X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

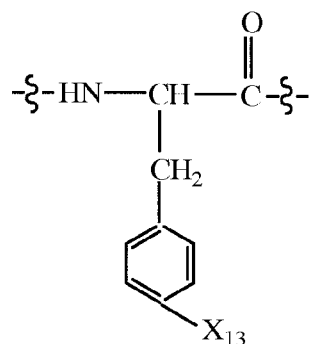
10 X<sub>7</sub> is an amino acid of the general structure



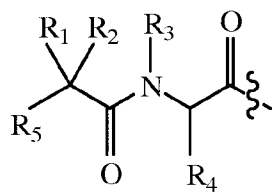
wherein X<sub>12</sub> is selected from the group consisting of OH, OCH<sub>3</sub> and NHR<sub>11</sub>, wherein R<sub>11</sub> is H or a dipeptide comprising the general structure:



15 X<sub>8</sub> is an amino acid of the general structure



wherein  $\text{X}_{13}$  is selected from the group consisting of H, OH,  $\text{OCH}_3$  and  $\text{NHR}_{12}$ , wherein  $\text{R}_{12}$  is H or a dipeptide comprising the general structure:



5                      wherein  $\text{R}_1$  is selected from the group consisting of H and  $\text{C}_1$ - $\text{C}_8$  alkyl;  
and

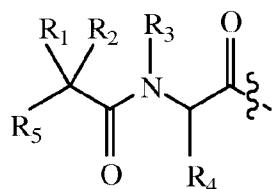
$\text{R}_2$  and  $\text{R}_4$  are independently selected from the group consisting of H,  $\text{C}_1$ - $\text{C}_8$  alkyl,  $\text{C}_2$ - $\text{C}_8$  alkenyl,  $(\text{C}_1$ - $\text{C}_4$  alkyl)OH,  $(\text{C}_1$ - $\text{C}_4$  alkyl)SH,  $(\text{C}_2$ - $\text{C}_3$  alkyl) $\text{SCH}_3$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{CONH}_2$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{COOH}$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NH}_2$ ,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NHC}(\text{NH}_2^+) \text{NH}_2$ ,  $(\text{C}_0$ - $\text{C}_4$  alkyl)( $\text{C}_3$ - $\text{C}_6$  cycloalkyl),  $(\text{C}_0$ - $\text{C}_4$  alkyl)( $\text{C}_6$ - $\text{C}_{10}$  aryl) $\text{R}_7$ ,  
10                      and  $\text{CH}_2(\text{C}_5$ - $\text{C}_9$  heteroaryl);

$\text{R}_3$  is selected from the group consisting of  $\text{C}_1$ - $\text{C}_8$  alkyl,  $(\text{C}_1$ - $\text{C}_4$  alkyl)OH,  $(\text{C}_1$ - $\text{C}_4$  alkyl)SH,  $(\text{C}_1$ - $\text{C}_4$  alkyl) $\text{NH}_2$ ,  $(\text{C}_3$ - $\text{C}_6$ )cycloalkyl or  $\text{R}_4$  and  $\text{R}_3$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic  
15                      ring;

$\text{R}_5$  is  $\text{NHR}_6$  or OH;

$\text{R}_6$  is H, or  $\text{R}_6$  and  $\text{R}_2$  together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

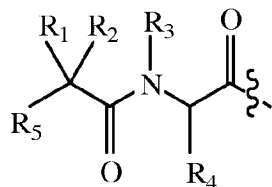
$\text{R}_7$  is selected from the group consisting of H and OH, with the proviso  
20                      that one and only one of Z, J,  $\text{R}_{10}$ ,  $\text{R}_{11}$  or  $\text{R}_{12}$  is a dipeptide comprising the general structure:



(i.e., only one dipeptide prodrug element is

attached to the insulin peptide).

In one embodiment J is a dipeptide comprising the general structure:



, and X and X<sub>12</sub> are each OH and X<sub>13</sub> is H, with

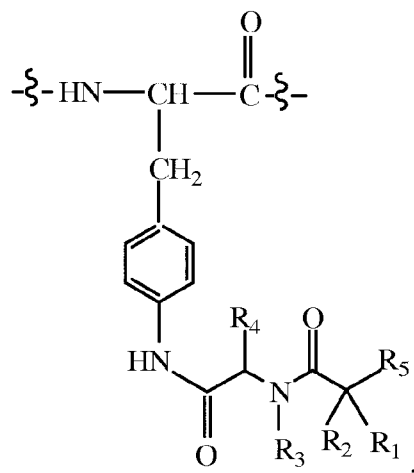
- 5 the proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring, both R<sub>1</sub> and R<sub>2</sub> are other than H. In an alternative embodiment X<sub>12</sub> comprises the dipeptide of Formula I, J and X<sub>13</sub> are each H and X is OH. In another alternative embodiment X<sub>13</sub> is comprises the dipeptide of Formula I, X and X<sub>12</sub> are each OH and J is H. In another embodiment X comprises the dipeptide
- 10 of Formula I, J and X<sub>13</sub> are each H and X<sub>12</sub> is OH. In one embodiment the B chain comprises the sequence J-X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTPKT (SEQ ID NO: 15) or J-X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTKPT (SEQ ID NO: 16), wherein J, X<sub>4</sub>, X<sub>5</sub>, X<sub>6</sub>, X<sub>7</sub>, X<sub>8</sub> and X<sub>9</sub> are defined as immediately above. In a further embodiment R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl and R<sub>4</sub> is selected from the group
- 15 consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl, or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring. In another further embodiment X<sub>4</sub> is histidine, X<sub>5</sub> is serine and X<sub>6</sub> is histidine.

In one embodiment the single chain insulin analog comprises a compound of the formula: B-P-A, wherein:

- 20 B represents a B chain sequence comprising a sequence of X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFF (SEQ ID NO: 4) or a functional analog thereof,

A represents an A chain sequence comprising a sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) or a functional analog thereof, wherein

- 25 X<sub>1</sub> is selected from the group consisting of threonine and histidine;  
X<sub>2</sub> is an amino acid of the general structure



X<sub>3</sub> is asparagine or glycine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

5 X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid.

P represents a linker, including a peptide linker, that covalently joins the amino-terminus of the A chain to the carboxy- terminus of the B chain. In an alternative embodiment the single-chain insulin analog comprises a compound of the formula: A-P-B, wherein: A represents a human insulin A chain, or a functional analog thereof, B represents a human insulin B chain, or a functional analog thereof, and P represents a linker, including a peptide linker, that covalently joins the amino-terminus of the B chain to the carboxy- terminus of the A chain. In one embodiment the peptide linker comprises 4 to 8 amino acids.

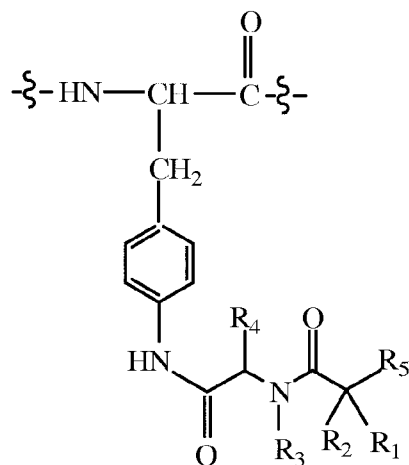
15 In accordance with one embodiment the peptide linker is 5 to 18 amino acids in length and comprises a sequence selected from the group consisting of: Gly-Gly-Gly-Pro-Gly-Lys-Arg (SEQ ID NO: 22), Gly-Tyr-Gly-Ser-Ser-Ser-Arg-Arg-Ala-Pro-Gln-Thr (SEQ ID NO: 23), Arg-Arg-Gly-Pro-Gly-Gly-Gly (SEQ ID NO: 32), Gly-Gly-Gly-Gly-Gly-Lys-Arg (SEQ ID NO: 24), Arg-Arg-Gly-Gly-Gly-Gly-Gly (SEQ ID NO: 25), Gly-Gly-Ala-Pro-Gly-Asp-Val-Lys-Arg (SEQ ID NO: 26), Arg-Arg-Ala-Pro-Gly-Asp-Val-Gly-Gly (SEQ ID NO: 27), Gly-Gly-Tyr-Pro-Gly-Asp-Val-Lys-Arg (SEQ ID NO: 28), Arg-Arg-Tyr-Pro-Gly-Asp-Val-Gly-Gly (SEQ ID NO: 29), Gly-Gly-His-Pro-Gly-Asp-Val-Lys-Arg (SEQ ID NO: 30) and Arg-Arg-His-Pro-Gly-Asp-Val-Gly-Gly (SEQ ID NO: 31). In one embodiment the peptide linker is 7 to 12

amino acids in length and comprises the sequence Gly-Gly-Gly-Pro-Gly-Lys-Arg (SEQ ID NO: 22) or Gly-Tyr-Gly-Ser-Ser-Ser-Arg-Arg-Ala-Pro-Gln-Thr (SEQ ID NO: 23).

In a further embodiment the peptide linker comprises a sequence selected from  
 5 the group consisting of AGRGSGK (SEQ ID NO: 35), AGLGSGK (SEQ ID NO: 36),  
 AGMGSGK (SEQ ID NO: 37), ASWGSGK (SEQ ID NO: 38), TGLGSGQ (SEQ ID  
 NO: 39), TGLGRGK (SEQ ID NO: 40), TGLGSGK (SEQ ID NO: 41), HGLYSGK  
 (SEQ ID NO: 42), KGLGSGQ (SEQ ID NO: 43), VGLMSGK (SEQ ID NO: 44),  
 VGLSSGQ (SEQ ID NO: 45), VGLYSGK (SEQ ID NO: 46), VGLSSGK (SEQ ID  
 10 NO: 47), VGMSSGK (SEQ ID NO: 48), VWSSSGK (SEQ ID NO: 49), VGSSSGK  
 (SEQ ID NO: 50), VGMSSGK (SEQ ID NO: 51), TGLGSGR (SEQ ID NO: 52),  
 TGLGKGQ (SEQ ID NO: 53), KGLSSGQ (SEQ ID NO: 54), VKLSSGQ (SEQ ID  
 NO: 55), VGLKSGQ (SEQ ID NO: 56), TGLGKGQ (SEQ ID NO: 57) SRVSRRSR  
 (SEQ ID NO: 65), GYGSSSRRAPQT (SEQ ID NO: 66) and VGLSKGQ (SEQ ID  
 15 NO: 58). In one embodiment the linker comprises GSSSRRAP (SEQ ID NO: 67) or  
 SRVSRRSR (SEQ ID NO: 65)

In one embodiment the single-chain insulin analog has the amino acid  
 sequence:

Phe-Val-Asn-Gln-His-Leu-Cys-Gly-Ser-His-Leu-Val-Glu- Ala-Leu-Tyr-Leu-Val-  
 20 Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr- Thr-Pro-Lys-Thr-Gly-Ile-Val-Glu-Gln-Cys-Cys-  
 Thr-Ser-Ile-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Xaa-Cys-Asn (SEQ ID NO: 33) or  
 Phe-Val-Asn-Gln-His-Leu-Cys-Gly-Ser-His-Leu-Val-Glu- Ala-Leu-Tyr-Leu-Val-  
 Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr- Thr-Pro-Lys-Thr-Gln-Pro-Leu-Ala-Leu-Glu-  
 Gly-Ser-Leu- Gln-Lys-Arg-Gly-Ile-Val-Glu-Gln-Cys-Cys-Thr-Ser-Ile-Cys-Ser-Leu-  
 25 Tyr-Gln-Leu-Glu-Asn-Xaa-Cys-Asn (SEQ ID NO: 34) wherein Xaa is an amino acid  
 of the general structure



The insulin peptides disclosed herein may be part of a dimer, trimer or higher order multimer comprising at least two, three, or more peptides bound via a linker, wherein at least one or both peptides comprises a dipeptide prodrug element linked to the insulin peptide. The dimer may be a homodimer or heterodimer, comprising peptides selected from the group consisting of native insulin, native IGF-1, native IGF-II and insulin analog peptides as disclosed herein. In one embodiment, the linker is selected from the group consisting of a bifunctional thiol crosslinker and a bi-functional amine crosslinker. In certain embodiments, the linker is PEG, e.g., a 5 kDa PEG, 20 kDa PEG. In some embodiments, the linker is a disulfide bond.

For example, each monomer of the dimer may comprise a Cys residue (e.g., a terminal or internally positioned Cys) and the sulfur atom of each Cys residue participates in the formation of the disulfide bond. Each monomer of the dimer represents a heterodimer of an A and B chain. The A and B chain are either linked via disulfide bonds or are prepared as single chain peptides. In some aspects of the invention, the monomers are connected via terminal amino acids (e.g., N-terminal or C-terminal), via internal amino acids, or via a terminal amino acid of at least one monomer and an internal amino acid of at least one other monomer. In specific aspects, the monomers are not connected via an N-terminal amino acid. In some aspects, the monomers of the multimer are attached together in a “tail-to-tail” orientation in which the C-terminal amino acids of each monomer are attached together. A conjugate moiety may be covalently linked to any of the insulin peptides described herein, including a dimer, trimer or higher order multimer.



The prodrugs disclosed herein can be further modified to improve the peptide's solubility in aqueous solutions at physiological pH, while enhancing the effective duration of the peptide by preventing renal clearance of the peptide. Peptides are easily cleared because of their relatively small molecular size when compared to  
5 plasma proteins. Increasing the molecular weight of a peptide above 40 kDa exceeds the renal threshold and significantly extends duration in the plasma. Accordingly, in one embodiment the peptide prodrugs are further modified to comprise a covalently linked hydrophilic moiety. In one embodiment the hydrophilic moiety is a plasma protein polyethylene oxide chain or the Fc portion of an immunoglobulin. Therefore, in  
10 one embodiment the presently disclosed prodrugs are further modified to comprise one or more hydrophilic groups covalently linked to the side chains of amino acids.

In accordance with one embodiment the insulin prodrugs disclosed herein are further modified by linking a hydrophilic moiety to either the N-terminal amino acid of the B chain or to the side chain of a lysine amino acid located at the carboxy  
15 terminus of the B chain, including for example, at position 28 of SEQ ID NO: 9/SEQ ID NO: 13 or at position 29 of SEQ ID NO: 8/SEQ ID NO: 12. In one embodiment a single-chain insulin prodrug analog is provided wherein one of the amino acids of the peptide linker is modified by linking a hydrophilic moiety to the side chain of the peptide linker. In one embodiment the modified amino acid is cysteine, lysine or  
20 acetyl phenylalanine. In one embodiment the peptide linker is selected from the group consisting of TGLGSGQ (SEQ ID NO: 39), VGLSSGQ (SEQ ID NO: 45), VGLSSGK (SEQ ID NO: 47), TGLGSGR (SEQ ID NO: 52), TGLGKGQ (SEQ ID NO: 53), KGLSSGQ (SEQ ID NO: 54), VKLSSGQ (SEQ ID NO: 55), VGLKSGQ (SEQ ID NO: 56), TGLGKGQ (SEQ ID NO: 57) and VGLSKGQ (SEQ ID NO: 58)  
25 and the hydrophilic moiety (e.g., polyethylene glycol) is linked to the lysine side chain of the peptide linker.

In another embodiment the insulin prodrug analogs disclosed herein are further modified by the addition of a modified amino acid to the carboxy terminus of the B chain of the insulin prodrug, wherein the C-terminally added amino acid is  
30 modified to comprise a hydrophilic moiety linked to the amino acid. In one embodiment the amino acid added to the C-terminus is a modified cysteine, lysine or acetyl phenylalanine. In one embodiment the hydrophilic moiety is selected from the

group consisting of a plasma protein, polyethylene oxide chain and an Fc portion of an immunoglobulin.

In one embodiment the hydrophilic group is a polyethylene oxide chain, and in one embodiment two or more polyethylene oxide chains are covalently attached to two or more amino acid side chains of the insulin prodrug analog. In accordance with one embodiment the hydrophilic moiety is covalently attached to an amino acid side chain of an insulin prodrug analog disclosed herein at a position selected from the group consisting of A9, A14, A15, B22, B28, B29 and the C-terminus or N-terminus of the B chain. For insulin prodrug analogs having multiple polyethylene oxide chains, the polyethylene oxide chains can be attached at the N-terminal amino acid of the B chain or to the side chain of a lysine amino acid located at the carboxy terminus of the B chain, or by the addition of a single amino acid at the C-terminus of the peptide wherein the added amino acid has a polyethylene oxide chain linked to its side chain. In accordance with one embodiment the polyethylene oxide chain or other hydrophilic moiety is linked to the side chain of one of the two amino acids comprising the dipeptide prodrug element. In one embodiment the dipeptide prodrug element comprises a lysine (in the D or L configuration) with a polyethylene oxide chain attached to the side chain amine of the lysine.

#### *Linkage of hydrophilic moieties*

In another embodiment the solubility of the insulin analogs disclosed herein are enhanced by the covalent linkage of a hydrophilic moiety to the peptide. Hydrophilic moieties can be attached to the insulin analogs under any suitable conditions used to react a protein with an activated polymer molecule. Any means known in the art can be used, including via acylation, reductive alkylation, Michael addition, thiol alkylation or other chemoselective conjugation/ligation methods through a reactive group on the PEG moiety (e.g., an aldehyde, amino, ester, thiol,  $\alpha$ -haloacetyl, maleimido or hydrazino group) to a reactive group on the target compound (e.g., an aldehyde, amino, ester, thiol,  $\alpha$ -haloacetyl, maleimido or hydrazino group). Activating groups which can be used to link the water soluble polymer to one or more proteins include without limitation sulfone, maleimide, sulfhydryl, thiol, triflate, tresylate, aziridine, oxirane and 5-pyridyl. If attached to the peptide by reductive alkylation, the polymer selected should have a single reactive aldehyde so that the

degree of polymerization is controlled. See, for example, Kinstler et al., *Adv. Drug Delivery Rev.* 54: 477-485 (2002); Roberts et al., *Adv. Drug Delivery Rev.* 54: 459-476 (2002); and Zalipsky et al., *Adv. Drug Delivery Rev.* 16: 157-182 (1995).

Suitable hydrophilic moieties include polyethylene glycol (PEG),  
5 polypropylene glycol, polyoxyethylated polyols (e.g., POG), polyoxyethylated sorbitol, polyoxyethylated glucose, polyoxyethylated glycerol (POG), polyoxyalkylenes, polyethylene glycol propionaldehyde, copolymers of ethylene glycol/propylene glycol, monomethoxy-polyethylene glycol, mono-(C1-C10) alkoxy- or aryloxy-polyethylene glycol, carboxymethylcellulose, polyacetals, polyvinyl  
10 alcohol (PVA), polyvinyl pyrrolidone, poly-1, 3-dioxolane, poly-1,3,6-trioxane, ethylene/maleic anhydride copolymer, poly (.beta.-amino acids) (either homopolymers or random copolymers), poly(n-vinyl pyrrolidone)polyethylene glycol, propylene glycol homopolymers (PPG) and other polyalkylene oxides, polypropylene oxide/ethylene oxide copolymers, colonic acids or other  
15 polysaccharide polymers, Ficoll or dextran and mixtures thereof.

The hydrophilic moiety, e.g., polyethylene glycol chain in accordance with one embodiment has a molecular weight selected from the range of about 500 to about 40,000 Daltons. In one embodiment the hydrophilic moiety, e.g. PEG, has a  
molecular weight selected from the range of about 500 to about 5,000 Daltons, or  
20 about 1,000 to about 5,000 Daltons. In another embodiment the hydrophilic moiety, e.g., PEG, has a molecular weight of about 10,000 to about 20,000 Daltons. In yet other exemplary embodiments the hydrophilic moiety, e.g., PEG, has a molecular weight of about 20,000 to about 40,000 Daltons.

In one embodiment dextrans are used as the hydrophilic moiety. Dextrans are  
25 polysaccharide polymers of glucose subunits, predominantly linked by  $\alpha$ 1-6 linkages. Dextran is available in many molecular weight ranges, e.g., about 1 kD to about 100 kD, or from about 5, 10, 15 or 20 kD to about 20, 30, 40, 50, 60, 70, 80 or 90 kD.

Linear or branched polymers are contemplated. Resulting preparations of conjugates may be essentially monodisperse or polydisperse, and may have about 0.5,  
30 0.7, 1, 1.2, 1.5 or 2 polymer moieties per peptide.

In accordance with one embodiment, the insulin prodrug analogs disclosed herein are further modified by amino acid substitutions, wherein the substituting amino acid comprises a side chain suitable for crosslinking with hydrophilic moieties, including for example, polyethylene glycol. In one embodiment the amino acid at the position of the insulin prodrug analog where the hydrophilic moiety is to be linked is substituted (or added at the C-terminus) with a natural or synthetic amino acid to introduce, or allow for ease in attaching, the hydrophilic moiety. For example, in one embodiment a native amino acid at position selected from A5, A8, A9, A10, A12, A14, A15, A17, A18, B1, B2, B3, B4, B5, B13, B14, B17, B21, B22, B26, B27, B28, B29 and B30 is substituted with a lysine, cysteine or acetyl phenylalanine residue (or a lysine, cysteine or acetyl phenylalanine residue is added to the C-terminus) to allow for the covalent attachment of a polyethylene glycol chain.

In one embodiment the insulin prodrug analog has a single cysteine residue added to the carboxy terminus of the B chain, or the insulin prodrug analog is substituted with at least one cysteine residue, wherein the side chain of the cysteine residue is further modified with a thiol reactive reagent, including for example, maleimido, vinyl sulfone, 2-pyridylthio, haloalkyl, and haloacyl. These thiol reactive reagents may contain carboxy, keto, hydroxyl, and ether groups as well as other hydrophilic moieties such as polyethylene glycol units. In an alternative embodiment, the insulin prodrug analog has a single lysine residue added to the carboxy terminus of the B chain, or the insulin prodrug analog is substituted with lysine, and the side chain of the substituting lysine residue is further modified using amine reactive reagents such as active esters (succinimido, anhydride, etc) of carboxylic acids or aldehydes of hydrophilic moieties such as polyethylene glycol.

In those embodiments wherein the insulin prodrug analog comprises a polyethylene glycol chain, the polyethylene glycol chain may be in the form of a straight chain or it may be branched. In accordance with one embodiment the polyethylene glycol chain has an average molecular weight selected from the range of about 20,000 to about 60,000 Daltons. Multiple polyethylene glycol chains can be linked to the insulin prodrug analog to provide an insulin prodrug analog with optimal solubility and blood clearance properties. In one embodiment the insulin prodrug analog is linked to a single polyethylene glycol chain that has an average molecular weight selected from the range of about 20,000 to about 60,000 Daltons. In another

embodiment the insulin prodrug analog is linked to two polyethylene glycol chains wherein the combined average molecular weight of the two chains is selected from the range of about 40,000 to about 80,000 Daltons. In one embodiment a single polyethylene glycol chain having an average molecular weight of 20,000 or 60,000  
5 Daltons is linked to the insulin prodrug analog. In another embodiment a single polyethylene glycol chain is linked to the insulin prodrug analog and has an average molecular weight selected from the range of about 40,000 to about 50,000 Daltons. In one embodiment two polyethylene glycol chains are linked to the insulin prodrug analog wherein the first and second polyethylene glycol chains each have an average  
10 molecular weight of 20,000 Daltons. In another embodiment two polyethylene glycol chains are linked to the insulin prodrug analog wherein the first and second polyethylene glycol chains each have an average molecular weight of 40,000 Daltons.

In a further embodiment an insulin prodrug analog comprising two or more polyethylene glycol chains covalently bound to the peptide is provided, wherein the  
15 total molecular weight of the polyethylene glycol chains is about 40,000 to about 60,000 Daltons. In one embodiment the pegylated insulin prodrug analog comprises a polyethylene glycol chain linked to one or more amino acids selected from the N-terminus of the B chain and/or position 28 of SEQ ID NO: 9 or at position 29 of SEQ ID NO: 8, wherein the combined molecular weight of the PEG chain(s) is about  
20 40,000 to about 80,000 Daltons.

In accordance with one embodiment, an insulin peptide, or prodrug/depot derivative thereof, is fused to an accessory peptide which is capable of forming an extended conformation similar to chemical PEG (e.g., a recombinant PEG (rPEG) molecule), such as those described in International Patent Application Publication No.  
25 WO2009/023270 and U.S. Patent Application Publication No. US2008/0286808. The rPEG molecule is not polyethylene glycol. The rPEG molecule in some aspects is a polypeptide comprising one or more of glycine, serine, glutamic acid, aspartic acid, alanine, or proline. In some aspects, the rPEG is a homopolymer, e.g., poly-glycine, poly-serine, poly-glutamic acid, poly-aspartic acid, poly-alanine, or poly-proline. In  
30 other embodiments, the rPEG comprises two types of amino acids repeated, e.g., poly(Gly-Ser), poly(Gly-Glu), poly(Gly-Ala), poly(Gly-Asp), poly(Gly-Pro), poly(Ser-Glu), etc. In some aspects, the rPEG comprises three different types of amino acids, e.g., poly(Gly-Ser-Glu). In specific aspects, the rPEG increases the half-

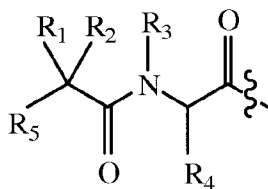
life of the insulin peptide. In some aspects, the rPEG comprises a net positive or net negative charge. The rPEG in some aspects lacks secondary structure. In one embodiment, the rPEG is greater than or equal to 10 amino acids in length, and in one embodiment is about 40 to about 50 amino acids in length. The accessory peptide in  
5 some aspects is fused to the N- or C- terminus of the peptide of the invention through a peptide bond or a proteinase cleavage site, or is inserted into the loops of the peptide of the invention. The rPEG in some aspects comprises an affinity tag or is linked to a PEG that is greater than 5 kDa. In one embodiment, the rPEG confers the peptide of the invention with an increased hydrodynamic radius, serum half-life, protease  
10 resistance, or solubility and in some aspects confers the peptide with decreased immunogenicity.

In accordance with one embodiment, an insulin prodrug analog is provided wherein a plasma protein has been covalently linked to an amino acid side chain of the peptide to improve the solubility, stability and/or pharmacokinetics of the insulin  
15 prodrug analog. For example, serum albumin can be covalently bound to the insulin prodrug analogs presented herein. In one embodiment the plasma protein is covalently bound to the N-terminus of the B chain and/or to an amino acid corresponding to position 28 of SEQ ID NO: 9 or at position 29 of SEQ ID NO: 8.

In accordance with one embodiment, an insulin prodrug analog is provided  
20 wherein a linear amino acid sequence representing the Fc portion of an immunoglobulin molecule has been covalently linked to an amino acid side chain of an insulin prodrug analog disclosed herein to improve the solubility, stability and/or pharmacokinetics of the insulin prodrug analog. For example, the amino acid sequence representing the Fc portion of an immunoglobulin molecule can be covalently bound to the N-terminus of  
25 the B chain or the C-terminus of the A or B chain, or the C-terminus of an A or B chain that has been terminally extended. For example, the amino acid sequence representing the Fc portion of an immunoglobulin molecule can be covalently bound to the C-terminus of the B chain, including for example linkage to an amino acid corresponding to position 28 of SEQ ID NO: 9 or at position 29 of SEQ ID NO: 8.  
30 The Fc portion is typically one isolated from IgG, but the Fc peptide fragment from any immunoglobulin should function equivalently.

In a specific aspect of the invention, the insulin prodrug analog is modified to comprise an alkyl or acyl group by direct alkylation or acylation of an amine, hydroxyl, or thiol of a side chain of an amino acid of the insulin prodrug analog. In one embodiment, the insulin prodrug analog is directly acylated through the side chain amine, hydroxyl, or thiol of an amino acid. In one embodiment, acylation is at one or more positions selected from A9, A14, A15, B22, B28 or B29. In this regard, the acylated insulin prodrug analog can comprise an A chain amino acid sequence of SEQ ID NO: 3 and a B chain of SEQ ID NO: 5, or a modified amino acid sequence of SEQ ID NO: 3 and/or SEQ ID NO: 5 with at least one of the amino acids at positions A9, A14, A15, B22, B28 or B29 modified to any amino acid comprising a side chain amine, hydroxyl, or thiol. In some specific embodiments, the direct acylation of the insulin prodrug analog occurs through the side chain amine, hydroxyl, or thiol of the amino acid at position B28 or B29. In one further embodiment the insulin prodrug analog comprises an acyl group of a carboxylic acid with 1-24 carbon atoms bound to the epsilon-amino group of a Lys present at position B28 or B29. In one embodiment a single-chain insulin prodrug analog is provided wherein one of the amino acids of the peptide linker is modified to comprise an acyl group by direct acylation of an amine, hydroxyl, or thiol of a side chain of an amino acid of the peptide linker. In accordance with one embodiment the peptide linker of the single-chain insulin analog is selected from the group consisting of AGRGSGK (SEQ ID NO: 35), AGLGSGK (SEQ ID NO: 36), AGMSGK (SEQ ID NO: 37), ASWGSGK (SEQ ID NO: 38), TGLGSGQ (SEQ ID NO: 39), TGLGRGK (SEQ ID NO: 40), TGLGSGK (SEQ ID NO: 41), HGLYSGK (SEQ ID NO: 42), KGLGSGQ (SEQ ID NO: 43), VGLMSGK (SEQ ID NO: 44), VGLSSGQ (SEQ ID NO: 45), VGLYSGK (SEQ ID NO: 46), VGLSSGK (SEQ ID NO: 47), VGMSSGK (SEQ ID NO: 48), VWSSSGK (SEQ ID NO: 49), VGSSSGK (SEQ ID NO: 50), VGMSSGK (SEQ ID NO: 51), TGLGSGR (SEQ ID NO: 52), TGLGKGQ (SEQ ID NO: 53), KGLSSGQ (SEQ ID NO: 54), VKLSSGQ (SEQ ID NO: 55), VGLKSGQ (SEQ ID NO: 56), TGLGKGQ (SEQ ID NO: 57) and VGLSKGQ (SEQ ID NO: 58) wherein at least one lysine residue in the A-chain, in the B-chain or in the connecting peptide has been chemically modified by acylation. In one embodiment the acylating group comprises a 1-5, 10-12 or 12-24 carbon chain.

In accordance with one embodiment the insulin prodrug analogs as disclosed herein are further modified to link an additional compound to the prodrug dipeptide moiety of the analog. In one embodiment the side chain of an amino acid comprising the dipeptide prodrug element is pegylated, acylated or alkylated. In one embodiment the dipeptide is acylated with a group comprising a 1-5, 10-12 or 12-24 carbon chain. In one embodiment the dipeptide is pegylated with a 40-80 KDa polyethylene glycol chain. In one embodiment the dipeptide prodrug element is pegylated and the insulin peptide linked to the dipeptide is acylated, including, for example, acylation at the C-terminal lysine of the B chain. In accordance with one embodiment a hydrophilic moiety or a sequestering macromolecule is covalently linked to the R<sub>2</sub> side chain of the dipeptide comprising the general structure:



, wherein R<sub>2</sub> is selected from the group consisting of (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, and (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>. In one embodiment R<sub>2</sub> is (C<sub>3</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>. Sequestering macromolecules are known to those skilled in the art and include dextrans and large molecular weight polyethylene glycol (i.e., greater than or equal to 80 KDa). By linking the sequestering macromolecule to the dipeptide moiety, the prodrug will remain sequestered, while the active insulin peptide is slowly released based on the kinetics of the cleavage of the dipeptide amide bond.

The present disclosure also encompasses other conjugates in which insulin prodrug analogs of the invention are linked, optionally via covalent bonding, and optionally via a linker, to a conjugate. Linkage can be accomplished by covalent chemical bonds, physical forces such as electrostatic, hydrogen, ionic, van der Waals, or hydrophobic or hydrophilic interactions. A variety of non-covalent coupling systems may be used, including biotin-avidin, ligand/receptor, enzyme/substrate, nucleic acid/nucleic acid binding protein, lipid/lipid binding protein, cellular adhesion molecule partners; or any binding partners or fragments thereof which have affinity for each other.

Exemplary conjugates include but are not limited to a heterologous peptide or polypeptide (including for example, a plasma protein), a targeting agent, an



immunoglobulin or portion thereof (e.g. variable region, CDR, or Fc region), a diagnostic label such as a radioisotope, fluorophore or enzymatic label, a polymer including water soluble polymers, or other therapeutic or diagnostic agents. In one embodiment a conjugate is provided comprising an insulin prodrug analog of the present disclosure and a plasma protein, wherein the plasma protein is selected from the group consisting of albumin, transferin and fibrinogen. In one embodiment the plasma protein moiety of the conjugate is albumin or transferin. In one embodiment, the linker comprises a chain of atoms from 1 to about 60, or 1 to 30 atoms or longer, 2 to 5 atoms, 2 to 10 atoms, 5 to 10 atoms, or 10 to 20 atoms long. In one embodiment, the chain atoms are all carbon atoms. In one embodiment, the chain atoms in the backbone of the linker are selected from the group consisting of C, O, N, and S. Chain atoms and linkers may be selected according to their expected solubility (hydrophilicity) so as to provide a more soluble conjugate. In one embodiment, the linker provides a functional group that is subject to cleavage by an enzyme or other catalyst or hydrolytic conditions found in the target tissue or organ or cell. In one embodiment, the length of the linker is long enough to reduce the potential for steric hindrance. If the linker is a covalent bond or a peptidyl bond and the conjugate is a polypeptide, the entire conjugate can be a fusion protein. Such peptidyl linkers may be any length. Exemplary linkers are from about 1 to 50 amino acids in length, 5 to 50, 3 to 5, 5 to 10, 5 to 15, or 10 to 30 amino acids in length. Such fusion proteins may alternatively be produced by recombinant genetic engineering methods known to one of ordinary skill in the art.

#### *Conjugates and fusions*

The present disclosure also encompasses other conjugates in which the insulin analogs disclosed herein are linked, optionally via covalent bonding and optionally via a linker, to a conjugate moiety. Linkage can be accomplished by covalent chemical bonds, physical forces such electrostatic, hydrogen, ionic, van der Waals, or hydrophobic or hydrophilic interactions. A variety of non-covalent coupling systems may be used, including biotin-avidin, ligand/receptor, enzyme/substrate, nucleic acid/nucleic acid binding protein, lipid/lipid binding protein, cellular adhesion molecule partners; or any binding partners or fragments thereof which have affinity for each other.

The peptide can be linked to conjugate moieties via direct covalent linkage by reacting targeted amino acid residues of the peptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of these targeted amino acids. Reactive groups on the peptide or conjugate include, e.g., an aldehyde, amino, ester, thiol,  $\alpha$ -haloacetyl, maleimido or hydrazino group. Derivatizing agents include, for example, maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride or other agents known in the art.

Alternatively, the conjugate moieties can be linked to the peptide indirectly through intermediate carriers, such as polysaccharide or polypeptide carriers. Examples of polysaccharide carriers include aminodextran. Examples of suitable polypeptide carriers include polylysine, polyglutamic acid, polyaspartic acid, co-polymers thereof, and mixed polymers of these amino acids and others, e.g., serines, to confer desirable solubility properties on the resultant loaded carrier.

Cysteinyll residues most commonly are reacted with  $\alpha$ -haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyll residues also are derivatized by reaction with bromotrifluoroacetone,  $\alpha$ -bromo- $\beta$ -(5-imidozoyl)propionic acid, chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl disulfide, p-chloromercuribenzoate, 2-chloromercuri-4-nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction with diethylpyrocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Para-bromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysinyll and amino-terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect of reversing the charge of the lysinyll residues. Other suitable reagents for derivatizing  $\alpha$ -amino-containing residues include imidoesters such as methyl picolinimate, pyridoxal phosphate, pyridoxal, chloroborohydride, trinitrobenzenesulfonic acid, O-

methylisourea, 2,4-pentanedione, and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues are modified by reaction with one or several conventional reagents, among them phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and  
5 ninhydrin. Derivatization of arginine residues requires that the reaction be performed in alkaline conditions because of the high  $pK_a$  of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine epsilon-amino group.

The specific modification of tyrosyl residues may be made, with particular  
10 interest in introducing spectral labels into tyrosyl residues by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidazole and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by  
15 reaction with carbodiimides ( $R-N=C=N-R'$ ), where R and R' are different alkyl groups, such as 1-cyclohexyl-3-(2-morpholinyl-4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residues are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

20 Other modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (T. E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), deamidation of asparagines or glutamine, acetylation of the N-  
25 terminal amine, and/or amidation or esterification of the C-terminal carboxylic acid group.

Another type of covalent modification involves chemically or enzymatically coupling glycosides to the peptide. Sugar(s) may be attached to (a) arginine and histidine, (b) free carboxyl groups, (c) free sulfhydryl groups such as those of  
30 cysteine, (d) free hydroxyl groups such as those of serine, threonine, or hydroxyproline, (e) aromatic residues such as those of tyrosine, or tryptophan, or (f)

the amide group of glutamine. These methods are described in WO87/05330 published 11 Sep. 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Exemplary conjugate moieties that can be linked to any of the insulin analogs described herein include but are not limited to a heterologous peptide or polypeptide (including for example, a plasma protein), a targeting agent, an immunoglobulin or portion thereof (e.g. variable region, CDR, or Fc region), a diagnostic label such as a radioisotope, fluorophore or enzymatic label, a polymer including water soluble polymers, or other therapeutic or diagnostic agents. In one embodiment a conjugate is provided comprising an insulin analog disclosed herein and a plasma protein, wherein the plasma protein is selected from the group consisting of albumin, transferrin, fibrinogen and globulins.

In one embodiment, the linker comprises a chain of atoms from 1 to about 60, or 1 to 30 atoms or longer, 2 to 5 atoms, 2 to 10 atoms, 5 to 10 atoms, or 10 to 20 atoms long. In one embodiment, the chain atoms are all carbon atoms. In one embodiment, the chain atoms in the backbone of the linker are selected from the group consisting of C, O, N, and S. Chain atoms and linkers may be selected according to their expected solubility (hydrophilicity) so as to provide a more soluble conjugate. In one embodiment, the linker provides a functional group that is subject to cleavage by an enzyme or other catalyst or hydrolytic conditions found in the target tissue or organ or cell. In one embodiment, the length of the linker is long enough to reduce the potential for steric hindrance. If the linker is a covalent bond or a peptidyl bond and the conjugate is a polypeptide, the entire conjugate can be a fusion protein. Such peptidyl linkers may be any length. Exemplary linkers are from about 1 to 50 amino acids in length, 5 to 50, 3 to 5, 5 to 10, 5 to 15, or 10 to 30 amino acids in length. Such fusion proteins may alternatively be produced by recombinant genetic engineering methods known to one of ordinary skill in the art.

As noted above, in one embodiment, the insulin analogs are conjugated, e.g., fused to an immunoglobulin or portion thereof (e.g. variable region, CDR, or Fc region). Known types of immunoglobulins (Ig) include IgG, IgA, IgE, IgD or IgM. The Fc region is a C-terminal region of an Ig heavy chain, which is responsible for binding to Fc receptors that carry out activities such as recycling (which results in

prolonged half-life), antibody dependent cell-mediated cytotoxicity (ADCC), and complement dependent cytotoxicity (CDC).

For example, according to some definitions the human IgG heavy chain Fc region stretches from Cys226 to the C-terminus of the heavy chain. The "hinge region" generally extends from Glu216 to Pro230 of human IgG1 (hinge regions of other IgG isotypes may be aligned with the IgG1 sequence by aligning the cysteines involved in cysteine bonding). The Fc region of an IgG includes two constant domains, CH2 and CH3. The CH2 domain of a human IgG Fc region usually extends from amino acids 231 to amino acid 341. The CH3 domain of a human IgG Fc region usually extends from amino acids 342 to 447. References made to amino acid numbering of immunoglobulins or immunoglobulin fragments, or regions, are all based on Kabat et al. 1991, Sequences of Proteins of Immunological Interest, U.S. Department of Public Health, Bethesda, Md. In a related embodiments, the Fc region may comprise one or more native or modified constant regions from an immunoglobulin heavy chain, other than CH1, for example, the CH2 and CH3 regions of IgG and IgA, or the CH3 and CH4 regions of IgE.

Suitable conjugate moieties include portions of immunoglobulin sequence that include the FcRn binding site. FcRn, a salvage receptor, is responsible for recycling immunoglobulins and returning them to circulation in blood. The region of the Fc portion of IgG that binds to the FcRn receptor has been described based on X-ray crystallography (Burmeister et al. 1994, Nature 372:379). The major contact area of the Fc with the FcRn is near the junction of the CH2 and CH3 domains. Fc-FcRn contacts are all within a single Ig heavy chain. The major contact sites include amino acid residues 248, 250-257, 272, 285, 288, 290-291, 308-311, and 314 of the CH2 domain and amino acid residues 385-387, 428, and 433-436 of the CH3 domain.

Some conjugate moieties may or may not include FcγR binding site(s). FcγR are responsible for ADCC and CDC. Examples of positions within the Fc region that make a direct contact with FcγR are amino acids 234-239 (lower hinge region), amino acids 265-269 (B/C loop), amino acids 297-299 (C'/E loop), and amino acids 327-332 (F/G) loop (Sondermann et al., Nature 406: 267-273, 2000). The lower hinge region of IgE has also been implicated in the FcRI binding (Henry, et al., Biochemistry 36, 15568-15578, 1997). Residues involved in IgA receptor binding are described in

Lewis et al., (J Immunol. 175:6694-701, 2005). Amino acid residues involved in IgE receptor binding are described in Sayers et al. (J Biol Chem. 279(34):35320-5, 2004).

Amino acid modifications may be made to the Fc region of an

immunoglobulin. Such variant Fc regions comprise at least one amino acid

5 modification in the CH3 domain of the Fc region (residues 342-447) and/or at least one amino acid modification in the CH2 domain of the Fc region (residues 231-341). Mutations believed to impart an increased affinity for FcRn include T256A, T307A, E380A, and N434A (Shields et al. 2001, J. Biol. Chem. 276:6591). Other mutations may reduce binding of the Fc region to FcγRI, FcγRIIA, FcγRIIB, and/or FcγRIIIA  
10 without significantly reducing affinity for FcRn. For example, substitution of the Asn at position 297 of the Fc region with Ala or another amino acid removes a highly conserved N-glycosylation site and may result in reduced immunogenicity with concomitant prolonged half-life of the Fc region, as well as reduced binding to FcγRs (Routledge et al. 1995, Transplantation 60:847; Friend et al. 1999, Transplantation  
15 68:1632; Shields et al. 1995, J. Biol. Chem. 276:6591). Amino acid modifications at positions 233-236 of IgG1 have been made that reduce binding to FcγRs (Ward and Ghetie 1995, Therapeutic Immunology 2:77 and Armour et al. 1999, Eur. J. Immunol. 29:2613). Some exemplary amino acid substitutions are described in US Patents 7,355,008 and 7,381,408, each incorporated by reference herein in its entirety.

#### *Acylation and alkylation*

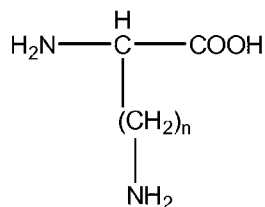
In accordance with one embodiment, the insulin analogs disclosed herein are modified to comprise an acyl group or alkyl group. Acylation or alkylation can increase the half-life of the insulin analogs in circulation. Acylation or alkylation can  
25 advantageously delay the onset of action and/or extend the duration of action at the insulin and/or IGF-1 receptors and/or improve resistance to proteases such as DPP-IV and/or improve solubility. Insulin analogs may be acylated or alkylated at the same amino acid position where a hydrophilic moiety is linked, or at a different amino acid position.

30 In one embodiment, the invention provides an insulin analog modified to comprise an acyl group or alkyl group covalently linked to the amino acid at a position corresponding to A10, B28, B29 of native insulin, or at the C-terminus or N-terminus of the A or B chain. The Insulin analog may further comprise a spacer

between the Insulin analog amino acid and the acyl group or alkyl group. In one embodiment, the acyl group is a fatty acid or bile acid, or salt thereof, e.g. a C4 to C30 fatty acid, a C8 to C24 fatty acid, cholic acid, a C4 to C30 alkyl, a C8 to C24 alkyl, or an alkyl comprising a steroid moiety of a bile acid. The spacer is any moiety with suitable reactive groups for attaching acyl or alkyl groups. In exemplary  
5       embodiments, the spacer comprises an amino acid, a dipeptide, or a tripeptide, or a hydrophilic bifunctional spacer. In one embodiment, the spacer is selected from the group consisting of: Trp, Glu, Asp, Cys and a spacer comprising  
10        $\text{NH}_2(\text{CH}_2\text{CH}_2\text{O})_n(\text{CH}_2)_m\text{COOH}$ , wherein m is any integer from 1 to 6 and n is any integer from 2 to 12. Such acylated or alkylated insulin peptides may also further comprise a hydrophilic moiety, optionally a polyethylene glycol. Any of the foregoing insulin analogs may comprise two acyl groups or two alkyl groups, or a combination thereof.

Acylation can be carried out at any positions within the insulin analog,  
15       provided that insulin analog insulin agonist activity is retained. The acyl group can be covalently linked directly to an amino acid of the insulin analog, or indirectly to an amino acid of the insulin analog via a spacer, wherein the spacer is positioned between the amino acid of the insulin peptide and the acyl group. In a specific aspect of the invention, the insulin analog is modified to comprise an acyl group by direct  
20       acylation of an amine, hydroxyl, or thiol of a side chain of an amino acid of the insulin peptide. In one embodiment, the insulin analog is directly acylated through the side chain amine, hydroxyl, or thiol of an amino acid. In one embodiment, acylation is at a position corresponding to A10, B28, B29 of native insulin, or at the C-terminus or N-terminus of the A or B chain. In this regard, the acylated insulin  
25       analog can comprise the amino acid sequence of SEQ ID NO: 9 and SEQ ID NO: 10, or a modified amino acid sequence thereof comprising one or more of the amino acid modifications described herein, with at least one of the amino acids at a position corresponding to A10, B28, B29 of native insulin, or at the C-terminus or N-terminus of the A or B chain modified to any amino acid comprising a side chain amine,  
30       hydroxyl, or thiol. In some specific embodiments, the direct acylation of the insulin peptide occurs through the side chain amine, hydroxyl, or thiol of the amino acid at a position corresponding to A10, B28, B29 of native insulin. In accordance with one embodiment one of the amino acid side chains of the dipeptide element is acylated.

In one embodiment, the amino acid to be acylated is an amino acid of Formula IV:



5

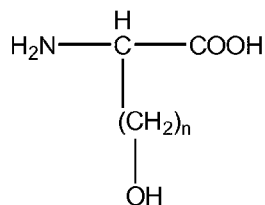
wherein n = 1 to 4

[Formula IV]

In some exemplary embodiments, the amino acid of Formula IV, is the amino acid wherein n is 4 (Lys) or n is 3 (Orn).

10

In other embodiments, the amino acid comprising a side chain hydroxyl is an amino acid of Formula V:



15

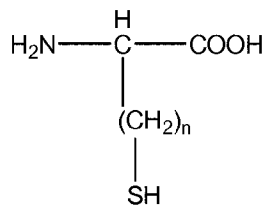
wherein n = 1 to 4

[Formula V]

In some exemplary embodiments, the amino acid of Formula V is the amino acid wherein n is 1 (Ser).

In yet other embodiments, the amino acid comprising a side chain thiol is an amino acid of Formula VI:

20



wherein n = 1 to 4



## [Formula VI]

In some exemplary embodiments, the amino acid of Formula VI is the amino acid wherein n is 1 (Cys).

5           In some exemplary embodiments, the insulin analog is modified to comprise an acyl group by acylation of an amine, hydroxyl, or thiol of a spacer, which spacer is attached to a side chain of an amino acid at position A10, B28 or B29 (according to the amino acid numbering of wild type insulin). The amino acid to which the spacer is attached can be any amino acid comprising a moiety which permits linkage to the  
10           spacer. For example, an amino acid comprising a side chain NH<sub>2</sub>, –OH, or –COOH (e.g., Lys, Orn, Ser, Asp, or Glu) is suitable. In one embodiment, the spacer is an amino acid comprising a side chain amine, hydroxyl, or a dipeptide or tripeptide comprising an amino acid comprising a side chain amine, hydroxyl, or thiol.

15           When acylation occurs through an amine group of a spacer the acylation can occur through the alpha amine of the amino acid or a side chain amine. In the instance in which the alpha amine is acylated, the spacer amino acid can be any amino acid. For example, the spacer amino acid can be a hydrophobic amino acid, e.g., Gly, Ala, Val, Leu, Ile, Trp, Met, Phe, Tyr. Alternatively, the spacer amino acid can be an  
20           acidic residue, e.g., Asp and Glu. In the instance in which the side chain amine of the spacer amino acid is acylated, the spacer amino acid is an amino acid comprising a side chain amine, e.g., an amino acid of Formula I (e.g., Lys or Orn). In this instance, it is possible for both the alpha amine and the side chain amine of the spacer amino acid to be acylated, such that the insulin peptide is diacylated. The present disclosure  
25           further contemplates diacylated insulin analogs.

          When acylation occurs through a hydroxyl group of a spacer, the amino acid or one of the amino acids of the dipeptide or tripeptide can be an amino acid of Formula II. In a specific exemplary embodiment, the amino acid is Ser.

          When acylation occurs through a thiol group of a spacer, the amino acid or  
30           one of the amino acids of the dipeptide or tripeptide can be an amino acid of Formula III. In a specific exemplary embodiment, the amino acid is Cys.

          In one embodiment, the spacer comprises a hydrophilic bifunctional spacer. In a specific embodiment, the spacer comprises an amino poly(alkyloxy)carboxylate.

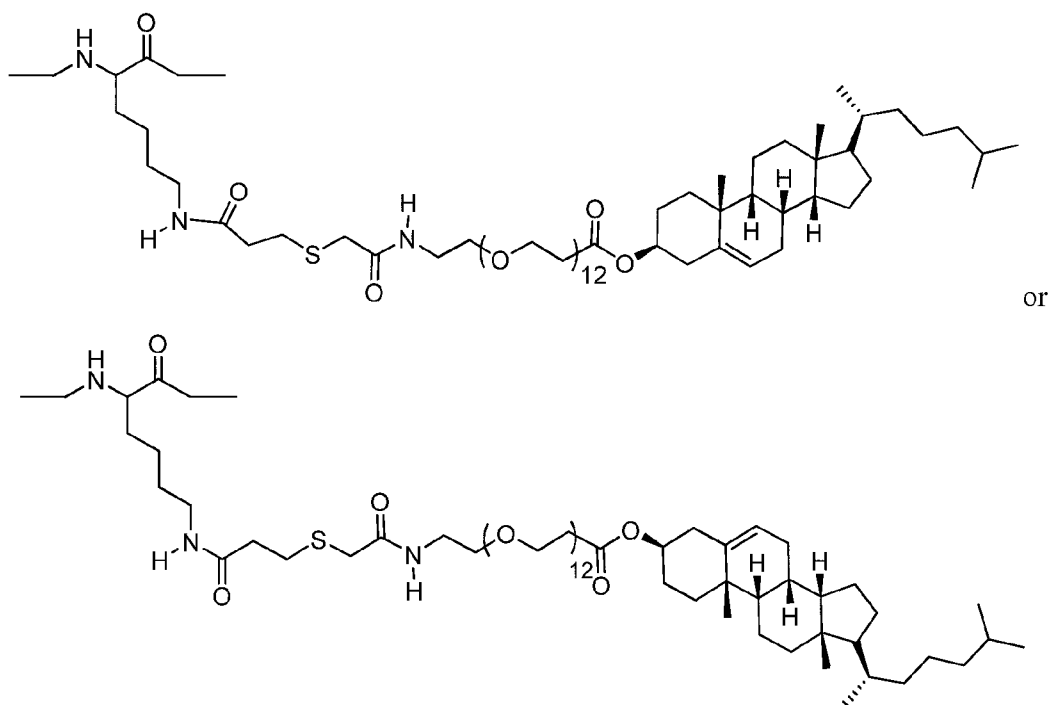
In this regard, the spacer can comprise, for example,  $\text{NH}_2(\text{CH}_2\text{CH}_2\text{O})_n(\text{CH}_2)_m\text{COOH}$ , wherein  $m$  is any integer from 1 to 6 and  $n$  is any integer from 2 to 12, such as, e.g., 8-amino-3,6-dioxaoctanoic acid, which is commercially available from Peptides International, Inc. (Louisville, KY).

5           Suitable methods of peptide acylation via amines, hydroxyls, and thiols are known in the art. See, for example, Miller, *Biochem Biophys Res Commun* 218: 377-382 (1996); Shimohigashi and Stammer, *Int J Pept Protein Res* 19: 54-62 (1982); and Previero et al., *Biochim Biophys Acta* 263: 7-13 (1972) (for methods of acylating through a hydroxyl); and San and Silviu, *J Pept Res* 66: 169-180 (2005) (for methods  
10 of acylating through a thiol); *Bioconjugate Chem.* "Chemical Modifications of Proteins: History and Applications" pages 1, 2-12 (1990); Hashimoto et al., *Pharmaceutical Res.* "Synthesis of Palmitoyl Derivatives of Insulin and their Biological Activity" Vol. 6, No: 2 pp.171-176 (1989)..

          The acyl group of the acylated insulin peptide can be of any size, e.g., any  
15 length carbon chain, and can be linear or branched. In some specific embodiments of the invention, the acyl group is a C4 to C30 fatty acid. For example, the acyl group can be any of a C4 fatty acid, C6 fatty acid, C8 fatty acid, C10 fatty acid, C12 fatty acid, C14 fatty acid, C16 fatty acid, C18 fatty acid, C20 fatty acid, C22 fatty acid, C24 fatty acid, C26 fatty acid, C28 fatty acid, or a C30 fatty acid. In one  
20 embodiment, the acyl group is a C8 to C20 fatty acid, e.g., a C14 fatty acid or a C16 fatty acid.

          In an alternative embodiment, the acyl group is a bile acid. The bile acid can be any suitable bile acid, including, but not limited to, cholic acid, chenodeoxycholic acid, deoxycholic acid, lithocholic acid, taurocholic acid, glycocholic acid, and  
25 cholesterol acid.

          In a specific embodiment, the insulin analog comprises a cholesterol acid, which is linked to a Lys residue of the insulin analog through an alkylated des-amino Cys spacer, i.e., an alkylated 3-mercaptopropionic acid spacer. The alkylated des-amino Cys spacer can be, for example, a des-amino-Cys spacer comprising a  
30 dodecaethylene glycol moiety. In one embodiment, the insulin analog comprises the structure:



The acylated insulin analogs described herein can be further modified to  
 5 comprise a hydrophilic moiety. In some specific embodiments the hydrophilic moiety  
 can comprise a polyethylene glycol (PEG) chain. The incorporation of a hydrophilic  
 moiety can be accomplished through any suitable means, such as any of the methods  
 described herein.

Alternatively, the acylated insulin peptide can comprise a spacer, wherein the  
 10 spacer is both acylated and modified to comprise the hydrophilic moiety. Nonlimiting  
 examples of suitable spacers include a spacer comprising one or more amino acids  
 selected from the group consisting of Cys, Lys, Orn, homo-Cys, and Ac-Phe.

In accordance with one embodiment, the insulin analog is modified to  
 comprise an alkyl group which is attached to the insulin analog via an ester, ether,  
 15 thioether, amide, or alkyl amine linkage for purposes of prolonging half-life in  
 circulation and/or delaying the onset of and/or extending the duration of action and/or  
 improving resistance to proteases such as DPP-IV.

The alkyl group of the alkylated insulin peptide can be of any size, e.g., any  
 length carbon chain, and can be linear or branched. In one embodiment of the  
 20 invention, the alkyl group is a C1 to C30 alkyl. For example, the alkyl group can be  
 any of a C1 alkyl, C2 alkyl, C3 alkyl, C4 alkyl, C6 alkyl, C8 alkyl, C10 alkyl, C12

alkyl, C14 alkyl, C16 alkyl, C18 alkyl, C20 alkyl, C22 alkyl, C24 alkyl, C26 alkyl, C28 alkyl, or a C30 alkyl. In one embodiment, the alkyl group is a C8 to C20 alkyl, e.g., a C14 alkyl or a C16 alkyl.

In some specific embodiments, the alkyl group comprises a steroid moiety of a bile acid, e.g., cholic acid, chenodeoxycholic acid, deoxycholic acid, lithocholic acid, taurocholic acid, glycocholic acid, and cholesterol acid.

In accordance with one embodiment a pharmaceutical composition is provided comprising any of the novel insulin prodrug analogs disclosed herein, preferably at a purity level of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%, and a pharmaceutically acceptable diluent, carrier or excipient. Such compositions may contain an A19 insulin analog as disclosed herein at a concentration of at least 0.5 mg/ml, 1 mg/ml, 2 mg/ml, 3 mg/ml, 4 mg/ml, 5 mg/ml, 6 mg/ml, 7 mg/ml, 8 mg/ml, 9 mg/ml, 10 mg/ml, 11 mg/ml, 12 mg/ml, 13 mg/ml, 14 mg/ml, 15 mg/ml, 16 mg/ml, 17 mg/ml, 18 mg/ml, 19 mg/ml, 20 mg/ml, 21 mg/ml, 22 mg/ml, 23 mg/ml, 24 mg/ml, 25 mg/ml or higher. In one embodiment the pharmaceutical compositions comprise aqueous solutions that are sterilized and optionally stored contained within various package containers. In other embodiments the pharmaceutical compositions comprise a lyophilized powder. The pharmaceutical compositions can be further packaged as part of a kit that includes a disposable device for administering the composition to a patient. The containers or kits may be labeled for storage at ambient room temperature or at refrigerated temperature.

In one embodiment, a composition is provided comprising a mixture of a first and second insulin prodrug analog, wherein the first and second insulin prodrug analogs differ from one another based on the structure of the prodrug element. More particularly, the first insulin prodrug analog may comprise a dipeptide prodrug element that has a half life substantially different from the dipeptide prodrug element of the second insulin prodrug analog. Accordingly, selection of different combinations of substituents on the dipeptide element will allow for the preparation of compositions that comprise a mixture of insulin prodrug analogs that are activated in a controlled manner over a desired time frame and at specific time intervals. For example, the compositions can be formulated to release active insulin at mealtimes followed by a subsequent activation of insulin during nighttime with suitable dosages being released based on time of activation. In another embodiment the

pharmaceutical composition comprises a mixture of an insulin prodrug analog disclosed herein and native insulin, or a known bioactive derivative of insulin.

The disclosed insulin prodrug analogs are believed to be suitable for any use that has previously been described for insulin peptides. Accordingly, the insulin  
5 prodrug analogs described herein can be used to treat hyperglycemia, or treat other metabolic diseases that result from high blood glucose levels. Accordingly, the present invention encompasses pharmaceutical compositions comprising an insulin prodrug analog of the present disclosure, and a pharmaceutically acceptable carrier for use in treating a patient suffering from high blood glucose levels. In accordance with  
10 one embodiment the patient to be treated using the insulin prodrug analogs disclosed herein is a domesticated animal, and in another embodiment the patient to be treated is a human.

One method of treating hyperglycemia in accordance with the present disclosure comprises the steps of administering the presently disclosed insulin  
15 prodrug analog to a patient using any standard route of administration, including parenterally, such as intravenously, intraperitoneally, subcutaneously or intramuscularly, intrathecally, transdermally, rectally, orally, nasally or by inhalation. In one embodiment the composition is administered subcutaneously or intramuscularly. In one embodiment, the composition is administered parenterally  
20 and the insulin prodrug analog composition is prepackaged in a syringe.

The insulin prodrug analogs of the invention may be administered alone or in combination with other anti-diabetic agents. Anti-diabetic agents known in the art or under investigation include native insulin, native glucagon and functional derivatives thereof, sulfonylureas, such as tolbutamide (Orinase), acetohexamide (Dymelor),  
25 tolazamide (Tolinase), chlorpropamide (Diabinese), glipizide (Glucotrol), glyburide (Diabeta, Micronase, Glynase), glimepiride (Amaryl), or gliclazide (Diamicron); meglitinides, such as repaglinide (Prandin) or nateglinide (Starlix); biguanides such as metformin (Glucophage) or phenformin; thiazolidinediones such as rosiglitazone (Avandia), pioglitazone (Actos), or troglitazone (Rezulin), or other PPAR $\gamma$  inhibitors;  
30 alpha glucosidase inhibitors that inhibit carbohydrate digestion, such as miglitol (Glyset), acarbose (Precose/Glucobay); exenatide (Byetta) or pramlintide; Dipeptidyl peptidase-4 (DPP-4) inhibitors such as vildagliptin or sitagliptin; SGLT (sodium-

dependent glucose transporter 1) inhibitors; or FBPase (fructose 1,6-bisphosphatase) inhibitors.

Pharmaceutical compositions comprising the insulin prodrug analogs disclosed herein can be formulated and administered to patients using standard  
5 pharmaceutically acceptable carriers and routes of administration known to those skilled in the art. Accordingly, the present disclosure also encompasses pharmaceutical compositions comprising one or more of the insulin prodrug analogs disclosed herein, or a pharmaceutically acceptable salt thereof, in combination with a pharmaceutically acceptable carrier. In one embodiment the pharmaceutical  
10 composition comprises a 1mg/ml concentration of the insulin prodrug analog at pH of about 4.0 to about 7.0 in a phosphate buffer system. The pharmaceutical compositions may comprise the insulin prodrug analog as the sole pharmaceutically active component, or the insulin prodrug analog can be combined with one or more additional active agents. In accordance with one embodiment a pharmaceutical  
15 composition is provided comprising one of the insulin prodrug analogs disclosed herein, preferably sterile and preferably at a purity level of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%, and a pharmaceutically acceptable diluent, carrier or excipient. Such compositions may contain an insulin prodrug analog wherein the resulting active peptide is present at a concentration of at least 0.5 mg/ml,  
20 1 mg/ml, 2 mg/ml, 3 mg/ml, 4 mg/ml, 5 mg/ml, 6 mg/ml, 7 mg/ml, 8 mg/ml, 9 mg/ml, 10 mg/ml, 11 mg/ml, 12 mg/ml, 13 mg/ml, 14 mg/ml, 15 mg/ml, 16 mg/ml, 17 mg/ml, 18 mg/ml, 19 mg/ml, 20 mg/ml, 21 mg/ml, 22 mg/ml, 23 mg/ml, 24 mg/ml, 25 mg/ml or higher. In one embodiment the pharmaceutical compositions comprise aqueous solutions that are sterilized and optionally stored within various containers.  
25 The compounds of the present invention can be used in accordance with one embodiment to prepare pre-formulated solutions ready for injection. In other embodiments the pharmaceutical compositions comprise a lyophilized powder. The pharmaceutical compositions can be further packaged as part of a kit that includes a disposable device for administering the composition to a patient. The containers or  
30 kits may be labeled for storage at ambient room temperature or at refrigerated temperature.

All therapeutic methods, pharmaceutical compositions, kits and other similar embodiments described herein contemplate that insulin prodrug analogs include all pharmaceutically acceptable salts thereof.

In one embodiment the kit is provided with a device for administering the  
5 insulin prodrug analog composition to a patient. The kit may further include a variety of containers, *e.g.*, vials, tubes, bottles, and the like. Preferably, the kits will also include instructions for use. In accordance with one embodiment the device of the kit is an aerosol dispensing device, wherein the composition is prepackaged within the aerosol device. In another embodiment the kit comprises a syringe and a needle,  
10 and in one embodiment the insulin analog composition is prepackaged within the syringe.

The compounds of this invention may be prepared by standard synthetic methods, recombinant DNA techniques, or any other methods of preparing peptides and fusion proteins. Although certain non-natural amino acids cannot be expressed  
15 by standard recombinant DNA techniques, techniques for their preparation are known in the art. Compounds of this invention that encompass non-peptide portions may be synthesized by standard organic chemistry reactions, in addition to standard peptide chemistry reactions when applicable.

## 20 EXAMPLE 1

### Synthesis of Insulin A & B Chains

Insulin A & B chains were synthesized on 4-methylbenzhyryl amine (MBHA) resin or 4-Hydroxymethyl-phenylacetamidomethyl (PAM) resin using Boc chemistry. The peptides were cleaved from the resin using HF/p-cresol 95:5 for 1 hour at 0°C.  
25 Following HF removal and ether precipitation, peptides were dissolved into 50% aqueous acetic acid and lyophilized. Alternatively, peptides were synthesized using Fmoc chemistry. The peptides were cleaved from the resin using Trifluoroacetic acid (TFA)/ Triisopropylsilane (TIS)/ H<sub>2</sub>O (95:2.5:2.5), for 2 hour at room temperature. The peptide was precipitated through the addition of an excessive amount of diethyl  
30 ether and the pellet solubilized in aqueous acidic buffer. The quality of peptides were monitored by RP-HPLC and confirmed by Mass Spectrometry (ESI or MALDI).

Insulin A chains were synthesized with a single free cysteine at amino acid 7 and all other cysteines protected as acetamidomethyl A-(SH)<sup>7</sup>(Acm)<sup>6,11,20</sup>. Insulin B

chains were synthesized with a single free cysteine at position 7 and the other cysteine protected as acetamidomethyl B-(SH)<sup>7</sup>(Acm)<sup>19</sup>. The crude peptides were purified by conventional RP-HPLC.

The synthesized A and B chains were linked to one another through their native disulfide bond linkage in accordance with the general procedure outlined in Fig. 1. The respective B chain was activated to the Cys<sup>7</sup>-Npys derivative through dissolution in DMF or DMSO and reacted with 2,2'-Dithiobis (5-nitropyridine) (Npys) at a 1:1 molar ratio, at room temperature. The activation was monitored by RP-HPLC and the product was confirmed by ESI-MS.

The first B7-A7 disulfide bond was formed by dissolution of the respective A-(SH)<sup>7</sup>(Acm)<sup>6,11,20</sup> and B-(Npys)<sup>7</sup>(Acm)<sup>19</sup> at 1:1 molar ratio to a total peptide concentration of 10 mg/ml. When the chain combination reaction was complete the mixture was diluted to a concentration of 50% aqueous acetic acid. The last two disulfide bonds were formed simultaneously through the addition of iodine. A 40 fold molar excess of iodine was added to the solution and the mixture was stirred at room temperature for an additional hour. The reaction was terminated by the addition of an aqueous ascorbic acid solution. The mixture was purified by RP-HPLC and the final compound was confirmed by MALDI-MS. As shown in Fig. 2 and the data in Table 1, the synthetic insulin prepared in accordance with this procedure compares well with purified insulin for insulin receptor binding.

Insulin peptides comprising a modified amino acid (such as 4-amino phenylalanine at position A19) can also be synthesized *in vivo* using a system that allows for incorporation of non-coded amino acids into proteins, including for example, the system taught in US Patent Nos. 7,045,337 and 7,083,970.

Table 1: Activity of synthesized insulin relative to native insulin

	Insulin Standard		A7-B7 Insulin	
	AVER.	STDEV	AVER.	STDEV
IC <sub>50</sub> (nM)	0.24	0.07	0.13	0.08
% of Insulin Activity	100		176.9	



## EXAMPLE 2

## Pegylation of Amine Groups (N-Terminus and Lysine) by Reductive Alkylation

## a. Synthesis

5           Insulin (or an insulin analog), mPEG20k-Aldhyde, and  $\text{NaBH}_3\text{CN}$ , in a molar ratio of 1:2:30, were dissolved in acetic acid buffer at a pH of 4.1-4.4. The reaction solution was composed of 0.1 N NaCl, 0.2 N acetic acid and 0.1 N  $\text{Na}_2\text{CO}_3$ . The insulin peptide concentration was approximately 0.5 mg/ml. The reaction occurs over six hours at room temperature. The degree of reaction was monitored by RP-HPLC  
10 and the yield of the reaction was approximately 50%.

## b. Purification

          The reaction mixture was diluted 2-5 fold with 0.1% TFA and applied to a preparative RP-HPLC column. HPLC condition: C4 column; flow rate 10 ml/min; A buffer 10% ACN and 0.1% TFA in water; B buffer 0.1% TFA in ACN; A linear  
15 gradient B% from 0-40% (0-80 min); PEG-insulin or analogues was eluted at approximately 35% buffer B. The desired compounds were verified by MALDI-TOF, following chemical modification through sulftolysis or trypsin degradation.

## Pegylation of Amine Groups (N-Terminus and Lysine) by N-Hydroxysuccinimide

## 20   Acylation.

## a. Synthesis

          Insulin (or an insulin analog) along with mPEG20k-NHS were dissolved in 0.1 N Bicine buffer (pH 8.0) at a molar ratio of 1:1. The insulin peptide concentration was approximately 0.5 mg/ml. Reaction progress was monitored by HPLC. The  
25 yield of the reaction is approximately 90% after 2 hours at room temperature.

## b. Purification

          The reaction mixture was diluted 2-5 fold and loaded to RP-HPLC. HPLC condition: C4 column; flow rate 10 ml/min; A buffer 10% ACN and 0.1% TFA in water; B buffer 0.1% TFA in ACN; A linear gradient B% from 0-40% (0-80 min);  
30 PEG-insulin or analogues was collected at approximately 35% B. . The desired compounds were verified by MAIDI-TOF, following chemical modification through sulftolysis or trypsin degradation.

## Reductive Aminated Pegylation of Acetyl Group on the Aromatic Ring Of The Phenylalanine

### a. Synthesis

Insulin (or an insulin analogue), mPEG20k-Hydrazide, and NaBH<sub>3</sub>CN in a molar ratio of 1:2:20 were dissolved in acetic acid buffer (pH of 4.1 to 4.4). The reaction solution was composed of 0.1 N NaCl, 0.2 N acetic acid and 0.1 N Na<sub>2</sub>CO<sub>3</sub>. Insulin or insulin analogue concentration was approximately 0.5 mg/ml. at room temperature for 24h. The reaction process was monitored by HPLC. The conversion of the reaction was approximately 50%. (calculated by HPLC)

### 10 b. Purification

The reaction mixture was diluted 2-5 fold and loaded to RP-HPLC. HPLC condition: C4 column; flow rate 10 ml/min; A buffer 10% ACN and 0.1% TFA in water; B buffer 0.1% TFA in ACN; A linear gradient B% from 0-40% (0-80 min); PEG-insulin, or the PEG-insulin analogue was collected at approximately 35%B. . The desired compounds were verified by MAIDI-TOF, following chemical modification through sulftolysis or trypsin degradation.

## EXAMPLE 3

### Insulin Receptor Binding Assay:

The affinity of each peptide for the insulin or IGF-1 receptor was measured in a competition binding assay utilizing scintillation proximity technology. Serial 3-fold dilutions of the peptides were made in Tris-Cl buffer (0.05 M Tris-HCl, pH 7.5, 0.15 M NaCl, 0.1% w/v bovine serum albumin) and mixed in 96 well plates (Corning Inc., Acton, MA) with 0.05 nM (3-[125I]-iodotyrosyl) A TyrA14 insulin or (3-[125I]-iodotyrosyl) IGF-1 (Amersham Biosciences, Piscataway, NJ). An aliquot of 1-6 micrograms of plasma membrane fragments prepared from cells over-expressing the human insulin or IGF-1 receptors were present in each well and 0.25 mg/well polyethylene imine-treated wheat germ agglutinin type A scintillation proximity assay beads (Amersham Biosciences, Piscataway, NJ) were added. After five minutes of shaking at 800 rpm the plate was incubated for 12h at room temperature and radioactivity was measured with MicroBeta1450 liquid scintillation counter (Perkin-Elmer, Wellesley, MA). Non-specifically bound (NSB) radioactivity was measured in the wells with a four-fold concentration excess of "cold" native ligand than the

highest concentration in test samples. Total bound radioactivity was detected in the wells with no competitor. Percent specific binding was calculated as following: % Specific Binding = (Bound-NSB / Total bound-NSB) x 100. IC50 values were determined by using Origin software (OriginLab, Northampton, MA).

5

#### EXAMPLE 4

##### Insulin Receptor Phosphorylation Assay:

To measure receptor phosphorylation of insulin or insulin analog, receptor transfected HEK293 cells were plated in 96 well tissue culture plates (Costar #3596, Cambridge, MA) and cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 100 IU/ml penicillin, 100 µg/ml streptomycin, 10 mM HEPES and 0.25% bovine growth serum (HyClone SH30541, Logan, UT) for 16-20 hrs at 37 °C, 5% CO<sub>2</sub> and 90% humidity. Serial dilutions of insulin or insulin analogs were prepared in DMEM supplemented with 0.5% bovine serum albumin (Roche Applied Science #100350, Indianapolis, IN) and added to the wells with adhered cells. After 15 min incubation at 37 °C in humidified atmosphere with 5% CO<sub>2</sub> the cells were fixed with 5% paraformaldehyde for 20 min at room temperature, washed twice with phosphate buffered saline pH 7.4 and blocked with 2% bovine serum albumin in PBS for 1 hr. The plate was then washed three times and filled with horseradish peroxidase-conjugated antibody against phosphotyrosine (Upstate biotechnology #16-105, Temecula, CA) reconstituted in PBS with 2% bovine serum albumin per manufacturer's recommendation. After 3 hrs incubation at room temperature the plate was washed 4 times and 0.1 ml of TMB single solution substrate (Invitrogen, #00-2023, Carlsbad, CA) was added to each well. Color development was stopped 5 min later by adding 0.05 ml 1 N HCl. Absorbance at 450 nm was measured on Titertek Multiscan MCC340 (ThermoFisher, Pittsburgh, PA). Absorbance vs. peptide concentration dose response curves were plotted and EC<sub>50</sub> values were determined by using Origin software (OriginLab, Northampton, MA).

#### 30 EXAMPLE 5

##### **Determination of rate of model dipeptide cleavage (in PBS)**

A specific hexapeptide (HSRGTF-NH<sub>2</sub>; SEQ ID NO: 59) was used as a model peptide upon which the rate of cleavage of dipeptide N-terminal extensions could be

studied. The dipeptide-extended model peptides were prepared Boc-protected sarcosine and lysine were successively added to the model peptide-bound resin to produce peptide A (Lys-Sar-HSRGTF-NH<sub>2</sub>; SEQ ID NO: 60). Peptide A was cleaved by HF and purified by preparative HPLC.

5           The rate of cleavage was determined for the respective propeptides. The concentrations of the propeptides and the model parent peptide were determined by their respective peak areas. The first order dissociation rate constants of the prodrugs were determined by plotting the logarithm of the concentration of the prodrug at various time intervals. The slope of this plot provides the rate constant 'k'. The half  
10       lives for cleavage of the various prodrugs were calculated by using the formula  $t_{1/2} = .693/k$ . The half life of the Lys-Sar extension to this model peptide HSRGTF-NH<sub>2</sub> (SEQ ID NO: 59) was determined to be 14.0h.

#### EXAMPLE 6

15       Rate of dipeptide cleavage half time in plasma as determined with an all d-isofom model peptide

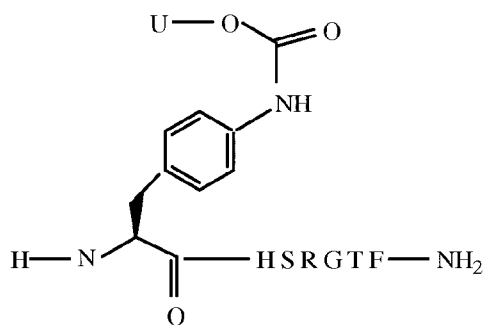
          An additional model hexapeptide (dHdTdRGdTdF-NH<sub>2</sub> SEQIDNO: 63) was used to determine the rate of dipeptide cleavage in plasma. The d-isomer of each amino acid was used to prevent enzymatic cleavage of the model peptide, with the  
20       exception of the prodrug extension. This model d-isomer hexapeptide was synthesized in an analogous fashion to the l-isomer. The sarcosine and lysine were successively added to the N-terminus as reported previously for peptide A to prepare peptide B (dLys-dSar-dHdTdRGdTdF-NH<sub>2</sub> SEQ ID NO: 64)

          The rate of cleavage was determined for the respective propeptides. The  
25       concentrations of the propeptides and the model parent peptide were determined by their respective peak areas. The first order dissociation rate constants of the prodrugs were determined by plotting the logarithm of the concentration of the prodrug at various time intervals. The slope of this plot provides the rate constant 'k'. The half life of the Lys-Sar extension to this model peptide dHdTdRGdTdF-NH<sub>2</sub> (SEQ ID NO:  
30       63) was determined to be 18.6h.

#### EXAMPLE 7

The rate of cleavage for additional dipeptides linked to the model hexapeptide (HSRGTF-NH<sub>2</sub>; SEQ ID NO: 59) were determined using the procedures described in Example 5. The results generated in these experiments are presented in Tables 2 and 3.

- 5 Table 2: Cleavage of the Dipeptides O-U that are linked to the side chain of an N-terminal para-amino-Phe from the Model Hexapeptide (HSRGTF-NH<sub>2</sub>; SEQ ID NO: 59) in PBS



10

Compounds	U (amino acid)	O (amino acid)	t <sub>1/2</sub>
1	F	P	58h
2	Hydroxyl-F	P	327h
3	d-F	P	20h
4	d-F	d-P	39h
5	G	P	72h
6	Hydroxyl-G	P	603h
7	L	P	62h
8	tert-L	P	200h
9	S	P	34h
10	P	P	97h
11	K	P	33h
12	dK	P	11h
13	E	P	85h
14	Sar	P	≈1000h
15	Aib	P	69min
16	Hydroxyl-Aib	P	33h
17	cyclohexane	P	6min
18	G	G	No cleavage
19	Hydroxyl-G	G	No cleavage
20	S	N-Methyl-Gly	4.3h
21	K	N-Methyl-Gly	5.2h
22	Aib	N-Methyl-Gly	7.1min
23	Hydroxyl-Aib	N-Methyl-Gly	1.0h

Table 3: Cleavage of the Dipeptides U-O linked to histidine (or histidine derivative) at position 1 (X) from the Model Hexapeptide (XSRGTF-NH<sub>2</sub>; SEQ ID NO: 59) in PBS

5

NH<sub>2</sub>-U-O-XSRGTF-NH<sub>2</sub>

Cmp	U (amino acid)	O (amino acid)	X (amino acid)	t <sub>1/2</sub>
1	F	P	H	No cleavage
2	Hydroxyl-F	P	H	No cleavage
3	G	P	H	No cleavage
4	Hydroxyl-G	P	H	No cleavage
5	A	P	H	No cleavage
6	C	P	H	No cleavage
7	S	P	H	No cleavage
8	P	P	H	No cleavage
9	K	P	H	No cleavage
10	E	P	H	No cleavage
11	Dehydro V	P	H	No cleavage
12	P	d-P	H	No cleavage
13	d-P	P	H	No cleavage
14	Aib	P	H	32h
15	Aib	d-P	H	20h
16	Aib	P	d-H	16h
17	Cyclohexyl-	P	H	5h
18	Cyclopropyl-	P	H	10h
19	N-Me-Aib	P	H	>500h
20	$\alpha$ , $\alpha$ -diethyl-Gly	P	H	46h
21	Hydroxyl-Aib	P	H	61
22	Aib	P	A	58
23	Aib	P	N-Methyl-His	30h
24	Aib	N-Methyl-Gly	H	49min
25	Aib	N-Hexyl-Gly	H	10min
26	Aib	Azetidine-2-carboxylic acid	H	>500h
27	G	N-Methyl-Gly	H	104h
28	Hydroxyl-G	N-Methyl-Gly	H	149h
29	G	N-Hexyl-Gly	H	70h
30	dK	N-Methyl-Gly	H	27h
31	dK	N-Methyl-Ala	H	14h
32	dK	N-Methyl-Phe	H	57h
33	K	N-Methyl-Gly	H	14h
34	F	N-Methyl-Gly	H	29h
35	S	N-Methyl-Gly	H	17h
36	P	N-Methyl-Gly	H	181h

## EXAMPLE 8

## Identification of an Insulin Analog with Structure Suitable for Prodrug Construction

Position 19 of the A chain is known to be an important site for insulin activity. Modification at this site to allow the attachment of a prodrug element is therefore  
5 desirable. Specific analogs of insulin at A19 have been synthesized and characterized for their activity at the insulin receptors. Two highly active structural analogs have been identified at A19, wherein comparable structural changes at a second active site aromatic residue (B24) were not successful in identification of similarly full activity insulin analogs.

10       Tables 4 and 5 illustrate the high structural conservation at position A19 for full activity at the insulin receptor (receptor binding determined using the assay described in Example 3). Table 4 demonstrates that only two insulin analogs with modifications at A19 have receptor binding activities similar to native insulin. For the 4-amino insulin analog, data from three separate experiments is provided. The  
15 column labeled "Activity (in test)" compares the percent binding of the insulin analog relative to native insulin for two separate experiments conducted simultaneously. The column labeled "Activity (0.60 nM)" is the relative percent binding of the insulin analog relative to the historical average value obtained for insulin binding using this assay. Under either analysis, two A19 insulin analogs ( 4-amino phenylalanine and 4-  
20 methoxy phenylalanine) demonstrate receptor binding approximately equivalent to native insulin. Fig. 3 represents a graph demonstrating the respective specific binding of native insulin and the A19 insulin analog to the insulin receptor. Table 5 presents data showing that the two A19 insulin analogs (4-amino and 4-methoxy) that demonstrate equivalent binding activities as native insulin also demonstrate  
25 equivalent activity at the insulin receptor (receptor activity determined using the assay described in Example 4).

Table 4: Insulin Receptor Binding Activity of A19 Insulin Analogs

Analogue	Insulin Receptor			
			% native ligand	% native ligand
	IC <sub>50</sub>	STDev	Activity (in test)	Activity (0.60 nM)
4-OH (native insulin)	0.64	0.15	100.0	100.0
4-COCH <sub>3</sub>	31.9	9.47	0.6	1.9
4-NH <sub>2</sub>	0.31	0.12	203.0	193.5
	0.83	0.15	103.0	72.3
	0.8	0.1	94.0	75.0
4-NO <sub>2</sub>	215.7	108.01	0.3	1.3
3,4,5-3F	123.29	31.10	0.5	0.5
4-OCH <sub>3</sub>	0.5	0.50	173.0	120.0
3-OCH <sub>3</sub>	4.74	1.09	28.0	12.7
	5.16	3.88	18.0	11.6
4-OH, 3,5-2Br	1807.17	849.72	0.0	0.0
4-OH, 3,5-2 NO <sub>2</sub>	2346.2	338.93	0.0	0.0

Table 5: Insulin Receptor Phosphorylation Activity of A19 Insulin Analogs

Analogue	Insulin Receptor		
	EC <sub>50</sub>	STDev	% native ligand Activity (in test)
4-OH (native insulin)	1.22	0.4	100.0
4-NH <sub>2</sub>	0.31	0.14	393.5
4-OCH <sub>3</sub>	0.94	0.34	129.8

5

## EXAMPLE 9

## Insulin like Growth Factor (IGF) Analog

Typical purification schemes for isolating the insulin A-chain use an  
 10 NH<sub>4</sub>HCO<sub>3</sub> buffer (pH=7.8). Under these conditions the dipeptide prodrug element is  
 rapidly cleaved from the A-chain. To simplify purification of prodrugs to investigate



their activities at the insulin receptor, applicants conducted such studies using an IGF analog that demonstrates similar activity at the insulin receptor as native insulin. More particularly, the IGF analog (IGF1 (Y<sup>B16</sup>L<sup>B17</sup>)) comprises the native IGF A and B chain (SEQ ID NO: 61 and SEQ ID NO: 62, respectively), wherein the native

5 glutamine and phenylalanine at positions 15 and 16 of the native IGF B-chain (corresponding to positions 16 and 17 of native insulin B-chain, respectively) have been replaced with tyrosine and leucine residues, respectively. As shown in Fig. 4 and Table 6 below the binding activities of IGF1 (Y<sup>B16</sup>L<sup>B17</sup>) demonstrate the compound is a highly potent insulin analog.

10

Table 6

	Insulin Standard		IGF1(Y <sup>B16</sup> L <sup>B17</sup> )	
	AVER.	STDEV	AVER.	STDEV
IC <sub>50</sub> (nM)	1.32	0.19	0.51	0.18
% of Insulin Activity	100		262	

## EXAMPLE 10

## IGF Prodrug Derivatives

15 Based on the activity of the A19 insulin analog (see Example 5), a similar modification was made to the IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>) analog and its ability to bind and stimulate insulin receptor activity was investigated. Fig. 5 provides the general synthetic scheme for preparing IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>) wherein the native tyrosine is replace with a 4-amino phenylalanine [IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>)(p-NH<sub>2</sub>-F)<sup>A19</sup>amide] as

20 well as the preparation of its dipeptide extended derivative [IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>)<sup>A19</sup>-AiBAla amide], wherein a dipeptide comprising AiB and Ala are linked to the peptide through an amide linkage to the A19 4-amino phenylalanine. As shown in Fig. 6 and Table 7, the IGF analog, IGF1 (Y<sup>B16</sup>L<sup>B17</sup>) A(p-NH<sub>2</sub>-F)<sup>19</sup> specifically binds to the insulin receptor wherein the dipeptide extended derivative of that analog fails to

25 specifically bind the insulin receptor. Note the dipeptide extension lacks the proper structure to allow for spontaneous cleavage of the dipeptide (absence of an N-alkylated amino acid at the second position of the dipeptide) and therefore there is no restoration of insulin receptor binding.

Table 7

	Insulin Standard		IGF1(Y <sup>B16</sup> L <sup>B17</sup> ) (p-NH <sub>2</sub> -F) <sup>A19</sup> amide		IGF1(Y <sup>B16</sup> L <sup>B17</sup> ) (AiBAIa) <sup>A19</sup> amide	
	AVER.	STDEV	AVER.	STDEV.	AVER.	STDEV
IC <sub>50</sub> (nM)	0.24	0.07	1.08	.075	No Activity	
% of Insulin Activity	100		22			

- 5 A further prodrug derivative of an IGF<sup>B16B17</sup> derivative peptide was prepared wherein the dipeptide prodrug element (alanine-proline) was linked via an amide bond to the amino terminus of the A chain (IGF1(Y<sup>B16</sup>L<sup>B17</sup>)(AlaPro)<sup>A-1,0</sup>). As shown in Table 8, the IGF1(Y<sup>B16</sup>L<sup>B17</sup>)(AlaPro)<sup>A-1,0</sup> has substantially reduced affinity for the insulin receptor. Note, based on the data of Table 3, the dipeptide prodrug element
- 10 lacks the proper structure to allow for spontaneous cleavage of the dipeptide prodrug element, and therefore the detected insulin receptor binding is not the result of cleavage of the prodrug element.

Table 8

	Insulin Standard		IGF1(YL) <sup>B16,17</sup> (AlaPro) <sup>A-1,0</sup>	
	AVER.	STDEV	AVER.	STDEV.
IC <sub>50</sub> (nM)	0.72	0.09	1.93	.96
% of Insulin Activity	100		37.12	

15

## EXAMPLE 11

## Additional IGF Insulin Analogs.

- Further modifications of the IGF1 (Y<sup>B16</sup>L<sup>B17</sup>)(YL)B<sup>16</sup>B<sup>17</sup> peptide sequence
- 20 reveal additional IGF insulin analogs that vary in their potency at the insulin and IGF-1 receptor. Binding data is presented in Table 9 for each of these analogs (using the

assay of Example 3), wherein the position of the modification is designated based on the corresponding position in the native insulin peptide (DPI = des B26-30). For example, a reference herein to "position B28" absent any further elaboration would mean the corresponding position B27 of the B chain of an insulin analog in which the  
5 first amino acid of SEQ ID NO: 2 has been deleted. Thus a generic reference to "B(Y16)" refers to a substitution of a tyrosine residue at position 15 of the B chain of the native IGF-1 sequence. Data regarding the relative receptor binding of insulin and IGF analogs is provided in Table 9, and data regarding IGF analog stimulated phosphorylation (using the assay of Example 4) is provided in Table 10.

10

Table 9 Receptor Binding Affinity of Insulin and IGF Analogues

Analogue	Insulin Receptor					IGF-1 Receptor					Ratio
	nM		% insulin		% native insulin activity (0.6 nM)	% IGF-1					
	IC <sub>50</sub>	STDev	Date	(in test)	IC <sub>50</sub>	STDev	Date	(in test)	activity (0.55 nM)		
IGF-1 A:B	10.41	1.65	9/4/2007	5.8	5.8						
IGF-1 A:B(E10Y16L17)	0.66	0.36	5/22/2007	58.7	90.9	7.85	1.98	6/4/2007	6.8	7.0	11.9
	0.51	0.18	5/29/2007	98.8	117.6	12.19	2.17	9/18/2007	5.0	4.5	
IGF-1 A:B(E10 Y16L17)-E31E3 2B-COOH	1.22	0.30	3/20/2008	36.5	50.0	17.50	2.25	4/4/2007	3.0	3.1	14.3
IGF-1 A:B(D10Y16L17) DPI A-COOH	0.26	0.02	11/9/2007	301.0	231.0	6.79	1.50	4/4/2008	7.7	8.1	
	0.2	0.02	12/4/2007	380.1	300.0						
	0.42	0.06	6/5/2008	174.1	144.1						
IGF-1 A:B (E10Y16L17) DPI	0.38	0.08	8/10/2007	51.1	157.9	22.89	5.26	9/18/2007	3.3	2.4	60.2
IGF-1 A:B (H5D10Y16L17) DPI	0.16	0.07	11/9/2007	479.0		4.66	0.77	4/4/2008	11.2	11.8	29.1
IGF-1 A:B (H5D10Y16L17) (S=O)DPI	0.25	0.04	11/9/2007	316.0							
IGF-1 A (H8 A9 N21): B(H5D10Y16L17) DPI A-COOH	0.05	0.01	12/4/2007	1576.7		4.03	0.50	4/4/2008	12.9	13.6	80.6
	0.09	0.02	12/14/2007	1667.0							
IGF-1 A (H8 A9 N21): B(H5D10Y16L17 A22) DPI A-COOH	0.12	0.02	12/14/2007	1171.4		22.83	3.53	4/4/2008	2.3	2.4	190.3

Analogue	Insulin Receptor					IGF-1 Receptor					Ratio
	nM		% insulin		% native insulin activity (0.6 nM)	% IGF-1		% native IGF-1 activity (0.55 nM)			
	IC <sub>50</sub>	STDev	Date	(in test)	IC <sub>50</sub>	STDev	Date	(in test)			
IGF-1 A (H8 A9 N21) :B(H5D10Y16L17A22) (S=O) DPI A-COOH	0.36	0.10	12/14/2007	400.7							
IGF-1 A: IGF-1 B(1-8)-In (9-17)- IGF-1 B(18-30)	1.59	0.62	5/22/2007	19.1	37.7	131.30	58.05	6/4/2007	0.3	0.4	82.6
IGF-1 A: In (1-17)- IGF-1 B (18- 30)	2.77	1.19	5/22/2007	14.0	21.7	62.50	30.28	6/4/2007	0.9	0.9	22.6
	2.67	0.67	5/18/2007	11.3	22.5						
	2.48	1.35	5/29/2007	20.1	24.2						
IGF-1 A: In B(1-5)- IGF-1 B(YL)(6- 30)	0.31	0.19	8/10/2007	62.4	193.5	27.54	6.57	9/25/2007	3.6	2	88.8
IGF-2 native						13.33	1.85	9/25/2007	7.5	4.5	
IGF-2 AB											
IGF-2 AB(YL)	6.81	3.81	10/10/2007	8.4	8.8						
In A: IGF-1 B(YL)	82.62	31.75	9/4/2007	0.9	0.7						
	107.24	65.38	9/4/2007	0.7	0.6						
In A- IGF-2 D : In B- IGF-2 C	0.53	0.11	9/4/2007	141.0	113.0	1.59	0.34	9/18/2007	47.6	34.6	
	0.37	0.05	10/13/2007	179.1	162.2	14.69	3.02	9/25/2007	6.8	3.7	39.7
**All C terminals are amides (DPI) unless specified otherwise											

Table 10: Total Phosphorylation by IGF-1 &amp; IGF-2 Analogues

Analogue	Insulin Receptor			IGF-1 Receptor			Selective Ratio
	EC50	STDev	Date	% Insulin	EC50	STDev	
Insulin	1.26	0.098	12/14/2007		114.88	46.66	
	1.43	0.72	4/1/2008		86.02	29.35	90.89
	1.12	0.11	3/31/2008				
	1.53	0.13	4/11/2008				
	2.70	0.71	4/16/2008				
	1.22	0.40	5/20/2008				
IGF-1	54.39	21.102	12/14/2007	2.3	0.87	0.16	100
					0.49	0.13	0.02
					0.97	0.48	
IGF-1 AB							
IGF-1 A: B(E10Y16L17)	2.57	0.59	3/31/2008	49.2	7.42	5.59	13
IGF-1 A:B(E10 Y16L17)-E31E32 B-COOH	7.00	2.82	3/31/2008	18.1			
	8.52	4.34	4/16/2008	31.7			
IGF-1 AB(D10Y16L17) DPI A-COOH	0.08	0.006	12/14/2007	1575	0.78	0.17	111.538
	4.38	2.98	4/16/2008	??			9.75
IGF-1 AB (E10Y16L17) DPI							
IGF-1 AB (H5D10Y16L17) DPI					12.22	5.46	7.1
IGF-1 AB (H5D10Y16L17) (S=O)DPI							

Analogue	Insulin Receptor				IGF-1 Receptor				Selective Ratio
	EC50:	STDev	Date	% Insulin	EC50:	STDev	Date	% IGF	
IGF-1 A (H8 A9 N21) B(H5D10Y16L17) DPI A-COOH	0.15	0.054	12/14/2007	840	0.43	0.44	1/23/2008	181.395	2.81
	0.25	0.2	4/16/2008	1080					
IGF-1 A (H8 A9 N21) B(H5D10Y16L17A22) DPI A-COOH	0.35	0.064	12/14/2007	360	11.26	2.55	1/23/2008	7.7	32.54
	0.44	0.17	4/16/2008	614					
IGF-1 A (H8 A9 N21) B(H5D10Y16L17A22) (S=O) DPI A-COOH	0.72	0.098	12/14/2007						
*All C-terminals are amides unless specified otherwise.									

## EXAMPLE 12

## Dipeptide Cleavage from Prodrug forms of IGFB16B17 derivative peptides

- 5           The cleavage of an (pNH<sub>2</sub>-Phe) amide linked dipeptide AibPro from various IGF-1 peptides was measured to determine the impact of the peptide sequence or heteroduplex on the dipeptide cleavage. Results for the tested peptides is shown in Table 12 and the data reveals that the IGF1-A chain alone represents a good model for the study of prodrug half life for IGF1 B:A (YL)<sup>B16,17</sup> peptides.

10

Table 12

Parent Peptide	Half Life (hr)
IGF1A(Ala) <sup>6,11,20</sup> (pNH <sub>2</sub> -Phe) <sup>A19</sup>	2.2
IGF1A(Acm) <sup>6,11,20</sup> (pNH <sub>2</sub> -Phe) <sup>A19</sup>	1.8
IGF1 B:A(S-S) <sup>A7,B7</sup> (Acm) <sup>A6,11,20,B19</sup> (pNH <sub>2</sub> -Phe) <sup>A19</sup>	1.8
IGF1 B:A(pNH <sub>2</sub> -Phe) <sup>A19</sup>	1.6

- Comparison of prodrug derivatives of the IGF A-chain relative to the disulfide bound A chain and B chain construct (IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>)) revealed the two compounds had similar half lives for the prodrug form. Accordingly, the IGF1A chain alone was determined to be a good model for the study of pro-drug half life on IGF1 B:A (Y<sup>B16</sup>L<sup>B17</sup>). Note the AibAla derivative does not cleave and thus is not a prodrug, but serves to show the modification can inactivate the insulin analog IGF1 A:B(Y<sup>B16</sup>L<sup>B17</sup>)(p-NH<sub>2</sub>-F)<sup>A19</sup>amide. For simplicity, prodrug half lives were determined using only the IGF1 A chain in the absence of the B chain. The half lives of each propeptide was determined as described in Example 5. The data is presented in Table 13:



Table 13: Dipeptide half life on IGF1 dipeptide extended (p-NH<sub>2</sub>-F)<sup>A19</sup> amide

Dipeptide		Half Life (hr)
AiB	Pro	2.2
AiBOH	Pro	165.0
AiB	dPro	1.9
AiBOH	Sar	2.3
dK(acetyl)	Sar	16.3
K	Sar	21.8
K(acetyl)	N-methyl Ala	23.6
dK(acetyl)	N-methyl Ala	35.3

The data shows that by altering the substituents on the dipeptide prodrug element  
 5 that the half life of prodrug can be varied from 2 hrs to >100 hrs.

Additional prodrug derivative peptides were prepared using an IGF1-A(pNH<sub>2</sub>-  
 F)<sup>19</sup> base peptide and alting the amino acid composition of the dipeptide prodrug element  
 linked through the 4-amino phenylalanine at position A19. Dipeptide half lives were  
 measured for different constructs both in PBS and in 20% plasma/PBS (i.e. in the  
 10 presence of serum enzymes. The results are provided in Table 14. The results indicate  
 that three of the four peptides tested were not impacted by serum enzymes.

Table 14: Dipeptide half life on IGF1-A(pNH<sub>2</sub>-F)<sup>19</sup>

		Half Life (hr)	
		PBS	20% Plasma/PBS
AiB	Pro	2.2	2.1
AiB	dPro	2.1	2.2
AiBOH	Sar	2.3	
dK	N-isobutyl Gly	4.4	4.1
dK	N-hexyl Gly	10.6	
dK(acetyl)	Sar	17.2	
K	Sar	21.8	5.9
K(acetyl)	N-methyl Ala	23.6	
dK(acetyl)	N-methyl Ala	35.3	
AiBOH	Pro	165.0	
K(acetyl)	Azetidine-2-carboxylic acid	Not cleavable	
dK(acetyl)	Azetidine-2-carboxylic acid	Not cleavable	

## EXAMPLE 13

Receptor Binding of IGF<sup>B16B17</sup> Derivative Peptides over Time

- 5 Prodrug formulations of IGF<sup>B16B17</sup> Derivative Peptides were prepared and their degradation over time was measured using the insulin receptor binding assay of Example 3. Peptides used in the assay were prepared as follows:

## Dipeptide-IGF1A analogs

- 10 If not specified, Boc-chemistry was applied in the synthesis of designed peptide analogs. Selected dipeptide H<sub>2</sub>N-AA1-AA2-COOH was added to (pNH<sub>2</sub>-Phe)<sup>19</sup> on IGF1A (Ala)<sup>6,7,11,20</sup>. The IGF-1 A chain C-terminal tripeptide Boc(Fmoc-pNH-Phe)-Ala-Ala was synthesized on MBHA resin. After removal of Fmoc by the treatment with 20% piperidine/DMF at room temperature for 30 minutes, Fmoc-AA2 was coupled to the p-
- 15 amino benzyl side chain at A19 by using a threefold excess of amino acid, PyBop, DIEA and catalytic amount of pyridine. The Boc-synthesis of the remaining IGF-1 A chain (Ala)<sup>6,7,11,20</sup> sequence was completed using an Applied Biosystems 430A Peptide

Synthesizer, yielding IGF-1 A chain  $(\text{Boc})^0(\text{Ala})^{6,7,11,20}(\text{Fmoc-AA2-pNH-Phe})^{19}\text{-MBHA}$ . After the Fmoc group was removed from the N-terminus of AA2, Boc-AA1 was then coupled to the amine using threefold excess of amino acid, DEPBT and DIEA. Removal of the two Boc groups remaining on the A chain by TFA was followed by HF cleavage, yielding IGF-1 A-chain  $(\text{Ala})^{6,7,11,20}(\text{H}_2\text{N-AA1-AA2-pNH-Phe})^{19}$  amide. In the case of AA1 being d-lysine, acetylation on the  $\epsilon$ -amine was performed prior to Boc removal. Dipeptide-IGF-1 A chain analogs were purified by semi-preparative RP-HPLC and characterized by analytical RP-HPLC and MALDI mass spectrometry.

#### Dipeptide-IGF-1 (YL) analogs

A selected dipeptide  $\text{H}_2\text{N-AA1-AA2-COOH}$  was added to  $(\text{pNH}_2\text{-Phe})^{19}$  on IGF-1 A chain  $(\text{Acm})^{6,11,20}$  as described immediately above except PAM resin was used for the synthesis of IGF-1 A chain to yield a C terminal acid upon HF-cleavage. IGF-1 B chain  $(\text{YL})^{16,17}(\text{Acm})^{19}$  was synthesized on MBHA resin to yield a C terminal amide. The free thiol on Cys<sup>B7</sup> was modified by Npys through reaction with DTNP at a 1:1 molar ratio in 100% DMSO. Purified dipeptide-IGF-1 A chain and IGF-1 B chain  $(\text{YL})^{16,17}$  derivatives were assembled using the "1+2" two step chain combination strategy illustrated in Scheme 1. Intermediate and final purifications were performed on semi-preparative RP-HPLC and characterized by analytical RP-HPLC and MALDI mass spectrometry.

The IGF<sup>B16B17</sup> derivative peptide prodrugs were incubated in PBS, pH 7.4 at 37°C and at predetermined time intervals an aliquot was taken and further degradation was quenched with 0.1% TFA and the aliquot was subjected to analytical HPLC analysis. Peaks a and b, representing the prodrug and active forms of the IGF<sup>B16B17</sup> derivative peptide were identified with LC-MS and quantified by integration of peak area on HPLC. Figs 8A-8C show the output of an HPLC analysis of the degradation of the IGF<sup>B16B17</sup> derivative peptide prodrug: IGF1A $(\text{Ala})^{6,7,11,20}(\text{Aib-Pro-pNH-F})^{19}$ . Aliquots were taken at 20 minutes (Fig. 8A), 81 minutes (Fig 8B) and 120 minutes (Fig. 8C) after beginning the incubation of the prodrug in PBS. The data indicate the spontaneous, non-enzymatic conversion of IGF1A $(\text{Ala})^{6,7,11,20}(\text{Aib-Pro-pNH-F})^{19}$  amide to IGF1A $(\text{Ala})^{6,7,11,20}(\text{pNH}_2\text{-F})^1$  amide over time.

The degradation of the prodrug forms of IGF<sup>B16B17</sup> derivative peptides to there active form was also measured based on the compounds ability to bind to the insulin receptor as measured using the *in vitro* assay of Example 3. Fig. 9A & 9B are graphs depicting the *in vitro* activity of the prodrug Aib,dPro-IGF1YL (dipeptide linked throught the A19 4-aminoPhe). Fig 9A is a graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at 4°C) and the A19 IGF prodrug analog (Aib,dPro-IGF1YL) over time (0 hours, 2.5 hours and 10.6 hours) incubated in PBS. Fig 9B is a graph comparing relative insulin receptor binding of native insulin (measured at 1.5 hour at 4°C) and the A19 IGF prodrug analog (Aib,dPro-IGF1YL) over time (0 hours, 1.5 hours and 24.8 hours) incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered form the A19 IGF prodrug analog sample as the prodrug form is converted to the active IGF1YL peptide. The activity of the IGF<sup>B16B17</sup> derivative peptides was measured relative to insulin receptor binding, and since the underlying IGF<sup>B16B17</sup> derivative peptides have more activity than native insulin, activity of greater than 100% relative to insulin is possible.

Fig. 10A & 10B are graphs depicting the *in vitro* activity of the prodrug dK,(N-isobutylG)-IGF1YL (dipeptide linked throught the A19 4-aminoPhe). Fig 10A is a graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK,(N-isobutylG) over time (0 hours, 5 hours and 52 hours) incubated in PBS. Fig 10B is a graph comparing relative insulin receptor binding of native insulin (measured at 1.5 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK,(N-isobutylG) over time (0 hours, 3.6 hours and 24.8 hours) incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered form the A19 IGF prodrug analog sample as the prodrug form is converted to the active IGF1YL peptide.

Fig. 11A & 11B are graphs depicting the *in vitro* activity of the prodrug dK(e-acetyl),Sar)-IGF1YL (dipeptide linked throught the A19 4-aminoPhe). Fig 11A is a graph comparing relative insulin receptor binding of native insulin (measured at 1 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK(e-acetyl),Sar) over time (0 hours, 7.2 hours and 91.6 hours) incubated in PBS. Fig 11B is a graph comparing relative

- insulin receptor binding of native insulin (measured at 1.5 hour at 4°C) and the A19 IGF prodrug analog (IGF1YL: dK(e-acetyl),Sar) over time (0 hours, 9 hours and 95 hours) incubated in 20% plasma/PBS. As indicated by the data presented in the graph, increased activity is recovered from the A19 IGF prodrug analog sample as the prodrug
- 5 form is converted to the active IGF1YL peptide

**Claims:**

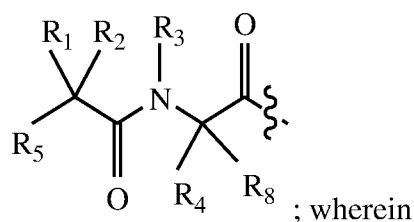
1. An insulin analog comprising an A chain, a B chain and a dipeptide element, wherein

said A chain has the general formula Z-peptide A, wherein peptide A comprises the sequence of GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3), or an analog thereof comprising a sequence that differs from SEQ ID NO: 3 by 1 to 3 amino acid modifications, selected from positions A5, A8, A9, A10, A14, A15, A17, and A18; and

said B chain has the general formula J-X<sub>14</sub>-peptide B, wherein peptide B comprises the sequence of X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub> (SEQ ID NO: 14), or an analog thereof comprising a sequence that differs from SEQ ID NO: 14 by 1 to 3 amino acid modifications, selected from positions B1, B2, B3, B4, B5, B13, B14, B17, B20, B22, and B23;

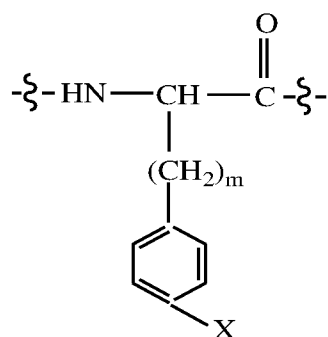
said B chain being linked to said A chain through intermolecular disulfide linkages, wherein

Z and J are independently H or a dipeptide element of the structure of Formula I:



X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the structure of:



wherein m is an integer selected from the group consisting of 0-3, and X is selected from the group consisting of OH, OCH<sub>3</sub>, NH<sub>2</sub> and NHR<sub>10</sub>, wherein R<sub>10</sub> is a dipeptide element comprising the general structure of Formula I;

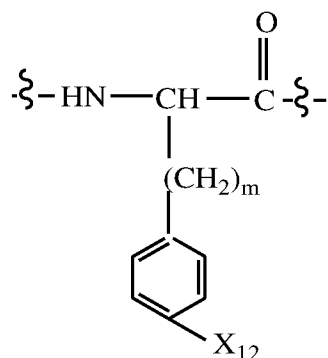
X<sub>3</sub> is selected from the group consisting of asparagine, glycine and alanine, threonine and serine;

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

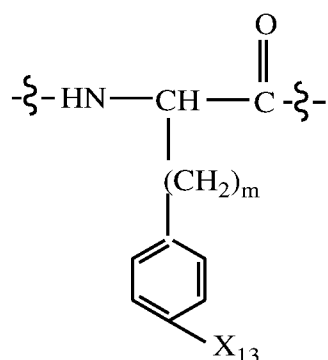
X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

X<sub>7</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X<sub>12</sub> is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>11</sub> and OCH<sub>3</sub>, wherein R<sub>11</sub> is a dipeptide element comprising the structure of Formula I;

X<sub>8</sub> is an amino acid of the general structure



wherein m is an integer selected from 0-3 and X<sub>13</sub> is selected from the group consisting of H, OH, NH<sub>2</sub>, NHR<sub>12</sub> and OCH<sub>3</sub>, wherein R<sub>12</sub> is a dipeptide element of the structure of Formula I;

X<sub>14</sub> is selected from the group consisting of a bond, X<sub>9</sub>VNQ (SEQ ID NO: 21), VNQ, NQ and Q, wherein X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine;

R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W<sub>1</sub>)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W<sub>1</sub> is a heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>1</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen and OH; and

R<sub>13</sub> is COOH or CONH<sub>2</sub>, with the proviso that one or more of X, X<sub>12</sub>, X<sub>13</sub>, J and Z is a dipeptide element of the structure of Formula I and at least one of R<sub>1</sub> and R<sub>2</sub> is other than H;

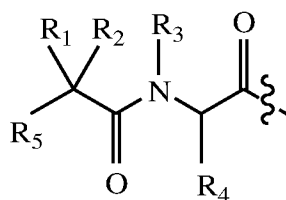
said dipeptide element being linked to said A or B chain through an amide bond wherein the chemical cleavage half-life (t<sub>1/2</sub>) of said dipeptide from said A or B chain is at least about 1 hour to about 1 week in PBS under physiological conditions.

2. The insulin analog of claim 1 wherein

Z and X<sub>13</sub> are each H;

X and X<sub>12</sub> are each OH; and

J is a dipeptide element comprising the general structure:



, with the further proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, R<sub>1</sub> and R<sub>2</sub> are both other than H; optionally wherein

i) R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl;

R<sub>2</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>2</sub> and R<sub>6</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring;

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl;



R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of hydrogen and OH; or

5 ii) R<sub>2</sub> comprises (C<sub>4</sub> alkyl)NH<sub>2</sub> that is acylated at the amino group of R<sub>2</sub> with an acyl group of sufficient size to bind serum albumin;

R<sub>1</sub> and R<sub>4</sub> are each hydrogen;

R<sub>3</sub> is CH<sub>3</sub>; and

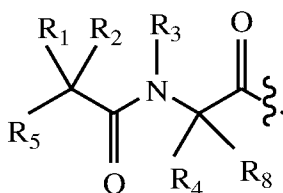
R<sub>5</sub> is NH<sub>2</sub>.

10 3. The insulin analog of claim 1 or claim 2 wherein

X<sub>14</sub>-peptide B comprises the sequence X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTPKT (SEQ ID NO: 15) or X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTKPT (SEQ ID NO: 16).

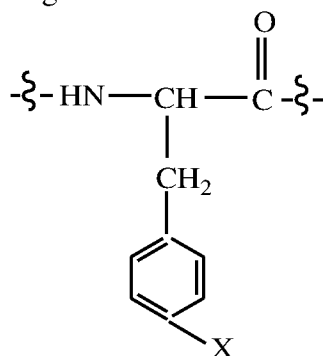
15 4. The insulin analog of claim 1 wherein said A chain comprises a sequence of Z-GIVEQCCX<sub>1</sub>SICSLYQLENX<sub>2</sub>CX<sub>3</sub> (SEQ ID NO: 3) and said B chain sequence comprises a sequence of X<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFF (SEQ ID NO: 4) wherein

Z is H or a dipeptide element comprising the general structure:

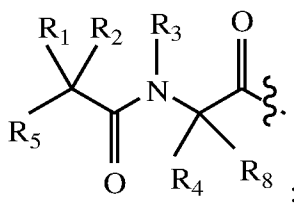


20 X<sub>1</sub> is selected from the group consisting of threonine and histidine;

X<sub>2</sub> is an amino acid of the general structure



wherein X is selected from the group consisting of OH, NH<sub>2</sub>, NHR<sub>10</sub> and OCH<sub>3</sub>, wherein R<sub>10</sub> is a dipeptide element comprising the general structure:



X<sub>3</sub> is selected from the group consisting of asparagine and glycine;

5 X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid;

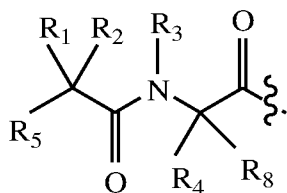
10 R<sub>1</sub>, R<sub>2</sub>, R<sub>4</sub> and R<sub>8</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>18</sub> alkyl, C<sub>2</sub>-C<sub>18</sub> alkenyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl), and C<sub>1</sub>-C<sub>12</sub> alkyl(W)C<sub>1</sub>-C<sub>12</sub> alkyl, wherein W is a  
15 heteroatom selected from the group consisting of N, S and O, or R<sub>1</sub> and R<sub>2</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>12</sub> cycloalkyl or aryl; or R<sub>4</sub> and R<sub>8</sub> together with the atoms to which they are attached form a C<sub>3</sub>-C<sub>6</sub> cycloalkyl;

R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>18</sub> alkyl, (C<sub>1</sub>-C<sub>18</sub> alkyl)OH, (C<sub>1</sub>-C<sub>18</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>18</sub> alkyl)SH, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub>)cycloalkyl,  
20 (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>2</sub>-C<sub>5</sub> heterocyclic), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and (C<sub>1</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>9</sub> heteroaryl) or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring;

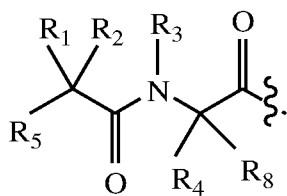
R<sub>5</sub> is NHR<sub>6</sub> or OH;

25 R<sub>6</sub> is H, C<sub>1</sub>-C<sub>8</sub> alkyl or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 4, 5 or 6 member heterocyclic ring; and

R<sub>7</sub> is selected from the group consisting of H and OH, with the proviso that Z is H when X is NHR<sub>10</sub>, and Z is



, when X is OH, NH<sub>2</sub>, or OCH<sub>3</sub> and when Z is



and R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, then R<sub>1</sub> and R<sub>2</sub> are each other than H.

5. The insulin analog of claim 4 wherein  
 5 the B chain sequence comprises the sequence  
 X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTPKT (SEQ ID NO: 12) or  
 X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTKPT (SEQ ID NO: 13) wherein

a. X<sub>4</sub> is selected from the group consisting of histidine and threonine;  
 X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;  
 10 X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid,  
 glutamic acid, homocysteic acid and cysteic acid; and

X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine; or

b. X<sub>4</sub> is histidine;  
 15 X<sub>5</sub> is serine; and  
 X<sub>6</sub> is histidine; and  
 X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine; optionally wherein

Z is H;  
 20 X is NHR<sub>10</sub>;  
 R<sub>1</sub> is selected from the group consisting of H and  
 C<sub>1</sub>-C<sub>4</sub> alkyl;

R<sub>2</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl),  
 25 (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>2</sub> and R<sub>6</sub> together with the  
 atoms to which they are attached form a 5 member heterocyclic ring;

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl;

R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl or R<sub>3</sub> and  
 R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring.

6. The insulin analog of claim 85 wherein  
R<sub>3</sub> is CH<sub>3</sub>;  
R<sub>4</sub> is H; and  
R<sub>5</sub> is NHR<sub>6</sub>.
- 5 7. The insulin analog of claim 5 wherein  
a. Z is H;  
X is NHR<sub>10</sub>;  
R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl; and  
R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl, or R<sub>3</sub> and  
10 R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic  
ring; or  
b. X is OH;  
R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl; and  
R<sub>4</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>4</sub> alkyl,  
15 (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH and (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub> aryl)R<sub>7</sub>, or  
R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member  
heterocyclic ring.
8. The insulin analog of claim 7 wherein the B chain sequence comprises  
the sequence X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTX<sub>10</sub> X<sub>11</sub> (SEQ ID NO: 20)  
20 wherein  
Z is H;  
X is NHR<sub>10</sub>;  
X<sub>4</sub> is selected from the group consisting of histidine and threonine;  
X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;  
25 X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid,  
homocysteic acid and cysteic acid;  
X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-  
phenylalanine;  
X<sub>10</sub> is aspartate-lysine dipeptide, a lysine-proline dipeptide, or a proline-lysine  
30 dipeptide;

X<sub>11</sub> is threonine, alanine, or a threonine-arginine-arginine tripeptide

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl; and

R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl, or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring.

- 5            9.            The insulin analog of any one of claims 1-4 further comprising
- a.            a hydrophilic moiety, optionally, polyethylene glycol, linked to the insulin analog; or
- b.            when the B chain sequence comprises the sequence of SEQ ID NO: 8 or SEQ ID NO: 9 a hydrophilic moiety, optionally, polyethylene glycol, is linked to the
- 10          insulin analog at a position selected from the group consisting of the N-terminal amino acid of the B chain, the lysine at position 29 of SEQ ID NO: 8, and the lysine at position 28 of SEQ ID NO: 9; or
- c.            the B chain sequence comprises the sequence of SEQ ID NO: 8 or SEQ ID NO: 9.

- 15           10.           The insulin analog of claim 1, 2 or 4 wherein R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>8</sub> alkyl; and

R<sub>2</sub> and R<sub>4</sub> are independently selected from the group consisting of H, C<sub>1</sub>-C<sub>8</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>2</sub>-C<sub>3</sub> alkyl)SCH<sub>3</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)CONH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)COOH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>1</sub>-C<sub>4</sub> alkyl)NHC(NH<sub>2</sub><sup>+</sup>)

20          NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl)

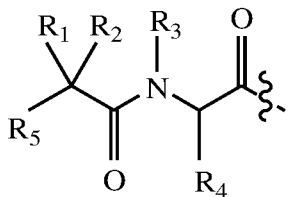
R<sub>3</sub> is selected from the group consisting of C<sub>1</sub>-C<sub>8</sub> alkyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>3</sub>-C<sub>6</sub>)cycloalkyl or R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic

25          ring;

R<sub>5</sub> is NHR<sub>6</sub> or OH;

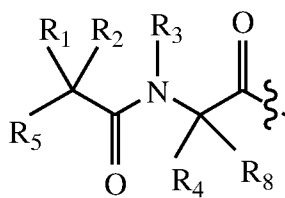
R<sub>6</sub> is H, or R<sub>6</sub> and R<sub>2</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring; and

$R_7$  is selected from the group consisting of H and OH, with the proviso that one and only one of  $R_{10}$ ,  $R_{11}$ ,  $R_{12}$ , J and Z comprises a dipeptide element of the general structure:



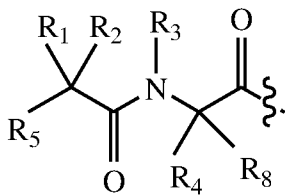
5 11. The insulin analog of claim 10 wherein

a. Z comprises a dipeptide element of the general structure:



, and said dipeptide element is acylated or  
pegylated;

b. J comprises a dipeptide element of the general structure:

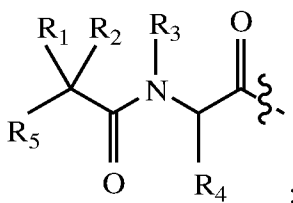


10 , and said dipeptide element is acylated or  
pegylated;

c. Z, J and  $X_{13}$  are each H;

X is OH;

$X_{12}$  is  $NHR_{11}$ , wherein  $R_{11}$  is a dipeptide element comprising the general structure:

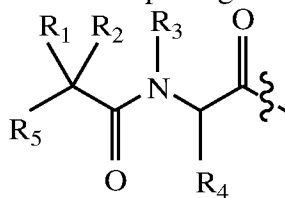


15

d. Z and  $X_{13}$  are each H;

X and  $X_{12}$  are each OH;

J is a dipeptide element comprising the general structure:

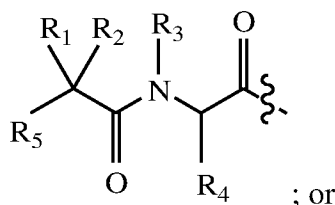


, with the further proviso that when R<sub>4</sub> and R<sub>3</sub> together with the atoms to which they are attached form a 5 or 6 member heterocyclic ring, R<sub>1</sub> and R<sub>2</sub> are both other than H;

5 e. Z and J are both H;

X and X<sub>12</sub> are each OH;

X<sub>13</sub> is NHR<sub>12</sub>, wherein R<sub>12</sub> is a dipeptide element comprising the general structure:



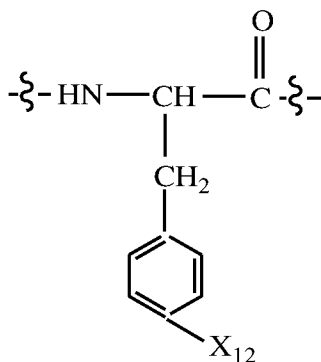
f. Z and J are each H;

10 X is OH; and

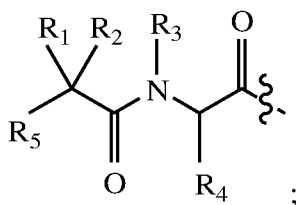
the B chain sequence comprises the sequence X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTPKT (SEQ ID NO: 15) or X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTKPT (SEQ ID NO: 16) wherein

15 X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine, and

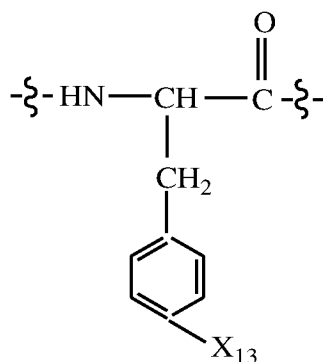
X<sub>7</sub> is an amino acid of the general structure



wherein  $X_{12}$  is selected from the group consisting of OH,  $NH_2$ , and  $NHR_{11}$ , wherein  $R_{11}$  is a dipeptide element comprising the general structure:

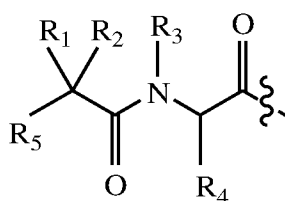


$X_8$  is an amino acid of the general structure



5

wherein  $X_{13}$  is selected from the group consisting of H,  $NH_2$ ,  $OCH_3$  and  $NHR_{12}$ , wherein  $R_{12}$  is a dipeptide element comprising the general structure:



with the proviso that  $X_{12}$  is  $NHR_{11}$  when  $X_{13}$  is H,  $NH_2$  or  $OCH_3$  and  $X_{12}$  is OH,  $NH_2$  or  $OCH_3$  when  $X_{13}$  is  $NHR_{12}$ ; optionally wherein

10

- i)  $X_1$  is threonine;  
 $X_4$  is histidine;  
 $X_5$  is serine; and

$X_6$  is histidine, or

15

- ii)  $X_1$  is threonine;  
 $X_3$  is asparagine or glycine  
 $X_4$  is histidine;  
 $X_5$  is serine; and  
 $X_6$  is histidine.



12. The insulin analog of claim 11 wherein

a. said dipeptide element is pegylated with one or two polyethylene glycol chains wherein the combined molecular weight of the polyethylene glycol chains ranges from about 20,000 to about 80,000 Daltons; or

5 b. said dipeptide element is acylated with an acyl group comprising 16 to 30 carbon atoms.

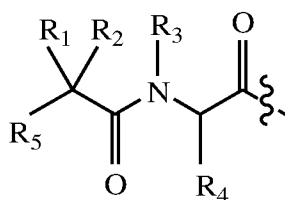
13. The insulin analog of claim 11 wherein the B chain sequence comprises the sequence J-X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTPKT (SEQ ID NO: 15) or J-X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALX<sub>7</sub>LVCGERGFX<sub>8</sub>YTKPT (SEQ ID NO: 16) wherein

10 X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine.

14. The insulin analog of claim 13 wherein

a. a hydrophilic moiety, optionally polyethylene glycol, is linked to an amino acid at a position selected from the group consisting of the N-terminal amino acid of the B chain, the lysine at position 29 of SEQ ID NO: 15, or the lysine at position 28 of SEQ ID NO: 16; or

b. a hydrophilic moiety or a sequestering macromolecule is covalently linked to the R<sub>2</sub> side chain of the dipeptide element comprising the general structure:



20 , wherein R<sub>2</sub> represents (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)SH, or (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>.

15. The insulin analog of claim 1 further comprising a linking peptide and formed as a single chain compound of the formula: B<sub>0</sub>-P-A<sub>0</sub>, wherein:

B<sub>0</sub> represents said B chain having the general formula J-X<sub>14</sub>-peptide B;

A<sub>0</sub> represents said A chain having the general formula Z-peptide A; and

25 P represents the linking peptide, wherein said linking peptide consisting of about 4 to about 14 amino acids.

16. The insulin analog of claim 15 wherein

a. the linking peptide comprises the sequence: Gly-Gly-Gly-Pro-Gly-Lys-Arg (SEQ ID NO: 22), AGRGSGK (SEQ ID NO: 35), AGLGSGK (SEQ ID NO: 36), AGMGSGK (SEQ ID NO: 37), ASWGSGK (SEQ ID NO: 38), TGLGSGQ (SEQ ID NO: 39), TGLGRGK (SEQ ID NO: 40), TGLGSGK (SEQ ID NO: 41), HGLYSGK (SEQ ID NO: 42), KGLGSGQ (SEQ ID NO: 43), VGLMSGK (SEQ ID NO: 44), VGLSSGQ (SEQ ID NO: 45), VGLYSGK (SEQ ID NO: 46), VGLSSGK (SEQ ID NO: 47), VGMSSGK (SEQ ID NO: 48), VWSSSGK (SEQ ID NO: 49), VGSSSGK (SEQ ID NO: 50), VGMSSGK (SEQ ID NO: 51), TGLGSGR (SEQ ID NO: 52), TGLGKGQ (SEQ ID NO: 53), KGLSSGQ (SEQ ID NO: 54), VKLSSGQ (SEQ ID NO: 55), VGLKSGQ (SEQ ID NO: 56), TGLGKGQ (SEQ ID NO: 57) and VGLSKGQ (SEQ ID NO: 58); or

b. the B chain sequence comprises the sequence X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTPKT (SEQ ID NO: 12) or X<sub>9</sub>VNQX<sub>4</sub>LCGX<sub>5</sub>X<sub>6</sub>LVEALYLVCGERGFFYTKPT (SEQ ID NO: 13) wherein

X<sub>4</sub> is selected from the group consisting of histidine and threonine;

X<sub>5</sub> is selected from the group consisting of alanine, glycine and serine;

X<sub>6</sub> is selected from the group consisting of histidine, aspartic acid, glutamic acid, homocysteic acid and cysteic acid; and

X<sub>9</sub> is selected from the group consisting of phenylalanine and desamino-phenylalanine; or

c. Z is H;

X is NHR<sub>10</sub>;

X<sub>4</sub> is histidine;

X<sub>5</sub> is serine;

X<sub>6</sub> is histidine;

R<sub>1</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl;

R<sub>2</sub> is selected from the group consisting of H, C<sub>1</sub>-C<sub>6</sub> alkyl, C<sub>2</sub>-C<sub>8</sub> alkenyl, (C<sub>1</sub>-C<sub>4</sub> alkyl)OH, (C<sub>1</sub>-C<sub>4</sub> alkyl)NH<sub>2</sub>, (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>3</sub>-C<sub>6</sub> cycloalkyl), (C<sub>0</sub>-C<sub>4</sub> alkyl)(C<sub>6</sub>-C<sub>10</sub> aryl)R<sub>7</sub>, and CH<sub>2</sub>(C<sub>5</sub>-C<sub>9</sub> heteroaryl) or R<sub>2</sub> and R<sub>6</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring;

R<sub>3</sub> is C<sub>1</sub>-C<sub>6</sub> alkyl; and

R<sub>4</sub> is selected from the group consisting of H and C<sub>1</sub>-C<sub>4</sub> alkyl or R<sub>3</sub> and R<sub>4</sub> together with the atoms to which they are attached form a 5 member heterocyclic ring; or

d. any combination of a, b or c.

17. The insulin analog of any of one of claims 1-14 wherein said analog is acylated at one or more positions selected from A9, A14, A15, B22, B28 or B29.

18. A dimer or multimer comprising an insulin analog of any of claims 1 to 17.

5 19. The insulin analog of any one of claims 1-18, wherein the first amino acid of the dipeptide element is an amino acid in the D-stereochemical configuration.

20. A pharmaceutical composition comprising the insulin analog of any one of claims 1-19, and a pharmaceutically acceptable carrier.

10 21. The use of a compound of any one of claims 1 to 19, in the manufacture of a medicament for the treatment of hyperglycemia or diabetes.

**Indiana University Research and Technology Corporation**

**Patent Attorneys for the Applicant/Nominated Person  
SPRUSON & FERGUSON**

Two-Step Insulin  
Chain Combination

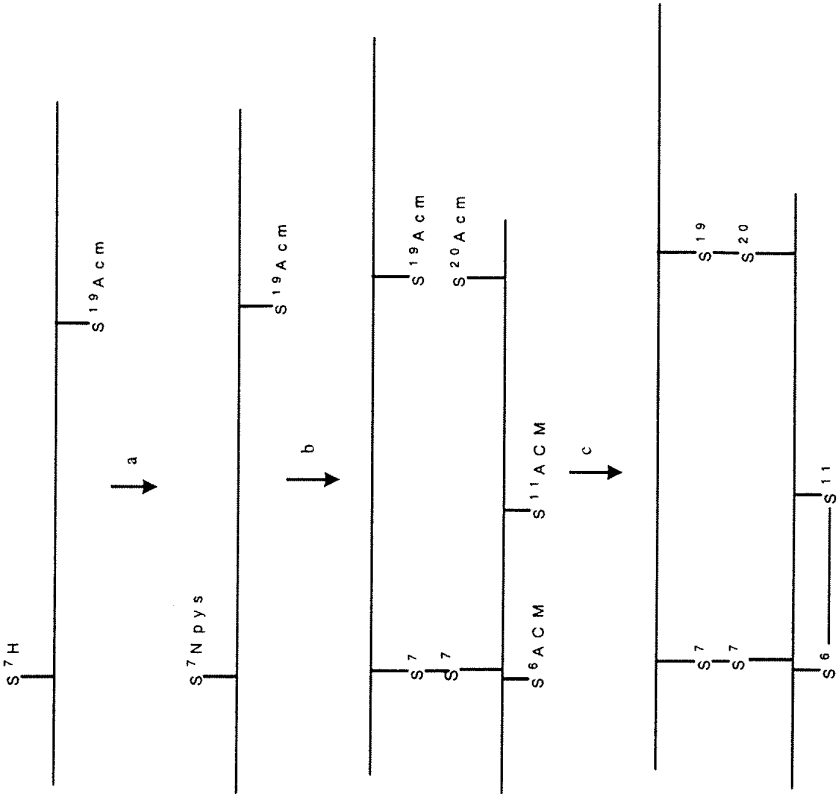
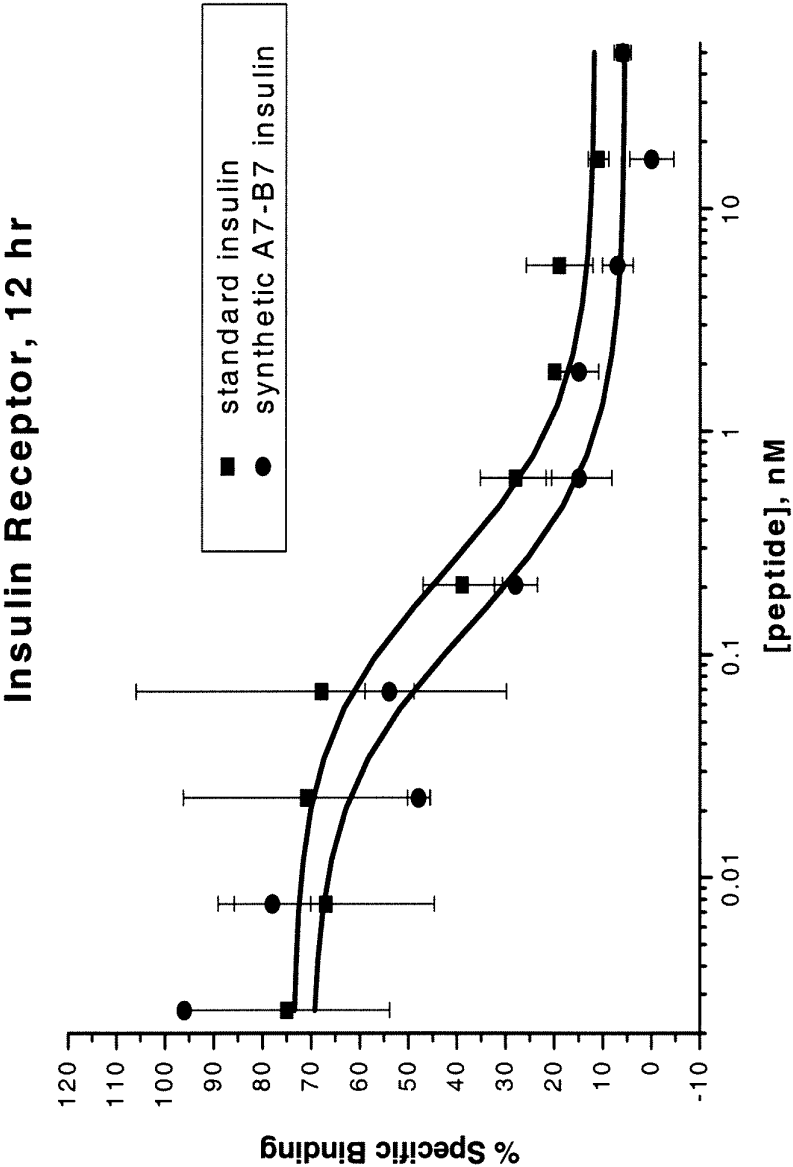


Fig. 1

**Fig. 2**  
**Synthetic “A<sup>7</sup>-B<sup>7</sup>”-derived Insulin**  
**Receptor Binding**



3/16

Fig. 3

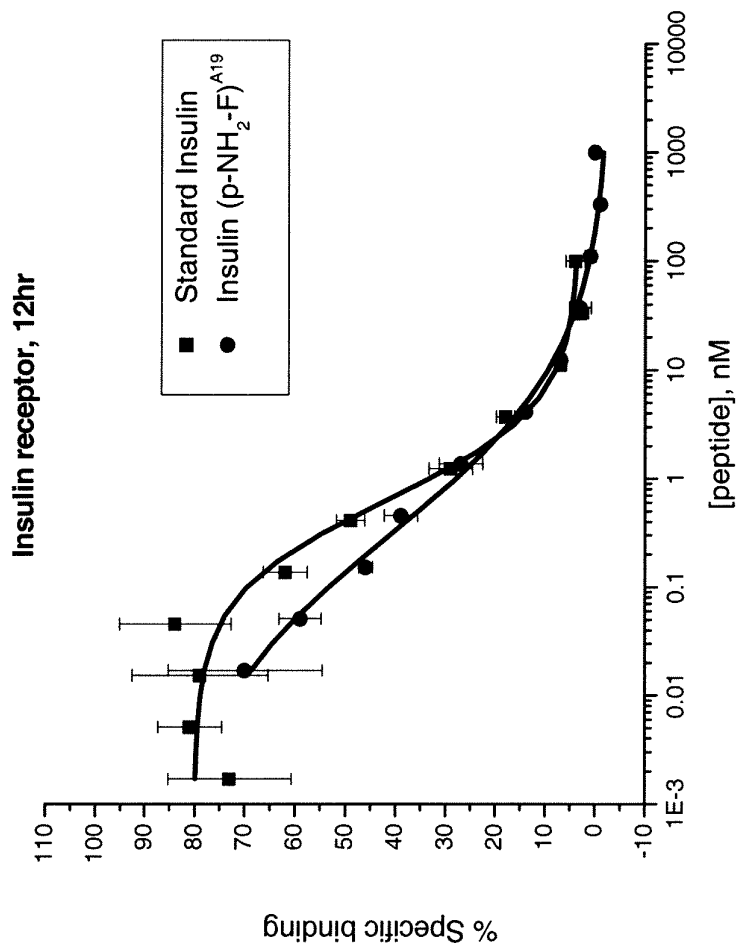
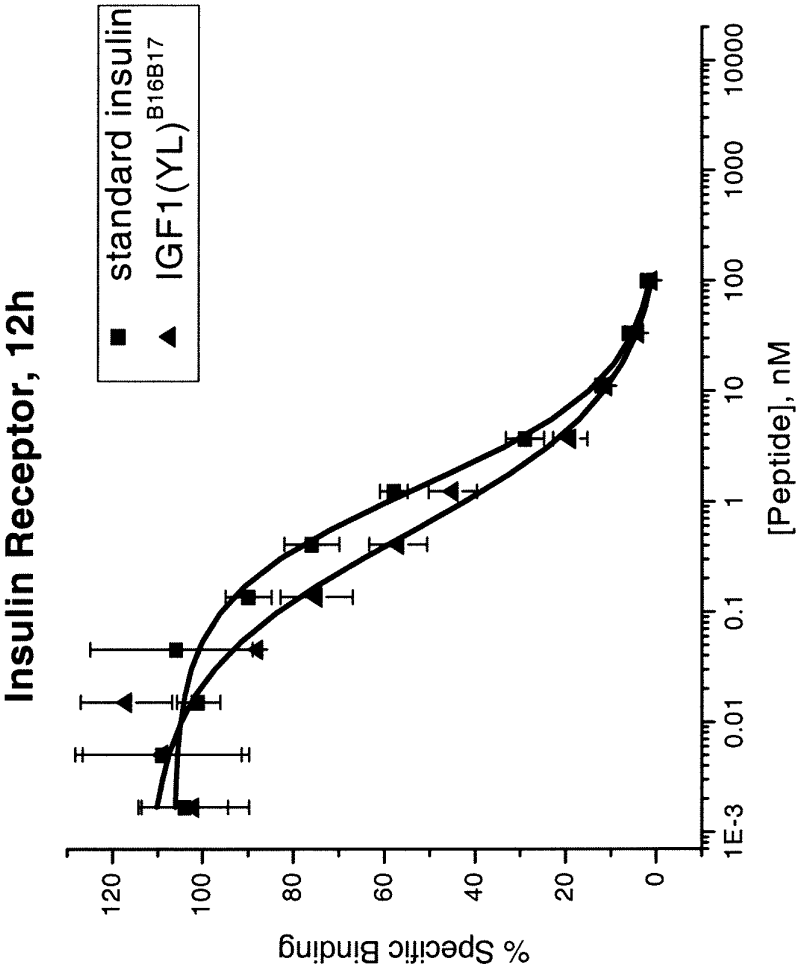


Fig. 4



5/16

Fig. 5

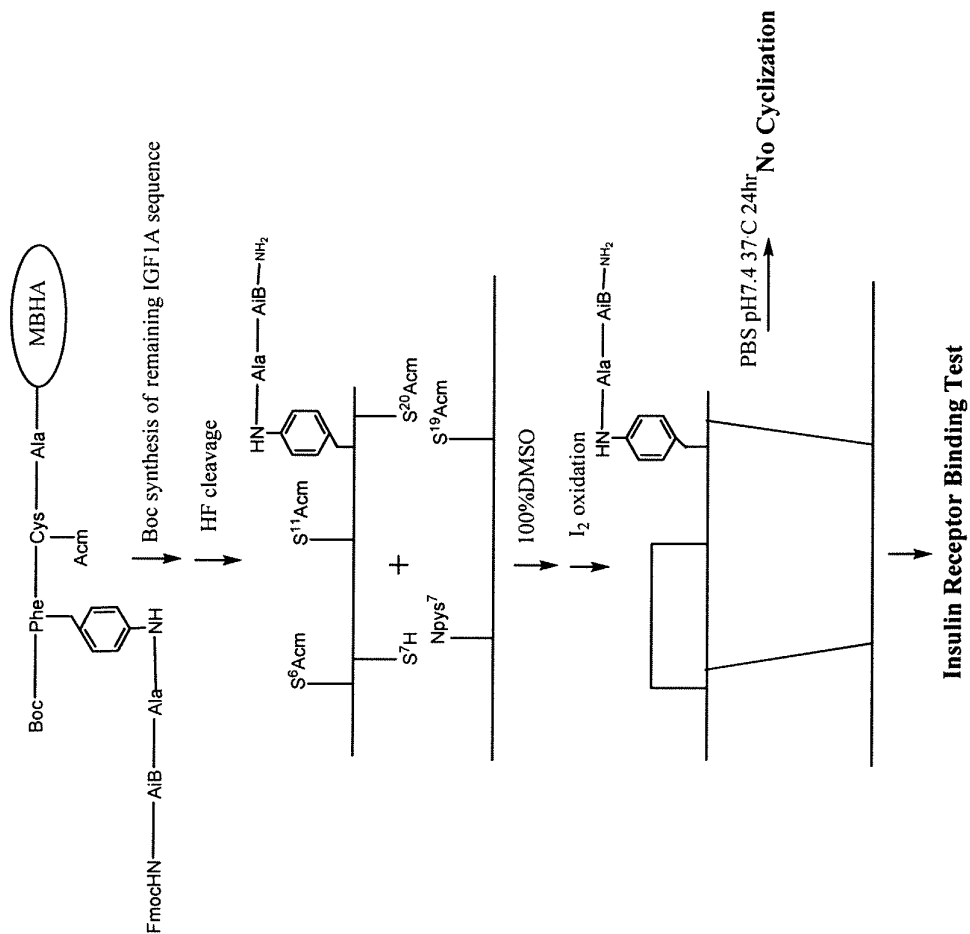
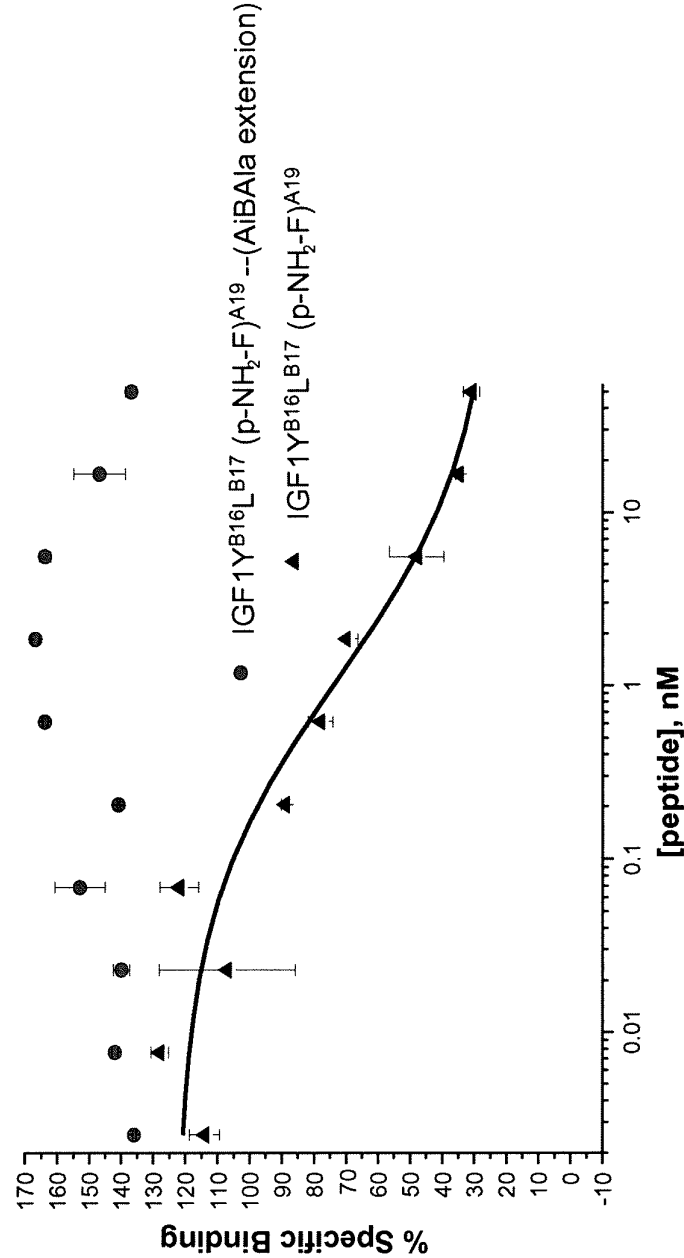




Fig. 6

Insulin Receptor, 12 hr



7/16

IGF-1B[C<sup>0</sup>H<sup>5</sup>Y<sup>16</sup>L<sup>17</sup>O<sup>22</sup>] - A[O<sup>9,14,15</sup>N<sup>18,21</sup>] Dimer

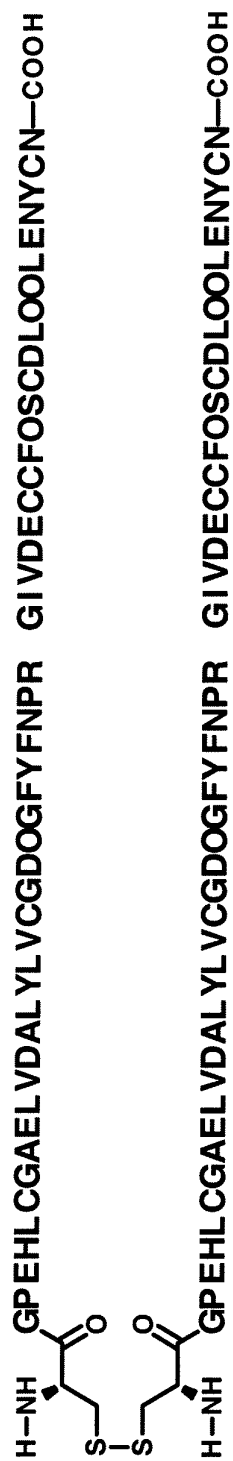


Fig. 7A

8/16

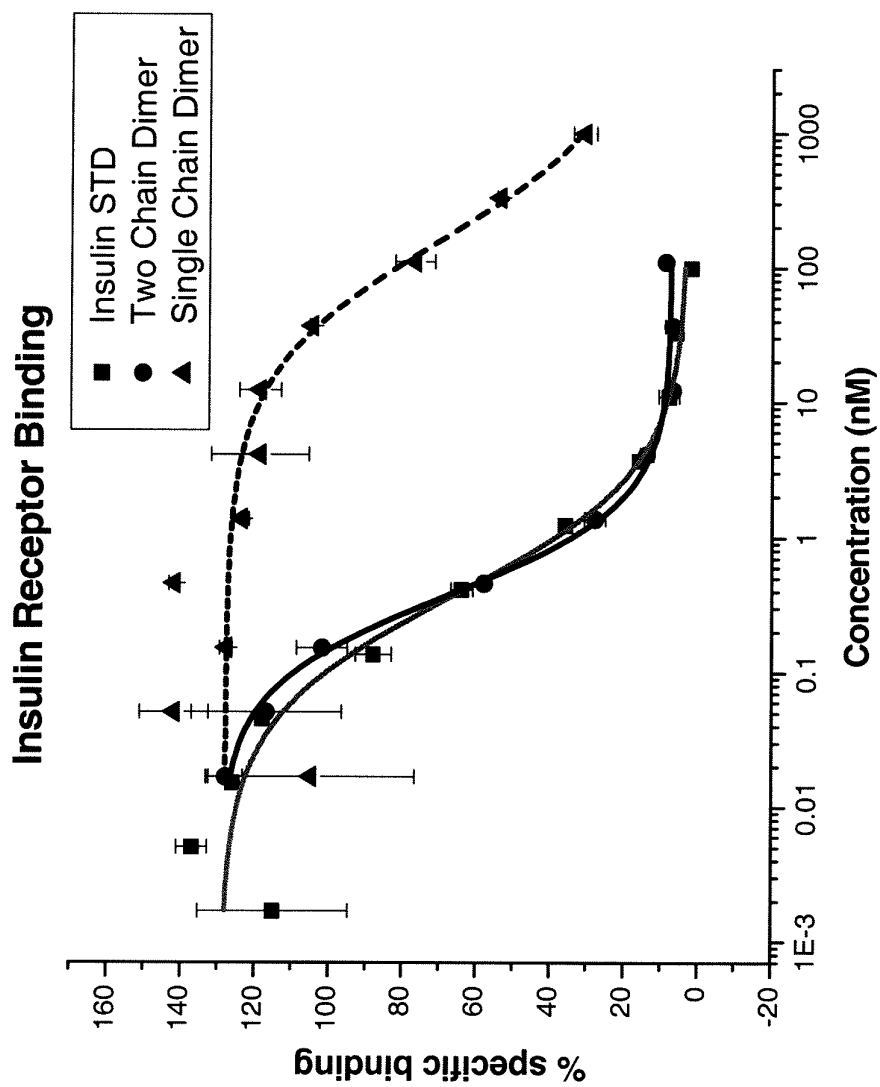


Fig. 7B

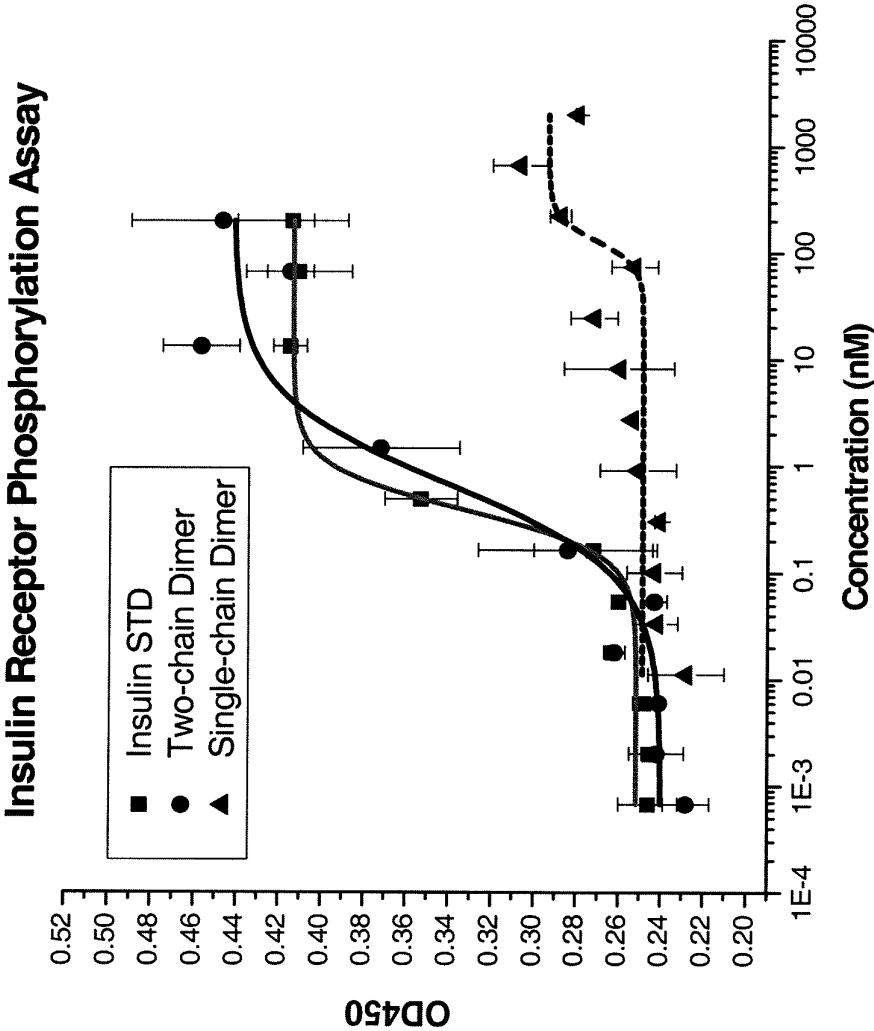


Fig. 7C

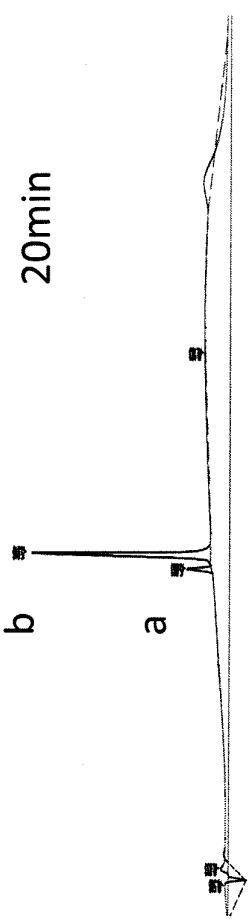


Fig. 8A

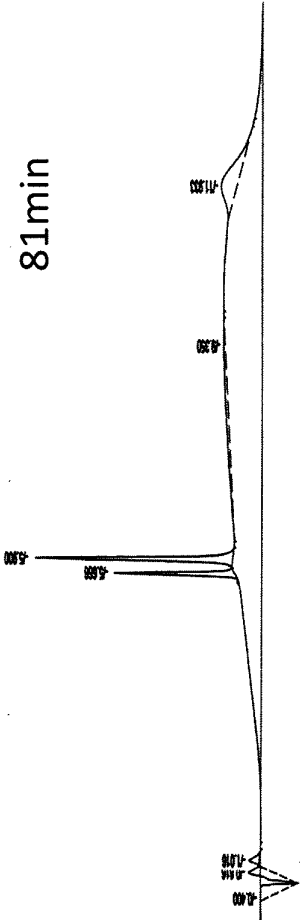


Fig. 8B

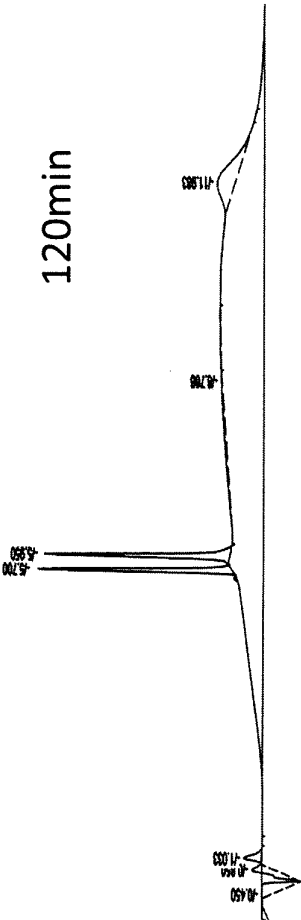


Fig. 8C

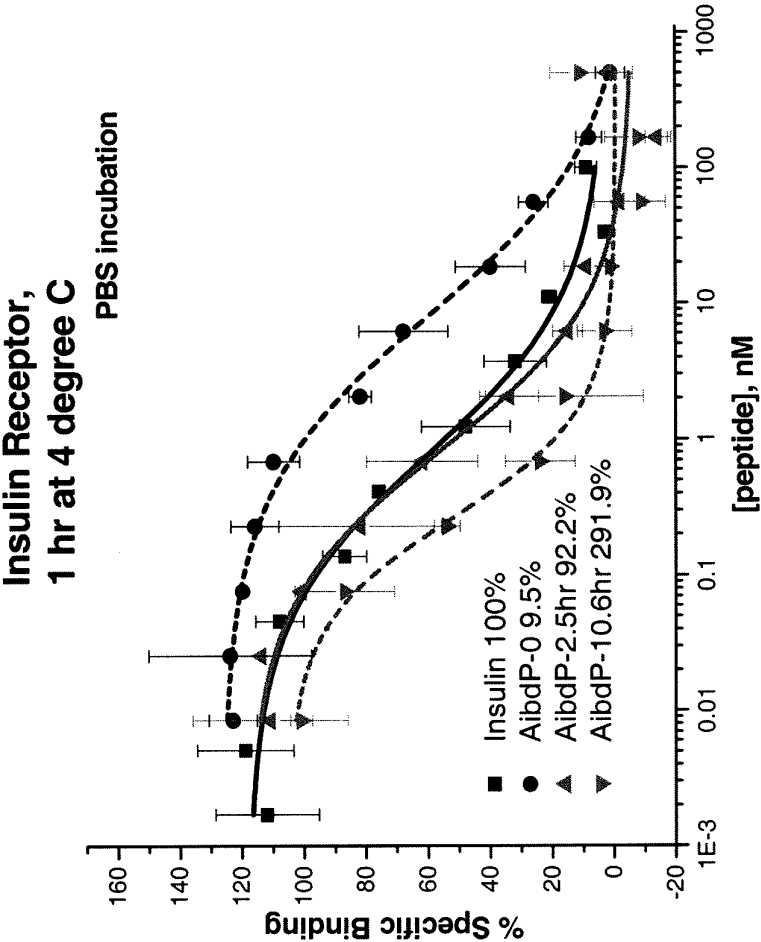


Fig. 9A

12/16

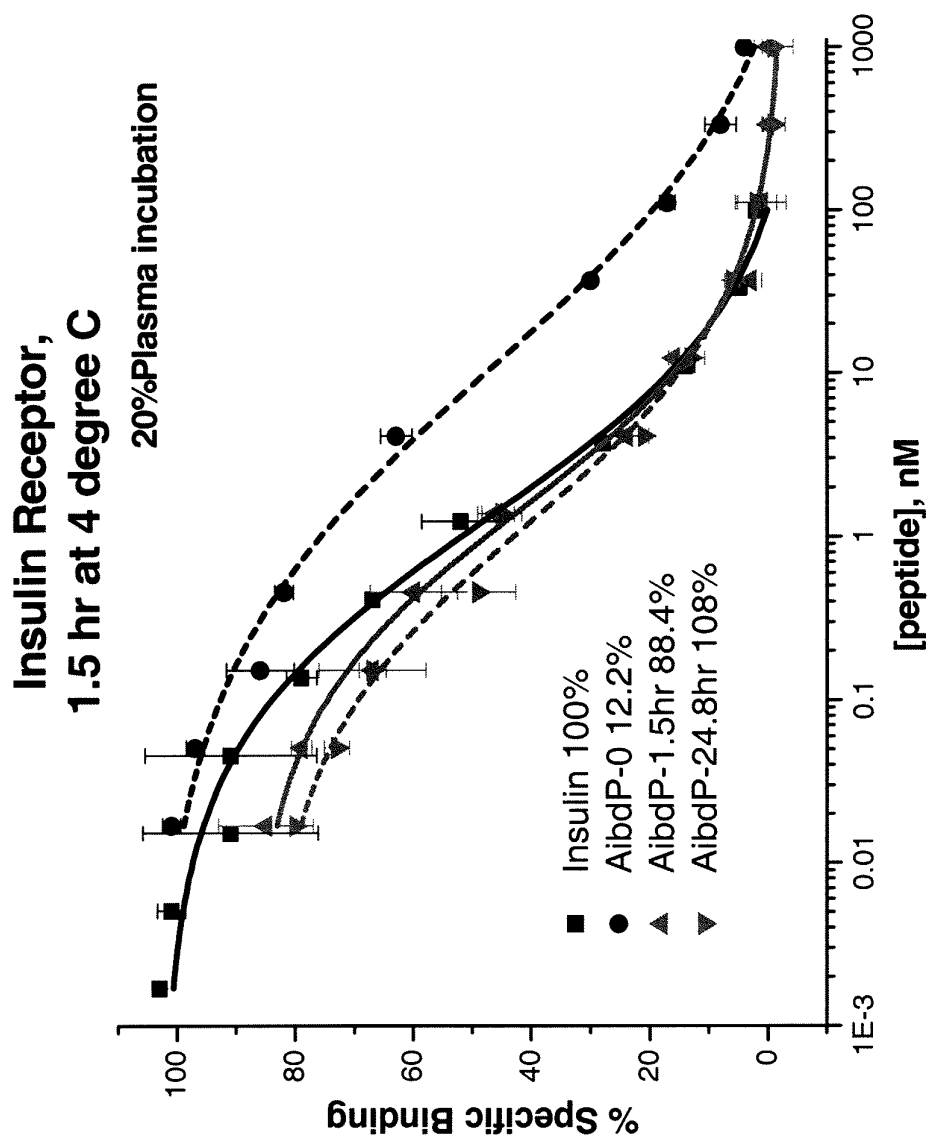


Fig. 9B

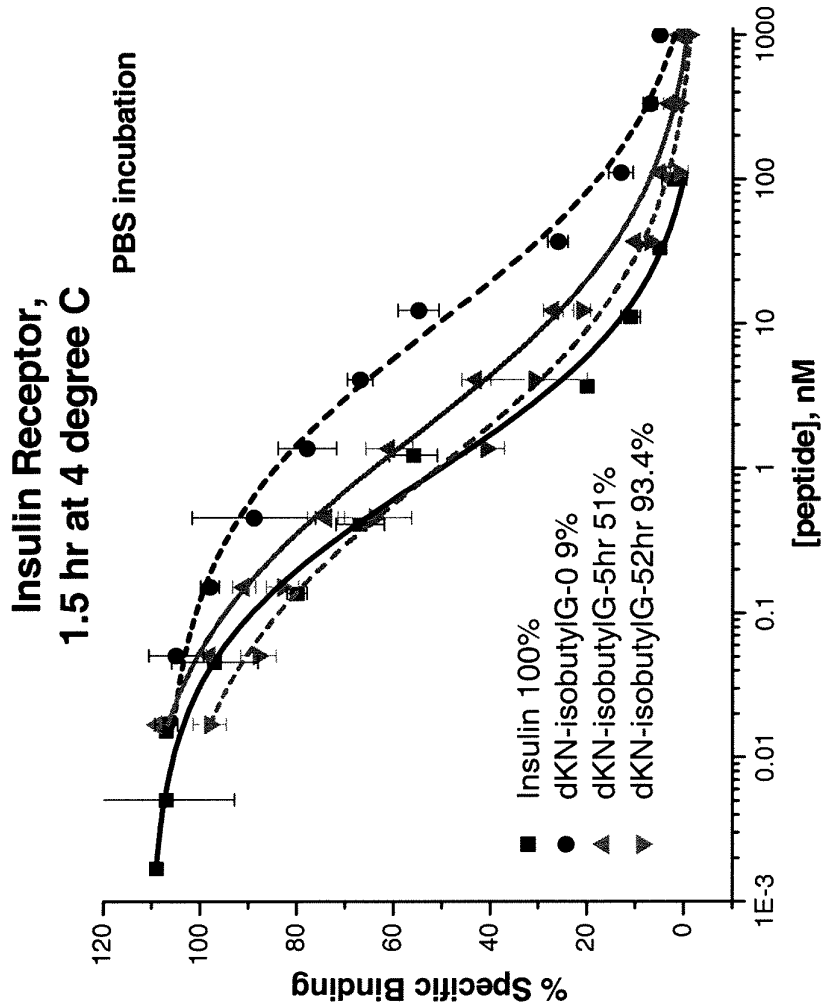


Fig. 10A



14/16

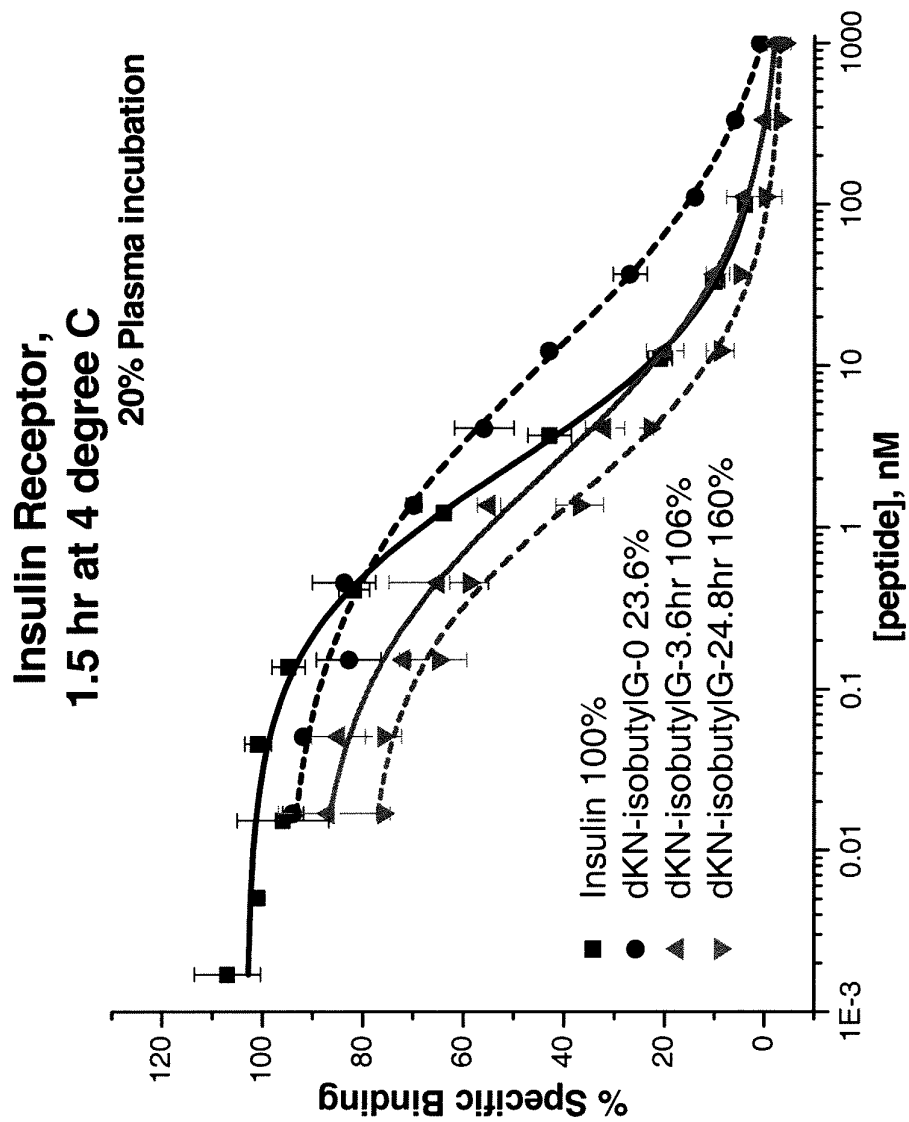


Fig. 10B

15/16

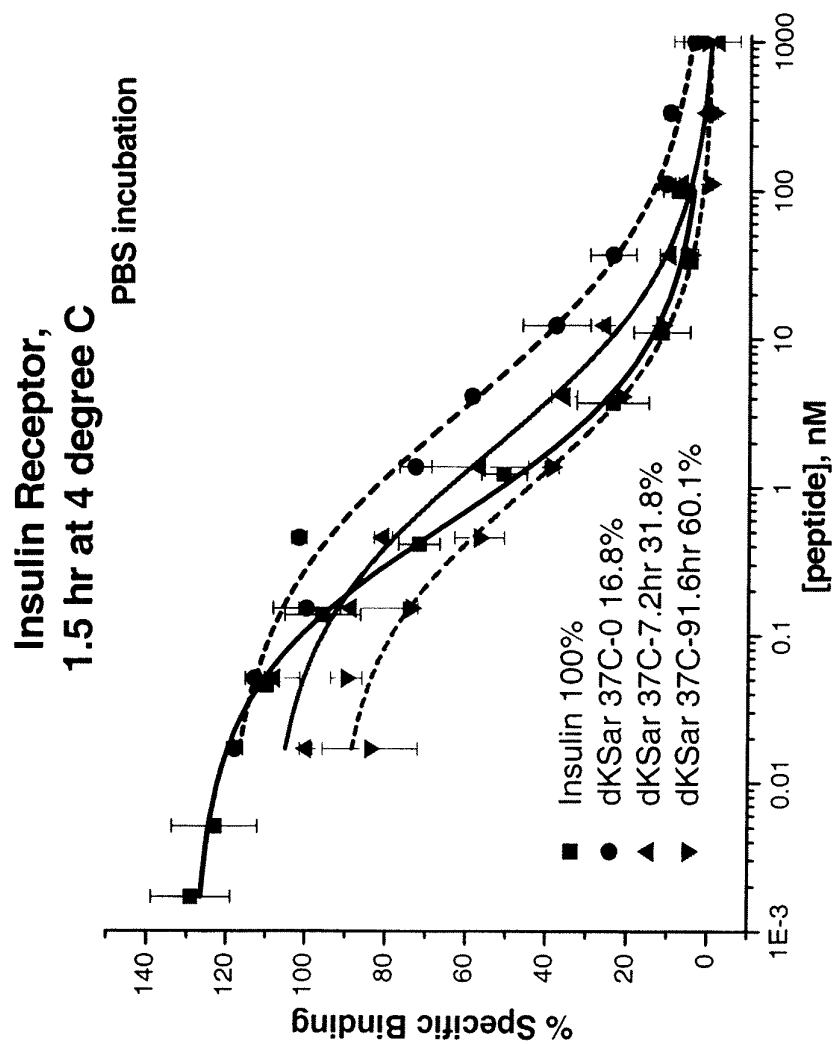


Fig. 11A

16/16

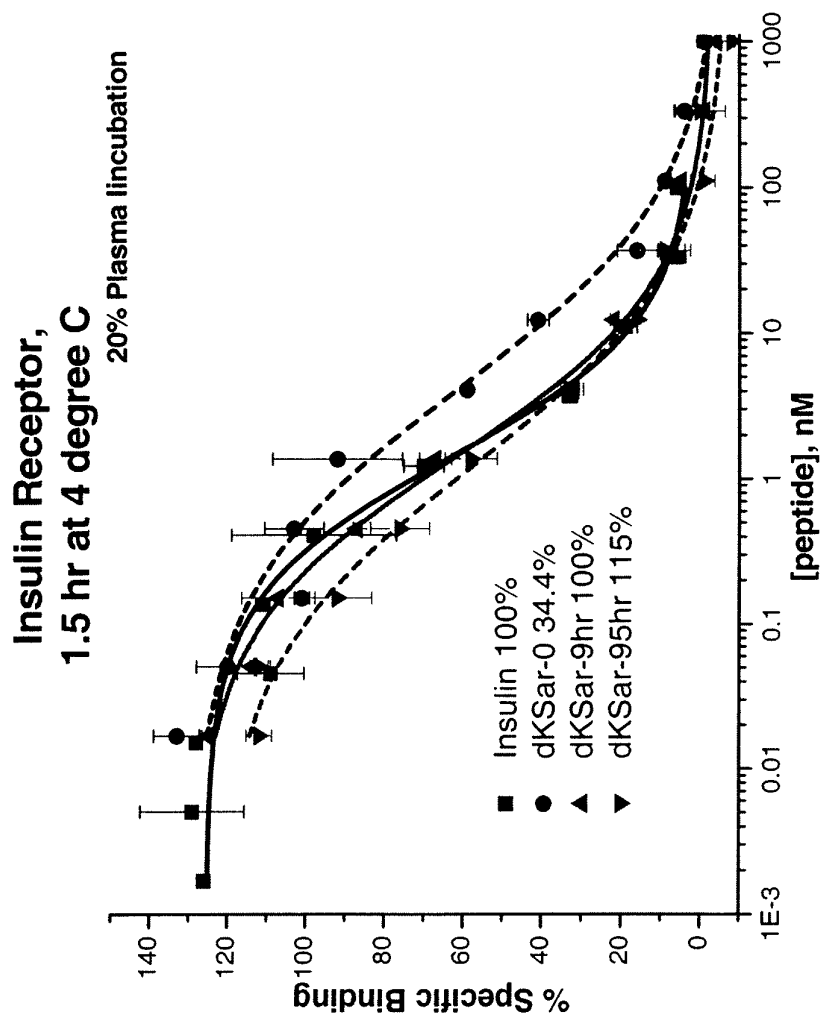


Fig. 11B

997054Seq. TXT  
SEQUENCE LISTING

<110> Di Marchi , Richard  
Kou, Binbin  
Cheng, Shujia

<120> AMIDE-BASED INSULIN PRODRUGS

<130> 29920-204933

<150> 61/139,218

<151> 2008-12-19

<160> 68

<170> PatentIn version 3.3

<210> 1

<211> 21

<212> PRT

<213> Homo sapiens

<400> 1

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	

Glu	Asn	Tyr	Cys	Asn
				20

<210> 2

<211> 30

<212> PRT

<213> Homo sapiens

<400> 2

Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5					10					15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr
			20					25					30

<210> 3

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Analog of the insulin A chain

<220>

<221> MISC\_FEATURE

<222> (8)..(8)

<223> Xaa at position 8 is threonine, histidine, arginine or lysine

<220>

<221> MISC\_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is tyrosine, 4-methoxy phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<220>  
 <221> MI SC\_FEATURE  
 <222> (21)..(21)  
 <223> Xaa at position 21 is asparagine, glycine, alanine glutamine,  
 glutamate, threonine, or serine

<400> 3

Gly Ile Val Glu Gln Cys Cys Xaa Ser Ile Cys Ser Leu Tyr Gln Leu  
 1 5 10 15

Glu Asn Xaa Cys Xaa  
 20

<210> 4  
 <211> 21  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of the insulin B chain

<220>  
 <221> MI SC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is histidine or threonine

<220>  
 <221> MI SC\_FEATURE  
 <222> (5)..(5)  
 <223> Xaa at position 5 is alanine, glycine or serine

<220>  
 <221> MI SC\_FEATURE  
 <222> (6)..(6)  
 <223> Xaa at position 6 is histidine, aspartic acid, glutamic acid,  
 homocysteic acid or cysteic acid

<400> 4

Xaa Leu Cys Gly Xaa Xaa Leu Val Glu Ala Leu Tyr Leu Val Cys Gly  
 1 5 10 15

Glu Arg Gly Phe Phe  
 20

<210> 5  
 <211> 29  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of insulin B chain

<220>  
 <221> MI SC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is phenylalanine or desamino-phenylalanine.

<220>  
 <221> MI SC\_FEATURE  
 <222> (5)..(5)

<223> Xaa at position 5 is histidine, threonine or serine

<220>

<221> MISC\_FEATURE

<222> (9)..(9)

<223> Xaa at position 9 is threonine or serine

<220>

<221> MISC\_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<220>

<221> MISC\_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is tyrosine or 4-hydroxymethyl phenylalanine

<220>

<221> MISC\_FEATURE

<222> (25)..(25)

<223> Xaa at position 25 is tyrosine, phenylalanine or 4-hydroxymethyl phenylalanine

<220>

<221> MISC\_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is aspartate, a lysine-proline dipeptide, or a proline-lysine dipeptide

<220>

<221> MISC\_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is threonine or an arginine-arginine dipeptide

<400> 5

Xaa Val Asn Gln Xaa Leu Cys Gly Xaa Xaa Leu Val Glu Ala Leu Xaa  
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Xaa Tyr Thr Xaa Xaa  
20 25

<210> 6

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Analog of insulin a chain

<220>

<221> MISC\_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is 4-amino phenylalanine

<400> 6

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
1 5 10 15

Glu Asn Xaa Cys Asn  
20

997054Seq. TXT

<210> 7  
 <211> 21  
 <212> PRT  
 <213> Homo sapiens

<400> 7

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly  
 1 5 10 15

Glu Arg Gly Phe Phe  
 20

<210> 8  
 <211> 30  
 <212> PRT  
 <213> Homo sapiens

<400> 8

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr  
 20 25 30

<210> 9  
 <211> 30  
 <212> PRT  
 <213> Homo sapiens

<400> 9

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Lys Pro Thr  
 20 25 30

<210> 10  
 <211> 32  
 <212> PRT  
 <213> Homo sapiens

<400> 10

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg  
 20 25 30

<210> 11  
 <211> 4  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> N-terminus of the B chain of an insulin analog

&lt;400&gt; 11

Phe Val Asn Gln  
1

&lt;210&gt; 12

&lt;211&gt; 30

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Analog of the insulin B chain

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (1)..(1)

&lt;223&gt; Xaa at position 1 is phenyl alanine or desamino-phenyl alanine

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (5)..(5)

&lt;223&gt; Xaa at position 5 is histidine or threonine

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (9)..(9)

&lt;223&gt; Xaa at position 9 is alanine, glucine or serine

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (10)..(10)

&lt;223&gt; Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

&lt;400&gt; 12

Xaa Val Asn Gln Xaa Leu Cys Gly Xaa Xaa Leu Val Glu Ala Leu Tyr  
1 5 10 15Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr  
20 25 30

&lt;210&gt; 13

&lt;211&gt; 30

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Analog of the insulin B chain

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (1)..(1)

&lt;223&gt; Xaa at position 1 is phenyl alanine or desamino-phenyl alanine

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (5)..(5)

&lt;223&gt; Xaa at position 5 is histidine or threonine

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE



997054Seq.TXT

<222> (9)..(9)  
 <223> Xaa at position 9 is alanine, glycine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (10)..(10)  
 <223> Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<400> 13

Xaa	Val	Asn	Gln	Xaa	Leu	Cys	Gly	Xaa	Xaa	Leu	Val	Glu	Ala	Leu	Tyr
1				5				10						15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Lys	Pro	Thr
			20					25					30

<210> 14  
 <211> 21  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of the insulin B chain

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is histidine, threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (5)..(5)  
 <223> Xaa at position 5 is threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (6)..(6)  
 <223> Xaa at position 6 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<220>  
 <221> MISC\_FEATURE  
 <222> (12)..(12)  
 <223> Xaa at position 12 is tyrosine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<220>  
 <221> MISC\_FEATURE  
 <222> (21)..(21)  
 <223> Xaa at position 21 is tyrosine, phenylalanine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<400> 14

Xaa	Leu	Cys	Gly	Xaa	Xaa	Leu	Val	Glu	Ala	Leu	Xaa	Leu	Val	Cys	Gly
1				5				10						15	

Glu	Arg	Gly	Phe	Xaa
			20	

<210> 15  
 <211> 30  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of the insulin B chain

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is phenylalanine or desamino-phenylalanine

<220>  
 <221> MISC\_FEATURE  
 <222> (5)..(5)  
 <223> Xaa at position 5 is histidine, threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (9)..(9)  
 <223> Xaa at position 9 is threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (10)..(10)  
 <223> Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<220>  
 <221> MISC\_FEATURE  
 <222> (16)..(16)  
 <223> Xaa at position 16 is tyrosine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic amine)

<220>  
 <221> MISC\_FEATURE  
 <222> (25)..(25)  
 <223> Xaa at position 25 is tyrosine, phenylalanine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic amine)

<220>  
 <221> MISC\_FEATURE  
 <222> (25)..(25)  
 <223> Xaa at position 25 is tyrosine, phenylalanine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<400> 15

Xaa	Val	Asn	Gln	Xaa	Leu	Cys	Gly	Xaa	Xaa	Leu	Val	Glu	Ala	Leu	Xaa
1				5				10						15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Xaa	Tyr	Thr	Pro	Lys	Thr
			20					25					30

<210> 16  
 <211> 30  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of the insulin B chain

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is phenylalanine or desamino-phenylalanine

<220>  
 <221> MISC\_FEATURE  
 <222> (5)..(5)  
 <223> Xaa at position 5 is histidine, threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (9)..(9)  
 <223> Xaa at position 9 is threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (10)..(10)  
 <223> Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<220>  
 <221> MISC\_FEATURE  
 <222> (16)..(16)  
 <223> Xaa at position 16 is tyrosine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<220>  
 <221> MISC\_FEATURE  
 <222> (25)..(25)  
 <223> Xaa at position 25 is tyrosine, phenylalanine, 4-hydroxymethyl phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<400> 16  
 Xaa Val Asn Gln Xaa Leu Cys Gly Xaa Xaa Leu Val Glu Ala Leu Xaa  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Xaa Tyr Thr Lys Pro Thr  
 20 25 30

<210> 17  
 <211> 21  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of the insulin A chain

<220>  
 <221> MISC\_FEATURE  
 <222> (8)..(8)  
 <223> Xaa at position 8 is threonine, histidine, arginine or lysine

<220>  
 <221> MISC\_FEATURE  
 <222> (21)..(21)

997054Seq.TXT

<223> Xaa at position 21 is asparagine, glycine, alanine glutamine, glutamate, threonine, or serine

<400> 17

Gly Ile Val Glu Gln Cys Cys Xaa Ser Ile Cys Ser Leu Tyr Gln Leu  
1 5 10 15

Glu Asn Tyr Cys Xaa  
20

<210> 18

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Analog of the insulin A chain

<220>

<221> MISC\_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is tyrosine, 4-methoxy phenylalanine, 4-amino phenylalanine or a modified 4-amino phenylalanine (having a dipeptide linked via an amide bond at the aromatic ring amine)

<220>

<221> MISC\_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is asparagine, glycine, alanine glutamine, glutamate, threonine, or serine

<400> 18

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
1 5 10 15

Glu Asn Xaa Cys Xaa  
20

<210> 19

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Analog of the insulin B chain

<220>

<221> MISC\_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is histidine, threonine or serine

<400> 19

Xaa Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly  
1 5 10 15

Glu Arg Gly Phe Phe  
20

997054Seq. TXT

<210> 20  
 <211> 29  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Analog of insulin B chain

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is phenylalanine or desamino-phenylalanine.

<220>  
 <221> MISC\_FEATURE  
 <222> (5)..(5)  
 <223> Xaa at position 5 is histidine, threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (9)..(9)  
 <223> Xaa at position 9 is threonine or serine

<220>  
 <221> MISC\_FEATURE  
 <222> (10)..(10)  
 <223> Xaa at position 10 is histidine, aspartic acid, glutamic acid, homocysteic acid or cysteic acid

<220>  
 <221> MISC\_FEATURE  
 <222> (16)..(16)  
 <223> Xaa at position 16 is tyrosine or 4-hydroxymethyl phenylalanine

<220>  
 <221> MISC\_FEATURE  
 <222> (25)..(25)  
 <223> Xaa at position 25 is tyrosine, phenylalanine or 4-hydroxymethyl phenylalanine

<220>  
 <221> MISC\_FEATURE  
 <222> (28)..(28)  
 <223> Xaa at position 28 is aspartate, a lysine-proline dipeptide, or a proline-lysine dipeptide

<220>  
 <221> MISC\_FEATURE  
 <222> (29)..(29)  
 <223> Xaa at position 29 is threonine or an arginine-arginine dipeptide

<400> 20

Xaa Val Asn Gln Xaa Leu Cys Gly Xaa Xaa Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Xaa Xaa  
 20 25

<210> 21  
 <211> 4  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> N-terminus of the B chain of an insulin analog

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(1)  
 <223> Xaa at position 1 is phenylalanine or desamino-phenylalanine.

<400> 21

Xaa Val Asn Gln  
 1

<210> 22  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker

<400> 22

Gly Gly Gly Pro Gly Lys Arg  
 1 5

<210> 23  
 <211> 12  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker

<400> 23

Gly Tyr Gly Ser Ser Ser Arg Arg Ala Pro Gln Thr  
 1 5 10

<210> 24  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker

<400> 24

Gly Gly Gly Gly Gly Lys Arg  
 1 5

<210> 25  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker

<400> 25

Arg Arg Gly Gly Gly Gly Gly

Gly Gly Ala Pro Gly Asp Val Lys Arg  
1 5

Arg Arg Ala Pro Gly Asp Val Gly Gly  
1 5

Gly Gly Tyr Pro Gly Asp Val Lys Arg  
1 5

Arg Arg Tyr Pro Gly Asp Val Gly Gly  
1 5

<400> 30

Gly Gly His Pro Gly Asp Val Lys Arg  
1 5

<210> 31  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 31

Arg Arg His Pro Gly Asp Val Gly Gly  
1 5

<210> 32  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 32

Arg Arg Gly Pro Gly Gly Gly  
1 5

<210> 33  
<211> 51  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> single chain human insulin analog

<220>  
<221> MISC\_FEATURE  
<222> (49)..(49)  
<223> Xaa at position 49 is 4-amino phenyl alanine or 4-methoxy  
phenyl alanine

<400> 33

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Gly Ile  
20 25 30

Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn  
35 40 45

Xaa Cys Asn  
50

<210> 34  
<211> 63



<212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> single-chain human insulin analog

<220>  
 <221> MI SC\_FEATURE  
 <222> (61)..(61)  
 <223> Xaa at position 61 is 4-amino phenylal anine or 4-methoxy phenyl al anine

<400> 34

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Gln Pro  
 20 25 30

Leu Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys  
 35 40 45

Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Xaa Cys Asn  
 50 55 60

<210> 35  
 <211> 7  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> peptide linker

<400> 35

Ala Gly Arg Gly Ser Gly Lys  
 1 5

<210> 36  
 <211> 7  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> peptide linker

<400> 36

Ala Gly Leu Gly Ser Gly Lys  
 1 5

<210> 37  
 <211> 7  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> peptide linker

<400> 37

Ala Gly Met Gly Ser Gly Lys  
1 5

<210> 38  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 38

Ala Ser Trp Gly Ser Gly Lys  
1 5

<210> 39  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 39

Thr Gly Leu Gly Ser Gly Gln  
1 5

<210> 40  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 40

Thr Gly Leu Gly Arg Gly Lys  
1 5

<210> 41  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

<400> 41

Thr Gly Leu Gly Ser Gly Lys  
1 5

<210> 42  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> peptide linker

&lt;400&gt; 42

His Gly Leu Tyr Ser Gly Lys  
 1 5

&lt;210&gt; 43

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; peptide linker

&lt;400&gt; 43

Lys Gly Leu Gly Ser Gly Gln  
 1 5

&lt;210&gt; 44

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; peptide linker

&lt;400&gt; 44

Val Gly Leu Met Ser Gly Lys  
 1 5

&lt;210&gt; 45

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; peptide linker

&lt;400&gt; 45

Val Gly Leu Ser Ser Gly Gln  
 1 5

&lt;210&gt; 46

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; peptide linker

&lt;400&gt; 46

Val Gly Leu Tyr Ser Gly Lys  
 1 5

&lt;210&gt; 47

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
 <223> peptide linker  
 <400> 47  
 Val Gly Leu Ser Ser Gly Lys  
 1 5

<210> 48  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker  
 <400> 48

Val Gly Met Ser Ser Gly Lys  
 1 5

<210> 49  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker  
 <400> 49

Val Trp Ser Ser Ser Gly Lys  
 1 5

<210> 50  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker  
 <400> 50

Val Gly Ser Ser Ser Gly Lys  
 1 5

<210> 51  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> peptide linker  
 <400> 51

Val Gly Met Ser Ser Gly Lys  
 1 5

<210> 52  
 <211> 7  
 <212> PRT

<213>    Arti f i c i a l    Sequence

 $\langle 220 \rangle$ 

<223> pepti de l i nker

<400> 52

Thr Gly Leu Gly Ser Gly Arg  
1 5

<210> 53

$\langle 211 \rangle$

<212> PRT

<213> Arti fi ci al Sequence

 $\langle 220 \rangle$ 

<223> peptide linker

<400> 53

Thr Gly Leu Gly Lys Gly Gln  
1 5

<210> 54

$\langle 211 \rangle$

<212> PRT

<213> Arti fi ci al Sequence

 $\langle 220 \rangle$ 

<223> pepti de l i nker

<400> 54

Lys Gly Leu Ser Ser Gly Gln  
1 5

<210> 55

 $\langle 211 \rangle$ 

<212> PRT

<213> Arti fi ci al Sequence

 $\langle 220 \rangle$ 

<223> peptide linker

<400> 55

Val Lys Leu Ser Ser Gly Gln  
1 5

<210> 56

 $\langle 211 \rangle$ 

<212> PRT

<213> Arti fi ci al Sequence

 $\langle 220 \rangle$ 

<223> pepti de linker

<400> 56

Val Lys Leu Ser Ser Gly Gln  
1 5

<210> 57

<211> 7  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> peptide linker

<400> 57

Thr Gly Leu Gly Lys Gly Gln  
 1 5

<210> 58  
 <211> 7  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> peptide linker

<400> 58

Val Gly Leu Ser Lys Gly Gln  
 1 5

<210> 59  
 <211> 6  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> synthetic peptide for testing cleavage of dipeptide prodrug  
 element

<400> 59

His Ser Arg Gly Thr Phe  
 1 5

<210> 60  
 <211> 8  
 <212> PRT  
 <213> Arti fi ci al Sequence

<220>  
 <223> synthetic peptide for testing cleavage of dipeptide prodrug  
 element

<220>  
 <221> MI SC\_FEATURE  
 <222> (2).(2)  
 <223> Xaa at position 2 is Sarcosine

<400> 60

Lys Xaa His Ser Thr Gly Thr Phe  
 1 5

<210> 61  
 <211> 21  
 <212> PRT  
 <213> Homo sapiens

&lt;400&gt; 61

Gly Ile Val Asp Glu Cys Cys Phe Arg Ser Cys Asp Leu Arg Arg Leu  
 1 5 10 15

Glu Met Tyr Cys Ala  
 20

&lt;210&gt; 62

&lt;211&gt; 29

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 62

Gly Pro Glu Thr Leu Cys Gly Ala Glu Leu Val Asp Ala Leu Gln Phe  
 1 5 10 15

Val Cys Gly Asp Arg Gly Phe Tyr Phe Asn Lys Pro Thr  
 20 25

&lt;210&gt; 63

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; glucagon analogue

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; (1)..(1)

&lt;223&gt; d-histidine

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; (2)..(2)

&lt;223&gt; d-threonine

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; (3)..(3)

&lt;223&gt; d-arginine

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; (5)..(5)

&lt;223&gt; d-threonine

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; (6)..(6)

&lt;223&gt; d-phenylalanine

&lt;400&gt; 63

Xaa Xaa Xaa Gly Xaa Xaa  
 1 5

&lt;210&gt; 64

&lt;211&gt; 8

&lt;212&gt; PRT

<213> Arti fi ci al Sequence

<220>

<223> gl ucagon anal ogue

<220>

<221> MOD\_RES

<222> (1)..(1)

<223> d-l ysi ne

<220>

<221> MOD\_RES

<222> (2)..(2)

<223> d-sarcosi ne

<220>

<221> MOD\_RES

<222> (3)..(3)

<223> d-hi sti di ne

<220>

<221> MOD\_RES

<222> (4)..(4)

<223> d-threoni ne

<220>

<221> MOD\_RES

<222> (5)..(5)

<223> d-argi ni ne

<220>

<221> MOD\_RES

<222> (7)..(7)

<223> d-threoni ne

<220>

<221> MOD\_RES

<222> (8)..(8)

<223> d-phenyl al ani ne

<400> 64

Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa  
1 5

<210> 65

<211> 8

<212> PRT

<213> Homo sapi ens

<400> 65

Ser Arg Val Ser Arg Arg Ser Arg  
1 5

<210> 66

<211> 12

<212> PRT

<213> Homo sapi ens

<400> 66

Gly Tyr Gly Ser Ser Ser Arg Arg Ala Pro Gl n Thr  
1 5 10



<210> 67  
 <211> 8  
 <212> PRT  
 <213> Homo sapiens

<400> 67

Gly Ser Ser Ser Arg Arg Ala Pro  
 1 5

<210> 68  
 <211> 48  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> IGF1 B and A chain analogs

<220>  
 <221> MISC\_FEATURE  
 <222> (22)..(22)  
 <223> Xaa at position 22 is ornithine

<220>  
 <221> MISC\_FEATURE  
 <222> (36)..(36)  
 <223> Xaa at position 36 is ornithine

<220>  
 <221> MISC\_FEATURE  
 <222> (41)..(41)  
 <223> Xaa at position 41 is ornithine

<220>  
 <221> MISC\_FEATURE  
 <222> (42)..(42)  
 <223> Xaa at position 42 is ornithine

<400> 68

Cys Gly Pro Glu His Leu Cys Gly Ala Glu Leu Val Asp Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Asp Xaa Gly Phe Tyr Phe Asn Gly Ile Val Asp Glu  
 20 25 30

Cys Cys Phe Xaa Ser Cys Asp Leu Xaa Xaa Leu Glu Asn Tyr Cys Asn  
 35 40 45