

# Insulin analog with additional disulfide bond has increased stability and preserved activity

Tine N. Vinther,<sup>1,2</sup> Mathias Norrman,<sup>1</sup> Ulla Ribel,<sup>1</sup> Kasper Huus,<sup>1</sup> Morten Schlein,<sup>1</sup> Dorte B. Steensgaard,<sup>1</sup> Thomas Å. Pedersen,<sup>1</sup> Ingrid Pettersson,<sup>1</sup> Svend Ludvigsen,<sup>1</sup> Thomas Kjeldsen,<sup>1</sup> Knud J. Jensen,<sup>2</sup> and František Hubálek<sup>1\*</sup>

<sup>1</sup>Diabetes Research Unit, Novo Nordisk A/S, Novo Nordisk Park, Måløv DK-2760, Denmark

<sup>2</sup>Department of Chemistry, University of Copenhagen, Faculty of Science, Frederiksberg DK-1871, Denmark

Received 4 October 2012; Revised 4 December 2012; Accepted 7 December 2012

DOI: 10.1002/pro.2211

Published online 26 December 2012 proteinscience.org

**Abstract:** Insulin is a key hormone controlling glucose homeostasis. All known vertebrate insulin analogs have a classical structure with three 100% conserved disulfide bonds that are essential for structural stability and thus the function of insulin. It might be hypothesized that an additional disulfide bond may enhance insulin structural stability which would be highly desirable in a pharmaceutical use. To address this hypothesis, we designed insulin with an additional interchain disulfide bond in positions A10/B4 based on C $\alpha$ -C $\alpha$  distances, solvent exposure, and side-chain orientation in human insulin (HI) structure. This insulin analog had increased affinity for the insulin receptor and apparently augmented glucodynamic potency in a normal rat model compared with HI. Addition of the disulfide bond also resulted in a 34.6°C increase in melting temperature and prevented insulin fibril formation under high physical stress even though the C-terminus of the B-chain thought to be directly involved in fibril formation was not modified. Importantly, this analog was capable of forming hexamer upon Zn addition as typical for wild-type insulin and its crystal structure showed only minor deviations from the classical insulin structure. Furthermore, the additional disulfide bond prevented this insulin analog from adopting the R-state conformation and thus showing that the R-state conformation is not a prerequisite for binding to insulin receptor as previously suggested. In summary, this is the first example of an insulin analog featuring a fourth disulfide bond with increased structural stability and retained function.

**Keywords:** disulfide; evolution; insulin; protein design; stability; yeast

*Abbreviations:* CD, circular dichroism; DSC, differential scanning calorimetry; HI, human insulin; IGF, insulin-like growth factor; SEC, size-exclusion chromatography; SPA, scintillation proximity assay.

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: All authors employed at Novo Nordisk own Novo Nordisk stocks.

Grant sponsor: Danish Ministry of Science, Technology and Innovation and the Novo Nordisk STAR program.

\*Correspondence to: František Hubálek, Diabetes Research Unit, Novo Nordisk A/S, Novo Nordisk Park, Måløv DK-2760, Denmark. E-mail: fhub@novonordisk.com

## Introduction

Insulin is a small peptide hormone which main function is to regulate blood glucose levels and lack in insulin secretion and/or signaling leads to the severe metabolic disorder, diabetes mellitus.<sup>1</sup> Human insulin (HI) consists of 51 amino acids divided into two chains, the A-chain (amino acids A1-A21) and the B-chain (amino acids B1-B30). The classical insulin fold which is shared by members of the insulin super-family includes three  $\alpha$ -helices, two in the A-chain and one in the B-chain, and a hydrophobic core of non-polar residues which are important for folding and maintaining the structure.<sup>2,3</sup> The two

chains are linked by two interchain disulfide bonds between residues A7-B7 and A20-B19. The A-chain also contains an intra-chain disulfide bond between residues A6-A11.<sup>4,5</sup> In the presence of zinc, the monomers assemble into hexamers where each of the two central zinc ions is coordinated by three HisB10 residues. In the hexameric form the HI molecule exists in two distinct conformations called the T-state and the R-state deviating in the length of the central  $\alpha$ -helix in the B-chain.<sup>2,6,7</sup> In the R-state there is a continuous  $\alpha$ -helix from residue B1 to B19, whilst in the T-state the first eight residues are found in an extended conformation.

In all vertebrate insulin homologues including the hagfish,<sup>8</sup> which is believed to be the earliest vertebrate ancestor, the pattern with three disulfide bonds is highly conserved.<sup>9</sup> The pattern is also found in insulin-like growth factor (IGF-1) and in a common ancestor of insulin and IGF-1, the insulin-like peptides from the primitive chordate amphioxus, and it is therefore believed that the insulin structure originates more than 1 billion years ago.<sup>10</sup> The disulfide bonds in insulin are crucial for maintaining a functional structure and removal of any of the three disulfide bonds, results in loss of function.<sup>11,12</sup> Even though insulin contains three stabilizing disulfide bonds, the insulin molecule is still prone to formation of amyloid fibrils and the stability is also challenged when stored at elevated temperatures. Stabilization of insulin in pharmaceutical formulations is therefore often achieved by addition of zinc ions as the hexameric form of insulin is more stable compared with the monomeric form. The hexameric form can be further stabilized by addition of small aromatic alcohols (phenol, m-cresol) and salt which further stabilizes the R-state.<sup>7</sup> Despite the stabilisation through formulation there is still a desire to further increase insulin's stability. Such insulin analogs would likely have longer in use time, not require refrigeration and could be used in continuous insulin pumps.

Addition of disulfide bonds has previously been used in attempts to increase stability of proteins.<sup>13–18</sup> The reason for disulfide stabilization has been attributed to both entropic and enthalpic effects. One theory suggests that a decrease in entropy of the unfolded state is causing destabilisation of this state compared with the folded state.<sup>19</sup> Another theory suggests, it is a result of an increase in enthalpy for the folded state compared with the unfolded state as a result of burial of non-polar groups.<sup>20</sup> No clear evidence of either of the theories has been found and the answer to the question is more likely a combination of the two depending on the sequence of the protein.<sup>21,22</sup>

Different strategies of disulfide bond introduction have been pursued with both intra- and intermolecular locations, for overview and examples see

reference.<sup>14</sup> Many attempts resulted in protein destabilization. Most successful efforts involved intramolecular disulfide bonds between subunits but there have also been some successful attempts with intermolecular disulfide bonds with some retaining their activities. Selecting new positions for introduction of an additional stabilizing disulfide bond is not trivial as there are strict stereochemical requirements for the relative position and orientation of the two involved Cys.<sup>14,23</sup>

The introduction of a fourth disulfide into the relatively small insulin molecule poses a challenge to the yeast expression system. In the yeast expression system insulin is folded and the three existing disulfide bonds are formed during expression. Introduction of two new Cys residues requires that the fourth disulfide bond is also easily formed during expression to avoid disulfide bond scrambling and misfolding of the protein.

A few examples of an additional disulfide bonds in members of the insulin super-family have been found in genomes of invertebrates,<sup>24–26</sup> however, these insulin analogs have limited affinity towards HI receptor.<sup>27</sup>

In this article, we demonstrate that it is possible to introduce a fourth disulfide bond into the insulin structure and express the analog in yeast with correctly formed disulfide bonds. The structure of this new analog resembles that of HI. This analog appears to be more active than HI, is also highly stabilized in terms of thermodynamic stability and resistance to fibrillation as a result of the additional disulfide bond. Moreover, this analog provides important clues concerning insulin binding to its receptor and the process of insulin fibril formation.

## Results

We selected positions for introduction of the two Cys residues based on  $C\alpha$ – $C\alpha$  distances, solvent exposure and side-chain orientation in the HI structure PDB 1MSO.<sup>28</sup> The known receptor binding positions were avoided<sup>29</sup> and the information from our recently established ability to express analogs with single Cys substitutions<sup>30</sup> was used.

Based on these criteria, one position, A10/B4, was selected for introduction of an additional interchain disulfide bond. The insulin precursor was cloned and expressed in yeast. The expression yield (RP-HPLC) of the CysA10/CysB4 precursor was ~70% relative to HI using the standard expression system when analyzed directly in the yeast culture supernatant without refolding. The precursor was converted to a desB30 mature analog by enzymatic removal of the extension and mini C-peptide and purified by RP-HPLC. Insulin lacking the B30 threonine has previously been shown to have full activity.<sup>31</sup>

The structure of the CysA10/CysB4/desB30 analog (4SS-insulin) was solved by X-ray crystallography to a resolution of 2.0 Å (see Table I for data collection and refinement statistics). The crystals belong to the rhombohedral crystal system with space group R3 with two insulin molecules in the asymmetric unit. By applying appropriate symmetry operations an insulin hexamer could be formed with an overall structure reminiscent of a HI hexamer in a T-state [Fig. 1(A)]. The fourth disulfide bond was clearly seen in the electron density [Fig. 1(B)] as were the three classical disulfide bonds, showing that the yeast cells were capable of formation of the fourth disulfide bond together with the three existing disulfide bonds. A comparison with a HI structure 1MSO.PDB gave an RMS distance between C $\alpha$  atoms of 0.62 Å. A chain-wise comparison between the 4SS-insulin and 17 crystal structures of insulin, all residing in the rhombohedral crystal system, space group R3, with the B chain in an extended T-state conformation is shown in Figure 1(C). Interesting differences were found in the N-terminus of the B-chain of both molecules in the asymmetric unit and in the N-terminus of the A-chain in one of the molecules. From numerous structures of insulin available, the N-terminal part of the B-chain is known to be rather flexible and susceptible to crystal packing- and/or intra hexamer interactions when it is in the T-state conformation. In the 4SS-insulin structure, the additional disulfide bond kept the residues B1-B4 in a constrained conformation with a positive B8  $\Phi$  dihedral angle of 46.2° and 42.4° for

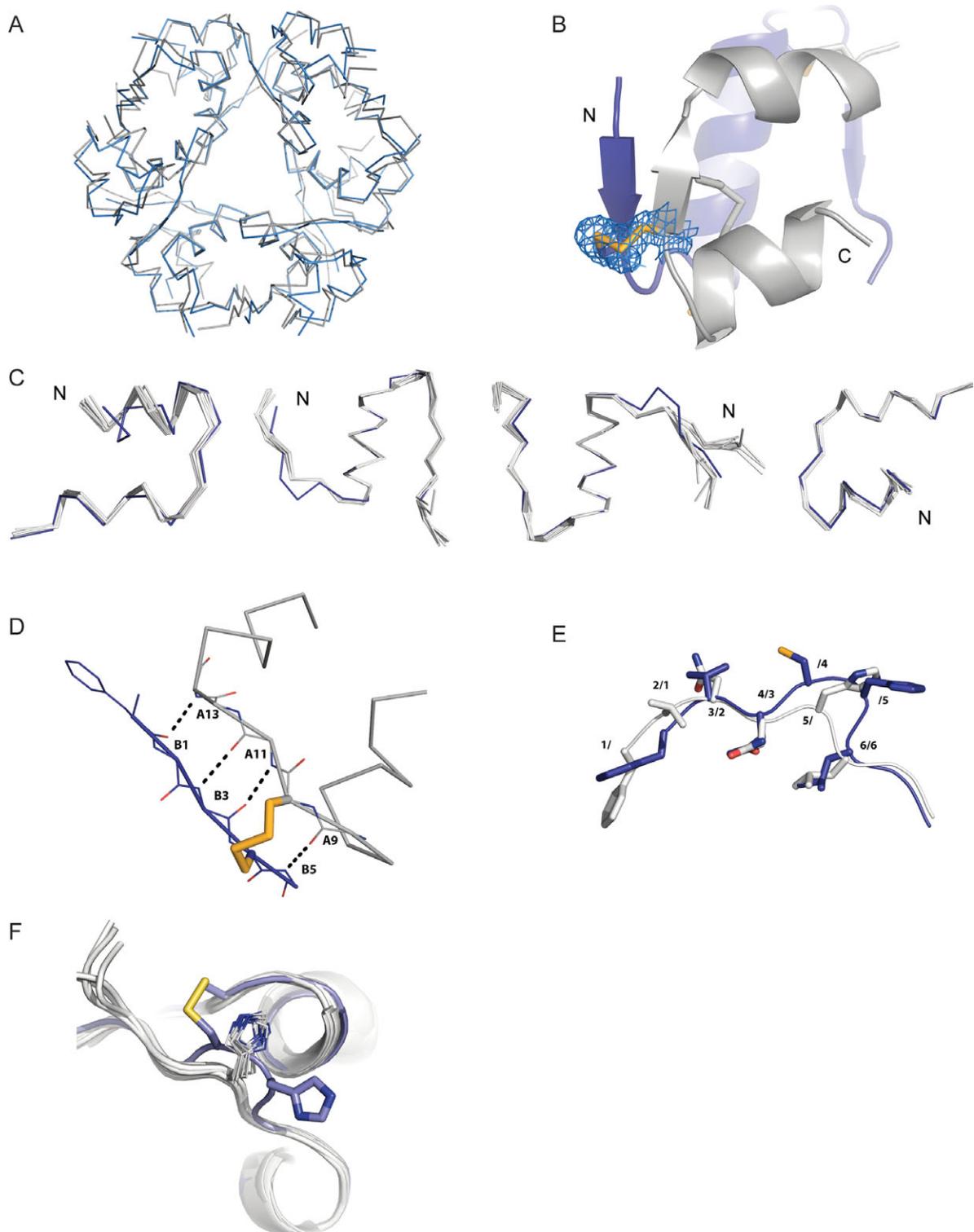
the two molecules in the asymmetric unit. The B8  $\Phi$  dihedral angles in HI are found to be positive in the T-state (56.4°  $\pm$  4.1) and negative in the R-state (-63.0°  $\pm$  3.12).<sup>32</sup> Four hydrogen bonds are formed between residues B1-A13, B3-A11, and A9-B5 [Fig. 1(D)]. Noteworthy, in one of the monomers the additional disulfide bond resulted in the formation of a novel  $\beta$ -sheet involving residues A10-A12 and B2-B4. A secondary structure at this position has only been observed before with extension of the  $\alpha$ -helix in the R-state, see below. Average distance between adjacent C $\alpha$  atoms in the N-terminus of the B-chain and the central part of A-chain was 4.9 Å ( $\pm$ 0.7) in the 4SS-insulin structure whereas corresponding distances in the 1MSO.PDB structure was 6.3 Å ( $\pm$ 1.7). Further comparison with the superposed crystal structures in Figure 1(C) revealed a shift in the positions of the residues in the N-terminus of the B-chain. The locations of the first five residues are shifted one position relative to HI [Fig. 1(D)]. This is probably a necessity to facilitate a conformation where the disulfide bond could be formed. Likely due to this registry shift, the hydrogen bond often found between the imidazole ring of HisB5 and carbonyl oxygen of either CysA7 or SerA9 is replaced by a hydrogen bond involving the amino nitrogen of HisB5 and the carbonyl oxygen of SerA9. This difference could be the reason for the structural difference also found in the N-terminus of the A-chain. As seen in Figure 1(C), the A-chain N-terminus in one of the monomers clearly deviates structurally from the 17 compared structures. The reason could be a loss of helix capping of the first helical fragment in the A-chain due to the relocation of HisB5 side chain [Fig. 1(F)].

When analyzed by RP-HPLC the retention time of the 4SS-insulin analog indicated a remarkable change in hydrophobicity as it eluted >7 min earlier relative to desB30 HI during a 25 min gradient (Supporting Information Fig. S1). The change in retention time may be a result of several factors. One factor is the substitution of the hydrophobic Ile residue with a Cys. Another factor is shielding of the hydrophobic residues (LeuA13, LeuA18, LeuB6, AlaB14, and ValB18) in the  $\alpha$ -helix of the B-chain, when the N-terminal of the B-chain is locked in place by the newly introduced disulfide bond.

The 4SS-insulin analog was able to form zinc-induced hexamers as shown by size-exclusion chromatography (SEC, Supporting Information Fig. S2). The 4SS-insulin analog in Zn-free solution gave a broad peak with apparent hydrodynamic radius smaller than monomeric insulin and a large shoulder indicating its ability to form higher molecular weight species. This indicated that an equilibrium between monomers, dimers, tetramers and hexamers is formed under Zn-free conditions. The monomeric peak eluted after the monomeric

**Table I.** Data Collection and Refinement Statistics

	4SS-insulin
Data collection	
Space group	R 3
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.37, 79.37, 34.23
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 120.00
Resolution (Å)	50–1.98 (2.01–1.98)
<i>R</i> <sub>sym</sub>	0.049 (0.202)
<i>I</i> / $\sigma$ <i>I</i>	31.6 (2.7)
Completeness (%)	97.0 (57.4)
Redundancy	5.3 (1.6)
Refinement	
Resolution (Å)	39.7–1.98
No. reflections	5284
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.21/0.28
No. atoms	
Protein	762
Ligand/ion	2
Water	33
<i>B</i> -factors	
Protein	25.1
Ligand/ion	25.0
Water	32.3
R.m.s. deviations	
Bond lengths (Å)	0.015
Bond angles (°)	2.023



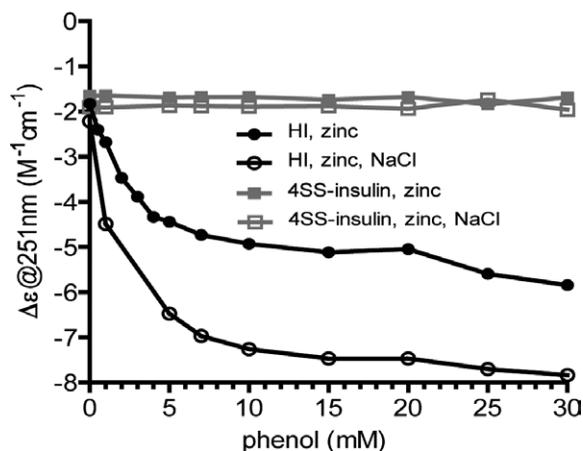
**Figure 1.** Crystal structure of the 4SS-insulin analog. (A) Overlay of the crystal structure of the 4SS-insulin analog (blue) and HI (grey) (pdb code 1MSO.PDB), in a hexamer assembly. (B) The structure of the 4SS-insulin analog (pdb code 4EFX) with the A-chain in grey and B-chain in blue. The additional disulphide bond between B4C and A10C is shown by yellow stick representation. The 2fo-Fc electron density from the additional disulphide bond is rendered in blue at  $1.0\sigma$  level. The N- and C-terminus of the B-chain are labeled. (C) Ribbon representation of the 4SS-insulin analog (blue) compared with 17 insulin structure obtained from the Protein Data Bank codes 1IZB, 1M5A, 1MSO, 1OS3, 2A3G, 2INS, 2ZP6, 3EXX, 3FHP, 3FQ9, 3ILG, 3INC, 3INS, 3RTO, 3TT8, 4E7T, and 4INS. The chains are depicted separately for clarity. The two B-chains are shown in the middle, flanked by their respective A-chain. (D) The hydrogen bonding between the A (grey) and B (blue) chains formed by introduction of the additional disulphide bond. (E) Comparison of the N-terminal part of the B-chain with that of HI (1MSO.PDB). The positions of the residues in the 4SS-insulin analog are shifted one position relative to the HI. Labels refer to sequence numbering of the HI/4SS-insulin structures, respectively. (F) The orientation of the HisB5 side chain in the 4SS-insulin analog deviates from the 17 structures compared. The HisB5 in the 4SS-insulin analog point away from the N-terminal  $\alpha$ -helix and is not active in the helical capping which may exist in the other insulin structures. The additional disulphide bond between CysB4 and CysA10 is shown by yellow stick representation.

reference indicating a smaller hydrodynamic radius probably as a result of a more compact structure caused by the additional disulfide bond. On addition of zinc, the 4SS-insulin analog eluted just before the Co(III)hexamer reference indicating the formation of insulin hexamer with a slightly smaller hydrodynamic radius.

Introduction of the additional disulfide bond between the A10C/B4C positions appears to restrict the flexibility of the N-terminal of the B-chain, thus not allowing the R-state to be formed. The ability of 4SS-insulin to adopt the T-and R-state in solution was probed by phenol titration and analyzed by circular dichroism (CD) at 251 nm (Fig. 2).<sup>33</sup> HI readily adopted the R-state when titrated with phenol in the presence of zinc, and sodium chloride. In contrast 4SS-insulin did not undergo conformational change and remained in the T-state under the experimental conditions known to promote the R-state formation.

The insulin receptor affinity of the 4SS-insulin was measured relative to HI in a receptor binding competition assay. Surprisingly, 4SS-insulin had an increased receptor affinity of 156% ( $n = 3$ , SD = 15.9%) relative to HI (Fig. 3). The ability of the 4SS-insulin to elicit a metabolic response was tested in a lipogenesis assay and the results reflected those seen in the receptor assay with EC50 values 132% (relative to HI in the same plate,  $n = 3$ , SD = 35%).

The *in vivo* activity was evaluated by intravenous administration of two doses to rats. Remarkably, it appeared that the 4SS-insulin had not only increased affinity for the insulin receptor but there was also an indication that it had increased ability to lower the blood glucose levels [Fig. 4(A,B)].

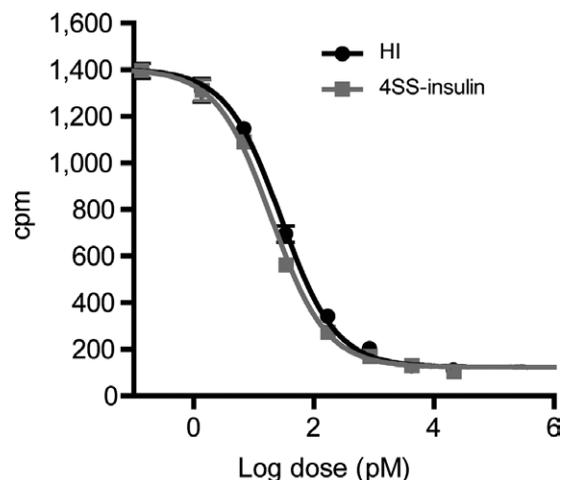


**Figure 2.** Probing phenol induced changes of the conformational state of insulin in the presence of zinc and sodium chloride. Phenol titration of HI with zinc (●), HI with zinc and sodium chloride (○), 4SS-insulin with zinc (■), and 4SS-insulin with zinc and sodium chloride (□). The molar ellipticity is indicative of the conformational state such that an amplitude at  $-2.0 M^{-1} cm^{-1}$  and  $-8.0 M^{-1} cm^{-1}$  is indicative of T-state and R-state, respectively.

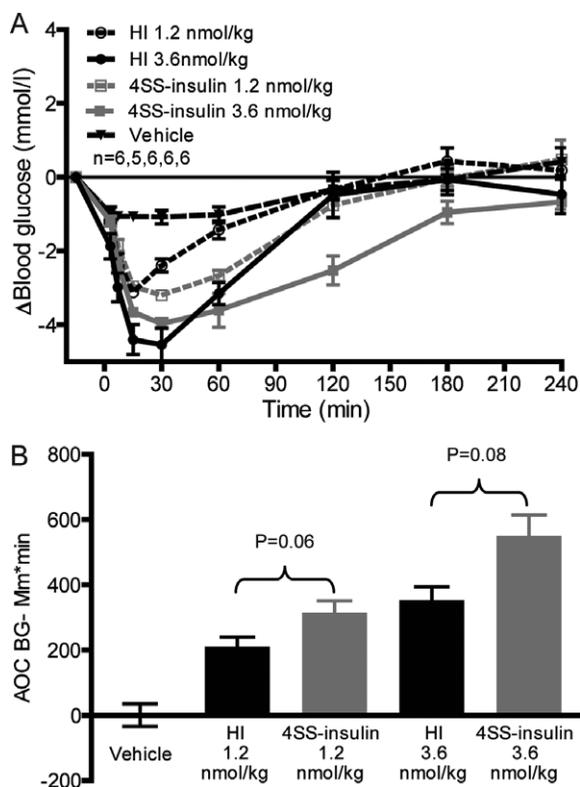
Although the main objective of this experiment was to show that 4SS-insulin is capable of lowering blood glucose levels in an animal model, one could speculate that an increased stability and/or decreased non-receptor mediated clearance of 4SS-insulin could explain this apparent higher *in vivo* activity. The final explanation of these results, however, requires a separate, detailed pharmacological investigation for example in a hyperinsulinemic euglycemic clamp experiment.

Introduction of an extra intermolecular disulfide bond could in principle result in either a decrease in stability because of strain in the structure<sup>34</sup> or an increase in stability by a reduction in entropy.<sup>21</sup> The thermodynamic stability of 4SS-insulin was assessed by differential scanning calorimetry (DSC) and compared with HI [Fig. 5(A)]. The 4SS-insulin had a  $T_m$  of 98.8°C which was a remarkable 34.6°C higher than HI. For comparison, the classic insulin stabilization by Zn-induced formation of an insulin hexamer, results only in a 19°C increase in  $T_m$ .<sup>35</sup> A similar stabilization has only recently been observed for an insulin dimer, which was covalently linked by a disulfide bond in position B25, though this dimer was biologically inactive.<sup>36</sup> As the endothermic unfolding transition had not been completed before a sharp exotherm occurred due to precipitation/aggregation, the unfolding of the 4SS-insulin was irreversible and it was therefore not possible to calculate thermodynamic parameters.

One of the main issues with the physical stability of insulin is fibrillation as a result of physical stress, low pH, and/or elevated temperatures.<sup>37</sup> The 4SS-insulin was tested in a Thioflavin T (ThT) assay for 45 h with vigorous shaking [Fig. 5(B)]. This assay was optimized to show differences between different insulin analogs. HI fibrillated rapidly within the first hour of the experiment and HPLC analyses



**Figure 3.** Representative insulin receptor binding curves. Representative curves for HI (●) and the 4SS-insulin (■) with number of assays,  $n = 3$ . Each point on the graph represents the mean  $\pm$  SD,  $n = 4$  within one assay.

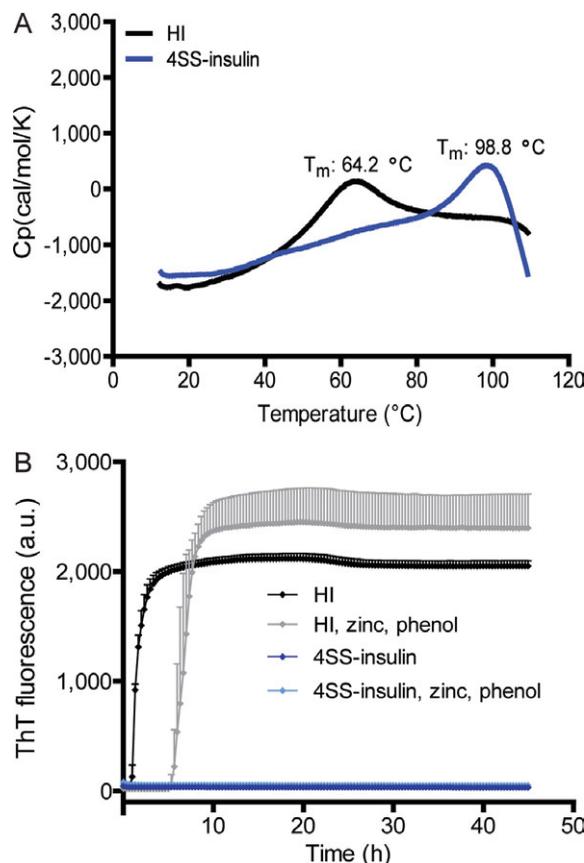


**Figure 4.** *In vivo* activity of the 4SS-insulin analog. (A) Blood glucose profiles following i.v. administration of HI and 4SS-insulin. The two analogs were administered in two doses in anaesthetized wistar rats; 1.2 nmol/kg HI (○), 3.6 nmol/kg HI (●), 1.2 nmol/kg 4SS-insulin (□), 3.6 nmol/kg 4SS-insulin (■), and vehicle (▼) mean values  $\pm$  SEM,  $n = 6, 5, 6, 6, 6$ . (B) Blood glucose lowering effect (area over curve (AOC) with the background subtracted) after i.v. administration of 4SS-insulin, HI or vehicle in fed anaesthetized wistar rats. The  $P$  value from a student two-tailed  $t$ -test between HI and the 4SS-insulin is indicated for each dose.

showed no remaining insulin in the solution after 45 h. The 4SS-insulin showed no indications of fibrillation throughout the experiment and HPLC analyses showed full recovery of the analog. Addition of zinc together with phenol increased the lag time before fibrillation of HI to just around 5 h but the 4SS-insulin still showed no indications of fibrillation during 45 h. A stability similar to the 4SS-insulin analog was only demonstrated before with the non-active insulin B25C-dimer which involves an inter-chain disulfide bond.<sup>36</sup>

## Discussion

The classical insulin structure with three helices and three disulfide bonds has been conserved throughout insulin's evolution and originated more than 500 million years ago.<sup>38</sup> Our finding that an insulin analog with four disulfide bonds, 4SS-insulin, has not only greater physical stability but also improved receptor affinity and activity, provides a



**Figure 5.** Stability of the 4SS-insulin analog. DSC thermographs for HI (black) and the 4SS-insulin analog (grey) recorded from  $10^{-1}$  to  $10^2$  °C with an increase of  $1^\circ\text{C}/\text{min}$  (B) ThT assay for HI (black), HI with 3Zn/hexamer and 30 mM phenol (grey), the 4SS-insulin analog (dark blue), and the 4SS-insulin analog with 3Zn/hexamer and 30 mM phenol (light blue). SD is indicated for each measure point,  $n = 4$ . The analog was subjected to continuously shaking for 45 h at  $37^\circ\text{C}$ .

new perspective on the evolutionary biology of insulin. Why has the structure of insulin and its ancestors, with three disulfide bonds, remained largely unchanged for more than 1 billion years? The answer could simply be a lack of evolutionary pressure due to the efficiency of this finely tuned system. Alternatively, the challenge of introducing two extra Cys in this small Cys-rich protein may have been the evolutionary stumbling block. Introduction of a single Cys leads to reduced stability due to an increased likelihood of disulfide scrambling.<sup>30</sup> Also, introduction of a single Cys has been associated with mutant INS gene-induced diabetes of youth<sup>39</sup> underlining the negative evolutionary selection for single Cys mutations. Consequently, introduction of the two Cys would need to occur in a single evolutionary step—a very improbable event—to avoid a deleterious single Cys substitution mutation.

4SS-insulin was shown to bind to the insulin receptor with a higher affinity than HI. Mutations in

the A10 position were previously shown to decrease receptor affinity.<sup>40,41</sup> Especially introduction of the flexible Gly mutation decreases the binding affinity by around two orders of magnitude indicating a negative effect of increased flexibility in the loop region in the A-chain.<sup>40</sup> In the 4SS-insulin, the disulfide bond decreases the flexibility of this region, which could explain the positive effect of the introduced disulphide bond on receptor binding. Previously, it was also suggested that reorganization of the N-terminal end of the B-chain, such as the transition from T-state to R-state, is necessary for the insulin receptor binding.<sup>32</sup> Our results demonstrate for the first time that structural flexibility of the N-terminal end of the B-chain of insulin is not a prerequisite for binding and that an insulin analog not capable of forming the classical R-state is fully active. Since the first four amino acids of the B-chain do not have a large effect on insulin interaction with the insulin receptor,<sup>42,43</sup> the “functional” difference between the T-state and the R-state lies around the GlyB8 “hinge” region. Dihedral angles of GlyB8 range from positive in the T-state ( $56.4^\circ \pm 4.1$ ) to negative in the R-state ( $-63.0^\circ \pm 3.12$ ).<sup>32</sup> Dihedral angles of GlyB8 in 4SS-insulin were found to be positive and similar to those observed in the T-state ( $46.2^\circ$  and  $42.4^\circ$  for the two molecules in the asymmetric unit). Furthermore, the CD spectroscopy results show no changes in CD spectra upon phenol titration, suggesting no or only very limited changes in the “hinge” region. Therefore, we believe that a transition from the positive  $\Phi$  angle of GlyB8 in the 4SS-insulin analog to a negative angle found in the R-state is unlikely.

Previous attempts to stabilize the insulin structure have mainly involved introduction of helix capping mutations (e.g., HisA8 and AspB10)<sup>44</sup> or construction of single-chain analogs where the C-terminal of the B-chain is linked by a short connecting peptide to the N-terminal of the A-chain.<sup>45</sup> Single-chain insulin analogs and the 4SS-insulin analog both introduce an extra constraint between the two chains. In single-chain analogs, the length and nature of the connecting peptide plays an important role in determining receptor binding properties<sup>46,47</sup> as flexibility of the C-terminal end of the B-chain is believed to be important for receptor binding.<sup>48</sup> In contrast, the disulfide bond in the 4SS-insulin analog only decreases the flexibility of the N-terminus of the B-chain and the analog retains full receptor binding affinity and activity.

We showed that introduction of a fourth disulfide bond had a profound effect on insulin stability, resulting in total absence of fibril formation in our assay. Fibril formation proceeds by partial unfolding of the insulin monomer and the C-terminus of the B-chain is thought to play an important role in this process.<sup>37</sup> Our results demonstrate that stabilization

of the N-terminal part of the insulin molecule is sufficient to prevent fibrillation even with a flexible B-chain C-terminus, likely by stabilization of the monomeric structure, which prevents its unfolding.

We designed an insulin analog containing an additional disulfide and demonstrated its ability to bind to the insulin receptor and facilitate lowering of blood glucose levels. The indications of increased potency would mean that reduced amount of insulin needs to be dosed to obtain the desired decrease in blood glucose levels, thereby lowering the cost. Also introduction of the additional disulfide bond resulted in high thermodynamic stability and prevention of fibril formation. All of these characteristics would be highly desirable in future insulin analogs for the treatment of diabetes by requiring less stringent storage and transport conditions, longer in use time and so forth, and thus providing extra benefits over current situation, especially in developing countries where cost should be kept at a minimum and storage facilities are far from optimal.

## Materials and Methods

### Plasmids construction and expression

Material, vectors, strains, and construction were as previously described.<sup>42,49,50</sup> Shortly, the mutation to cysteine residues were introduced in selected positions in the insulin coding sequence by overlapping PCRs. The insulin precursors were expressed in *Saccharomyces cerevisiae* as proinsulin-like single-chains consisting of a spacer GluGluAlaGluAlaGluAlaProLys (EEAEAEAPK) followed by the B-chain (B1-B29), a mini C-peptide Ala Ala Lys (AAK) and the A-chain (A1-A21). The expression yields of the insulin precursors were determined by reversed-phase HPLC (RP-HPLC) based on peak area using HI as external standard. A gradient going from 17% acetonitrile to 27% within the first 20 min and from 27% to 37% acetonitrile within the next 5 min was used. The mass was determined by LC/MS on a Dionex Ultimate 3000 liquid chromatography system coupled to a mass spectrometer, Ultra Performance High Capacity Ion Trap MS (HCT Ultra) (Bruker Daltonics, Bremen, Germany) with electrospray ionization.

### Purification

The CysA10-CysB4 precursor was fermented in 200 mL cultures in shaking flasks. The cells were removed by centrifugation and the culture supernatant was acidified. The precursor was partially purified and concentrated by a capture step using cation exchange chromatography. The eluate was diluted 1:1 in MilliQ, and 15 mM sodium glutamate was added and the pH adjusted to 9.7 with 10N NaOH. The precursor was converted to the mature insulin

analog by removal of the spacer and C-peptide using ALP (Novo Nordisk A/S) in solution overnight. The CysA10/CysB4 analog (4SS-insulin) was further purified by RP-HPLC using a formic acid buffer system and lyophilized.

### Crystallography

Crystals belonging to the R3 space group were obtained by vapour diffusion. To the protein solution with an insulin concentration of 5 mg/mL was added zinc acetate to a concentration corresponding to 3 Zn<sup>2+</sup> per six insulin molecules to obtain hexameric insulin. Diffraction data to a resolution of 2.0 Å were collected in house using a rotating anode (Rigaku MicroMax-007HF, Cu/K $\alpha$  radiation,  $\lambda = 1.5418$  Å) and a Rayonix SX-165 CCD detector (Mar Research, Hamburg) at a temperature of 100 K. The data was processed and scaled using the HKL2000 software package.<sup>51</sup> The structure was solved by molecular replacement using Phaser<sup>52</sup> with an in house insulin monomer as search model. Residues B1-B5 of the search model were removed to reduce model bias. Atom coordinates were manually adjusted in Coot<sup>53</sup> and the structure was refined in Refmac.<sup>54</sup> 96.6% of the residues were found in the most favoured regions of the Ramachandran plot and 3.3% in additional allowed regions. Data collection details and refinement statistics are summarized in Table I. The X-ray crystallographic coordinates have been deposited in the protein data bank with accession ID 4EFX.

### Size-exclusion chromatography

References for monomeric insulin was the stable monomeric insulin AspB9, GluB27 0.6 mM,<sup>55</sup> for hexameric insulin it was the stable Co(III) hexamer 0.6 mM<sup>56</sup> and included was also albumin (HSA) and NaNO<sub>3</sub> 0.1M. The 4SS-insulin analog was formulated in a 2 mM phosphate buffer in either 0.2 mM or 0.6 mM with or without three zinc per hexamer (zinc acetate) a slight excess of zinc to enforce hexamer formation. As the analog is restricted from forming the R-state this amount of zinc will not induce the formation of T<sub>3</sub>R<sub>3</sub> form.<sup>33</sup> The chromatography was run on an Acquity UPLC BEH125 SEC column (1.7  $\mu$ M, 150 mm) at 37°C, with a flow rate of 0.3 mL/min. The eluent was 20 mM NaCl, 10 mM Tris, 3% v/v 2-propanol, and 0.1% Na azide, pH 7.3.

### Circular dichroism spectroscopy

Recording of CD spectral data were carried out as described in Ref.<sup>57</sup> except that a Jasco J-815 instrument was used. The conditions were 0.6 mM HI, 0.3 mM zinc acetate, 10 mM tris/CIO<sub>4</sub><sup>-</sup> pH 7.4,  $\pm$ 100 mM sodium chloride or 0.6 mM 4SS-Insulin, 0.3 mM zinc acetate, 10 mM tris/CIO<sub>4</sub>-pH 7.6,  $\pm$ 100 mM sodium chloride. Phenol was titrated to a final concentration of 30 mM.

### Receptor binding assay

Receptor binding was measured on the A isoform of the receptor in a scintillation proximity assay (SPA) as previously described.<sup>31</sup> Briefly, the binding receptor affinities were determined by competition of the 4SS-insulin analog and [125I]TyrA14-labeled insulin (Novo Nordisk A/S) in the SPA assay. The 4SS-insulin analog ( $n = 4$ ) was tested together with a human standard ( $n = 4$ ) in one plate. The data were analyzed according to a four-parameter logistic model<sup>58</sup> and the affinities were expressed relative to the HI standard [IC<sub>50</sub>(insulin)/IC<sub>50</sub>(analog)  $\times$  100%].

### Metabolic potency determination

The metabolic potency determinations were done by lipogenesis essentially as described before.<sup>36</sup> Shortly, primary rat adipocytes from SPRD rats were isolated and placed in a degradation buffer containing Hepes buffer, Krebs buffer, 0.1% HSA, collagenase, and glucose. The cells were shaken vigorously for 1 h at 37°C and the cell suspensions were filtered, washed twice, and resuspended in incubation buffer containing Hepes buffer, Krebs buffer and 0.1% HSA. 100  $\mu$ L aliquots were distributed in 96-well PicoPlates and incubated 2 h at 37°C with gentle shaking together with 10  $\mu$ L glucose solution containing D-[3-<sup>3</sup>H]glucose and glucose and of increasing concentration of 10  $\mu$ L insulin analog. The incubation was stopped by adding 100  $\mu$ L MicroScint E (Packard) and the plates were counted in a Top-Count NXT (PerkinElmer Life Science). The data were analyzed according to a four-parameter logistic model<sup>58</sup> and the metabolic potency were expressed relative to a HI standard [EC<sub>50</sub>(insulin)/EC<sub>50</sub>(analog)  $\times$  100%].

### In vivo effect of 4SS-insulin analog in Wistar rats

The *in vivo* effect was determined essentially as described before.<sup>30</sup> The Wistar (253-319 g) (Taconic, Hudson, NY) rats were allocated into five groups, six rats in each group being administrated HI or insulin analog, and five in the group being administrated vehicle. The animals were dosed with an intravenous injection in a tail vein (1 mL/kg) of either vehicle (5 mM phosphate buffer, 140 mM NaCl, 70 ppm polysorbate 20, pH 7.4) or HI or the 4SS-insulin analog (1.2 nmol/kg or 3.6 nmol/kg). Blood samples for the determination of whole blood glucose concentration were collected by puncture of the capillary vessels in the tail tip to time -15 min and 0 min before dosing, and to time 3, 7, 15, 30, 60, 120, 180, and 240 min after dosing. The blood glucose concentrations were measured by the immobilized glucose oxidase method using a Biosen auto-analyzer (EKF Diagnostic, Germany).

### Differential scanning calorimetry

The DSC measurements were performed essentially as described before.<sup>35</sup> The insulin analogs were formulated at 0.2 mM in a 0.2 mM phosphate buffer pH 7.5 which were also used as reference buffer. The samples were heated from 10°C to 110°C with a scan rate of 60°C per hour.

### Thioflavin T fibrillation assay

The Thioflavin T (ThT) assay was performed as previously described.<sup>36</sup> Briefly, the samples were prepared freshly. The experiment was performed at 37°C and the plate with the samples was incubated for 10 min before the first measurement and then measured every 20 min for up to 45 h. Between each measurement, the plate was continuously shaken (960 rpm) and heated. Each shown time point is the mean of the four replicas with standard deviation error bars. The peptide concentration in each of the tested formulations was measured by RP-HPLC both before application and after completion of the ThT fibrillation.

### Acknowledgments

The authors thank Svend Havelund for scientific input. Berit Bugge Nielsen, Kirsten Vestergaard, Susanne G. Kjeldsen, Charlotte Jensen, Louise Qvistgaard, Lene Villadsen, Rikke Lill Rønhede, and Anne Ahrensberg Jensen provided excellent technical assistance.

### References

1. World Health Organization. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications (1999) Part 1: Diagnosis and Classification of Diabetes Mellitus. WHO Consultation. Geneva: World Health Organization, Department of Noncommunicable Disease Surveillance.
2. Baker EN, Blundell TL, Cutfield JF, Cutfield SM, Dodson EJ, Dodson GG, Hodgkin DM, Hubbard RE, Isaacs NW, Reynolds CD, Sakabe K, Sakabe N, Vijayan NM (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos Trans R Soc Lond B Biol Sci* 319:369–456.
3. Blundell T, Dodson G, Hodgkin D, Mercola D (1972) Insulin: The structure in the crystal and its reflection in chemistry and biology. *Adv Protein Chem* 26:279–402.
4. Nicol DS, Smith LF (1960) Amino-acid sequence of human insulin. *Nature* 187:483–485.
5. Ryle AP, Sanger F, Smith LF, Kitai R (1955) The disulphide bonds of insulin. *Biochem J* 60:541–556.
6. Bentley G, Dodson E, Dodson G, Hodgkin D, Mercola D (1976) Structure of insulin in 4-zinc insulin. *Nature* 261:166–168.
7. Derewenda U, Derewenda Z, Dodson EJ, Dodson GG, Reynolds CD, Smith GD, Sparks C, Swenson D (1989) Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature* 338:594–596.
8. Peterson JD, Steiner DF (1975) The amino acid sequence of the insulin from a primitive vertebrate, the atlantic hagfish (*Myxine glutinosa*). *J Biol Chem* 250:5183–5191.

9. Steiner DF, Chan SJ, Welsh JM, Kwok SC (1985) Structure and evolution of the insulin gene. *Annu Rev Genet* 19:463–484.
10. Chan SJ, Cao QP, Steiner DF (1990) Evolution of the insulin superfamily: cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus. *Proc Natl Acad Sci U S A* 87:9319–9323.
11. Chang SG, Choi KD, Jang SH, Shin HC (2003) Role of disulfide bonds in the structure and activity of human insulin. *Mol Cells* 16:323–330.
12. Guo ZY, Feng YM (2001) Effects of cysteine to serine substitutions in the two inter-chain disulfide bonds of insulin. *Biol Chem* 382:443–448.
13. Clarke J, Henrick K, Fersht AR (1995) Disulfide mutants of barnase. I: Changes in stability and structure assessed by biophysical methods and X-ray crystallography. *J Mol Biol* 253:493–504.
14. Das M, Kobayashi M, Yamada Y, Sreeramulu S, Ramakrishnan C, Wakatsuki S, Kato R, Varadarajan R (2007) Design of disulfide-linked thioredoxin dimers and multimers through analysis of crystal contacts. *J Mol Biol* 372:1278–1292.
15. Guzzi R, Andolfi L, Cannistraro S, Verbeet MP, Canters GW, Sportelli L (2004) Thermal stability of wild type and disulfide bridge containing mutant of poplar plasmatocyanin. *Biophys Chem* 112:35–43.
16. Mansfeld J, Vriend G, Dijkstra BW, Veltman OR, Van den BB, Venema G, Ulbrich-Hofmann R, Eijssink VG (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J Biol Chem* 272:11152–11156.
17. McAuley A, Jacob J, Kolvenbach CG, Westland K, Lee HJ, Brych SR, Rehder D, Kleemann GR, Brems DN, Matsumura M (2008) Contributions of a disulfide bond to the structure, stability, and dimerization of human IgG1 antibody CH3 domain. *Protein Sci* 17:95–106.
18. Ray SS, Nowak RJ, Strokovich K, Brown RH, Jr., Walz T, Lansbury PT, Jr. (2004) An intersubunit disulfide bond prevents in vitro aggregation of a superoxide dismutase-1 mutant linked to familial amyotrophic lateral sclerosis. *Biochemistry* 43:4899–4905.
19. Flory PJ (1956) Theory of elastic mechanisms in fibrous proteins. *J Am Chem Soc* 78:5222–5235.
20. Doig AJ, Williams DH (1991) Is the hydrophobic effect stabilizing or destabilizing in proteins? The contribution of disulphide bonds to protein stability. *J Mol Biol* 217:389–398.
21. Betz SF (1993) Disulfide bonds and the stability of globular proteins. *Protein Sci* 2:1551–1558.
22. Harrison PM, Sternberg MJ (1994) Analysis and classification of disulphide connectivity in proteins. The entropic effect of cross-linkage. *J Mol Biol* 244:448–463.
23. Sowdhamini R, Srinivasan N, Shoichet B, Santi DV, Ramakrishnan C, Balaram P (1989) Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis. *Protein Eng* 3:95–103.
24. Duret L, Guex N, Peitsch MC, Bairoch A (1998) New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling. *Genome Res* 8:348–353.
25. Pierce SB, Costa M, Wisotzkey R, Devadhar S, Homberger SA, Buchman AR, Ferguson KC, Heller J, Platt DM, Pasquinelli AA, Liu LX, Doberstein SK, Ruvkun G (2001) Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev* 15:672–686.

26. Smit AB, Vreugdenhil E, Ebberink RH, Geraerts WP, Klootwijk J, Joosse J (1988) Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* 331:535–538.
27. Hua QX, Nakagawa SH, Wilken J, Ramos RR, Jia W, Bass J, Weiss MA (2003) A divergent INS protein in *Caenorhabditis elegans* structurally resembles human insulin and activates the human insulin receptor. *Genes Dev* 17:826–831.
28. Smith GD, Pangborn WA, Blessing RH (2003) The structure of T6 human insulin at 1.0 Å resolution. *Acta Crystallogr, Sect D: Biol Crystallogr* 59:474–482.
29. Pullen RA, Lindsay DG, Wood SP, Tickle IJ, Blundell TL, Wollmer A, Krail G, Brandenburg D, Zahn H, Gliemann J, Gammeltoft S (1976) Receptor-binding region of insulin. *Nature* 259:369–373.
30. Vinther TN, Ribøl U, Askov PT, Kjeldsen TB, Jensen KJ, Hubalek F (2011) Identification of anchor points for chemical modification of a small cysteine-rich protein by using a cysteine scan. *ChemBioChem* 12:2448–2455.
31. Glendorf T, Sørensen AR, Nishimura E, Pettersson I, Kjeldsen T (2008) Importance of the solvent-exposed residues of the insulin B chain alpha-helix for receptor binding. *Biochemistry* 47:4743–4751.
32. Nakagawa SH, Zhao M, Hua QX, Hu SQ, Wan ZL, Jia W, Weiss MA (2005) Chiral mutagenesis of insulin. Foldability and function are inversely regulated by a stereospecific switch in the B chain. *Biochemistry* 44:4984–4999.
33. Kaarsholm NC, Ko HC, Dunn MF (1989) Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin. *Biochemistry* 28:4427–4435.
34. Katz B, Kossiakoff AA (1990) Crystal structures of subtilisin BPN' variants containing disulfide bonds and cavities: concerted structural rearrangements induced by mutagenesis. *Proteins* 7:343–357.
35. Huus K, Havelund S, Olsen HB, van de Weert M, Frokjaer S (2005) Thermal dissociation and unfolding of insulin. *Biochemistry* 44:11171–11177.
36. Vinther TN, Norrman M, Strauss HM, Huus K, Schlein M, Pedersen TA, Kjeldsen T, Jensen KJ, Hubalek F (2012) Novel covalently linked insulin dimer engineered to investigate the function of insulin dimerization. *PLoS One* 7:e30882.
37. Brange J, Andersen L, Laursen ED, Meyn G, Rasmussen E (1997) Toward understanding insulin fibrillation. *J Pharm Sci* 86:517–525.
38. Cutfield JF, Cutfield SM, Dodson EJ, Dodson GG, Emdin SF, Reynolds CD (1979) Structure and biological activity of hagfish insulin. *J Mol Biol* 132:85–100.
39. Liu M, Haataja L, Wright J, Wickramasinghe NP, Hua QX, Phillips NF, Barbetti F, Weiss MA, Arvan P (2010) Mutant INS-gene induced diabetes of youth: proinsulin cysteine residues impose dominant-negative inhibition on wild-type proinsulin transport. *PLoS One* 5:e13333.
40. Bullesbach EE, Schwabe C (1994) Functional importance of the A chain loop in relaxin and insulin. *J Biol Chem* 269:13124–13128.
41. Sajid W, Holst PA, Kiselyov VV, Andersen AS, Conlon JM, Kristensen C, Kjeldsen T, Whittaker J, Chan SJ, De MP (2009) Structural basis of the aberrant receptor binding properties of hagfish and lamprey insulins. *Biochemistry* 48:11283–11295.
42. Kristensen C, Kjeldsen T, Wiberg FC, Schäffer L, Hach M, Havelund S, Bass J, Steiner DF, Andersen AS (1997) Alanine scanning mutagenesis of insulin. *J Biol Chem* 272:12978–12983.
43. Kerp L, Steinhilber S, Kasemir H, Han J, Henrichs HR, Geiger R (1974) Changes in immunospecificity and biologic activity of bovine insulin due to subsequent removal of the amino acids B1, B2, and B3. *Diabetes* 23:651–656.
44. Kaarsholm NC, Norris K, Jørgensen RJ, Mikkelsen J, Ludvigsen S, Olsen OH, Sørensen AR, Havelund S (1993) Engineering stability of the insulin monomer fold with application to structure-activity relationships. *Biochemistry* 32:10773–10778.
45. Hua QX, Nakagawa SH, Jia W, Huang K, Phillips NB, Hu SQ, Weiss MA (2008) Design of an active ultrastable single-chain insulin analog: synthesis, structure, and therapeutic implications. *J Biol Chem* 283:14703–14716.
46. Nakagawa SH, Tager HS (1989) Perturbation of insulin-receptor interactions by intramolecular hormone cross-linking. Analysis of relative movement among residues A1, B1, and B29. *J Biol Chem* 264:272–279.
47. Rajpal G, Liu M, Zhang Y, Arvan P (2009) Single-chain insulins as receptor agonists. *Mol Endocrinol* 23:679–688.
48. Hua QX, Shoelson SE, Kochoyan M, Weiss MA (1991) Receptor binding redefined by a structural switch in a mutant human insulin. *Nature* 354:238–241.
49. Kjeldsen T, Pettersson AF, Hach M (1999) The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. *J Biotechnol* 75:195–208.
50. Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Pettersson AF, Vad K (1996) A removable spacer peptide in an alpha-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*. *Gene* 170:107–112.
51. Otwinowski Z (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol Crystallogr Part A. Methods Enzymol* 276:307–326.
52. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
53. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
54. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr, Sect D: Biol Crystallogr* 53:240–255.
55. Brange J, Ribøl U, Hansen JF, Dodson G, Hansen MT, Havelund S, Melberg SG, Norris F, Norris K, Snel L (1988) Monomeric insulins obtained by protein engineering and their medical implications. *Nature* 333:679–682.
56. Kurtzhals P, Ribøl U (1995) Action profile of cobalt(III)-insulin. A novel principle of protraction of potential use for basal insulin delivery. *Diabetes* 44:1381–1385.
57. Olsen HB, Kaarsholm NC (2000) Structural effects of protein lipidation as revealed by LysB29-myristoyl, des(B30) insulin. *Biochemistry* 39:11893–11900.
58. Vølund A (1978) Application of the four-parameter logistic model to bioassay: comparison with slope ratio and parallel line models. *Biometrics* 34:357–365.