ORGANIC LETTERS

2013
Vol. 15, No. 4
960–963

Concise Synthetic Routes to Human Insulin

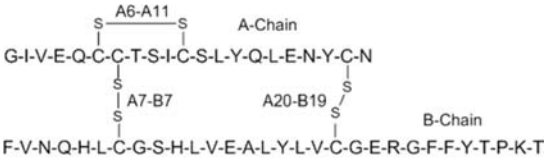
Fa Liu,* Ethan Y. Luo, David B. Flora, and John P. Mayer

Lilly Research Laboratories, Indianapolis, Indiana 46285, United States

liufx@lilly.com

Received January 18, 2013

ABSTRACT



Human Insulin (HI), **1**

We report a set of concise and efficient routes for the chemical synthesis of human insulin using a two- or three-step combination procedure that employs Trt, Acm, and *t*-Bu cysteine protection schemes. Starting with resin-bound assembled A and B chains, human insulin can be obtained within the span of a single work day in 5.4% overall yield based on the crude A or B chain.

Accessibility to desired molecules is always a key prerequisite to enable the validation of a biological hypothesis, and efficient material generation methods can potentially expedite this process. Current state of the art solid-phase peptide synthesis techniques have greatly advanced the ability to obtain most peptides of interest. However, the construction of multiple disulfide bridges within a peptide often remains challenging, mainly due to the tedious procedures and associated poor yields.[1,2] Naturally occurring cysteine rich peptides are structurally unique molecules with important bioactivities;[3−7] the insulin super family[8] represents a special subclass that can be distinguished by a characteristic two chain topology. Insulin is the most important member in its super family, which was identified in 1922 as a life saving molecule.[9] Despite nearly a century of investigation into its biology and therapeutic applications, insulin research remains a high priority, underscoring its indispensible role in diabetes treatment.[8] Human insulin (HI, **1**) consists of two individual chains (A and B) that are connected by two interchain disulfide bonds (A7−B7 and A20−B19). A third intra-A chain disulfide bond (A6-A11) further stabilizes its global conformation. Insulin was originally obtained by extraction of bovine and porcine pancreatic glands; it has also been made by semisynthesis, total chemical synthesis, and recombinant expression.[10] Currently, while all pharmaceutical insulin products are of recombinant origin, chemical synthesis still remains of great interest in discovery applications where a large number of analogs are generally required. For such a purpose, fast turnover times with reasonable yields become elements of practicality. Several elegant examples of insulin chemical synthesis have been reported, including cysteine S-sulfonate mediated chain combination[11−14] or refolding[15−17] and directed disulfide bond formation.[18−21] Considerable chemistry challenges have been overcome by existing methodology;[22] however, optimization for synthetic conciseness has been overlooked. Herein, we report a set of concise and efficient synthetic routes to human insulin.

(1) Andreu, D.; Albericio, F.; Sole, N. A.; Munson, M. C.; Ferrer, M.; Barany, G. *Methods Mol. Biol. (Totowa, N. J.)* **1994**, *35*, 91.
(2) Boulegue, C.; Musiol, H.-J.; Prasad, V.; Moroder, L. *Chim. Oggi* **2006**, *24*, 24.
(3) Craik, D. J.; Daly, N. L.; Mulvenna, J.; Plan, M. R.; Trabi, M. *Curr. Protein Pept. Sci.* **2004**, *5*, 297.
(4) Dimarcq, J. L.; Bulet, P.; Hetru, C.; Hoffmann, J. *Biopolymers* **1998**, *47*, 465.
(5) Garcia-Olmedo, F.; Molina, A.; Alamillo, J. M.; Rodriguez-Palenzuela, P. *Biopolymers* **1999**, *47*, 479.
(6) Loughnan, M. L.; Alewood, P. F. *Eur. J. Biochem.* **2004**, *271*, 2294.
(7) Terlau, H.; Olivera, B. M. *Physiol. Rev.* **2004**, *84*, 41.
(8) Litwack, G., Ed. *Vitamins and Hormones: Insulin and IGFs*; Elsevier: 2009.
(9) Banting, F. G.; Best, C. H. *J. Lab. Clin. Med.* **1922**, *7*, 251.
(10) Mayer, J. P.; Zhang, F.; DiMarchi, R. D. *Biopolymers* **2007**, *88*, 687.
(11) Katsoyannis, P. G.; Tometsko, A.; Zalut, C. *J. Am. Chem. Soc.* **1966**, *88*, 166.
(12) Kung, Y.-T. i.; et al. *Sci. Sin. (Engl. Ed.)* **1965**, *14*, 1710.
(13) Marglin, A.; Merrifield, R. B. *J. Am. Chem. Soc.* **1966**, *88*, 5051.
(14) Meienhofer, J.; Schnabel, E.; Bremer, H.; Brinkhoff, O.; Zabel, R.; Sroka, W.; Klostermayer, H.; Brandenburg, D.; Okuda, T.; Zahn, H. *Z. Naturforsch. B* **1963**, *18*, 1120.
(15) Sohma, Y.; Hua, Q.-X.; Whittaker, J.; Weiss, M. A.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 5489.
(16) Sohma, Y.; Kent, S. B. H. *J. Am. Chem. Soc.* **2009**, *131*, 16313.
(17) Tofteng, A. P.; Jensen, K. J.; Schaffer, L.; Hoeg-Jensen, T. *ChemBioChem* **2008**, *9*, 2989.

10.1021/ol400149j
Published on Web 02/07/2013

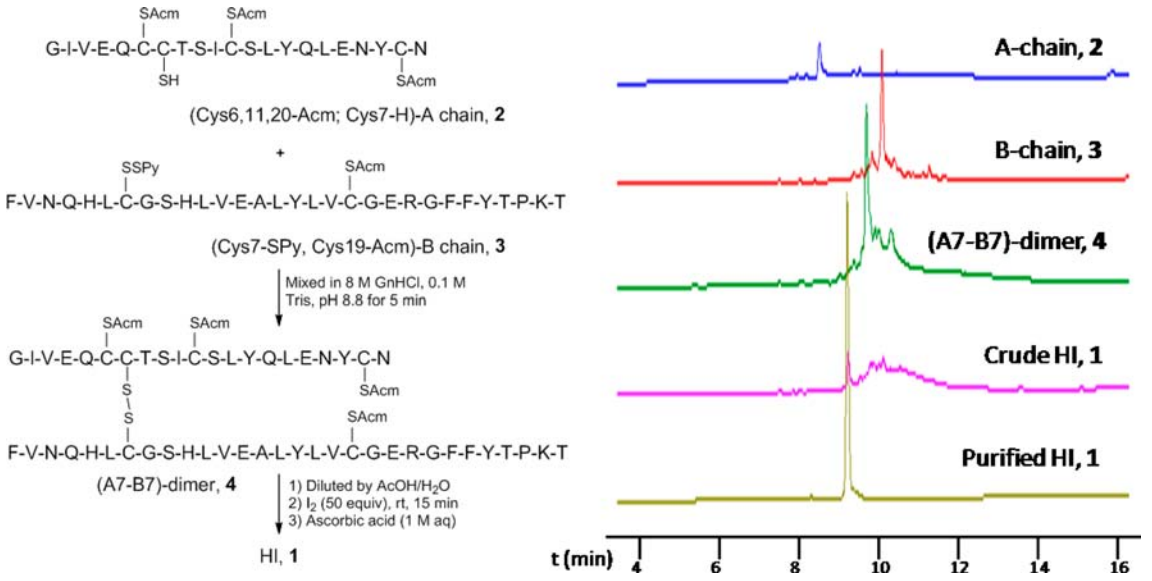


Figure 1. Two-step combination protocol through (A7-B7)-dimer.

Two key challenges need to be addressed in order to improve upon existing methodologies: first, the hydrophobicity of the A-chain and, second, the pairing of three native disulfide bonds.[10] Although the highly hydrophobic A-chain and A−B heterodimers (with one or two disulfide bonds) are problematic for separation by RP-HPLC (reversed phase high performance liquid chromatography),[19] the fully formed HI **1** behaves well on preparative RP-columns. Accordingly, we sought a one-pot procedure that would require a single purification step to obtain the final product **1**. In so doing, we explored two- and three-step disulfide bond construction protocols that utilized trityl (Trt),[23] acetamidomethyl (Acm),[24] and *tert*-butyl (*t*-Bu)[25] cysteine protection schemes.

Two-Step Combination through (A7-B7) or (A20-B19)-Dimers (Cys Protection Using 2 Trt and 4 Acm). A-chains **2**, **5** and B-chains **3**, **6** were assembled on PEG based or PEG grafted resins (ChemMatrix or NovaPEG resin) which were proven superior for the synthesis of a hydrophobic peptide.[26,27] Initially, the B-chains were cleaved from the resin first and then activated using DTDP (2,2′-dithiodipyridine) in AcOH (acetic acid)/isopropanol.[18] However, these conditions were found to be suboptimal due to the low solubility of the B-chain and the slow conversion rate. In addition, two separate steps were required for resin cleavage and cysteine activation. We found that neat TFA (trifluoroacetic acid) proved to be an excellent solvent allowing the DTDP activation to be completed within 30 min. The activated peptide was then precipitated using cold Et_2O (diethyl ether), and excess DTDP was easily removed by subsequent Et_2O washings. DTNP (2,2′-dithiobis(5-nitropyridine))[28] was also evaluated in this context. Although both DTDP and DTNP activated B-chains performed comparably in the thiolysis combination step, the low solubility of DTNP in Et_2O made it less preferable since excess DTNP during the B-chain activation could not be removed by Et_2O washings. In the final optimized procedure to obtain the activated B-chain, the resin cleavage and cysteine activation were conducted simultaneously by treating with a TFA/TIS (triisopropylsilane)/H_2O/DTDP cocktail for 1.5 h.[29] For both the (A7−B7) and (A20−B19)-dimer mediated routes (Figures 1 and 2 respectively), conditions for the thiolysis chain combination and the subsequent I_2 (iodine) oxidation were identical. In each case, the crude A-chain **2** or **5** (1.0 equiv) and its paired B-chain **3** or **6** (1.0 equiv) were mixed in 8 M GnHCl, 0.1 M tris buffer (pH 8.8), and the thiolysis chain combination was completed within 5 min as monitored by LC-MS. The resulting solution was then diluted by a 5-fold volume of AcOH/H_2O (4:1), followed by oxidation with I_2 (50 equiv based

(18) Akaji, K.; Fujino, K.; Tatsumi, T.; Kiso, Y. *J. Am. Chem. Soc.* **1993**, *115*, 11384.
(19) Hossain, M. A.; Belgi, A.; Lin, F.; Zhang, S.; Shabanpoor, F.; Chan, L.; Belyea, C.; Truong, H.-T.; Blair, A. R.; Andrikopoulos, S.; Tregear, G. W.; Wade, J. D. *Bioconjugate Chem.* **2009**, *20*, 1390.
(20) Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Riniker, B.; Rittel, W. *Helv. Chim. Acta* **1974**, *57*, 2617.
(21) Han, J.; Cheng, S.; DiMarchi, R. D. *Peptide Science* **2009**, *92* (P312), 371.
(22) Belgi, A.; Hossain, M. A.; Tregear, G. W.; Wade, J. D. *Immunol., Endocr. Metab. Agents Med. Chem.* **2011**, *11*, 40.
(23) McCurdy, S. N. *Pept. Res.* **1989**, *2*, 147.
(24) Veber, D. F.; Milkowski, J. D.; Varga, S. L.; Denkewalter, R. G.; Hirschmann, R. *J. Am. Chem. Soc.* **1972**, *94*, 5456.
(25) Callahan, F. M.; Anderson, G. W.; Paul, R.; Zimmerman, J. E. *J. Am. Chem. Soc.* **1963**, *85*, 201.
(26) de la Torre, B. G.; Jakab, A.; Andreu, D. *Int. J. Pept. Res. Ther.* **2007**, *13*, 265.
(27) Garcia-Martin, F.; Quintanar-Audelo, M.; Garcia-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Cote, S.; Tulla-Puche, J.; Albericio, F. *J. Comb. Chem.* **2006**, *8*, 213.
(28) Rabanal, F.; DeGrado, W. F.; Dutton, P. L. *Tetrahedron Lett.* **1996**, *37*, 1347.
(29) Ghosh, A. K.; Fan, E. *Tetrahedron Lett.* **2000**, *41*, 165.

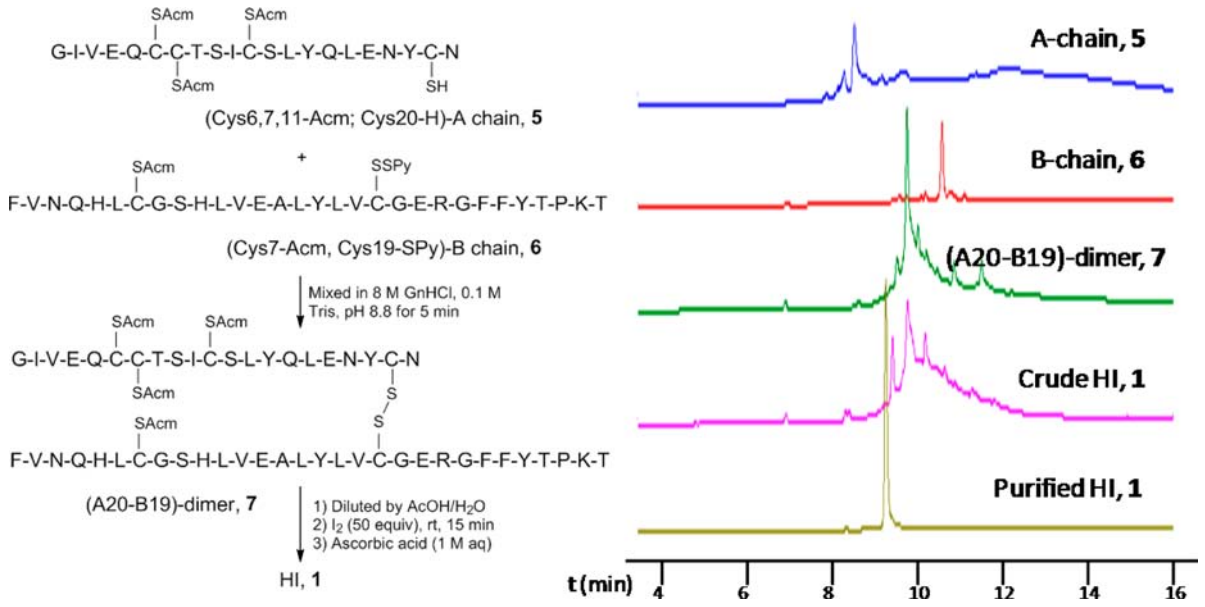


Figure 2. Two-step combination protocol through (A20-B19)-dimer.

on A- or B-chain). This removed the four Acm groups and simultaneously formed the remaining two disulfide bonds in a process that was generally completed within 15 min. More dilute solutions were also examined for the I_2 oxidation, but these were found not to improve the yield. The I_2 oxidation was quenched by the addition of ascorbic acid (1.0 M aq), and the resulting product was subjected to RP-HPLC purification to provide HI **1**.

Overall, the routes through the (A7-B7)-dimer **4** and the (A20-B19)-dimer **7** were largely comparable, although it appeared the former route provided a cleaner LC-MS profile following I_2 oxidation. The native disulfide pattern of synthetic **1** was unambiguously confirmed by endoproteinase Glu-C mapping, using authentic HI as a reference standard (Supporting Information: Figures S4, S6 and Table S1). Notably, the correctly folded HI elutes as the earliest fraction by RP-HPLC, which considerably simplifies purification. The early elution of native **1** can be explained by the normally buried hydrophobic core residues of the native protein being exposed to some degree in mis-folded isomers, leading to later elution on RP-HPLC.

Three-Step Combination (Cys Protections Using 2 Trt, 2 Acm, and 2 *t*-Bu). In the two-step combination procedures, the formation of the last two disulfide bonds during the I_2 oxidation did not proceed in a regioselective manner, the HI corresponding mass signals (5808 with four or five charges) were consistently detected in the regions eluted just after the correctly folded **1** on the LC-MS traces. These were assumed to be the mis-folded isomers (Supporting Information Figures S1 and S2). To achieve a more selective process, a stepwise disulfide bond construction strategy was employed. To facilitate an overall three-step sequential disulfide formation protocol, *t*-Bu protection was selected for Cys6 and Cys11 in A-chain **8** since *t*-Bu is considered to be orthogonal to the existing Trt and Acm groups. The B-chain **3** cleavage/activation and subsequent thiolysis combination with the A-chain **8** was conducted exactly as in the previously discussed two-step combination procedure. However, it was difficult to obtain a clear analytical RP-HPLC of A-chain **8** presumably due to its higher hydrophobicity. A partial purification or desalting of the (A7-B7)-dimer **9** was performed to facilitate solvent exchange in the subsequent steps. The recovery of **9** from the preparative C4, C8, and C18 RP-HPLC column was found to be extremely low. RP-HPLC using a C1 phase was eventually found to give an acceptable recovery. A C1 or cyano solid-phase extraction catridge served the same purpose, since desalting is the main objective of this step. The lyophilized dimer **9** was dissolved in AcOH/H_2O (95/5) with 25 equiv of I_2. This achieved Acm removal and resulted in complete A20−B19 disulfide bond formation within 15 min. The crude (A7-B7, A20-B19)-dimer **10** was precipitated by adding cold Et_2O, and the resulting solid was briefly air-dried before redissolving in TFA with 5% DMSO (45 min).[30] This cleanly cleaved the *t*-Bu groups and simultaneously formed the A6−A11 disulfide bridge (Figure 3). Crude HI was precipitated by adding cold Et_2O and purified by RP-HPLC. The native disulfide pattern of HI **1** prepared by this route was also confirmed by Glu-C mapping (Supporting Information: Figures S4, S6 and Table S1). The silyl chloride−sulfoxide method[31] was also attempted in the final step. However, it was found to

(30) Cuthbertson, A.; Indrevoll, B. *Org. Lett.* **2003**, *5*, 2955.

(31) Akaji, K.; Tatsumi, T.; Yoshida, M.; Kimura, T.; Fujiwara, Y.; Kiso, Y. *J. Am. Chem. Soc.* **1992**, *114*, 4137.

(32) One possible reason could be the starting material of this step in our procedure was relatively impure comparing to Kiso's experiment (ref 18), and these impurities might accelerate the disulfide reshuffling.

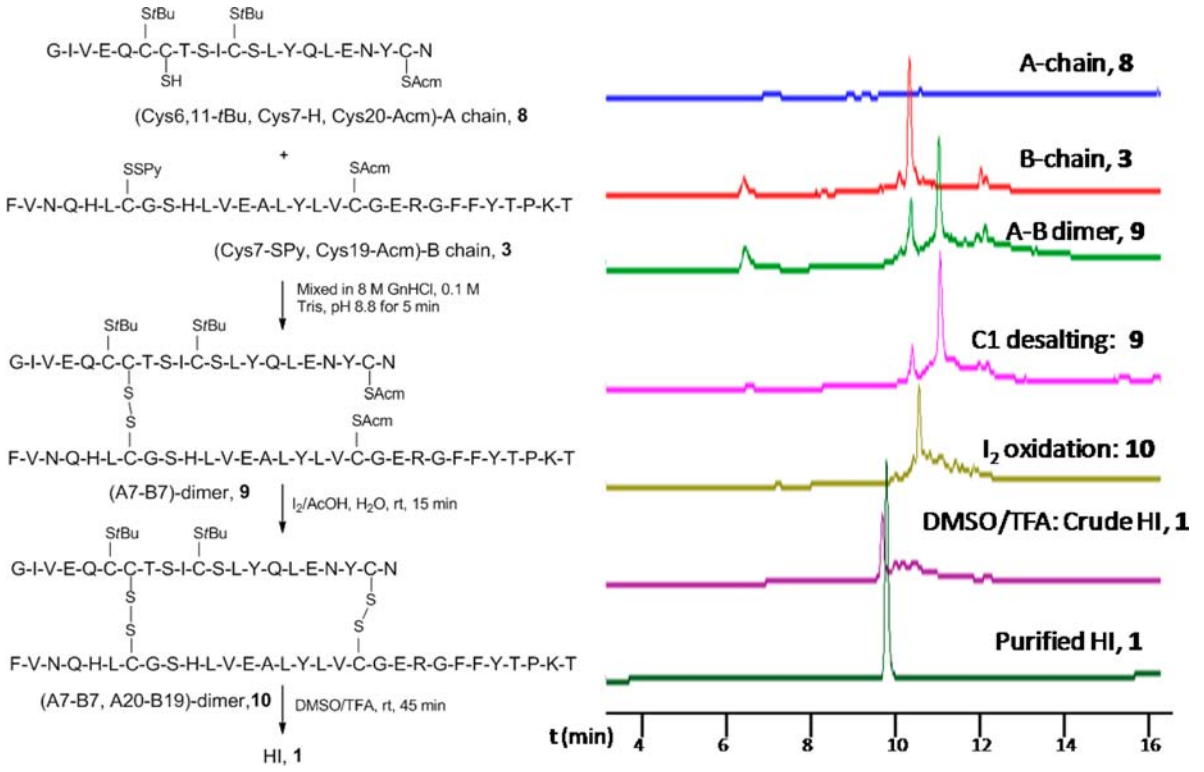


Figure 3. Three-step combination protocol.

degrade dimer **10** into the individual A and B chain, which was presumably caused by disulfide bonds reshuffling.[32]

In summary, we report herein a set of concise and efficient procedures for the total chemical synthesis of insulin. Starting with the resin-bound assembled A and B chains, the desired insulin can be obtained within the span of a single work day in reasonable yield. The two- and three-step combination protocols which we report are comparably practical, although the three-step approach appears slightly superior based on crude LC-MS profiles. The one-pot concept described here for the synthesis of human insulin could also be applied to other cysteine-rich peptides that do not contain methionine and tryptophan in the sequence.

Acknowledgment. We thank Dr. Adam R. Mezo, Dr. John M. Beals, and Dr. Scott D. Putney (Lilly Research Laboratories, Indianapolis) for their helpful discussions and thank Ms. Julia A. Cummins (Lilly Research Laboratories, Indianapolis) for her help with setting up solid phase peptide synthesis.

Supporting Information Available. The experimental procedures and spectroscopic data. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.