VIP Protein Folding

DOI: 10.1002/anie.201001151

Design and Folding of [Glu^{A4}(O$^{\beta}$Thr^{B30})]Insulin ("Ester Insulin"): A Minimal Proinsulin Surrogate that Can Be Chemically Converted into Human Insulin**

Youhei Sohma, Qing-Xin Hua, Jonathan Whittaker, Michael A. Weiss, and Stephen B. H. Kent**

Insulin biosynthesis involves the efficient folding of a single polypeptide-chain precursor with concomitant formation of three disulfide bonds to give proinsulin followed by enzymatic removal of the C peptide to give mature insulin.[1,2] A proinsulin- or miniproinsulin-based approach is currently used in the recombinant production of human insulin.[3,4] However, recombinant production of insulin analogues is effectively limited to the creation of mutants from the twenty genetically encoded amino acids. In contrast to this, total chemical synthesis of insulin would in principle enable the incorporation of a wide range of nonnatural amino acids and other chemical modifications into the molecule,[5] and would thus enable the full exploration of the medicinal chemistry of this important therapeutic molecule. Until now, however, we have lacked an efficient approach to the chemical synthesis of human insulin.[5] This deficiency has impeded the development of next-generation insulin analogues containing non-standard side chains, D-amino acids,[6,7] or other novel chemical structural features.

Current chemical methods for insulin synthesis are limited by chain combination to give the three native disulfide bonds. Early chemical syntheses of insulin relied on inefficient folding/disulfide bond formation from separate A and B chains, which were prepared by solution synthesis[8–10] or solid phase peptide synthesis.[11] More recently, it has been found that optimal folding/disulfide bond formation requires a two-to-threefold stoichiometric excess of A chain over the B chain, and gives only approximately 12% folding yield based on the limiting amount of B chain;[12] because of the excess A chain used, only about 7% of the total weight of A and B chains ends up as final insulin product. Total synthesis with chemically directed formation of the disulfide bonds has been reported,[13,14] but has not found widespread use.

An alternative approach to high-yield folding/disulfide bond formation in the insulin molecule is to use a chemical tether to mimic the effect of covalently linking the A and B chains as occurs in proinsulin and miniproinsulin precursors. Most of the previously studied chemically tethered insulin precursors have involved covalent linking of the N terminal of the insulin A chain to the side chain of Lys^{B29}, which is near the C terminal of the insulin B chain.[15–20] With the goal of achieving an efficient total synthesis of human insulin and analogues, a variety of different length chemical tethers between these two functionalities has been explored, using both noncleavable[15,16] and cleavable[17–20] tethers. The shortest tether reported to be effective in promoting high yield folding/disulfide bond formation contained eight carbon atoms.[15]

Recently, there have been attempts to extend the chemical-tether approach to provide a more effective total chemical synthesis of insulin.[21,22] Our research group has reported a proof-of-principle synthesis of human insulin through a chemically synthesized miniproinsulin prepared by oxime-forming ligation.[22] A temporary chemical tether that linked the N terminus of the A chain to Lys^{B28} near the C terminal of the B chain enabled us to fold/form disulfide bonds with high efficiency. However, our approach involved a relatively long and complex chemical tether, which made the synthesis laborious. In addition, it was necessary to remove the chemical tether enzymatically in a subsequent step, as was the case for a similar miniproinsulin chemical synthesis approach to insulin(desB30).[21] Thus, the strategy was not practical for the efficient generation of chemical analogues of insulin.

Ideal features of an optimal chemically tethered miniproinsulin would include: 1) straightforward preparation by existing synthetic methods, 2) efficient folding/disulfide bond formation, and 3) ready chemical conversion to mature insulin. A model target molecule is provided by *Insulin Lispro* ([Lys^{B28},Pro^{B29}]-human insulin (herein called "KP-insulin"), the active ingredient of Humalog (Eli Lilly and Co.), a first-generation, fast-acting insulin analogue widely used for the treatment of diabetes mellitus.[23] In examining the three-dimensional structure of KP-insulin,[24] we noticed that the β-hydroxyl group of Thr^{B30} was in close proximity to, and virtually in contact with, the side-chain carboxyl group of

[*] Prof. Dr. Y. Sohma,[+] Prof. Dr. S. B. H. Kent
Department of Biochemistry and Molecular Biology
Department of Chemistry, Institute for Biophysical Dynamics
The University of Chicago, Chicago, IL 60637 (USA)
E-mail: skent@uchicago.edu
Dr. Q.-X. Hua, Prof. Dr. J. Whittaker, Prof. Dr. M. A. Weiss
Department of Biochemistry, Case Western Reserve University
Cleveland, OH 44106 (USA)

[+] Present address: Department of Medicinal Chemistry
Kyoto Pharmaceutical University
Yamashina-ku, Kyoto 607-8412 (Japan)
E-mail: ysohma@mb.kyoto-phu.ac.jp

[**] This research was supported by the NIH grant RO1 GM075993 to S.B.H.K, and in part by grants from the American Diabetes Association (1-08-RA-218) and the NIH (DK040949 and DK074176) to M.A.W. Y.S. is grateful for a JSPS Postdoctoral Fellowship for Research Abroad. We thank Salih Özcubukcu (University of Chicago) for NMR characterization of new compounds.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201001151.

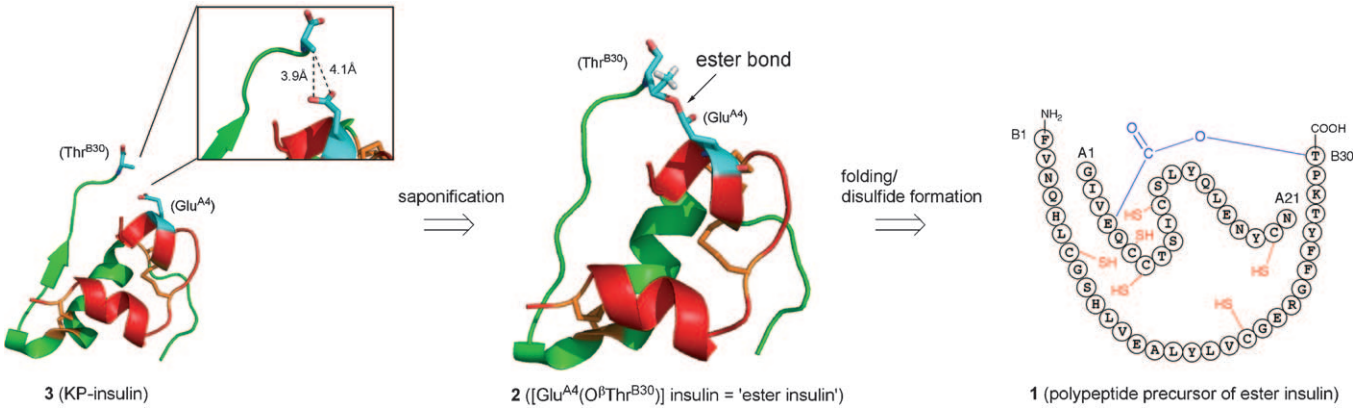

Figure 1. Retrosynthetic analysis for the preparation of human KP-insulin by the "ester insulin" strategy. KP-insulin is a fast acting form of human insulin in which the native Pro^{B28}-Lys^{B29} sequence is inverted to Lys^{B28}-Pro^{B29}.[23] KP-insulin coordinates are from Protein Databank entry 1 LPH.[24] Insulin A chain is shown in red, B chain in green, and Glu^{A4} and Thr^{B30} in cyan. The γ-CH_3 and β-OH of Thr^{B30} are not visible because of disorder in the crystal structure; however, the possible positions of the β-OH of Thr^{B30} can be accurately inferred because the position of the Thr^{B30} β-carbon is known. The desired ester linked molecule **1** was prepared by native chemical ligation[25] of [Phe^{B1}-Val^{B18}]-αthioester and the ester linked Cys^{B19}-[A1-Glu^{A4}(OβThr^{B30})-A21] peptide as described in the Supporting Information.

Glu^{A4} (Figure 1). This observation suggested to us that a covalently linked molecule, in which the A and B chains of insulin are directly connected through an ester bond (i.e. with no additional auxiliary moiety as a tether) between the β-hydroxyl group of Thr^{B30} and the γ-carboxyl group of Glu^{A4} (Figure 1), might serve as a surrogate proinsulin to promote efficient folding/disulfide bond formation in an insulin precursor molecule. We set out to make such a molecule and explore its folding properties and its chemical conversion into insulin.

A retrosynthetic analysis of a route to KP-insulin through the ester insulin precursor **1** is shown in Figure 1. The folding properties of purified **1** were examined under the following conditions: about 0.3 mg mL^{-1} of **1**, 20 mM of Tris, 8 mM of Cys, 1 mM of cystine, 1.5 M of GnHCl, pH 7.3 (Figure S3 in the Supporting Information). Folding was complete in 20 hours with the formation of folded ester insulin **2** as the predominant product (Figure 2a). The observed mass of **2** decreased by (5.8 ± 0.2) Da compared to that of the reduced polypeptide **1**, and is consistent with the formation of three disulfide bonds in **2**. We estimated that the yield of folded ester insulin **2** from **1** was approximately 70% as determined by HPLC analysis. The excellent folding profile of **2** demonstrates that the Thr^{B30}-Glu^{A4} ester linkage made the molecule as favorable for folding/disulfide formation as does the C peptide (35 amino acids long) in the proinsulin molecule.[26] The folding yield of **2** was also similar to that previously observed for oxime-linked miniproinsulin (ca. 60%)[22] where the N terminal of the A chain was connected to Lys^{B28} near the C terminal of B chain. Folded ester insulin **2** was isolated after purification by HPLC methods (Figure 2b).

To investigate the folded conformation of ester insulin, 2D ^{1}H NMR studies were conducted at pH 3.0 in 10 mM of deuterioacetic acid at 25 and 37 °C. Use of acidic pH conditions was chosen to retard the rate of hydrolysis of the ester during the 7 days of NMR data acquisition in H_2O and D_2O. Under these conditions the solution structure of KP-insulin[27] is essentially identical to its crystal structure (T-state protomer).[24] Because KP-insulin (like wild-type insulin)

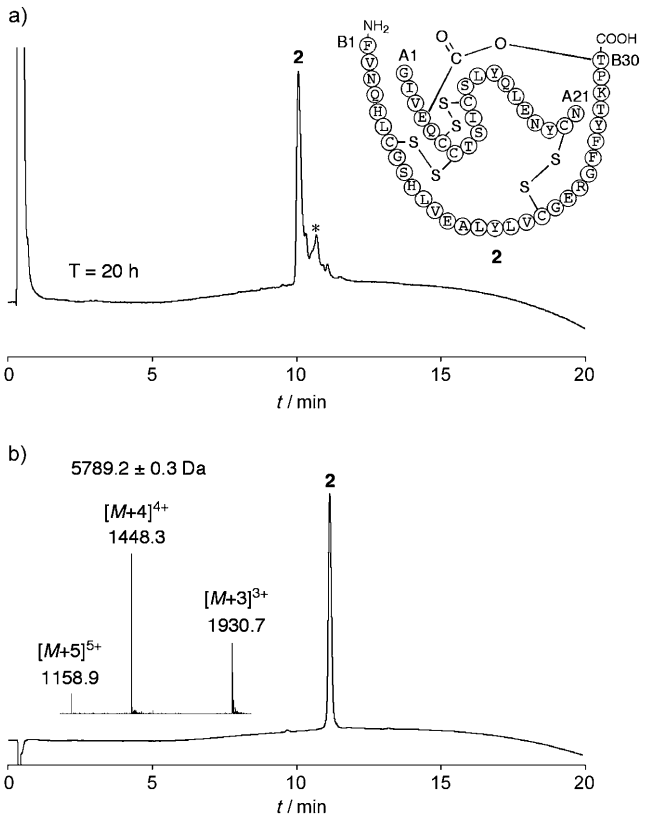

Figure 2. Folding/disulfide bond formation to give ester insulin. a) Folding of the ester insulin precursor **1** to form ester insulin **2** was monitored by LC analysis after 20 hours (UV absorbance profiles at 214 nm are shown). Essentially similar data were obtained at $T = 1$ hour. Folding reaction conditions were **1**: ca. 0.3 mg mL^{-1}, Tris: 20 mM, Cys: 8 mM, cystine: 1 mM, GnHCl: 1.5 M, pH 7.3, *: Cys adducts. b) Purified ester insulin **2**. (Inset: On-line ESI-MS spectra taken at the top of the main peak in the chromatogram.) The chromatographic separations were performed using a linear gradient (5–65%) of buffer B in buffer A over 15 minutes (buffer A = 0.1% TFA in water; buffer B = 0.08% TFA). Columns with different reverse-phase packings were used for (a) and (b). TFA = trifluoroacetic acid, Tris = tris(hydroxymethyl)aminomethane.

contains multiple aromatic side chains at key positions in the structure (4 Tyr, 3 Phe, and 2 His), the aromatic region of its NMR spectrum provides a fingerprint of the folded structure. Comparison of TOCSY ^{1}H NMR spectra of ester insulin and KP-insulin demonstrated retention of a nativelike pattern of aromatic chemical shifts (Figure 3a versus 3c). In particular,

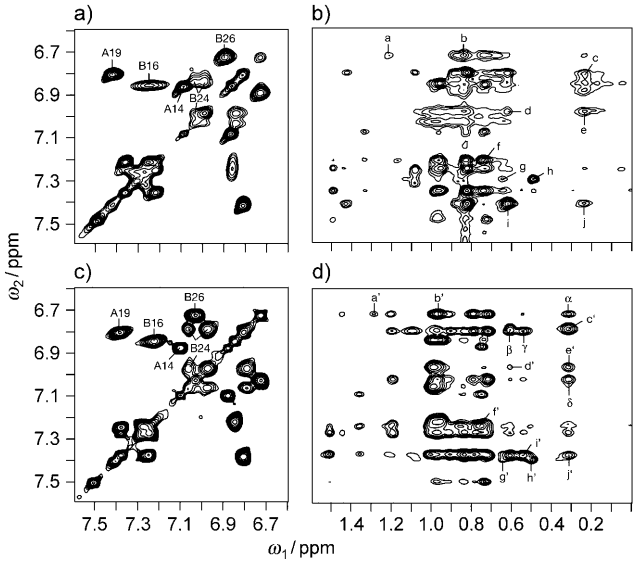

Figure 3. Tertiary structure of ester insulin and its relation to mature insulin by 2D ^{1}H NMR spectroscopy. Panels a) and b): Spectra of ester insulin. Aromatic TOCSY spin systems are shown in panel (a). Assignments of Tyr and Phe spin systems are as labeled; NOE interactions between aromatic protons (vertical axis) and aliphatic protons (horizontal axis) are shown in panel (b). Panels c) and d): Corresponding spectra of KP-insulin. Aromatic TOCSY spin systems are shown in panel (c) with related NOE interactions in panel (d). Assignment of NOE cross-peaks a–j (a′–j′) are reported in the Supporting Information. Samples were dissolved in 10 mM of deuterioacetic acid (pH 3.0). Spectra were acquired at 25°C and 700 MHz.

the large and corresponding secondary chemical shifts of Tyr^{A19} and Tyr^{B16} suggest that the α-helical moiety of the structure is retained, and the upfield shifts of Phe^{B24} and Tyr^{B26} suggest that ester insulin retains a nativelike αβ U-turn involving the C-terminal segment of the tethered B chain. The B24 spin system is anomalously broadened. Similar broadening is also observed involving non-aromatic resonances in the segments B27–B30 and A1–A5 (comparison of NOESY spectra in Figure 3b versus 3d). We ascribe such broadening to constrained millisecond motions in the ester insulin molecule, thus leading to incomplete averaging of chemical shifts.

Despite limitations of the NMR analysis near the ester moiety itself, complete resonance assignments were obtained elsewhere and enabled a detailed comparison between interresidue nuclear Overhauser effects (NOE) in ester insulin and the corresponding KP-insulin. Diagnostic long-range NOE interactions characteristic of the tertiary structure of insulin are retained in ester insulin. Of particular note are nativelike patterns of chemical shifts and NOE interactions involving the three cystine residues that gave evidence of native disulfide pairing. In addition, ester insulin exhibits selected interresidue NOE interactions in the A6-A12 segment that are in accord with insulin crystal structures but that are attenuated or not seen in the NMR spectrum of KP-insulin.

Interestingly, ester insulin itself had less than 1% activity (> 6 nM) in the insulin receptor binding assay as compared to KP-insulin (0.044 nM) in the same assay. It has previously been shown that an Ala replacement at either Thr^{B30} or Glu^{A4} did not cause any loss of the binding affinity,[5] thus suggesting that the inactive nature of ester insulin may arise from conformational restriction caused by the ester bond (see the NMR studies above) in accord with the low activities of chemically tethered insulin analogues[28] and single-chain analogues.[29,30]

Next, we investigated the chemical conversion of ester insulin into KP-insulin. Saponification of ester insulin **2** was performed under the following conditions: about 0.12 mg mL^{-1} of **2**, 25 mM of sodium hydroxide, 25% acetonitrile (in water), 4°C. As shown in Figure 4a, after 24 hours reaction time the desired KP-insulin molecule **3** was obtained in approximately 95% yield as determined by HPLC analysis. A few percent of individual A and B chains were observed as side products, probably because of disruption of disulfide bonds under these basic conditions. After saponification, the mass increased by (18.5 ± 0.5) Da compared to that of **2** which is consistent with the addition of the elements of water in the formation of **3**. We obtained pure KP-insulin **3** after final purification by HPLC (Figure 4b); the isolated yield of KP-insulin **3** from ester insulin **2** was 93%. The reverse-phase HPLC retention time of synthetic KP-insulin **3** was identical to that of an authentic sample of biosynthetic KP-insulin (extracted from Humalog). The synthetic KP-insulin **3** was also characterized by measurement of the relative binding affinity to the insulin receptor (Figure 4c). Within experimental uncertainty, the activity of synthetic KP- insulin was the same as that of an authentic sample of Humalog. This result further confirmed the formation of the correct disulfide bonds in the folded ester insulin molecule, and their retention in the isolated synthetic KP-insulin after saponification.

To evaluate the utility of ester insulin for the efficient chemical synthesis of insulin analogues, we prepared the ester-containing polypeptide precursor of [Gly23D-Ala]KP-insulin. The protein diastereomer [Gly23D-Ala]KP-insulin was designed to investigate the contribution of the Gly^{B23} residue to receptor recognition. As in ester insulin, the polypeptide precursor of [Gly23D-Ala]ester insulin was efficiently folded with concomitant formation of three disulfide bonds, and the resulting oxidized [Gly23D-Ala]ester insulin was saponified in a similar manner and gave the desired [Gly23D-Ala]KP-insulin (Figure S4 in the Supporting Information). The receptor binding affinity of the [Gly23D-Ala]KP-insulin was (0.021 ± 0.004) nM, a twofold higher activity than that of KP-insulin ((0.044 ± 0.007) nM). The [Gly23D-Ala]ester insulin had 100-fold lower activity ((2.1 ± 0.3) nM) compared with the [Gly23D-Ala]KP-insulin mature form (Figure S4C in the Supporting Information). The twofold higher activity of the [Gly23D-Ala] analogue suggests that the Gly^{B23} residue of the KP-insulin contributes to

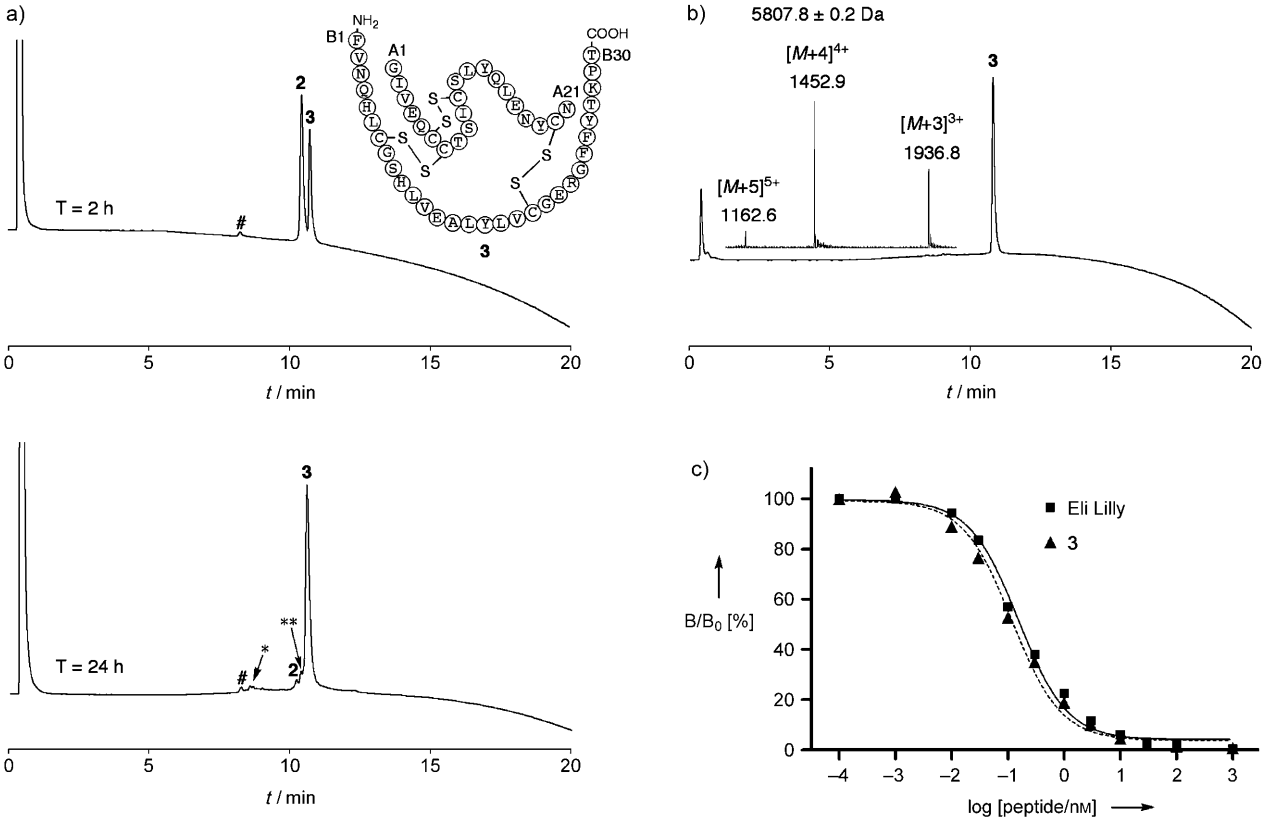

Figure 4. Conversion of ester insulin into native KP-insulin. a) Saponification of ester insulin **2** to give KP-insulin **3**. Reaction conditions were **2**: 0.12 mg mL^{-1}, NaOH: 25 mM, acetonitrile/H_2O (2.5:7.5), 4 °C. Reaction mixture at $T=2$ hours (upper panel) and $T=24$ hours (lower panel), #: unrelated column contaminant, *: derived from A chain, **: B chain; b) purified KP-insulin **3**. (Inset: On-line ESI-MS spectra taken at the top of the main peak in the chromatogram.) Chromatographic separations were performed as described in Figure 2 legend; c) binding affinities of synthetic KP-insulin **3** and authentic KP-insulin (purchased from Eli Lilly and Co.) to the insulin receptor.

receptor recognition by maintenance of the positive phi angle at B23, as was previously suggested by studies of DKP-insulin.[6]

In conclusion, we have designed and synthesized [Glu^{A4}-(OβThr^{B30})]insulin (ester insulin **2**) as a surrogate proinsulin with a "zero length" chemical tether moiety, and explored the potential utility of this novel molecule as an intermediate for the total chemical synthesis of human insulins. The reduced ester insulin precursor folded efficiently (ca. 70% yield based on HPLC analysis) under standard redox conditions with concomitant formation of the three native disulfide bonds. Thus, the Thr^{B30}-Glu^{A4} ester linkage made folding the precursor molecule as favorable as does the 35 residue C peptide in proinsulin. Finally, saponification of ester insulin gave the native folded insulin molecule in near-quantitative yield. Synthetic KP-insulin produced by the ester insulin route had full receptor-binding activity.

With suitable optimization of its preparation, ester insulin may prove to be the key to a simple and effective route to the total chemical synthesis of insulin.[5] The ester insulin precursor polypeptide could be made by any of several synthetic routes, including the hybrid solution/solid-phase method used for the cost-effective, large-scale manufacture of long peptides.[31] We believe that ester insulin will be a useful intermediate for the efficient generation of insulin analogues in the research laboratory and for cost-effective chemical manufacture of human insulins.

Received: February 25, 2010
Published online: May 27, 2010

Keywords: chemical ligation · chemical protein synthesis · esters · insulin · protein folding

[1] P. E. Oyer, S. Cho, J. D. Peterson, D. F. Steiner, *J. Biol. Chem.* **1971**, *246*, 1375–1386.
[2] W. Kemmler, J. D. Peterson, D. F. Steiner, *J. Biol. Chem.* **1971**, *246*, 6786–6791.
[3] B. H. Frank, R. E. Chance, *MMW Muench. Med. Wochenschr.* **1983**, *125(Suppl 1)*, 14–20.
[4] L. Thim, M. T. Hansen, K. Norris, I. Hoegh, E. Boel, J. Forstrom, G. Ammerer, N. P. Fiil, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 6766–6770.
[5] J. P. Mayer, F. Zhang, R. D. DiMarchi, *Biopolymers* **2007**, *88*, 687–713.
[6] S. H. Nakagawa, Q. X. Hua, S. Q. Hu, W. Jia, S. Wang, P. G. Katsoyannis, M. A. Weiss, *J. Biol. Chem.* **2006**, *281*, 22386–22396.
[7] M. Zhao, Z. L. Wan, L. Whittaker, B. Xu, N. B. Phillips, P. G. Katsoyannis, F. Ismail-Beigi, J. Whittaker, M. A. Weiss, *J. Biol. Chem.* **2009**, *284*, 32178–32187.

[8] J. Meienhofer, E. Schnabel, H. Bremer, O. Brinkhoff, R. Zabel, W. Sroka, H. Klostermeyer, D. Brandenburg, T. Okuda, H. Zahn, *Z. Naturforsch. B* **1963**, *18*, 1120 – 1121.
[9] P. G. Katsoyannis, K. Fukuda, A. Tometsko, K. Suzuki, M. Tilak, *J. Am. Chem. Soc.* **1964**, *86*, 930 – 932.
[10] Y. C. Du, R. Q. Jiang, C. L. Tsou, *Sci. Sin. (Engl. Ed.)* **1965**, *14*, 229 – 236.
[11] A. Marglin, R. B. Merrifield, *J. Am. Chem. Soc.* **1966**, *88*, 5051 – 5052.
[12] Q. Hua, Y. Chu, W. Jia, N. Phillips, R. Wang, P. Katsoyannis, M. Weiss, *J. Biol. Chem.* **2002**, *277*, 43443 – 43453.
[13] P. Sieber, B. Kamber, A. Hartmann, A. Jöhl, B. Riniker, W. Rittel, *Helv. Chim. Acta* **1977**, *60*, 27 – 37.
[14] K. Akaji, K. Fujino, T. Tatsumi, Y. Kiso, *J. Am. Chem. Soc.* **1993**, *115*, 11384 – 11392.
[15] D. Brandenburg, A. Wollmer, *Hoppe-Seyler's Z. Physiol. Chem.* **1973**, *354*, 613 – 627.
[16] A. Wollmer, D. Brandenburg, H.-P. Vogt, W. Schermutzki, *Hoppe-Seyler's Z. Physiol. Chem.* **1974**, *355*, 1471 – 1476.
[17] D. Brandenburg, W. Schermutzki, H. Zahn, *Hoppe-Seyler's Z. Physiol. Chem.* **1973**, *354*, 1521 – 1524.
[18] R. Geiger, R. Obermeier, *Biochem. Biophys. Res. Commun.* **1973**, *55*, 60 – 66.
[19] R. Obermeier, R. Geiger, *Hoppe-Seyler's Z. Physiol. Chem.* **1975**, *356*, 1631 – 1634.
[20] W-D. Busse, F. H. Carpenter, *Biochemistry* **1976**, *15*, 1649 – 1657.
[21] A. P. Tofteng, K. J. Jensen, L. Schäffer, T. Hoeg-Jensen, *ChemBioChem* **2008**, *9*, 2989 – 2996.
[22] Y. Sohma, S. B. H. Kent, *J. Am. Chem. Soc.* **2009**, *131*, 16313 – 16318.
[23] J. H. Anderson, Jr., R. L. Brunelle, P. Keohane, V. A. Koivisto, M. E. Trautmann, L. Vignati, R. DiMarchi, *Arch. Intern. Med.* **1997**, *157*, 1249 – 1255.
[24] PDB code 1LPH: E. Ciszak, J. M. Beals, B. H. Frank, J. C. Baker, N. D. Carter, G. D. Smith, *Structure* **1995**, *3*, 615 – 622.
[25] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776 – 779.
[26] J. Winter, H. Lilie, R. Rudolph, *Anal. Biochem.* **2002**, *310*, 148 – 155.
[27] Q. X. Hua, M. A. Weiss, unpublished results.
[28] S. H. Nakagawa, H. S. Tager, *J. Biol. Chem.* **1989**, *264*, 272 – 279.
[29] J. Markussen, K. H. Jørgensen, A. R. Sørensen, L. Thim, *Int. J. Peptide Protein Res.* **1985**, *26*, 70 – 77.
[30] Q. X. Hua, S. Q. Hu, W. Jia, Y. C. Chu, G. T. Burke, S. H. Wang, R. Y. Wang, P. G. Katsoyannis, M. A. Weiss, *J. Mol. Biol.* **1998**, *277*, 103 – 118.
[31] B. L. Bray, *Nat. Rev. Drug Discovery* **2003**, *2*, 587 – 593.

