



(43) International Publication Date  
15 November 2012 (15.11.2012)

- (51) International Patent Classification:  
C07K 14/62 (2006.01)
- (21) International Application Number:  
PCT/EP2012/001992
- (22) International Filing Date:  
9 May 2012 (09.05.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
11003823.9 10 May 2011 (10.05.2011) EP
- (71) Applicant (for all designated States except US): **GLUC-OMETRIX AG** [DE/DE]; Am Mühlenberg 11, 14476 Potsdam (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **ECKERT, Kelvin** [DE/DE]; Corinthstr. 56, 10245 Berlin (DE). **HERMANN, Stefan** [DE/DE]; Hegelallee 41, 14467 Potsdam (DE). **NOACK, Janina** [DE/DE]; Bertolt-Brecht-Str. 5, 14478 Potsdam (DE). **SCHWEDE, Matthias** [DE/DE]; Lise-Meitner-Str. 21, 14480 Potsdam (DE).
- (74) Agent: **BRITTINGER, Matthias**; Ter Meer Steinmeister & Partner GbR, Mauerkircherstraße 45, 81679 München (DE).

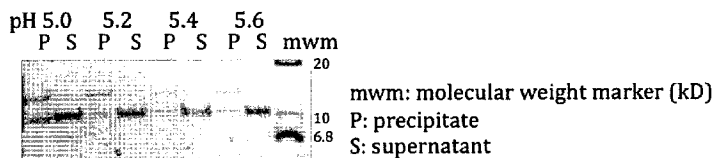
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: COMPREHENSIVE BUFFER SYSTEM FOR RENATURING HUMAN PROINSULIN OR PROINSULIN DERIVATIVES

Figure 1



(57) Abstract: The invention relates to an improved processing and purification method for the production of proinsulin or derivatives thereof starting from the solubilisation of inclusion bodies isolated from a prokaryotic expression system, such as *E. coli*, whereby all processing (i.e. refolding, precipitation, concentration) and purification steps are carried out in the same aqueous buffer system comprising an organic amphoteric substance containing at least two functional groups, which contain at least an amino and a carboxylic acid group, and which is characterized by a difference in the pKa values of its at least one amino group and at least one carboxylic acid group of at least 7. Furthermore, the organic amphoteric substance is an efficient scavenger of cyanate, thereby preventing irreversible carbamylation of primary amines present in proinsulin or the derivatives thereof, which has been observed in the presence of high concentrations of urea. Thus, the disadvantage associated with this widespread chaotropic auxiliary is obliterated, allowing the utilization of efficient and uncomplicated chromatography techniques for purification steps subsequent to solubilisation and refolding. The use of an efficient buffer substance in all processing steps from inclusion body solubilisation up to the re-folded, concentrated and purified precursor of insulin or a derivative thereof, which at the same time acts as an agent preventing protein modification by carbamylation, thereby abolishing the need to add an additional compound with scavenging activity, leads to significant advantages regarding the efficiency and efficacy of the procedure compared to previously employed production methods.



### **Comprehensive buffer system for renaturing human proinsulin or proinsulin derivatives**

The present invention relates to an improved method for obtaining a recombinantly expressed precursor of human insulin, human proinsulin, or a precursor of a human insulin derivative having correctly bonded cystine bridges by recombinant expression, encompassing the steps of solubilisation of isolated and purified inclusion bodies, purification of proinsulin or a proinsulin derivate from solubilised inclusion bodies using ion exchange chromatography, refolding of proinsulin or a proinsulin derivate comprising native disulfide bonds, precipitation of proinsulin or a proinsulin derivate containing non-native disulfide bonds, as well as the steps of concentration and purification of proinsulin or a proinsulin derivate comprising native disulfide bonds using ion exchange chromatography.

Human proinsulin is a protein of about 10 kD, consisting of an A chain, comprising 21 amino acid residues, and a B chain, comprising 30 amino acid residues, both of which are present in the biologically active, mature form of human insulin, and a linker region, the so called C peptide, which comprises 35 amino acid residues and which is absent from the mature human insulin. Six cysteine residues are present within the A and B chains at positions A6, A7, A11 and A20 as well as B7 and B19, whereby the letters represent the respective A and B chains and the numbers represent the position of the amino acid residue, counting from the amino to the carboxyl end of the respective amino acid chain. In the mature, biologically active human insulin, the A and B chains are linked by disulfide bridges between residues A7 and B7 and between residues A 20 and B 19. An intrachain disulfide bond is formed between residues A6 and A11. In the following, the term "correctly bonded cystine bridges" refers to these disulfide bonds, present in the biologically active human insulin.

Known processes for the preparation of human insulin are based on the recombinant expression of a precursor of human insulin (proinsulin) or a human proinsulin derivative in suitable microorganisms (e.g. EP 0 055 945). The formation of correct disulfide bonds in human proinsulin or proinsulin derivatives generated by recombinant expression depends, however, on the microorganism used for expressing the protein. In particular, proinsulins

prepared from genetically modified *Escherichia coli* (*E. coli*) cells do not exhibit correctly bonded cystine bridges. The use of *E. coli*, however, for the production of heterologous proteins, such as insulin, has been widespread for the past decades (Pines and Inouye, 1999, Mol. Biotechnol., 12, 25-34), due to a favourable cost-value relation. *E. coli* bacteria are easy to cultivate, rapidly generate biomass and are accessible to high density fermentation, reflecting the fact that, compared to eukaryotic expression in systems such as yeast cells, bacterial expression harbours the advantages of good scalability and far shorter fermentation runs. These advantages are counterbalanced by the fact that *E. coli* cells are unable to perform most of the posttranslational modifications of proteins, e.g. the formation of correct cystine bridges (see above), required for the functioning of most eukaryotic proteins. Moreover, high-yield expression in *E. coli* often results in protein aggregation in the form of inclusion bodies, which are intracellular aggregates of insoluble, biologically inactive proteins. While the target protein is present within these inclusion bodies in a relatively concentrated form, it is necessary to separate other inclusion body proteins from the protein of interest, which, after this initial purification, has to be renatured and converted to its biologically active form. Generally, this *in vitro* renaturation process involves a complex series of biochemical reactions, the nature and order of which is highly dependent on the structural requirements of the protein to be converted (Rudolph and Lilie, 1986, FASEB, 10: 49-56; DeBernadez-Clark, 2001, Curr. Opinion Biotech, 12: 202-207). Accordingly, starting from a fully denatured precursor isolated from inclusion bodies, the final yields of the biologically active, "native" form of a renatured protein often are very low. Hence, even a modest increase in the yield of biologically active, renatured protein is considered a significant improvement of procedure, often associated with a clear commercial advantage. Common problems encountered in the production process of recombinant proteins include misfolding of proteins, protein aggregation as well as substantial protein loss associated with a multi-step process. While the problems of misfolding and aggregation are tractable by way of providing optimal refolding conditions, possibly including folding promoters such as high or low molecular weight chaperones, and low protein concentrations, protein loss can be prevented by optimizing reactions such as reducing processing steps to a minimum, thereby preventing e.g. adsorption to solid phase materials during multiple purification steps.

Thus, in the case of recombinantly produced human insulin or a derivative thereof, starting from human proinsulin or human proinsulin derivative expressed in a prokaryotic system, e.g. *E. coli*, a considerably large number of process steps have to be performed for obtaining a biologically active human insulin or human insulin derivative. A method for producing human insulin from *E. coli* cells (EP 0 055 945) is based on the following process steps: fermentation - cell lysis - purification of the inclusion bodies - isolation of the fusion protein - cyanogen halide cleavage of the fusion protein - isolation of the proinsulin cleavage product - S-sulfonation of cysteine residues (sulfitolysis) – chromatographic purification of the sulfonated proinsulin- refolding of proinsulin for the formation of correctly bonded cystine bridges - chromatographic purification of the proinsulin comprising correctly bonded cystine bridges - concentration - chromatographic purification of concentrated proinsulin - enzymatic cleavage of proinsulin to generate mature human insulin - chromatographic purification. The larger the number of process steps required, the more significant is the loss of protein, which eventually leads to a low yield of insulin. In EP 0 055 945, a loss of up to 40% of the proinsulin product starting from the stage of the isolated fusion protein to the stage of purification of the proinsulin product is described, with further substantial losses occurring during subsequent processing steps before obtaining the final product.

Among the reasons for the loss of proinsulin are the multiple purification and desalting steps necessitated by the change of buffer systems during the protein production process. Different pH conditions are required for solubilisation of inclusion bodies (alkaline pH), reduction of pre-existing disulfide bonds (alkaline pH), shuffling (alkaline pH, narrow range), precipitation of wrongly folded intermediates (neutral pH), purification and concentration (acidic pH). Therefore, the use of different buffering compounds has been described in the patents (e.g. EP0489780, EP0347781, CA02315750) and in the scientific literature (e.g. Winter et al., 2002; Son et al., 2008). Patents EP0489780 and EP0347781 describe the use of a tris(hydroxymethyl) aminomethane (i.e. Tris)- buffered solution for solubilisation of inclusion bodies and sulfonation of the solubilised proteins, followed by anion exchange chromatography in either glycine buffer (EP0489780) or Tris buffer (EP0347781), followed by refolding in glycine buffer (EP0489780; EP0347781) and subsequent chromatographic purification by ion exchange in glycine buffer (EP0347781).

CA02315750 discloses a Tris buffered solution (20 mM) for the steps of solubilisation of inclusion bodies and protein sulfonation, while the refolding step is conducted in a buffer containing 50 mM glycine. Winter et al. (*Analytical Biochemistry*, 2002, 310, 148-155) disclose the denaturation of human proinsulin in a Tris buffered system, followed by re-  
5 folding in an equimolar Tris/glycine buffered solution. Son et al. (*J. Microbiol. Biotechnol.*, 2008, 18(5); 983-989) describe the solubilisation of inclusion bodies in a Tris/glycine buffered solution and refolding in 10 mM glycine buffer. Thus, in those protocols, the exchange of buffer solutions between the process steps is required.

10 Furthermore, in protein purification procedures starting from inclusion bodies, solubilisation and refolding steps are usually carried out in the presence of a chaotropic auxiliary. A chaotrope is a substance which disrupts the three-dimensional structure of proteins by interfering with non-covalent forces in proteins, such as hydrogen bonds and van der  
15 Waals interactions, occurring between amino acids within a chain and between chains of a protein. Examples of chaotropic substances are ammoniumsulfate, ethylenecarbonate, thiocyanate, dimethylsulfonate, thiourea, lithiumperchlorate, guanidinium hydrochloride and urea. Predominantly, two of these chaotropic auxiliaries are used for the denaturation of  
20 proteins: guanidinium hydrochloride (GdnHCl) and urea. A disadvantage of using urea is the generation of ammonium cyanate, which is in equilibrium with urea in highly concentrated solutions (such as 6-8 M) of the chaotrope in an aqueous buffer, and which is capable of modifying proteins via carbamylation. During carbamylation of proteins, primary  
25 amines are modified irreversibly (N termini and  $\epsilon$ -amino groups of lysines; see: Cole, 1961, *J. Biol.Chem.* 236, 2670-2671), thus possibly resulting in unwanted modifications of protein structure (Oimomi et al., 1987, *Nephron*, 46, 63-66). Moreover, modifications of  
30 hydroxyl, sulfhydryl or imidazole moieties also are possible, but at alkaline pH, this reaction is reversible. Carbamylation can be prevented by adding cyanate scavengers to the solution, such as ethylene diamines, which work in the range of neutral pH and slightly below (US 4,605,513) or non-ethylene diamine compounds, such as glycinamide, histidine, 4-hydroxyl proline and various dipeptides (US 7,459,425). Albeit the efficiency of carbamylation prevention of the added compounds is excellent and ranges from 80% to 100% protection depending on the assay used, the compounds have to be added to the aqueous medium and cleared from it in the subsequent purification steps.

Surprisingly, it has now been found that the yield of correctly folded proinsulins or proinsulin derivatives can be increased and the overall time for the production process can be reduced significantly by conducting all steps as described herein in the same aqueous buffer system comprising an organic amphoteric substance containing at least two functional groups, comprising at least one amino group and one carboxylic acid group, and wherein the functional groups exhibit pKa values which lie between 2.0 and 10.0 and wherein the difference between the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance is at least 7. This organic amphoteric substance is, in addition to its function as an advantageous buffering compound, capable of reducing harmful by-products associated with the presence of high concentrations of urea.

#### **Detailed description of the invention**

The present invention is a method for producing proinsulin or a proinsulin derivate by recombinant expression in a prokaryotic organism comprising the steps

- A) Solubilisation of inclusion bodies,
- B) Purification of proinsulin or a proinsulin derivate from solubilised inclusion bodies using ion exchange chromatography,
- 20 C) Refolding proinsulin or a proinsulin derivate comprising native disulfide bonds,
- D) Precipitation of proinsulin or a proinsulin derivate containing non-native disulfide bonds,
- E) Concentration and purification of proinsulin or a proinsulin derivate comprising native disulfide bonds using ion exchange chromatography,
- 25

wherein all steps are carried out in an aqueous buffer comprising an organic amphoteric substance at a concentration between 0.001 M and 0.5 M containing at least two functional groups, comprising at least one amino group and one carboxylic acid group, and wherein the functional groups exhibit pKa values which lie between 2.0 and 10.0 and wherein the difference between the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance is at least 7.

An amphoteric substance is one that can react as either an acid or a base. Preferably, the amphoteric substance of the present invention is an organic substance which contains at least two functional groups, comprising at least one amino group and one carboxylic acid group. The functional groups exhibit pKa values which lie between 2.0 and 10.0 and the difference between the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance is at least 7. A pKa spread of at least 7 was found to be essential for achieving high yield with good purity of a proinsulin or a proinsulin derivate comprising native disulfide bonds,

The term pKa is a constant referring to the acid-ionization constant,  $K_a$ , to which is related by the equation:

$$\text{pKa} = -\log_{10} K_a \quad (1).$$

$K_a$  is the equilibrium constant for a dissociation reaction in the context of acid base reactions, hence it is a measure of the strength of an acid in solution. For the present invention, the term "pKa value" refers to a value obtained under conditions relevant for standard biochemical procedures. It is well known that the highest buffering capacity  $\beta$  is observed at a constant buffer concentration when

$$\text{pH} = \text{pKa} \quad (2).$$

20

The concentration of the organic amphoteric substance in the aqueous medium is between 0.001 M and 0.5 M, preferably between 0.01 M and 0.1 M, most preferred is a concentration between 0.02 M and 0.05 M to provide for sufficiently high buffering capacity,  $\beta$ , to which the concentration ( $C_{\text{buf}}$ ) and the pKa (as  $K_a$ ) value of the organic amphoteric substance are related by the following equation:

25

$$\beta = 2.303 \left( \frac{K_w}{[H^+]} + [H^+] + \frac{C_{\text{buf}} K_a [H^+]}{(K_a + [H^+])} \right) \quad (3)$$

30

$K_w$  refers to the dissociation constant of pure water; at 25 °C  $K_w$  is approximately equal to  $1.0 \times 10^{-14}$ . Preferred are amphoteric substances with the pKa value of one carboxylic acid group ranging from 2.0 to 2.5 and the pKa value of amino group ranging from 9.5 to 10.5.

By exhibiting at least two pKa values falling within the indicated ranges, excellent buffer capacity is achieved in the pH ranges required for the purification of proinsulin and derivatives thereof from inclusion bodies. Specifically, the process steps of (A) solubilisation of inclusion bodies and (B) purification of proinsulin or proinsulin derivate from solubilised inclusion bodies using ion exchange chromatography as well as (C) refolding proinsulin or a proinsulin derivate comprising native disulfide bonds are conducted within the alkaline pH range. In particular, step A is carried out at a pH ranging from pH = 9.0 to pH = 11.5, especially preferred at pH = 9.5. The result of inclusion body solubilisation at high alkaline pH is soluble denatured protein without significant amounts of secondary structure. Step (B) is a chromatographic purification step required for separation of the fully denatured and reduced precursor protein from other inclusion body proteins. Bead- or membrane-based anion exchange chromatography, preferably employing quaternary amines as functional groups, is conducted at pH = 9.0 to pH = 11.5, especially preferred at pH = 9.5. In this step, proinsulin or proinsulin derivative are purified to some extent from the chaotropic denaturant as well as the reducing agent, which do not bind to the functional groups of the anion exchanger at the pH applied. The chaotropic denaturant as well as reducing agent have to be added back to keep the proinsulin or proinsulin derivative in a fully reduced and denatured state until initiation of the refolding reaction. Small amounts of both compounds are required in this subsequent reaction (step C). For the refolding reaction (C), a pH range of pH = 9.5 to pH = 10.5 is preferred, especially preferred at pH = 10.5. The reaction is stopped (D) by adjusting the pH to an acidic value, which is preferably between pH = 4.8 and pH = 5.6, most preferred to pH = 5.2, whereby proinsulin or proinsulin derivate containing non-native disulfide bonds are precipitated and can be isolated by a simple centrifugation or filtration step. In the subsequent step (E), bead- or membrane-based ion exchange chromatography is used to purify proinsulin or proinsulin derivate containing native disulfide bonds utilizing cation exchange at a pH ranging from pH = 2.5 to pH = 4.5, especially preferred at pH = 3.5, followed by further purification (F) using anion exchange at a pH ranging from pH = 8.5 to pH = 10.5, especially preferred at pH = 9.5.

30

The preferred organic amphoteric substances of the invention to be used as a buffer system are amino acids. Preferably, the organic amphoteric substance exhibits a suitable pKa



spread and an excellent solubility in a number of different solvents, in particular water, at a range of different temperatures, to yield a buffer compound with a sufficient buffer capacity b in the required pH ranges to be used in the context of the protein processing and purification procedure of the present invention. In particular, it is an amino acid exhibiting  
5 a water solubility of greater than 150 g/l at 20°C, which is required to achieve good buffer capacity.

Specifically suitable are the amino acids glycine, arginine, alanine, beta-alanine, serine or proline. In the present invention, it has been found that the amino acids glycine, alanine  
10 and serine are the most adequate organic amphoteric buffer substances in terms of buffering capacity, pKa spread and water solubility, most preferred is glycine. Especially, regarding buffering capacity, it was observed that glycine is a high strength compound in comparison to other amino acid buffers. Thus, concentration and pH-dependent tendency for aggregation, which proinsulin is prone to at neutral to slightly acidic pH even at low  
15 concentrations of protein, is diminished significantly using glycine as a buffering compound within the specified concentration range.

The most preferred organic amphoteric substance, glycine, exhibits high water solubility (25g/100g water at 25°C) and the most favourable pKa spread (2.35; 9.78). By comparison,  
20 other amino acids with high water solubility, such as arginine (15g/100g water at 25°C), alanine (16.5 g/100 g water at 25°C), serine (5 g/100 g water at 25°C), proline (162 g/100 g water at 25°C) and cysteine (>50 g/100 g water at 25°C) also can be used advantageously as amphoteric substances. Glycine, however, is the most preferred amino acid buffer substance in comparison, for the following reasons: for arginine and proline, the spread of  
25 their respective pKa values is not as suitable (arginine: 2.01, 9.4, 12.48; proline: 2.0, 10.6), for cysteine, the pKa spread (2.05, 10.25, 8.00) is more satisfactory, but it comprises a side chain with a functional group, which could react with the protein to be purified under the conditions utilized. Furthermore, considering also the ease and cost efficiency of preparation, glycine is considered inexpensive (unlike serine, alanine and beta-alanine).

30

In addition, the organic amphoteric substance of the invention efficiently acts as a cyanate scavenger, maintaining the cyanate concentration in an aqueous solution containing urea in

a concentration range of 5 M to 8 M at concentrations lower than 1.5 mM for a time period of at least 7 days. This is particularly advantageous, as primary amines (e.g. present at the N termini and the side chains of lysine and arginine) are prone to irreversible modification by carbamylation at the pH range required for solubilisation of inclusion bodies and refolding (i.e. high alkaline pH values). Carbamylation of these residues, resulting in N-terminal carbamine, epsilon-carbamyllysine and N-carbamylarginine, disrupts their capability to form native hydrogen bonds, which stabilize intra- and interchain interactions, thus resulting in modifications of protein structure, function and/or stability. While polyamines, especially ethylene diamines, have been described as efficient cyanate scavengers (US 4,605,513; US 76 459, 425 B2) here it is surprisingly found that the amino acid glycine can scavenge the decomposition products associated with high urea concentrations (i.e. ammoniumcyanate) with a comparable efficiency, thereby leading to an improved situation during the solubilisation and refolding step in comparison with other patented procedures, wherein an additional molecular scavenger has to be added and subsequently removed in the overall production process for e.g. proinsulin or a proinsulin derivative comprising native disulfide bonds.

The presence of an efficient cyanate scavenger in the buffer thus enables the use of urea as a chaotropic auxiliary for solubilisation of inclusion bodies containing human proinsulin at high alkaline pH values. As an alternative, among others, guanidinium hydrochloride could be utilized. Urea, however, is the most preferred chaotropic auxiliary, as it is an uncharged molecule (in contrast to guanidinium hydrochloride) and therefore compatible with chromatographic purification steps based on ion exchange, which are employed in the present invention in steps (B), (E) and (F). These chromatographic techniques bear relatively high economic (cost of material, flow volumes) and procedural (time) advantages compared to other concentration- and purification-techniques such as dia- and/or ultrafiltration.

While the amino acids glycine, alanine and serine are the most suitable organic amphoteric buffer substances in terms of buffering capacity, pKa spread and water solubility, with glycine being the most preferred compound, for the precipitation step (D), buffering capacity in the mildly acidic to neutral range is required. In the present invention, it was found surprisingly, that by adding small concentrations of a second buffer substance which

has at least one pKa value within the central 50% of the range defined by the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance, the pH could be stabilized efficiently to allow for quantitative precipitation of misfolded proteins while not affecting significantly proteins with correctly bonded cysteine bridges. The preferred pH range for (D) is between pH = 4.8 and pH = 5.6. Therefore, substances characterized by at least one pKa between 4.5 and 5.9 are best suited for buffering in this range. In particular, it was found that mono- and dicarboxylic acids of the general chemical formulae R-COOH and HOOC-R-COOH, respectively, including e.g. acetic, propanoic, butyric, valeric or benzoic acid and e.g. succinic, glutaric, adipic, pimelic, suberic, or azelaic acid, can efficiently stabilize the pH within the range most suitable for precipitation of misfolded proinsulin or derivatives thereof. Most preferred, a dicarboxylic acid is added to the aqueous medium comprising one first buffer substance at the end of the refolding reaction (step D). Dicarboxylic acids are preferred over monocarboxylic acids as they exhibit two pKa values which allow for a greater spread of buffering capacity in the indicated pH range and hence provide for more flexibility as to the pH at which precipitation is carried out. Moreover, water solubility of dicarboxylic acids in general is greater for comparable C- chain lengths, e.g. glutaric acid (HOOC-(CH<sub>2</sub>)<sub>3</sub>-COOH; 64.0 g/100ml) is more soluble than valeric acid (C<sub>4</sub>H<sub>9</sub>COOH; 4.97g/100ml). Especially preferred in the context of the invention are the dicarboxylic acids succinic acid and glutaric acid, with the highest preference given to succinic acid due to its most favourable spread of pKa values (pK<sub>a1</sub> = 4.2; pK<sub>a2</sub> = 5.6).

Thus, by using a comprehensive buffer system for renaturing human proinsulin or proinsulin derivatives, wherein the buffer provides excellent buffering capacity at the divergent pH ranges necessary for adequate processing of the protein in question, while at the same time acting as a scavenger for decomposition products of essential buffer components, a simple and shortened procedure can be used for the renaturation of human proinsulin with increased yield of the native biologically active form.

In the following, the present invention will be described with respect to working examples. Reference is made in this respect to the following figures:

FIG. 1: Optimization of the precipitation conditions: Semi-quantitative gel densitometry revealed that at pH=5.2 most of the misfolded proinsulin, but only minimal amounts of correctly folded proinsulin, was precipitated.

5 Fig. 2: Concentration and purification of the correctly folded proinsulin: Analytical RP-HPLC of purified human proinsulin after the final purification step.

### EXAMPLES

10 1. Scavenger activity of 50 mM glycine buffer, pH 9.5, calibration curve

Cyanate was added to a buffer solution containing 200 mM sodiumdiarsinate at a pH of 9, 9.5 or 11. Utilizing a modification of the method described by Warner (1942), copper-pyridine-cyanate complexes could be detected by measuring photometrically the absorbance at a wavelength  $\lambda = 700$  nm in a BioRad Smartspec Plus spectrophotometer. By using cyanate concentrations within a range relevant in the context of protein purification (i.e.: 0, 1.2, 2.5, 3.7, 4.9, 7.4 mM) a near linear relationship was obtained with a coefficient of determination,  $R^2$ , approximating 0.97.

20 The scavenging activity of 50 mM glycine buffer, pH 9.5, was determined in the presence of 5 M urea during the time course of 7 d at room temperature. For controls, no glycine was added.

	Cyanat [mM]
pH 9	0
pH 11	0
pH 9, 7 d	9
pH 11, 7 d	3
Glycin, pH 9,5, 0 d	0
Glycin, pH 9,5, 2 d	0
Glycin, pH 9,5, 3 d	0
Glycin, pH 9,5, 4 d	1

25 From this data a calibration curve can be generated by using the modified method according to Warner (J. Biol. Chem., 142, 705; 1942).

## 2. Solubilisation of inclusion bodies

Pre-purified inclusion bodies isolated from approximately  $1.5 \cdot 10^{13}$  *E. coli* cells expressing the appropriate construct (e.g. human proinsulin or a derivative thereof) were dissolved in 30 ml of a 50 mM glycine buffer at pH 9.5 and solubilised by adding solid urea to a final concentration of 7.5 M. Moreover, 100 mM beta-mercaptoethanol was added to the solution to provide a strongly reducing environment. Incubation was conducted for 1 hour at 20-37°C using slight agitation (stirring or shaking); subsequently, insoluble cellular debris was discarded using centrifugation or, as an alternative, microfiltration.

## 3. Purification of fully denatured and reduced proinsulin

The supernatant/filtrate was purified directly by column-based anion exchange chromatography, using membrane based exchangers exhibiting a high dynamic binding capacity (e.g. Mustang-Q, binding capacity: 34 mg/ml; 5 ml solidphase material). During the washing of the loaded column using 50 mM glycine buffer at pH 9.5, the concentrations of urea and beta-mercaptoethanol were lowered to 5M urea and 5 mM beta-mercaptoethanol. Fully denatured and reduced proinsulin was eluted using a linear gradient of NaCl (0-1000 mM, in 50 mM glycine buffer at pH 9.5 supplemented with 5 M urea and 5 mM [beta]-mercaptoethanol). Concentration of the precursor in the combined elution fractions was determined either by using an assay according to Bradford (Bradford, 1976) or by spectrophotometry (absorbance assay at  $\lambda = 280\text{nm}$ ; fully reduced denatured proinsulin has a molar extinction coefficient of  $[\epsilon] = 5960 \text{ l}/(\text{mol} \cdot \text{cm})$ ) and varied between 5 to 6 g/l.

## 4. Renaturation of the purified proinsulin: refolding reaction and precipitation of misfolded proinsulins

The elution fractions were diluted approximately 20-fold to a protein concentration of 0.1 – 1,5 g/l in refolding buffer, which contained 50 mM glycine, pH 10.5, and urea adjusted to a final concentration of 1 M. The reducing agent carried over from anion exchange chromatography (approximately 0.25 mM beta-mercaptoethanol) was complemented by adding 1 mM cystine (oxidized cysteine) for formation of a mixed redox system. The refolding mixture was incubated for 12-14 hours either at room temperature or at 4°C using slight agitation (stirring) under an air atmosphere. After the indicated time, succinic acid was added to the reaction mixture to a final concentration of 5 mM for pH stabilization

prior to stopping the refolding reaction by adjusting the pH to 5.2 using 1 M HCl. Precipitation of misfolded proinsulin molecules was allowed to proceed for 1 hour and RT under continuous stirring. Precipitates were separated by paper filtration using Whatman No. 1 filter paper. Precipitate can be denatured again in 50 mM glycine buffer, pH 9.5 containing  
5 urea at 5 M and beta-mercaptoethanol at 100 mM, upon which the fully denatured and reduced proinsulin can be subjected to an additional folding reaction to enhance overall yield.

#### 5. Optimization of the precipitation conditions.

10 DE3501641 disclosed the precipitation of misfolded proinsulins by acidifying to conditions slightly below their isoelectric point. However, the exact conditions vary with the particular amino acid sequence and have to be determined empirically.

Proinsulin folded according to example 4 was supplemented with 5 mM succinic acid. Aliquots were acidified with hydrochloric acid to pH=5.0, pH=5.2, pH=5.4 and pH=5.6, respectively, and the mixture allowed to incubate for another hour under gentle stirring. The precipitates formed were removed by centrifugation (10 min at 15000 g) and precipitates (P) and supernatants (S) were analyzed by SDS-PAGE.

20 According to Fig. 1, semi-quantitative gel densitometry revealed that at pH=5.2 most of the misfolded proinsulin, but only minimal amounts of correctly folded proinsulin, was precipitated.

#### 6. Concentration and purification of the correctly folded proinsulin

25 The pH of the filtrate obtained from the filtration step was adjusted to 3.5 using 1 M HCl and loaded directly onto a pre-equilibrated cation exchange column, e.g. S-Sepharose (GE Healthcare) with a bed volume of 10 ml. At this pH, proinsulin is positively charged and can be eluted by reversing the charges via a rapid pH change. Accordingly, the cation exchange column was washed using 20mM glycine buffer, pH 3.5, ( i.e. low strength), followed by elution in 100 ml of 50mM glycine buffer, pH 9.5 (high strength). The first 25 ml  
30 were discarded, as they contained no proinsulin. The combined proinsulin-containing fractions had a final concentration of 6-8 g/l. The elution fractions were subjected directly to

anion exchange chromatography (e.g. Mustang Q or Q Sepharose) for a final purification step. The column-bound proinsulin was washed with 20 mM Glycin pH=8.5 and eluted with 150 mM NaCl/20 mM Glycine pH=8.5.

5 At this point, the proinsulin was judged to be 99 % pure by analytical reverse phase high performance liquid chromatography (RP-HPLC, Fuji Silysia Chromatorex C18, pore size 10 nm, particle size 10 [micro]m, 4,6\*250 mm); see figure 2.

10 The eluted material can be directly subjected to enzymatic processing with trypsin and carboxypeptidase B.

The invention relates to an improved processing and purification method for the production of proinsulin or derivatives thereof starting from the solubilisation of inclusion bodies isolated from a prokaryotic expression system, such as *E. coli*, whereby all processing (i.e. refolding, precipitation, concentration) and purification steps are carried out in the same aqueous buffer system comprising an organic amphoteric substance containing at least two functional groups, which contain at least an amino and a carboxylic acid group, and which is characterized by a difference in the pKa values of its at least one amino group and at least one carboxylic acid group of at least 7. Furthermore, the organic amphoteric substance is an efficient scavenger of cyanate, thereby preventing irreversible carbamylation of primary amines present in proinsulin or the derivatives thereof, which has been observed in the presence of high concentrations of urea. Thus, the disadvantage associated with this widespread chaotropic auxiliary is obliterated, allowing the utilization of efficient and uncomplicated chromatography techniques for purification steps subsequent to solubilisation and refolding. The use of an efficient buffer substance in all processing steps from inclusion body solubilisation up to the re-folded, concentrated and purified precursor of insulin or a derivative thereof, which at the same time acts as an agent preventing protein modification by carbamylation, thereby abolishing the need to add an additional compound with scavenging activity, leads to significant advantages regarding the efficiency and efficacy of the procedure compared to previously employed production methods.

15  
20  
25  
30

### Claims

1. A method for producing proinsulin or a proinsulin derivate by recombinant expression in a prokaryotic organism comprising the steps

- 5           A) Solubilisation of inclusion bodies,  
          B) Purification of proinsulin or a proinsulin derivate from solubilised inclusion bodies using ion exchange chromatography,  
          C) Refolding proinsulin or a proinsulin derivate comprising native disulfide bonds,  
10          D) Precipitation of proinsulin or a proinsulin derivate containing non-native disulfide bonds,  
          E) Concentration of proinsulin or a proinsulin derivate comprising native disulfide bonds using ion exchange chromatography,  
          F) Purification of proinsulin or a proinsulin derivate comprising native disulfide bonds using ion exchange chromatography,  
15

          wherein all steps are carried out in an aqueous buffer comprising an organic amphoteric substance at a concentration between 0.001 M and 0.5 M containing at least two functional groups, comprising at least one amino group and one carboxylic acid group,  
20

          wherein the functional groups exhibit pKa values which lie between 2.0 and 10.0 and wherein the difference between the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance is at least 7.

25          2. A method according to claim 1, wherein steps A) and B) are carried out at a pH between pH = 9.0 and pH = 11.5.

3. A method according to any of the claims 1 or 2, wherein step C) is carried out at a pH between pH=9.5 and pH= 10.5.

30

4. A method according to any of the claims 1 to 3, wherein step D) is carried out at a pH between pH=4.8 and pH= 5.6.



5. A method according to any of the claims 1 to 4, wherein step E) is carried out at a pH between pH=3.5 and pH= 4.5.
- 5 6. A method according to any of the claims 1 to 5, wherein step F) is carried out at a pH between pH=8.5 and pH= 10.5.
7. A method according to claim 1 or 6, wherein the organic amphoteric substance is an amino acid with a water solubility of greater than 150 g/l at 20°C.
- 10 8. A method according to claim 1 or 7, wherein the organic amphoteric substance is in one or more of: glycine, arginine, alanine, beta-alanine, serine, cysteine or proline.
- 15 9. A method according to claim 1 or 8, wherein the organic amphoteric substance maintains the cyanate concentration in an aqueous solution containing urea in a concentration range of 5 M to 8 M at concentrations lower than 1.5 mM.
- 20 10. A method according to any of the claims 1 to 9, wherein the aqueous buffer comprises a second buffer substance which has at least one pKa value within the central 50% of the range defined by the pKa values of at least one amino group and at least one carboxylic acid group of the amphoteric substance.
- 25 11. A method according to claim 10, wherein at least one pKa value of the second buffer substance is between 4.5 and 5.9.
- 30 12. A method according to claim 10 or 11, wherein the second buffer substance is a monocarboxylic acid, in particular, acetic, propanoic, butyric, valeric or benzoic acid.

13. A method according to claim 10 to 12, wherein the second buffer substance is a dicarboxylic acid, in particular, succinic, glutaric, adipic, pimelic, suberic, or azelaic acid.
- 5 14. A method according to claim 10 to 12, wherein the second buffer substance is succinic acid.

Figure 1

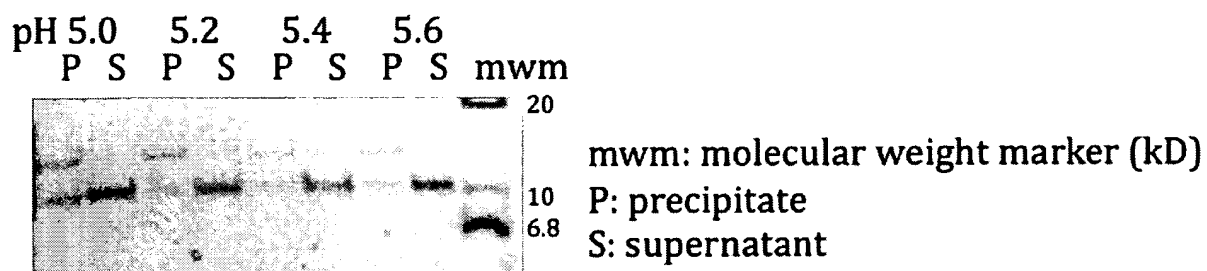
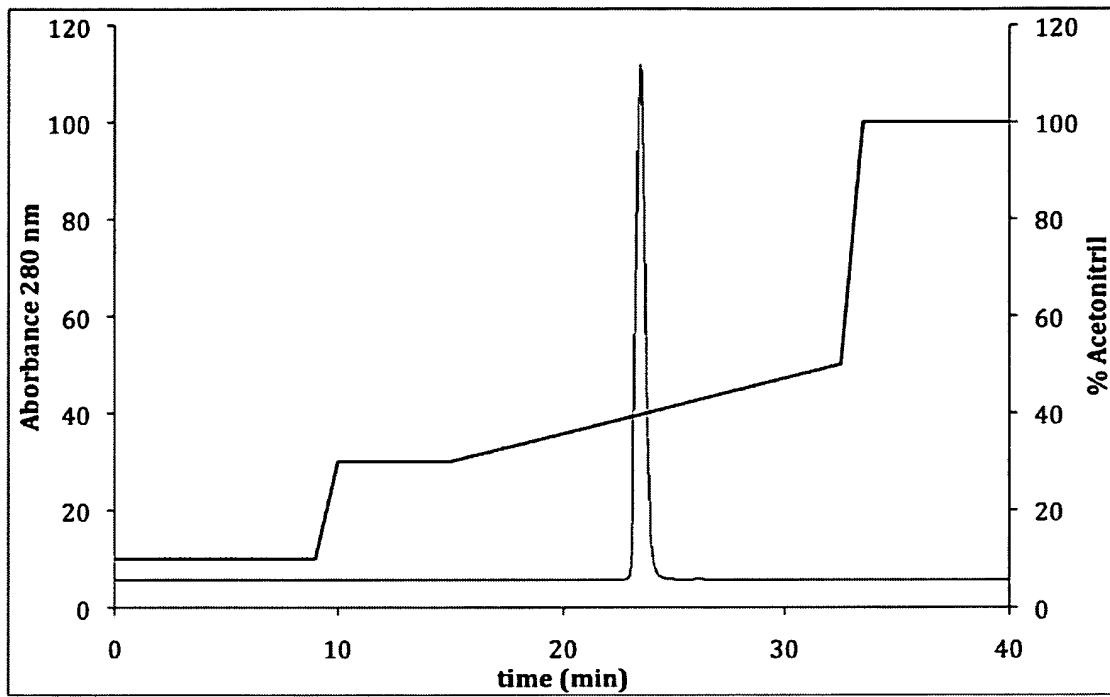


Figure 2



## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2012/001992

A. CLASSIFICATION OF SUBJECT MATTER INV. C07K14/62 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, BIOSIS, EMBASE, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	EP 2 374 888 A1 (LTD LIABILITY COMPANY MAKO [UA]) 12 October 2011 (2011-10-12) abstract paragraph [0027] - paragraph [0036]; examples 8-13	1-14
A	-& WO 2010/062279 A1 (MAKO LLC) 3 June 2010 (2010-06-03) abstract	1-14
A	----- US 5 514 646 A (CHANCE RONALD E [US] ET AL) 7 May 1996 (1996-05-07) abstract; example 28	1-14
A	----- EP 0 668 292 A2 (HOECHST AG [DE]) 23 August 1995 (1995-08-23) abstract page 5, line 30 - page 6, line 40; examples 1,2 ----- -/--	1-14
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search  18 July 2012	Date of mailing of the international search report  26/07/2012	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  Mossier, Birgit	

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2012/001992

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2009/133529 A2 (WOCKHARDT RESEARCH CENTER [IN]; EDUPUGANTI B RAJU [IN]; JAGIRDAR HASEE) 5 November 2009 (2009-11-05) abstract paragraph [0030] - paragraph [0043] -----	1-14
A	WO 99/33988 A1 (CHONG KUN DANG CORP [KR]; HONG CHUNG IL [US]; KIM JUNG WOO [KR]; LEE W) 8 July 1999 (1999-07-08) abstract page 5, line 32 - page 8, line 15; examples 1-13; table 1 -----	1-14
A	OLMOS J ET AL: "Production in Escherichia coli of a rat chimeric proinsulin polypeptide carrying human A and B chains and its preparative chromatography", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 38, no. 1, 30 November 1994 (1994-11-30), pages 89-96, XP023788198, ISSN: 0168-1656, DOI: 10.1016/0168-1656(94)90150-3 [retrieved on 1994-11-30] abstract page 92, column 1, paragraph 2 - page 93, column 1, paragraph 1 -----	1-14
A	CHANG S-G ET AL: "HUMAN INSULIN PRODUCTION FROM A NOVEL MINI-PROINSULIN WHICH HAS HIGH RECEPTOR-BINDING ACTIVITY", BIOCHEMICAL JOURNAL, THE BIOCHEMICAL SOCIETY, LONDON, GB, vol. 329, 1 January 1998 (1998-01-01), pages 631-635, XP001037629, ISSN: 0264-6021 abstract -----	1-14
A	VALLEJO LUIS FELIPE ET AL: "Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins", MICROBIAL CELL FACTORIES, BIOMED CENTRAL, LONDON, NL, vol. 3, no. 1, 2 September 2004 (2004-09-02), page 11, XP021007184, ISSN: 1475-2859, DOI: 10.1186/1475-2859-3-11 the whole document -----	1-14

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/EP2012/001992
---------------------------------------------------

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
EP 2374888	A1	12-10-2011	CA 2746984 A1	03-06-2010
			CN 102282260 A	14-12-2011
			EA 201100781 A1	30-01-2012
			EP 2374888 A1	12-10-2011
			US 2012058513 A1	08-03-2012
			WO 2010062279 A1	03-06-2010
			-----	
WO 2010062279	A1	03-06-2010	CA 2746984 A1	03-06-2010
			CN 102282260 A	14-12-2011
			EA 201100781 A1	30-01-2012
			EP 2374888 A1	12-10-2011
			US 2012058513 A1	08-03-2012
			WO 2010062279 A1	03-06-2010
			-----	
US 5514646	A	07-05-1996	GE P20074095 B	25-04-2007
			US 5514646 A	07-05-1996
			US 5700662 A	23-12-1997
-----				
EP 0668292	A2	23-08-1995	AT 166069 T	15-05-1998
			AU 699817 B2	17-12-1998
			AU 1228895 A	31-08-1995
			CA 2142780 A1	19-08-1995
			DE 4405179 A1	24-08-1995
			DK 668292 T3	23-11-1998
			EP 0668292 A2	23-08-1995
			ES 2119241 T3	01-10-1998
			FI 950699 A	19-08-1995
			HK 1010457 A1	17-06-1999
			JP 3863579 B2	27-12-2006
			JP 7265092 A	17-10-1995
			NO 950592 A	21-08-1995
			SG 46683 A1	20-02-1998
			US 5663291 A	02-09-1997
-----				
WO 2009133529	A2	05-11-2009	CA 2724510 A1	05-11-2009
			EP 2274318 A2	19-01-2011
			JP 2011519369 A	07-07-2011
			US 2011152507 A1	23-06-2011
			WO 2009133529 A2	05-11-2009
-----				
WO 9933988	A1	08-07-1999	AU 735480 B2	12-07-2001
			AU 1694199 A	19-07-1999
			BR 9814526 A	17-10-2000
			CA 2315750 A1	08-07-1999
			CN 1290299 A	04-04-2001
			EP 1042479 A1	11-10-2000
			JP 2002500013 A	08-01-2002
			RU 2203949 C2	10-05-2003
			US 5952461 A	14-09-1999
			WO 9933988 A1	08-07-1999
			-----	