Chemical Protein Synthesis

DOI: 10.1002/anie.200703521

Comparative Properties of Insulin-like Growth Factor 1 (IGF-1) and [Gly7D-Ala]IGF-1 Prepared by Total Chemical Synthesis**

*Youhei Sohma, Brad L. Pentelute, Jonathan Whittaker, Qin-xin Hua, Linda J. Whittaker, Michael A. Weiss, and Stephen B. H. Kent**

Dedicated to Professor Panayotis G. Katsoyannis

Human insulin-like growth factor 1 (IGF-1) is a 70-residue single-chain protein that mediates somatic growth.[1] Inborn mutations in the *igf-1* gene are associated with a childhood syndrome of impaired growth, mental retardation, and deafness.[2] Although such mutations are rare, broad interest in IGF-1 has been stimulated by its nonclassical roles in neoplasia and the regulation of metabolism. Ectopic expression of IGF-1 and autocrine activation of the IGF-1 signaling pathway contribute to several human cancers, including those of the breast, prostate, lung, and thyroid.[3] In addition, because of cross-talk between the insulin and IGF-1 signaling systems in metabolic target tissues, IGF-1 is under investigation as a possible treatment for diabetes mellitus.[4] Understanding the diverse functions of IGF-1 and its relation to insulin signaling is thus of great biomedical significance.

IGF-1 consists of a single polypeptide chain and (from the N to the C terminus) contains B, C, A, and D domains, which are defined by homology within the insulin-related superfamily (Scheme 1).[5] The B and A domains are similar to the two separate B and A chains of insulin (45–52% identity), whereas the C domain is unrelated to the connecting peptide of the single-chain insulin precursor (proinsulin). Like other members of the insulin-related superfamily, IGF-1 contains three disulfide bridges, which correspond to IGF-1 residues Cys6–Cys48, Cys18–Cys61, and Cys47–Cys52. Crystal structures of human IGF-1 resemble the classical T state of human insulin,[7] with novel features associated with the IGF-1 C and D domains.[8,9] Despite the similarities of their respective A and B regions, insulin and IGF-1 appear to differ in how they bind to their cognate receptors (the insulin receptor and IGF Type I receptor).[10,11] A promising approach to investigate functional similarities and differences between IGF-1 and insulin is the comparative study of analogous substitutions in their putative receptor-binding surfaces. The overall richness of the literature describing the synthesis of chemical analogues of insulin, which extends back 35 years,[12] has resulted in the existence of a trove of data that can act as a foundation for comparative studies of synthetic IGF-1 analogues. Particularly intriguing are the marked effects of chiral perturbations in insulin at specific sites in the main chain[13–16] or side chains.[17,18] To date, however, relatively few chemical analogues of IGF-1 have been prepared because of the inefficiency of stepwise solid-phase methods for the synthesis of a 70-residue polypeptide chain.[19–21] A modular synthetic approach would offer decisive advantages for the facile preparation of analogues by chemical means.

We set out to establish an efficient modular synthesis of IGF-1 and chemical analogues by using modern synthetic methods. Our synthetic strategy for the preparation of IGF-1 is based on the native chemical ligation of unprotected peptide segments in aqueous solution,[22–28] as shown in Scheme 1. Several challenges had to be overcome in this apparently straightforward synthesis: the ligation at Val17–Cys18 is sterically hindered and would not have been practical by standard native chemical ligation,[26] and solubility issues were encountered with the 1–17 peptide, which were overcome by using some of the findings from our recent work.[28] Furthermore, it proved important to use solid-phase extraction to purify the first ligation product for use in the next ligation reaction, rather than make use of the "one-pot" ligation strategy.[25] Together, these innovations enabled the efficient synthesis of this challenging target polypeptide chain.

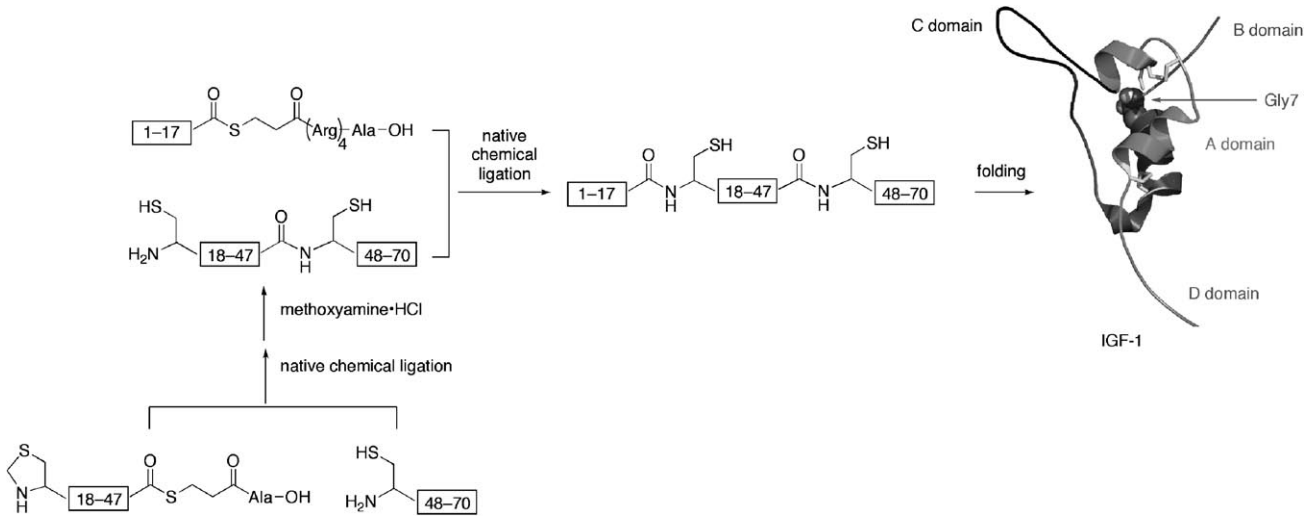
Ligation of IGF-1[Thz^{18}-47]$^{\alpha}$thioester and IGF-1[Cys^{48}-70] was complete in 14 h (see Figure S2(A, B) in the Supporting Information). The product IGF-1[Thz^{18}-70] was quantitatively converted into IGF-1[Cys^{18}-70] in 3 h by the addition of 0.2 M methoxylamine·HCl to the crude ligation

[*] Dr. Y. Sohma, B. L. Pentelute, Prof. 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. Y. Sohma
Department of Pharmaceutical Manufacturing Chemistry
Kyoto Pharmaceutical University
Yamashina-ku, Kyoto 607-8412 (Japan)
Dr. Q.-X. Hua, Prof. M. A. Weiss
Department of Biochemistry
Case Western Reserve University
Cleveland Ohio 44106 (USA)
Prof. J. Whittaker, L. J. Whittaker
Department of Nutrition
Case Western Reserve University
Cleveland Ohio 44106 (USA)

[**] This research was supported by the Office of Science (BER), the U.S. Department of Energy (grant no.DE-FG02-07ER64501 to S.B.H.K.), and by the NIH (grant no. DK065890 to J.W. and grant no. DK40949 to M.A.W.) Y.S. is grateful for a Research Fellowship from the Japan Society for the Promotion of Science for Young Scientists. We thank Wenhua Jia for technical assistance with spectroscopic studies.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

InterScience
© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 1. Strategy for the synthesis of IGF-1. The IGF-1 coordinates are from Protein Databank entry 1PMX.[6]

mixture at about pH 4 (see Figure S2(C) in the Supporting Information).[25] After solid-phase extraction and lyophilization, ligation of IGF-1[1-17]$^{\alpha}$thioester and crude IGF-1[Cys^{18}-70] was performed. This ligation was anticipated to be a slow reaction, because of the β-branched nature of the C-terminal Val residue of the IGF-1[1-17]$^{\alpha}$thioester,[24] so a greater concentration of the (4-carboxylmethyl)thiophenol[26] catalyst (MPAA, 200 mM) was used. During an exploratory synthesis we found that the peptide IGF-1[1-17]$^{\alpha}$COSCH$_2$CH$_2$CO-Ala-OH is hydrophobic and thus difficult to handle by reversed-phase HPLC. We therefore prepared IGF-1[1-17]$^{\alpha}$thioester-(Arg)$_4$-Ala-OH, which contains a solublizing "Arg tag" in the C-terminal thioester leaving group.[27,28] The Arg tag was observed to increase solubility, thereby conferring favorable handling properties and facilitating the use of this peptide in ligation. The reaction of IGF-1[1-17]$^{\alpha}$thioester-(Arg)$_4$-Ala-OH and IGF-1[Cys^{18}-70] (2 mM each) was approximately 70–80% complete in 18 h and yielded the final full-length reduced product (see Figure S3-(A–C) in the Supporting Information). Purification by HPLC provided pure, reduced full-length IGF-1(SH)$_6$ (see Figure S3(D) in the Supporting Information).

We aimed to optimize the folding reaction of IGF-1 because it is known that the folding of IGF-1 by air oxidation also results in the formation of non-native isomers;[29,30] under these conditions, the major non-native isomer (designated "IGF-swap"), which contains the disulfide pairings Cys6–Cys47, Cys18–Cys61, and Cys48–Cys52, is obtained at a similar rate and in similar yield to native IGF-1.[29] We explored different folding conditions, and found that a greater proportion of native IGF-1 was obtained under the following conditions: about 0.5 mg mL^{-1} IGF-1(SH)$_6$, 20 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl, pH 7.7), 8 mM cysteine, 1 mM cystine

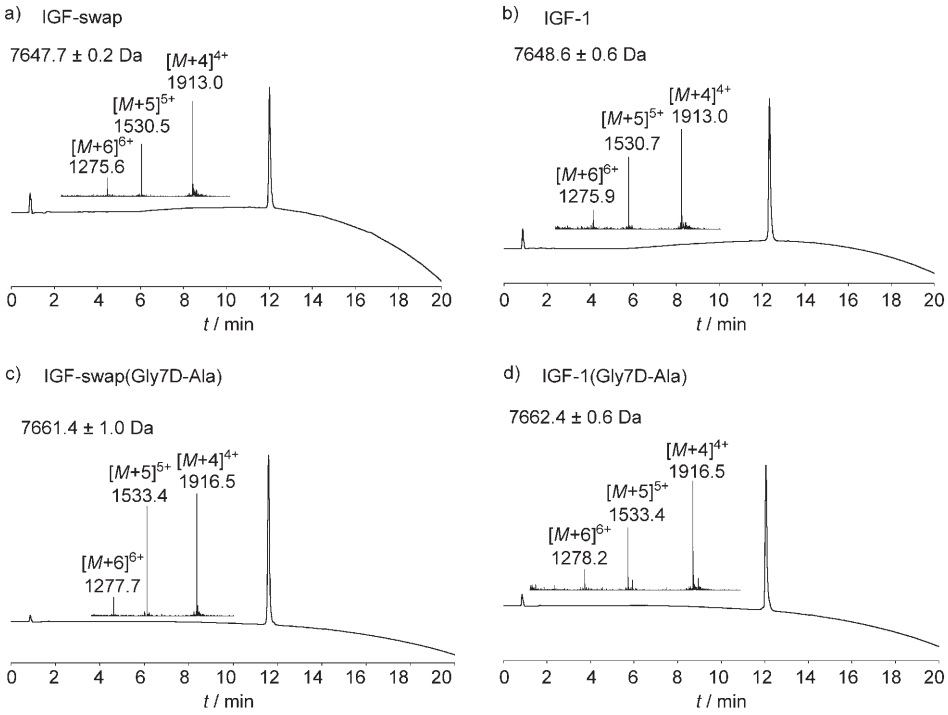


Figure 1. Folded and purified synthetic proteins. a) IGF-swap, b) IGF-1, c) [Gly7D-Ala]IGF-swap, and d) [Gly7D-Ala]IGF-1 (UV profiles at 214 nm). The chromatographic separations were performed using a linear gradient (5–65%) of buffer B in buffer A over 15 min (buffer A: 0.1% trifluoroacetic acid (TFA) in water; buffer B: 0.08% TFA in acetonitrile). Inset: On-line ESMS spectra taken at the top of the main peak in each chromatogram. The obtained mass is given for each product.

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

hydrochloride, and 0.5 M guanidine hydrochloride (GnHCl; see Figure S4 in the Supporting Information, IGF-swap:IGF-1 = 1:3). The amount of other minor isomers was also reduced, relative to the amounts reported in the literature.[30] Folding, as monitored by LCMS, was complete in one hour. Following oxidation, the mass decreased by 6.6 ± 0.7 Da, indicating the formation of three disulfide bonds. The principal components of the folding reaction were purified by preparative reversed-phase HPLC to give pure IGF-1 and IGF-swap (Figures 1a,b). The overall yield of IGF-1 was 6.7%.

Synthetic IGF-1 and IGF-swap were characterized by CD and ^{1}H NMR spectroscopy. Previous studies have established that native IGF-1 and IGF-swap exhibit different mean α-helix contents because of the segmental unfolding of helix 2 (residues 41–48 in IGF-1; canonical residues A1–A8) in the non-native isomer (see Figure S5 in the Supporting Information).[31,32] Structural differences between IGF-1 and IGF-swap are also associated with significant differences in the ^{1}H NMR chemical shifts. The ^{1}H NMR spectra of synthetic IGF-1 (or IGF-swap) obtained in the current study (Figure 2a) correspond to control spectra of biosynthetic IGF-1 (or IGF-swap) kindly provided by Eli Lilly and Co (see Figure S6 in the Supporting Information). These NMR data show that our synthetic IGF-1 and IGF-swap each had a defined tertiary structure, and also confirm the previously determined disulfide pairing schemes for the two protein isomers.[29] The synthetic proteins were further characterized by measurement of the relative binding affinities to the Type I IGF receptor (Table 1). The activity of synthetic IGF-1 is the same as that of biosynthetic IGF-1, and for both recombinant and synthetic proteins the activity of the disulfide "swap" isomer is reduced by about 30-fold.[29] The native chemical ligation route thus provides an efficient method for the synthesis and purification of a protein whose structure and activity are indistinguishable from biosynthetic IGF-1.

To illustrate the utility of this synthetic protocol in studies of nonstandard analogues, we investigated the substitution of Gly7 by D-Ala. This analogue was designed on the basis of corresponding studies on human insulin. Gly7 (canonical position B8 in insulin) participates in a Type 1′ β turn and exhibits absolute conservation among members of the vertebrate insulin-related superfamily. Although not part of the classical receptor-binding surface, the homologous glycine in insulin (Gly^{B8}) participates in a large-scale allosteric reorganization of zinc–insulin hexamers,[36,37] designated the TR transition.[38] In this transition, the dihedral angles of Gly^{B8} move from the right side of the Ramachandran plot (with a positive ϕ angle ordinarily forbidden to L-amino acids) to the left side (with a negative ϕ angle characteristic of L-amino acids). Substitution of Gly^{B8} by D-Ala in insulin markedly augments the thermodynamic stability of insulin (presumably by stabilizing the B8-related β turn) but impedes binding to the insulin receptor by 1000-fold.[15] [Gly7D-Ala]IGF-1 was obtained by using the same synthetic procedure as for native IGF-1 (Figure 1d). The overall synthetic yield of [Gly7D-Ala]IGF-1 was 6.7%. The observed folding rate of [Gly7D-Ala]IGF-1 was much slower than native

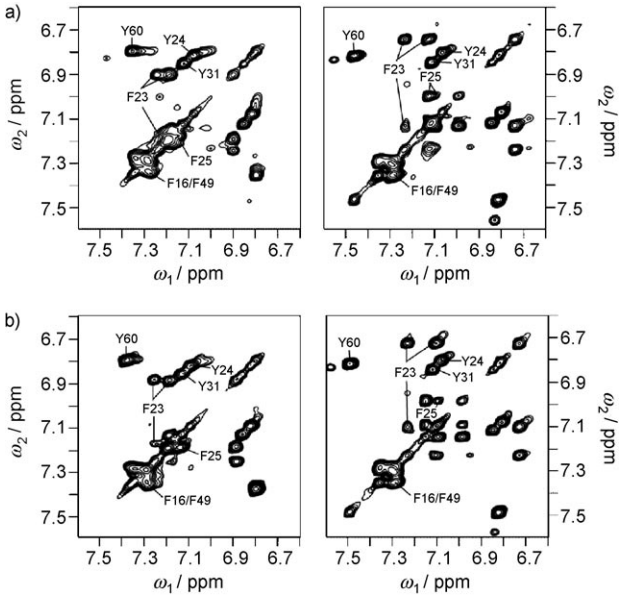


Figure 2. 2D TOCSY NMR aromatic proton spectra of the synthetic proteins. a) Folding of the polypeptide chain of wild-type IGF-1 leads to two disulfide isomers with distinct structures: native pairing scheme (left) and "swapped" pairing scheme (right).[29] Each isomer exhibits major and minor conformers; essentially identical spectra are obtained from authentic biosynthetic proteins (see Figure S6 in the Supporting Information). Resonance assignments of the major conformer are as indicated. Although signficant differences are observed in the chemical shifts of corresponding aromatic resonances, similar overall trends are observed in relative resonance positions (upfield or downfield of random-coil values), thus reflecting partial structural similarities between the native and swapped proteins. b) Corresponding 2D TOCSY spectra of D-Ala7 variants: native disulfide pairing (left) and "swapped" pairing (right). The similarity between the Gly7 and D-Ala7 spectra for each isomer provides evidence that the nonstandard D-Ala7 substitution is well accommodated. Spectra were obtained at 40°C and 700 MHz in 10% deuterioacetic acid and 90% D_2O at pD 2.0 (direct meter reading) as described.[29,33–35]

Table 1: Binding affinities of IGF-1, IGF-swap, [Gly7D-Ala]IGF-1, and [Gly7D-Ala]IGF-swap to the Type I IGF receptor.

	Dissociation constant [nM][a]	Relative affinity [%]
IGF-1	0.046 ± 0.0006	100
IGF-swap	6	0.8
[Gly7D-Ala]IGF-1	16.7 ± 0.3	0.3
[Gly7D-Ala]IGF-swap	N.A.[b]	< 0.05

[a] Mean S.E.M. values, $n = 3$ (S.E.M. = standard error of the mean). [b] Too high to measure; even at micromolar concentrations, less than 20% of the radiolabeled IGF-1 was displaced.

IGF-1. It is notable that the amount of [Gly7D-Ala]IGF-swap was greater in the folding of [Gly7D-Ala]IGF-1 than in native IGF-1 (see Figure S7 in the Supporting Information). The ^{1}H NMR spectra of synthetic [Gly7D-Ala]IGF-1 and [Gly7D-Ala]IGF-swap are shown in Figure 2b.

Characterization of the present D-Ala7 analogue of IGF-1 indicates augmented stability ($\Delta\Delta G_{unfolding}$ 0.6 ± 0.1 kcal mol^{-1} relative to native IGF-1) as probed by CD-monitored

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

guanidine denaturation experiments (see the Supporting Information for data and experimental details). Substitution of Gly7 by D-Ala augments the stability of the non-native disulfide "swap" isomer more substantially ($\Delta\Delta G_{unfolding} > 1.2$ kcal mol^{-1}) than it does the stability of the native isomer, which is consistent with the greater relative yield of the non-native isomer on oxidative folding of the D-Ala7 polypeptide. Although the structural basis for this selective stabilization is unclear, these findings demonstrate that single amino acid substitutions in IGF-1 can modulate the fidelity of disulfide pairing by favoring a preferred pairing scheme (even if non-native) rather than by destabilizing competing folds. Although the CD and ^{1}H NMR spectra of the D-Ala7 analogue are similar to those of native IGF-1 (see Figure S5 in the Supporting Information and Figure 2), its binding to the IGF receptor is reduced by 300-fold (Table 1). The similar properties of homologous D-Ala-substituted analogues of insulin and IGF-1 (namely, higher thermodynamic stability and lower receptor-binding affinity as compared to the native form) provide evidence that the N-terminal segments of these proteins play corresponding roles in their respective receptor-binding mechanisms. It is possible that the IGF-1 monomer, on binding to its receptor, undergoes a similar change in conformation near Gly7 as proposed for the insulin monomer. It would be of future interest to characterize sites elsewhere in the insulin and IGF-1 molecules at which corresponding substitutions yield divergent properties, as those sites would identify inequivalent receptor-binding mechanisms. Together, such comparative studies promise to provide a map of the functional similarities and differences between insulin and IGF-1. Such data could inform design of novel ligands with novel biological properties.

In conclusion, we have described herein an efficient synthetic scheme for the preparation of IGF-1 and the protein diastereomer analogue [Gly7D-Ala]IGF-1. Based on native chemical ligation and an optimized folding protocol, this scheme facilitates the introduction of nonstandard amino acids whose biosynthetic incorporation is difficult or not feasible. In addition to the preparation of IGF-1 analogues containing D-amino acids, this synthetic scheme will enable chemical modification of the polypeptide backbone and the introduction of nonstandard side chains, including photoactivatable derivatives that are useful in mapping the interface between IGF-1 and its receptor. Facile synthesis of chemical analogues of IGF-1 thus promises to lead to a new generation of structure–activity studies of potential interest in the treatment of cancer and diabetes mellitus.

Received: August 2, 2007
Revised: October 26, 2007
Published online: December 27, 2007

Keywords: growth factor · insulin · native chemical ligation · protein folding · protein synthesis

[1] E. Rinderknecht, R. E. Humbel, *J. Biol. Chem.* **1978**, *253*, 2769–2776.
[2] A. Denley, C. C. Wang, K. A. McNeil, M. J. Walenkamp, H. van Duyvenvoorde, J. M. Wit, J. C. Wallace, R. S. Norton, M. Karperien, B. E. Forbes, *Mol. Endocrinol.* **2005**, *19*, 711–721.
[3] M. Pollak, *Recent Results Cancer Res.* **2007**, *174*, 49–53.
[4] P. Pennisi, O. Gavrilova, J. Setser-Portas, W. Jou, S. Santopietro, D. Clemmons, S. Yakar, D. LeRoith, *Endocrinology* **2006**, *147*, 2619–2630.
[5] T. L. Blundell, R. E. Humbel, *Nature* **1980**, *287*, 781–787.
[6] M. L. Schaffer, K. Deshayes, G. Nakamura, S. Sidhu, N. J. Skelton, *Biochemistry* **2003**, *42*, 9324–9334.
[7] T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. C. Hodgkin, D. A. Mercola, M. Vijayan, *Nature* **1971**, *231*, 506–511.
[8] A. M. Brzozowski, E. J. Dodson, G. G. Dodson, G. N. Murshudov, C. Verma, J. P. Turkenburg, F. M. de Bree, Z. Dauter, *Biochemistry* **2002**, *41*, 9389–9397.
[9] F. F. Vajdos, M. Ultsch, M. L. Schaffer, K. D. Deshayes, J. Liu, N. J. Skelton, A. M. de Vos, *Biochemistry* **2001**, *40*, 11022–11029.
[10] P. De Meyts, J. Whittaker, *Nat. Rev. Drug Discovery* **2002**, *1*, 769–783.
[11] J. Whittaker, A. V. Groth, D. C. Mynarcik, L. Pluzek, V. L. Gadsboll, L. J. Whittaker, *J. Biol. Chem.* **2001**, *276*, 43980–43986.
[12] A. Cosmatos, P. G. Katsoyannis, *J. Biol. Chem.* **1973**, *248*, 7304–7309.
[13] M. Kobayashi, S. Ohgaku, M. Iwasaki, H. Maegawa, Y. Shigeta, K. Inouye, *Biochem. Biophys. Res. Commun.* **1982**, *107*, 329–336.
[14] R. G. Mirmira, S. H. Nakagawa, H. S. Tager, *J. Biol. Chem.* **1991**, *266*, 1428–1436.
[15] S. H. Nakagawa, M. Zhao, Q. X. Hua, S. Q. Hu, Z. L. Wan, W. Jia, M. A. Weiss, *Biochemistry* **2005**, *44*, 4984–4999.
[16] Q. X. Hua, S. Nakagawa, S. Q. Hu, W. Jia, S. Wang, M. A. Weiss, *J. Biol. Chem.* **2006**, *281*, 24900–24909.
[17] S. H. Nakagawa, H. S. Tager, *Biochemistry* **1992**, *31*, 3204–3214.
[18] B. Xu, Q. X. Hua, S. H. Nakagawa, W. Jia, Y. C. Chu, P. G. Katsoyannis, M. A. Weiss, *J. Mol. Biol.* **2002**, *316*, 435–441.
[19] C. H. Li, D. Yamashiro, D. Gospodarowicz, S. L. Kaplan, G. Van Vliet, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 2216.
[20] M. Iwai, M. Kobayashi, K. Tamura, Y. Ishii, H. Yamada, M. Niwa, *J. Biochem.* **1989**, *106*, 949–951.
[21] C. J. Bagley, K. M. Otteson, B. L. May, S. N. McCurdy, L. Pierce, F. J. Ballard, J. C. Wallace, *Int. J. Pept. Protein Res.* **1990**, *36*, 356–361.
[22] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
[23] P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
[24] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
[25] D. Bang, S. B. H. Kent, *Angew. Chem.* **2004**, *116*, 2588–2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538.
[26] E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
[27] T. Sato, Y. Saito, S. Aimoto, *J. Pept. Sci.* **2005**, *11*, 410–416.
[28] E. C. B. Johnson, S. B. H. Kent, *Tetrahedron Lett.* **2007**, *48*, 1795–1799.
[29] J. A. Miller, L. O. Narhi, Q.-X. Hua, R. Rosenfeld, T. Arakawa, M. Rohde, S. Prestrelski, S. Lauren, K. S. Stoney, L. Tsai, M. A. Weiss, *Biochemistry* **1993**, *32*, 5203–5213.
[30] M. Iwai, H. Yokoyama, H. Yamada, M. Niwa, M. Kobayashi, *Chem. Pharm. Bull.* **2000**, *48*, 1304–1309.
[31] Q. X. Hua, S. N. Gozani, R. E. Chance, J. A. Hoffmann, B. H. Frank, M. A. Weiss, *Nat. Struct. Biol.* **1995**, *2*, 129–138.
[32] Q. X. Hua, W. Jia, B. H. Frank, N. F. Phillips, M. A. Weiss, *Biochemistry* **2002**, *41*, 14700–14715.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[33] A. Sato, S. Nishimura, T. Ohkubo, Y. Kyogoku, S. Koyama, M. Kobayashi, T. Yasuda, Y. Kobayashi, *J. Biochem.* **1992**, *111*, 529 – 536.
[34] Q. X. Hua, L. Narhi, W. Jia, T. Arakawa, R. Rosenfeld, N. Hawkins, J. A. Miller, M. A. Weiss, *J. Mol. Biol.* **1996**, *259*, 297 – 313.
[35] Q. X. Hua, J. P. Mayer, W. Jia, J. Zhang, M. A. Weiss, *J. Biol. Chem.* **2006**, *281*, 28131 – 28142.
[36] G. Bentley, E. Dodson, G. Dodson, D. Hodgkin, D. Mercola, *Nature* **1976**, *261*, 166 – 168.
[37] U. Derewenda, Z. Derewenda, E. J. Dodson, G. G. Dodson, C. D. Reynolds, G. D. Smith, C. Sparks, D. Swenson, *Nature* **1989**, *338*, 594 – 596.
[38] M. L. Brader, M. F. Dunn, *Trends Biochem. Sci.* **1991**, *16*, 341 – 345.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim