Published in final edited form as:

Chemistry. 2017 January 31; 23(7): 1709–1716. doi:10.1002/chem.201605578.

Scope & Limitations of Fmoc Chemistry SPPS-Based Approaches to the Total Synthesis of Insulin Lispro via Ester Insulin

Balamurugan Dhayalan^[a], Kalyaneswar Mandal^[a], Nischay Rege^[b], Michael A. Weiss^[b], Simon H. Eitel^[c], Thomas Meier^[c], Ralph O. Schoenleber^[c], and Stephen B.H. Kent^{*,[a]}

^[a]Department of Chemistry, Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA ^[b]Department of Biochemistry, Case Western Reserve University, Cleveland, OH 44106, USA ^[c]Bachem AG, Hauptstrasse 144, Bubendorf 4416, Switzerland

Abstract

We have systematically explored three approaches based on Fmoc chemistry SPPS for the total chemical synthesis of the key depsipeptide intermediate for the efficient total chemical synthesis of insulin. The approaches used were: stepwise Fmoc chemistry SPPS; the ‘hybrid method’, in which maximally-protected peptide segments made by Fmoc chemistry SPPS are condensed in solution; and, native chemical ligation using peptide-thioester segments generated by Fmoc chemistry SPPS. A key building block in all three approaches was a Glu[Oβ(Thr)] ester-linked dipeptide equipped with a set of orthogonal protecting groups compatible with Fmoc chemistry SPPS. The most effective method for the preparation of the 51 residue ester-linked polypeptide chain of ester insulin was the use of unprotected peptide-thioester segments, prepared from peptide-hydrazides synthesized by Fmoc chemistry SPPS, and condensed by native chemical ligation. High resolution X-ray crystallography confirmed the disulfide pairings and three-dimensional structure of synthetic insulin lispro prepared from ester insulin lispro by this route. Further optimization of these pilot studies should yield an effective total chemical synthesis of insulin lispro (Humalog) based on peptide synthesis by Fmoc chemistry SPPS.

Keywords

native chemical ligation; Fmoc chemistry SPPS; hybrid synthesis; insulin lispro; ester insulin

Introduction

Insulin is a 51 amino acid residue protein hormone made up of two peptide chains which are connected by two inter-chain disulfide bonds. Insulin also has one intra-chain disulfide bond.[1] A number of different forms of insulin have been developed as human therapeutics. Insulin lispro is a form of human insulin in which the –Pro28–Lys29– amino acid sequence near the C-terminus of the B-chain is inverted to –Lys28–Pro29–.[2] This simple change

* skent@uchicago.edu.

inhibits formation of insulin dimers, while retaining full activity against the insulin receptor. Consequently insulin lispro is widely used as a rapid-acting form of insulin.[3]

Because of the poor handling properties of the A-chain peptide, and the inefficient formation of the folded insulin protein molecule from two individual peptide chains,[4] total chemical synthesis of insulin is still challenging despite the many strategies for its synthesis that have been developed over the decades since its covalent structure was elucidated.[5] We have previously reported the total synthesis of insulin lispro via a 51 residue depsipeptide chain, which contains an ester bond between the side chains of residues GluA4 and ThrB30. The covalent ester link facilitates correct folding to give the synthetic protein molecule 'ester insulin' (Figure 1), which is readily converted to fully active insulin by simple saponification at 4 °C.[6]

Results and Discussion

Here we report systematic exploration of three approaches based on Fmoc chemistry SPPS for the total chemical synthesis of insulin lispro via the ester insulin intermediate. The approaches used were: stepwise Fmoc chemistry SPPS;[7] the 'hybrid method',[8] in which protected peptide segments made by Fmoc chemistry SPPS are condensed in solution; and, native chemical ligation[9] using peptide-thioester segments generated by Fmoc chemistry SPPS.

Design & Synthesis of an Orthogonally-Protected Dipeptide Building Block

The key building block in all three synthetic approaches was an ester-linked dipeptide equipped with a set of mutually-compatible (i.e. 'orthogonal'[10]) protecting groups (Scheme 1). During Fmoc chemistry SPPS of the peptide segment containing the ester link, each amino-protecting group must be removable in the presence of the other, while the carboxyl group of ThrB30 must remain protected until the final global deprotection step. For the synthesis of the protected dipeptide itself, four different protecting groups – two for the α -amino groups and two for the α -carboxylic acid groups – were used. For the threonine (that would become ThrB30), we chose the $N\alpha$ -allyloxycarbonyl (Alloc) and α COOH tert-butyl ester protecting groups. For the glutamic acid (that would become GluA4), we chose the $N\alpha$ -Fmoc protecting group and, as the fourth orthogonal protecting group, a phenacyl (Pac) ester for the α -carboxyl group. The Pac group was ultimately removed using Zn/AcOH to give the desired Fmoc-Glu[O β (Alloc-Thr-OtBu)]-OH.

Synthesis of gram quantities of the target dipeptide Fmoc-Glu[O β (Alloc-Thr-OtBu)]-OH was achieved by this approach in a matter of few days. The overall yield was 65%.

Stepwise Fmoc chemistry SPPS

First, the orthogonally protected dipeptide building block Fmoc-Glu[O β (Alloc-Thr-OtBu)]-OH was used to make the full-length 51 amino acid ester insulin depsipeptide by using stepwise Fmoc chemistry SPPS (Scheme 2). A typical stepwise synthesis of the ester insulin 51 residue depsipeptide started from H-Asn(Trt)-2-Chlorotrityl-(S-DVB)resin. Stepwise Fmoc chemistry SPPS – including the incorporation of Fmoc-GluA4[O β (Alloc-ThrB30-OtBu)]-OH – was carried out through the incorporation of Boc-GlyA1. Then, after removing

the Na-Alloc protecting group from ThrB30, stepwise Fmoc chemistry SPPS was continued through the incorporation of Boc-PheB1. Full details of the synthetic protocols used are given in the Supporting Information.

Cleavage of the full-length peptide-resin with concomitant deprotection of the side chain protecting groups was performed with a TFA-scavenger cocktail, to give the crude peptide product (Figure 2).

Only a ~5% yield of the full-length polypeptide was evident by HPLC analysis after cleavage/deprotection, with principal byproducts arising from side reactions that included: an O/N-acyl shift after Alloc deprotection at ThrB30; and, H-GlyA1-GluA4(OEt)-AsnA21-OH, possibly generated by transesterification with Oxyma. Although undoubtedly stepwise Fmoc chemistry SPPS of this 51 residue depsipeptide could be substantially optimized, no such investigation was carried out in the current research. Rather, we set out to explore alternative Fmoc chemistry SPPS-based approaches.

Hybrid SPPS/Solution Condensation of Protected Peptide Segments

Next, we explored use of the ‘hybrid SPPS/solution condensation method’ for the convergent chemical synthesis of the 51 residue ester insulin polypeptide chain (Scheme 3). [8] We chose SerA9-IleA10 and GlyB23-PheB24 sites for condensation of the protected peptide segments. The protected peptide segments Boc-PheB1-GlyB23-OH, Boc-GlyA1-GluA4[Oβ(ThrB30-PheB24-Fmoc)]-SerA9-OH were prepared by Fmoc chemistry SPPS on a 2-chlorotrityl chloride-resin.[11] For H-IleA10-AsnA21-OtBu, we used side-chain attachment of the C-terminal Asn on a xanthylenyl linker-resin.[12] Details for the synthesis, purification, and characterization of the protected peptide segments are given in the Supporting Information.

For the reaction of Boc-GlyA1-GluA4[Oβ(ThrB30-PheB24-Fmoc)]-SerA9-OH with H-IleA10-AsnA21-OtBu, use of PyOxim as the coupling reagent in DMF gave a better conversion and more rapid reaction, with no detectable racemization, when compared to HBTU/HOBT as coupling reagent. For the reaction of Boc-PheB1-GlyB23-OH with Boc-GlyA1-GluA4[Oβ(ThrB30-PheB24-H)]-AsnA21-OtBu, only reaction using PyOxim/DIEA and the solvent combination 1:1 DCM/DMF gave a significant consumption of one of the starting peptides and detectable product (Figure 3). The large amounts of unreacted starting peptides left after the reaction was a major concern.

We changed various parameters such as stoichiometry of the reagents, order of addition of reagents, solvent combinations, and coupling reagents, without being able to improve this outcome. Because of the inability to obtain useful yields of the desired product, this hybrid Fmoc chemistry SPPS/solution condensation approach was also abandoned.

Native Chemical Ligation of Unprotected Peptide Segments

Finally, we explored the convergent synthesis of the ester insulin 51 residue depsipeptide by native chemical ligation of unprotected peptide segments.[9] The necessary peptide-thioester segments were prepared by oxidation of peptide-hydrazides prepared by Fmoc chemistry SPPS[13].

Because of the reported instability of the Thz-moiety to oxidation at reduced pH,[13] which we confirmed in a model study (see SI), we were unable to use our previously published C-to-N segment condensation strategy.[14] so a revised strategy was devised that made use of N-to-C native chemical ligation segment condensation (Scheme 4). This strategy obviated the need for Thz-protection of the N-terminal Cys-residue of the middle segment, while the hydrazide at the C-terminus of that peptide served as a cryptic thioester, facilitating the N → C ligation strategy.[13] The peptide-hydrazide intermediate product GlyA1-GluA4(Oβ(ThrB30-PheB1))-CysA6-αCO-NHNH2 is converted to the thioester by oxidation at pH 3.1, to give GlyA1-GluA4(Oβ(ThrB30-PheB1))-CysA6-αCOSR for the second (and final) native chemical ligation step.

Native chemical ligation of the peptide segments PheB1-ValB18-αCOSCH2CH2SO3H and GlyA1-GluA4[Oβ(ThrB30-CysB19))-CysA6-αCONHNH2 was carried out at pH7.0 in 6 M Gn.HCl under standard conditions. Experimental details and LC-MS data are presented in Supporting Information. It is important to note that, unlike in a C-to-N ligation strategy where consecutive one-pot reactions can be performed without intermediate purification steps, in an N-to-C strategy the product from first native chemical ligation must be purified in order to remove MPAA and TCEP, the presence of which may hamper the subsequent NaNO2 oxidation step).[15] After purification, 25.8 mg (72% yield, 6.3 μmol from 8.78 μmol of PheB1-ValB18-Mes thioester limiting peptide) of product peptide GlyA1-GluA4[Oβ(ThrB30-PheB1))-CysA6-αCONHNH2 was obtained.

Purified product peptide GlyA1-GluA4[Oβ(ThrB30-PheB1))-CysA6-αCONHNH2 was taken up in aqueous 6 M Gn.HCl, 0.1 M phosphate buffer, pH 3.1 and subjected to oxidation by 200mM NaNO2 at -15 °C. After 30 min, CysA7-AsnA21 dissolved in 6 M Gn.HCl, 0.1 M phosphate buffer at pH 6.8 containing 200 mM MPAA was added to the oxidation reaction solution at -15 °C. The reaction mixture was then allowed to warm to RT and adjusted to pH6.1 for the native chemical ligation reaction. For LC-MS analysis (Figure 4), aliquots were treated with TCEP.HCl before injection. After 4h, the reaction mixture was treated with 100mM TCEP.HCl, at pH 7.0, for 10 min, then diluted 10-fold with water to precipitate the full-length polypeptide, which was recovered by centrifugation. Final yield of desired branched depsipeptide H-GlyA1-GluA4[Oβ(ThrB30-PheB1))-AsnA21-COOH product after washing with water was 24.1 mg (67% yield, 4.16 μmol from 6.21 μmol of limiting peptide GlyA1-GluA4[Oβ(ThrB30-PheB1))-CysA6-αCONHNH2).

Folding & Saponification

Having established an effective synthesis of the 51 residue depsipeptide, it was folded with concomitant formation of disulfides to form ester insulin lispro.[14] Formation of the desired folded ester insulin lispro protein was indicated by the formation of a sharp, early eluting peak that had a mass lower by 5.5 ± 0.6 Da which corresponds to the formation of three disulfide bonds) (Figure 5). The earlier elution behavior of the folded ester insulin lispro protein compared with the unfolded polypeptide is consistent with the burying of hydrophobic side chains inside the protein tertiary structure stabilized by three disulfide bonds.

The ester insulin lispro was then saponified,[14] to give insulin lispro (Figure 6). After prep-HPLC purification, 3.39 mg insulin lispro was obtained (66% yield, 0.58 μ mol from 5.1 mg, 0.88 μ mol of ester insulin lispro).

Characterization of Synthetic Insulin Lispro

The synthetic protein was characterized by LCMS (Figure 7). Formation of a late-eluting peak with 18.1 ± 0.6 Da increased mass was consistent with saponification of the ester bond.

X-ray crystallography

The amino acid sequence, correct disulfide pairing, and three-dimensional structure of the synthetic human insulin lispro protein molecule were confirmed by X-ray crystallography. Data collected to 1.35 Å resolution and the structure was solved by molecular replacement. [16] A cartoon representation of the crystal structure of the final model is shown in Figure 7.

In the present crystal form, the B-chain of the insulin lispro molecule adapted the R6 state, forming an extended alpha-helix spanning residue PheB1-CysB19. It is worth mentioning that, in the first crystal structure of insulin lispro reported by Smith and co-workers in 1995, [17] the insulin lispro molecule was crystallized in presence of Zn²⁺ and phenol as a T3Rf3 hexamer, where the B-chain of three monomers had a T-conformation and of other three monomers had an Rf conformation. However, the R6 hexamer form was previously observed in several human insulin structures and in one analogue of insulin lispro structure.[16]

The 21 amino acid A-chain of the insulin lispro consisting of two short alpha-helices separated by a turn was packed against an extended alpha-helix of the 30 residue B-chain. In the crystalline state, the six insulin lispro protein monomers assembled into a hexamer mediated by two Zn²⁺ metal ions and six phenol ligands. Each Zn²⁺ ion formed a tetrahedral complex with the three neighboring imidazole side chains of the HisB10 residue from three monomers and the fourth ligand being the chloride ion located at the three-fold axis. The source of the chloride ion was Tris.HCl used as buffer in the crystallization conditions. Additional stabilization of the hexamer assembly was mediated by the six phenol ligands located at the interface of two neighboring B-chains. Each phenol molecule was hydrogen bonded with the C=O of CysA6 and HN- of CysA11, and was in weak van der Waals contact (3.5 Å to 4.0 Å) with the side chains of HisB5 and LeuB11 residues of the two neighboring insulin lispro molecules.

The C-terminal region of the B-chain formed a crystallographic dimer through a short antiparallel beta-sheet formation, involving residue PheB24-TyrB26. Significant destabilization of this dimer interface by partial disruption of the network of van der Waals contacts involving the side chains of Tyr-B26 and Pro-B28 was noticed at the C-terminal region of the B-chain, caused presumably by the inversion of the ProB28-LysB29 residues in the insulin lispro protein molecule compared to the native insulin. Thus, it is hypothesized that the destabilization of the dimer formation due to the LysB28 - ProB29 sequence inversion enables this molecule to act as a rapid-acting insulin for the treatment of diabetes. [2]

Receptor binding assay. The synthetic insulin lispro was fully active in an insulin receptor-binding assay,[18] with a K_d of 0.12 ± 0.05 nM

Conclusions

In the work reported here, we have explored three approaches based on Fmoc chemistry SPPS to the total chemical synthesis of the insulin lispro protein molecule, via 'ester insulin' as a key intermediate product that facilitates folding of the synthetic protein molecule. The goal of each synthetic approach was an efficient total chemical synthesis of the 51 residue ester-linked polypeptide chain of ester insulin. Neither a standard stepwise Fmoc chemistry SPPS approach nor the hybrid Fmoc chemistry SPPS synthesis of protected peptides in solution were effective methods for the preparation of the 51 residue ester-linked polypeptide chain of ester insulin. In the case of stepwise Fmoc chemistry SPPS, the primary shortcomings were the instability of the Glu[O β (Thr)] ester bond to the conditions used to remove the Fmoc protecting group, together with an O/N-acyl shift (side reaction after Alloc deprotection at ThrB30), and formation of H-GlyA1-GluA4(OEt)-AsnA21-OH possibly generated by transesterification with Oxyma. In the hybrid method, synthesis of protected peptide segments by Fmoc chemistry SPPS was satisfactory after some optimization. However, the second solution condensation step was problematic, and could not be further optimized in our hands.

We conclude that the most effective method for the total synthesis of insulin via ester insulin based on peptide synthesis by Fmoc chemistry SPPS is the covalent condensation of unprotected peptide segments by native chemical ligation. Further optimization of these pilot studies should yield an effective total chemical synthesis of insulin lispro (Humalog).

Experimental Section

General experimental procedures

Fmoc amino acids were from Peptide Institute, Osaka, Japan, and were purchased from Peptides International (Kentucky). Protecting groups used were: Arg(Pbf), Asn(Trt), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Tyr(tBu). These amino acid building blocks HCl.H-Thr-OtBu, Fmoc-Lys(Boc)-Pro-OH, Fmoc-Gly-Ser[Ψ (Me,Me)Pro]-OH, and resins used in the synthesis: H-Ser(tBu)-O-2-Chlorotrityl-(S-DVB)resin, H-Gly-O-2-Chlorotrityl-(S-DVB)resin, H-Cys(Trt)-O-2-Chlorotrityl-(S-DVB)resin, Fmoc-Xanthenyl linker-(S-DVB)resin, were obtained from Bachem, Switzerland. 4-Methylpiperidine and piperidine were purchased from Sigma-Aldrich. Trifluoroacetic acid (TFA) was purchased from Halocarbon (New Jersey). N,N-Diisopropylethylamine (DIEA) was obtained from Applied Biosystems. Diethyl ether, dichloromethane (DCM), N,N-dimethylformamide (DMF), acetonitrile (HPLC-grade), and guanidine hydrochloride were purchased from Fisher. All other reagents were purchased from Sigma-Aldrich and were of the purest grade available.

Analytical HPLC conditions

Analytical reversed phase HPLC and LC-MS were performed using an Agilent 1100 series HPLC system equipped with an online MSD ion trap. All chromatographic separations were

performed on either a C4 (2.1×50 mm) column or C4 (4.6×150 mm) column or C8 (4.6×150 mm) Phenomenex column at 40 °C, using a linear gradient (5–45%) of solvent B in solvent A over 40 min (solvent A = 0.1% TFA in water, solvent B = 0.08% TFA in acetonitrile) at a flow rate of 0.5 mL/min for 2.1×50 mm column and 1.0 mL/min for 4.6×150 mm column, with detection by UV absorption at 214 nm. Masses were obtained by online electrospray mass spectrometry. All MS data shown were collected across the entire principal UV absorbing peak in each chromatogram.

Preparative HPLC conditions

Crude peptides and ligation reaction products were dissolved in 6 M Gn.HCl, acidified to pH 2–3, and filtered (0.22μ). The clear solution was then loaded onto either a C4 (10×100 mm) in-house packed column or a C18 (9.4×250 mm) Agilent Zorbax column, and the peptide components eluted at a flow rate of 5 mL/min using a shallow gradient of solvent B in solvent A (solvent A = 0.1% TFA in water, solvent B = 0.08% TFA in acetonitrile). Fractions containing the desired purified peptides were identified by analytical LC and mass spectrometry, then combined and lyophilized.

Stepwise Synthesis by Fmoc chemistry SPPS

The detailed description of the methodology used for the stepwise Fmoc chemistry SPPS is provided in the Supporting Information.

Hybrid SPPS-solution segment condensation synthesis

The protocols followed for the Fmoc chemistry SPPS of the protected peptides, their isolation and procedures for the hybrid SPPS-solution segment condensation are described in the Supporting Information.

Convergent synthesis by native chemical ligation

The procedures for the solid phase peptide synthesis of N-terminal Cys-containing peptides and peptide hydrazides-their conversion to peptide-αthioesters are provided in the Supporting Information.

Native chemical ligations to give H-GlyA1-GluA4(Oβ(ThrB30-PheB1))-AsnA21-COOH polypeptide chain—NCL-1. PheB1-ValB18-MES thioester (19.0 mg, 8.78 μmol, 4.18 mM) and GlyA1-GluA4[Oβ(ThrB30-CysB19)]-CysA6-αCONHNH₂ (20.0 mg, 9.75 μmol, 4.64 mM) were dissolved in 6 M Gn.HCl, 0.1 M Na₂HPO₄, 200mM MPAA, 50mM TCEP.HCl, pH 7.0. After analytical LC-MS of an aliquot indicated the completion of the reaction (Figure S17, Supporting Information), typically 36–48h, the ligation mixture was acidified to pH 3.0 and purified on C4(10×100 mm) semi-prep column. Isolated yield was 25.8 mg (6.3 μmol, 72% based on limiting segment PheB1-ValB18-MES thioester).

NCL-2. GlyA1-GluA4[Oβ(ThrB30-PheB1)]-CysA6-αCONHNH₂ (25.3 mg, 6.21 μmol, 4.35mM) was weighed into an 5 mL glass vial and dissolved in 1.43 mL aqueous 6 M Gn.HCl, 0.1 M phosphate buffer at pH 3.1. The peptide solution was cooled to –15 °C using an ice and salt mixture. Then 143 μL of 200mM NaNO₂ in water was added to the GlyA1-GluA4[Oβ(ThrB30-PheB1)]-CysA6-αCONHNH₂ solution at –15 °C. The reaction mixture

was vortexed gently and the oxidation reaction was allowed to proceed at -15°C for 30 min. During that time, a solution was prepared containing 200 mM MPAA, 6 M Gn.HCl, 0.1 M phosphate pH 6.8, and 1.43 mL of this solution was used to dissolve CysA7-AsnA21 (13.1 mg, 7.45 μmol). The above solution containing MPAA and Cys peptide was added to the oxidation reaction solution at -15°C . After 5 min, the reaction was brought to $0-5^{\circ}\text{C}$ for 10 min. The pH was adjusted to 6.1 and the reaction was then allowed to proceed at RT. For LC-MS analysis, a 2.1 μL aliquot was briefly treated with 10 μL 0.1 M TCEP.HCl, 6 M Gn.HCl, 0.1 M phosphate pH 7.0 solution for 5 min, then acidified with pH 3.0, 6 M Gn.HCl, 0.1 M phosphate buffer. Mass: calculated for GlyA1-GluA4[O β (ThrB30-PheB1)]-AsnA21, 5795.6 Da (average isotope composition); Obsd. 5795.4 ± 0.6 Da.

After 4h, the reaction mixture was treated with 100mM TCEP.HCl (86 mg), at pH 7.0, for 10 min and diluted 10-fold with water to precipitate the full length polypeptide and centrifuged. The clear supernatant was discarded and water wash of the precipitated product was repeated once again. Finally, the white precipitate was dissolved in 1:1 A/B HPLC buffer and analyzed by LC-MS before lyophilization. After lyophilization, 24.1 mg of the desired full length depsipeptide product was obtained (4.16 μmol , 67% yield based on limiting peptide GlyA1-GluA4[O β (ThrB30-PheB1)]-CysA6- α CONHNH₂).

Folding with concomitant formation of disulfide bonds to give ester insulin lispro—Ester insulin lispro polypeptide (23.6 mg, 4.07 μmol) was subjected to following folding reaction conditions: 1.5 M Gn.HCl, 20 mM Tris, 8 mM Cysteine, 1 mM Cystine. 2HCl at a polypeptide concentration of 0.05 mg/mL, pH 7.6. Typically, 23.6 mg of polypeptide was dissolved in 118 mL of 6 M Gn.HCl. Folding buffer exclusive of Gn.HCl was prepared by dissolving Tris (1.143 g, 9.44 mmol, 26.6 mM), Cysteine (457 mg, 3.77 mmol, 10.65 mM) and Cystine.2HCl (147.8 mg, 0.47 mmol, 1.33 mM) in 354 mL water, pH 7.6. Both the solutions were purged with helium for 15 min to remove any dissolved oxygen in the above solutions. Then the folding buffer was transferred to the 6 M Gn.HCl solution containing polypeptide and gently mixed, then left to stand airtight until LC-MS analysis showed completion of the folding reaction.

After acidification to pH 3.0, the folding reaction mixture was then subjected to solid phase extraction on a C18 cartridge to remove buffer salts. The product mixture obtained by elution with 1:1 HPLC solvents A/B was then lyophilized and purified by reverse phase HPLC on C18 (9.4 $\times$ 250 mm) semi-prep column. After purification, 5.5 mg (0.95 μmol , 23.3% isolated yield based on ester insulin lispro polypeptide) ester insulin lispro obtained. (Correcting for 6.8 mg of recovered starting ester-containing polypeptide, the calculated yield was 32.7%).

Conversion of folded ester insulin lispro to insulin lispro—Ester insulin lispro (5.1 mg, 0.88 μmol) was dissolved in 5.1 mL water and cooled to 4°C . To this solution was added 5.1 mL of 50 mM LiOH in water at 4°C . The reaction mixture was mixed gently and allowed to proceed at the same temperature until LC-MS showed the completion of reaction. After which time (5.5h), the reaction mixture was acidified to pH 3.0, and purified on a C18 (9.4 $\times$ 250mm) semi-prep column. Isolated yield 3.39 mg (0.58 μmol , 66%).

Receptor Binding Assays

Procedures for the receptor binding assays are described in the Supporting Information.

Crystal structures of synthetic insulin lispro

To determine the disulfide pairing and three-dimensional structure, the chemically synthesized human insulin lispro protein molecule was crystallized under the following conditions previously reported in literature.[16] Diffraction quality crystals were obtained from a (1:1) mixture of 10 mg/mL aqueous protein solution and the reservoir solution containing 0.5 M Na₂SO₄, 0.6 mM Zn(OAc)₂, 0.06% phenol and 0.3 M Tris at pH 8.2. X-ray diffraction data were acquired using synchrotron radiation at the Advance Photon Source, Argonne National Lab, from a crystal that diffracted to 1.35 Å resolution. The insulin lispro protein crystallized in the P2₁ space group. Matthews cell content analysis[19] revealed that there were six synthetic insulin lispro protein molecules in the asymmetric unit. The reported coordinates of the insulin lispro analog structure (PDB ID 3ZS2) were used as a search model to solve the crystal structure of synthetic insulin lispro by molecular replacement, using the program PHASER.[20]. The initial model was refined using PHENIX-REFINE[21] with iterative model building using COOT.[22] The electron density of the C-terminal few residues of the B-chain was poorly defined and had substantially increased thermal parameters (B-factors). The residue ThrB30 was not visible in any of the protein molecules in the asymmetric unit. The final model had an R-factor/R-free of 16.1%/19.4%.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was performed under a Sponsored Research Agreement between Bachem AG and The University of Chicago. Use of NE-CAT beamline 24-ID at the Advanced Photon Source is supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the Department of Energy, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357.

References

1. (a) Sanger F. *Science*. 1959; 129:1340–1344. [PubMed: 13658959] (b) Nicol DSHW, Smith LF. *Nature*. 1960; 187:483–485. [PubMed: 14426955] (c) Ryle AP, Sanger F, Smith LF, Kitai R. *Biochem J*. 1955; 60:541–556. [PubMed: 13249947]
2. DiMarchi, RD., Mayer, JP., Fan, L., Brems, DN., Frank, BH., Green, LK., Hoffmann, JA., Howey, DC., Long, HB., Shaw, WN., Shields, JE., Slieker, LJ., Su, KSE., Sundell, KL., Chance, RE. In: Smith, JA., River, JE., editors. *Peptides: Proceedings of the 12th American Peptide Symposium*; Leiden, The Netherlands: ESCOM Science Publisher B.V; 1992. p. 26-28.
3. Brange J, Volund A. *Adv Drug Delivery Rev*. 1999; 35:307–335.
4. (a) Hossain MA, Belgi A, Lin F, Zhang S, Shabanpoor F, Chan L, Belyea C, Truong HT, Blair AR, Andrikopoulos S, Tregear GW, Wade JD. *Bioconjugate Chem*. 2009; 20:1390–1396.(b) Marglin A, Merrifield RB. *J Am Chem Soc*. 1966; 88:5051–5052. [PubMed: 5978833] (c) Katsoyannis PG, Tometsko A, Zalut C. *J Am Chem Soc*. 1966; 88:166–167. [PubMed: 5900390] (d) Meienhofer J, Schnabel E, Bremer H, Brinkhoff O, Zabel R, Sroka W, Klostermeyer H, Brandenburg D, Okuda T, Zahn H. *Z Naturforsch B*. 1963; 18:1120–1121. [PubMed: 14117584] (e) Du YC, Jiang RQ, Tsou

- CL. Sci Sin (Engl Ed). 1965; 14:229–236.(f) Akaji K, Fujino K, Tatsumi T, Kiso Y. J Am Chem Soc. 1993; 115:11384–11392.(g) Liu F, Luo EY, Flora DB, Mayer JP. Org Lett. 2013; 15:960–963. [PubMed: 23391097]
5. Mayer JP, Zhang F, DiMarchi RD. Biopolymers (Peptide Science). 2007; 88:687–713. [PubMed: 17410596]
6. Sohma Y, Hua Q-X, Whittaker J, Weiss MA, Kent SBH. Angew Chem Int Ed. 2010; 49:5489–5493.
7. (a) Chan, WC., White, PD. Fmoc Solid Phase Peptide Synthesis. Chan, WC., White, PD., editors. Oxford University Press; Oxford: 2000. (b) Amblard M, Fehrentz J-A, Martinez J, Subra G. Mol Biotechnol. 2006; 33:239–254. [PubMed: 16946453]
8. Barlos K, Gatos D. Biopolymers (Peptide Science). 1999; 51:266–278. [PubMed: 10618595]
9. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Science. 1994; 266:776–779. [PubMed: 7973629]
10. Barany G, Merrifield RB. J Am Chem Soc. 1977; 99:7363–7365. [PubMed: 915158]
11. Barlos K, Gatos D, Kapolos S, Papaphotiu G, Schafer W, Wenqing Y. Tetrahedron Lett. 1989; 30:3947–3950.
12. Sieber P. Tetrahedron Lett. 1987; 28:2107–2110.
13. Fang GM, Wang JX, Liu L. Angew Chem Int Ed. 2012; 51:10347–10350.
14. (a) Avital-Shmilovici M, Mandal K, Gates ZP, Phillips NB, Weiss MA, Kent SBH. J Am Chem Soc. 2013; 135:3173–3185. [PubMed: 23343390] (b) Dhayalan B, Fitzpatrick A, Mandal K, Whittaker J, Weiss MA, Tokmakoff A, Kent SBH. ChemBioChem. 2016; 17:415–420. [PubMed: 26715336]
15. Zheng JS, Tang S, Qi YK, Wang ZP, Liu L. Nat Protoc. 2013; 8:2483–2495. [PubMed: 24232250]
16. Antoliová E, Žáková L, Turkenburg JP, Watson CJ, Hanlová I, Šanda M, Cooper A, Kraus T, Brzozowski AM, Jiráček J. J Biol Chem. 2011; 286:36968–36977. [PubMed: 21880708]
17. Ciszak E, Beals JM, Frank BH, Baker JC, Carter ND, Smith GD. Structure. 1995; 3:615–622. [PubMed: 8590022]
18. Pandeyarajan V, Phillips NB, Rege N, Lawrence MC, Whittaker J, Weiss MA. J Biol Chem. 2016; 291:12978–12990. [PubMed: 27129279]
19. Matthews BW. J Mol Biol. 1968; 33:491–497. [PubMed: 5700707]
20. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
21. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. Acta Crystallogr Sect D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
22. Emsley P, Cowtan K. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

Figure 1.

Amino acid sequence and covalent structure of ester insulin (CysA7-CysB7 inter-chain disulfide bond is not shown for the clarity).

Figure 2.

HPLC analysis of the crude material obtained from the stepwise Fmoc chemistry SPPS of the ester insulin 51 residue depsipeptide. * H-GlyA1-GluA4(OEt)-AsnA21-OH impurity, most likely from transesterification with Oxyma.

Figure 3.

LC-MS data for the segment condensation of protected peptides Boc-Phe^{B1}-Gly^{B23}-OH (a) and Boc-Gly^{A1}-Glu^{A4}[Oβ(Thr^{B30}-Phe^{B24}-H.HCl)]-Asn^{A21} (b), (Top panel) HPLC profile of the crude Boc-Gly^{A1}-Glu^{A4}[Oβ(Thr^{B30}-Phe^{B1})]-Asn^{A21}. (Bottom panel) Online ESI-MS spectrum taken across the whole of the UV signal for product peak 'c'. Mass: Observed for Boc-Gly^{A1}-Glu^{A4}[Oβ(Thr^{B30}-Phe^{B1})]-Asn^{A21}, 5795.0 ± 0.6 Da; Calculated, 5795.6 Da (average isotope composition).

Figure 4. (NCL-2). Oxidation, thioesterification of GlyA1-GluA4[Oβ(ThrB30-PheB1)]-CysA6-αCONHNH₂, and native chemical ligation with CysA7-AsnA21. (Top panel) Reverse phase HPLC. (Bottom panel) Online ESI-MS spectrum taken across the whole of the main UV peak. Mass: Observed for GlyA1-GluA4[Oβ(ThrB30-PheB1)]-AsnA21, 5795.6 ± 0.6 Da; Calculated, 5795.6 Da (average isotope composition).
 #, thiolactone; &, GlyA1-GluA4[Oβ(ThrB30-PheB1)]-CysA6-αCOOH;
 *, GlyA1-GluA4[Oβ(ThrB30-PheB1)]-CysA6-αCOSPhCH₂COOH;
 §, GlyA1-GluA4[Oβ(ThrB30-PheB1)]-AsnA21.

Figure 5. Folding/disulfide bond formation of ester insulin lispro. (Top panel) Reverse phase HPLC. (Bottom panel) Online ESI-MS spectrum taken across the whole of the UV peak. Mass: Observed for ester insulin lispro, 5790.1 ± 0.6 Da; Calculated, 5789.6 Da (average isotope composition).

Figure 6. LiOH saponification of ester insulin lispro. (Top panel) Reverse phase HPLC. (Bottom panel) Online ESI-MS spectrum taken across the whole of the main UV peak. Mass: Observed for insulin lispro, 5808.2 ± 0.6 Da; Calculated, 5807.6 Da (average isotope composition). a : Ester insulin lispro; b : insulin lispro; c : oxidized B-chain.

Figure 7.

Crystal structure of the synthetic insulin lispro obtained by X-ray diffraction. Cartoon representation of the insulin lispro hexamer (A) and insulin lispro monomer (B) (Colors: A-chain, magenta; B-chain, orange; Zn²⁺, blue; Cl⁻, green; phenol, cyan). Cys residues involved in disulfide bonds are numbered.

Scheme 1.

Synthesis of orthogonally protected Fmoc-Glu[Oβ(Alloc-Thr-OtBu)]-OH.

Scheme 2.

Synthesis of the 51 residue insulin lispro depsipeptide by stepwise Fmoc chemistry SPPS. SPPS was performed on H-Asn(Trt)-2CT-resin (AsnA21 attached to the resin via the carboxy group). All couplings were carried out with DIC activation.

Scheme 3.

'Hybrid synthesis' of the ester insulin 51 residue desipeptide. (A) Preparation of the three protected peptide segments by Fmoc chemistry SPPS; (B) segment condensation in solution.

Scheme 4.

The N → C native chemical ligation segment condensation strategy. Preparation of the three unprotected peptide segments. Phe^{B1}-Val^{B18}-αCOSCH₂CH₂SO₃H, Gly^{A1}-Glu^{A4}[Oβ (Thr^{B30}-Cys^{B19})]-Cys^{A6}-αCONHNH₂, and Cys^{A7}-Asn^{A21} by Fmoc chemistry SPPS is reported in the Supporting Information section.