Insulin

DOI: 10.1002/anie.201003018

Two-Chain Insulin from a Single-Chain Branched Depsipeptide Precursor: The End of a Long Journey**

*Hans-Jürgen Musiol and Luis Moroder**

hormones · insulin · peptides · protein folding · total synthesis

Diabetes was a deadly disease up to the first few decades of the twentieth century. Since the discovery of the pancreatic peptide hormone insulin in 1922 by Banting and Best from experiments on the canine pancreas and of its outstanding pharmacology in the treatment of diabetes with bovine pancreas extracts,[1] this larger peptide/miniprotein has been in the focus of medicinal, chemical, structural, immunological, biological and biotechnological research. These studies gave rise to a phenomenal number of milestones in the life sciences.[2] Despite all the advances insulin research has seen over the last century, this polypeptide is still a persisting challenge for further developments of novel insulin-based medicines with breakthrough efficacy, safety, and patient compliance.

Only one year after its discovery, an industrial insulin preparation based on bovine and porcine pancreas extracts was marketed by Hoechst in 1923 as a first treatment for human diabetes. The quality of this insulin preparation experienced continuous improvements in terms of purity and reproducibility, resulting in the first crystalline zinc complex of insulin in 1936 as a longer-acting drug. Due to the dramatic setback in the production of insulin during World War II and the slow and difficult post-war recovery, it was the disclosure of its primary structure by Sanger and co-workers in 1955[3] that paved the way towards the synthetic access of this molecule as additional or alternative source of the indispensable drug for the ever increasing number of insulin-dependent diabetes patients. In this context, the successful total synthesis of oxytocin and vasopressin by Du Vigneaud and co-workers in 1954[4] and the cysteine chemistry developed for this purpose was encouraging and promising.

The human insulin molecule comprises two peptide strands, a 21 amino acid A chain containing an intramolecular disulfide bridge, which is further linked by two interchain disulfide bonds to a 30 amino acid B chain (Figure 1).[3] Based

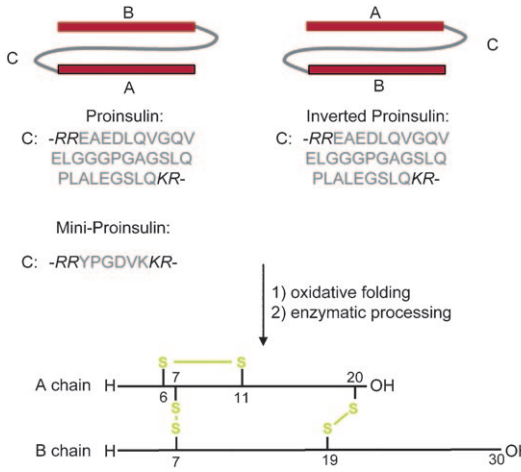

Figure 1. Conversion of native and non-native proinsulins into insulin.

on early experiments focusing on the recovery of insulin activity from naturally sourced A and B chains under oxidative conditions,[5] first total syntheses of insulin were reported across three continents in the early 1960s. The laboratories of Zahn (Germany),[17a] Katsoyannis (USA),[17b] and the Shanghai Insulin Research Group (China)[17c] utilized the oxidative assembly of synthetic A and B chains to produce native insulin, but at disappointingly low yields. This low-yield recovery of the native disulfide isomer defined the largest obstacle to insulin total syntheses by the two-chain approach. Although the method took advantage of S-sulfonate chain intermediates and of subsequent refinement of the mixed SSO_3H/SH combination strategy,[6] its limitations severely affected even the application in academic structure–function studies that are fundamental for improvements of the pharmacological performance of this molecule.

With the emergence of decisive advances in peptide chemistry in terms of orthogonal protection strategies, coupling procedures, and particularly of the thiol chemistry for synthesis of cystine-rich peptides both in solution and on solid supports, a successful bypass to the problematic A- and B-chain recombination method was realized. A regioselective disulfide pairing strategy was elaborated by Sieber and associates at Ciba-Geigy (Switzerland)[18a] on the basis of a stepwise assembly of fragments with concomitant inter- and intrachain disulfide formation. This synthetic route even allowed access to non-native disulfide isomers of insulin,

[*] H.-J. Musiol, Prof. L. Moroder
Max Planck Institute of Biochemistry
82152 Martinsried (Germany)
Fax: (+49) 898-578-2847
E-mail: moroder@biochem.mpg.de
Homepage: www.biochem.mpg.de/en/rg/moroder/

[**] The authors are grateful to Dr. Gerhard Müller, Proteros Fragments GmbH, Martinsried (Germany), for critical reading of the manuscript and for his help in preparing the figures.

which confirmed the importance of the native cystine connectivities for a properly folded, bioactive hormone. Finally, in the early 1990s, regioselective disulfide pairings between the fully synthetic A and B chain were achieved by Kiso and co-workers (Japan) by applying pairwise orthogonal thiol protection groups that led to the stepwise installation of the inter-/intrachain disulfide bonds for production of the native insulin isomer.[18b] Although both of these synthetic routes generated insulin at maximized purity of the final product, the yield of the total synthesis remained unsatisfactory. The low overall yields and the time and manpower demand of such syntheses, whilst brilliant from an academic point of view, is still the major obstacle to the development of insulin analogues in the quantities required for full biochemical, biological, and pharmacological analysis.

A feasible bypass to this bottleneck was expected from the knowledge gained on the biosynthesis of insulin and its biotechnological production by recombinant DNA technologies over the last few decades.[2a,c] The ribosomal synthesis of insulin occurs in pancreatic β cells as a single-chain preproinsulin consisting of a signal peptide located N-terminal to the B chain, which is connected via the C peptide to the N terminus of the A chain (Figure 1). Following the enzymatic removal of the signal peptide in the endoplasmatic reticulum, the B-C-A proinsulin folds with assistance of the competent folding enzyme catalysts, and the mature two-chain insulin is then produced by the action of prohormone convertases.[7a] Oxidative folding of the proinsulin into the native disulfide isomer proceeds with high efficiency (up to 80%) even in vitro, and its conversion to mature insulin can be achieved with trypsin and carboxypeptidase B (Figure 1).[7b] Inspired by the X-ray structure of insulin[8] that shows a spatial proximity of the N terminus of the A chain to the C terminus of the B chain, early cross-linking experiments of the two strands between the N^{ε}-Lys^{B29} and N^{α}-Gly^{A1} were performed with cleavable and non-cleavable bifunctional reagents to produce single-chain proinsulin mimics.[9] The high-yield recovery of the native insulin disulfide bonds clearly revealed that the C peptide merely acts as a tether of the two chains rather than to direct the oxidative folding process. Such simple tethering role of the C peptide was further substantiated by replacement of the large native connecting sequence with nonnative mini C peptides optimized in terms of oxidative folding yields of the linear precursors produced by secretory expression in *S. cerevisiae* and their subsequent conversion into mature insulin (Figure 1).[2c] Remarkably, even "inverted" proinsulin was found to fold correctly in high yield with the advantage of being significantly more active than the practically inactive proinsulin molecule (Figure 1).[10]

Although biotechnological approaches based on proinsulin and various optimized miniproinsulins serve nowadays for the commercial production of insulin in *E. coli* and *S. cerevisiae*, respectively,[2a,c] such linear peptidic precursors have found only sporadic application in the chemical synthesis of insulin because of the solubility problems that have been encountered. These could only be bypassed, for example, by N-terminal extensions with solubilizing sequences.[11] However, the resulting increased size of the precursors raises severe limitations to yields of highly purified precursors as required for efficient folding and enzymatic conversion processes. First attempts to resolve this peptide size complication were explored by Kent and Sohma[12] by utilizing a biomimetic design of mini-proinsulins assembled by oxime ligation from synthetic A- and B-chain derivatives containing the two peptidic extensions at selected positions with suitable reactive groups to cross-link N^{α}-Gly^{A1} with N^{ε}-Lys^{B28} or N^{α}-Phe^{B1} and alternatively C^{α}-Thr^{B30} with N^{α}-Gly^{A1}. Only the latter classical proinsulin-like tethering led to the correct oxidative A- and B-chain assembly and upon subsequent enzymatic excision to the insulin molecule in satisfactory yields.

The insulin variant containing an inverted Pro^{B28}-Lys^{B29} sequence, that is, Lys-Pro, was found to exhibit a significantly decreased aggregation tendency and rapid pharmacokinetics.[13] Because of these beneficial properties, the Lys^{B28},Pro^{B29}-insulin became the first commercially registered insulin analogue as active ingredient of Humalog (Eli Lilly & Co).[2a] As shown in Figure 2, the close proximity of the side-chain β-hydroxy group of Thr^{B30} and γ-carboxy group of Glu^{A4} in the X-ray structure of Lys^{B28},Pro^{B29}-insulin[14] inspired Kent

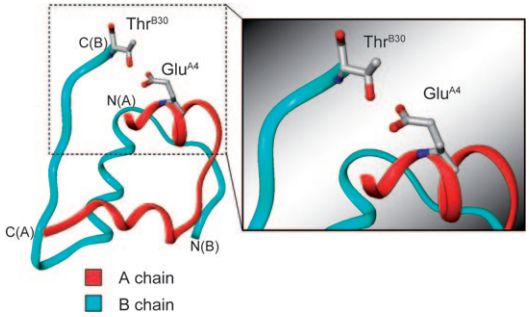

Figure 2. Left: X-ray structure of Lys^{B28},Pro^{B29}-insulin (PDB code: 1LPH) with the side chains of Glu^{A4} and Thr^{B30} depicted in the ribbon presentation of both chains. Right: Detailed picture of the C-terminal region of the B chain adjacent to the N-terminal region of the A chain, showing the close proximity of Glu^{A4} and Thr^{B30} side chains.

and co-workers to explore an ester linkage between the A- and B-chain to produce a new single-chain insulin precursor completely lacking a peptidic or nonpeptidic C-peptide-like tether.[15] Moreover, it had been well-established that the solubility of peptides with high aggregation propensities benefit significantly from incorporation of depsipeptide bonds.[16] Thus, by taking advantage of this property of ester bonds in peptides and of the chemistry developed for such purposes as well as of advanced chemical ligation methods, the single-chain side-chain branched peptide shown in Figure 3 was readily assembled without solubility problems, despite its significant size. In accordance with the underlying design principle applied by Kent and co-workers, it contains the depsipeptide unit between the C terminus of the B chain and the side-chain carboxy group of Glu^{A4}. Most surprisingly, oxidative folding into the correct insulin disulfide topology proceeded in yields comparable to those of proinsulin, and the last chemical saponification step with sodium hydroxide

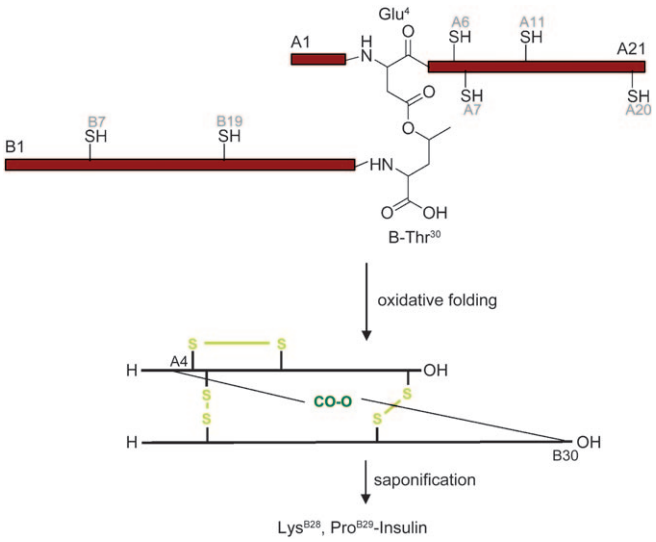

Figure 3. Conversion of the synthetic branched peptide precursor into Lys^{B28},Pro^{B29}-insulin.

occurred without significant side reactions, leading to the two-chain target insulin molecule in high yields and high purity. It is to be expected that even this last step could benefit from less drastic saponification procedures than the applied alkaline hydrolysis.

With this synthetic route of Kent and co-workers, a journey over several decades towards efficient total syntheses of insulin may have reached its final stage. Indeed, it offers a more facile and efficient synthetic access to insulin analogues with the unlimited diversity of non-natural amino acid residues. Such analogues may possibly allow increasing the potency of insulin-based life-saving medicines, to extend the time of action as well as to control it using prodrugs, and to enhance the bioavailability of this peptidic drug.

Received: May 18, 2010
Published online: August 16, 2010

[1] a) F. G. Banting, C. H. Best, *J. Clin. Med.* **1922**, *7*, 251 – 266; b) F. G. Banting, C. H. Best, *J. Clin. Med.* **1922**, *7*,464 – 472; c) F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell, A. A. Fletcher, *Can. Med. Assoc. J.* **1922**, *7*, 141 – 146.

[2] For reviews in synthesis, semisynthesis, and biotechnological preparation of insulin, see: a) J. P. Mayer, F. Zhang, R. D. DiMarchi, *Biopolymers* **2007**, *88*, 687 – 713; b) J. D. Wade in *Oxidative Folding of Peptides and Proteins* (Eds.: J. Buchner, L. Moroder) Royal Society of Chemistry, Cambridge, **2009**, pp. 345 – 366; c) T. Kjeldsen, *Appl. Microbiol. Biotechnol.* **2000**, *54*, 277 – 286.

[3] A. P. Ryle, F. Sanger, L. F. Smith, R. Kitai, *Biochem. J.* **1955**, *60*, 541 – 556.

[4] V. Du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, *J. Am. Chem. Soc.* **1954**, *76*, 3115 – 3121.

[5] H. Dixon, A. C. Wardlaw, *Nature* **1960**, *188*, 721 – 724.

[6] R. E. Chance, J. A. Hoffmann, E. P. Kroeff, M. G. Johnson, E. W. Schirmer, W. W. Bromer, M. J. Ross, R. Wetzel in *Peptides, Synthesis Structure and Function*, 7th American Peptide Symposium (Eds.: D. H. Rich, E. Gross), **1981**, pp. 721–728.

[7] a) D. F. Steiner, D. Cunningham, L. Spiegelman, B. Aten, *Science* **1967**, *157*, 697 – 700; b) D. F. Steiner, J. L. Clark, *Proc. Natl. Acad. Sci. USA* **1968**, *60*, 622 – 629.

[8] a) M. J. Adams, T. L. Blundell, E. J. Dodson, G. G. Dodson, M. Vijayan, E. N. Baker, M. M. Harding, D. C. Hodgkin, B. Rimmer, S. Sheat, *Nature* **1969**, *224*, 491 – 495; b) E. N. Baker, T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. M. C. Hodgkin, R. E. Hubbard, N. W. Isaacs, C. D. Reynolds, K. Sakabe, N. Sakabe, N. M. Vijayan, *Philos. Trans. R. Soc. London Ser. B* **1988**, *319*, 369 – 456.

[9] a) A. Wollmer, D. Brandenburg, H.-P. Vogt, W. Schermutzki, *Hoppe Seyler's Z. Physiol. Chem.* **1974**, *355*, 1471 – 1476; b) R. Obermeier, R. Geiger, *Hoppe Seyler's Z. Physiol. Chem.* **1975**, *356*, 1631 – 1634; c) W.-D. Busse, F. H. Carpenter, *Biochemistry* **1976**, *15*, 1649 – 1657.

[10] W. F. Heath, R. M. Belagaje, G. S. Brooke, R. E. Chance, J. A. Hoffmann, H. B. Long, S. G. Reams, C. Roundtree, W. N. Shaw, L. J. Slieker, K. L. Sundell, R. D. DiMarchi, *J. Biol. Chem.* **1992**, *267*, 419 – 425.

[11] A. P. Tofteng, K. J. Jensen, L. Schäffer, T. Hoeg-Jensen, *ChemBioChem* **2008**, *9*, 2989 – 2996.

[12] Y. Sohma, S. B. H. Kent, *J. Am. Chem. Soc.* **2009**, *131*, 16313 – 16318.

[13] D. N. Brems, L. A. Alter, M. J. Beckage, R. E. Chance, R. D. DiMarchi, L. K. Green, H. B. Long, A. H. Pekar, J. E. Shields, B. H. Frank, *Protein Eng.* **1992**, *6*, 527 – 533.

[14] E. Ciszak, J. M. Beals, B. H. Frank, J. C. Baker, N. D. Carter, G. D. Smith, *Structure* **1995**, *3*, 615 – 622.

[15] Y. Sohma, Q.-X. Hua, J. Whittaker, M. A. Weiss, S. B. H. Kent, *Angew. Chem.* **2010**, *122*, 5621—5625; *Angew. Chem. Int. Ed.* **2010**, *49*, 5489 – 5493.

[16] For a review, see: a) Y. Sohma, T. Yoshiya, A. Taniguchi, T. Kimura, Y. Hayashi, Y. Kiso, *Biopolymers* **2007**, *88*, 253 – 262; b) I. Coin, *J. Pept. Sci.* **2010**, *16*, 223 – 230.

[17] a) J. Meienhofer, E. Schnabel, H. Bremer, O. Brinkhoff, R. Zabel, W. Sroka, T. H. Klostermeyer, D. Brandenburg, T. Okuda, H. Zahn, *Z. Naturforsch. B* **1963**, *18*, 1120 – 1121; b) P. G. Katsoyannis, K. Fukuda, A. Tometsko, K. Suzuki, M. Tilak, *J. Am. Chem. Soc.* **1964**, *86*, 930—932; c) Y. T. Kung, Y. C. Du, W. T. Huang, C. C. Chen, L. T. Ke, S. C. Hu, R. Q. Jiang, S. Q. Chu, C. I. Niu, J. Z. Hsu, W. C. Chang, L. L. Cheng, H. S. Li, Y. Wang, T. P. Loh, A. H. Chi, C. H. Li, P. T. Shi, Y. H. Yieh, K. L. Tang, C. Y. Hsing, *Sci. Sinica* **1965**, *14*, 1710-1716.

[18] a) P. Sieber, B. Kamber, A. Hartmann, A. Jöhl, B. Riniker, W. Rittel, *Helv. Chim. Acta* **1974**, *57*, 2617 – 2621; b) K. Akaji, K. Fujino, T. Tatsumi, Y. Kiso, *J. Am. Chem. Soc.* **1993**, *115*, 11384 – 11392.

