*Drug Design*

# Site-Specific Introduction of Sialic Acid into Insulin**

*Masaaki Sato, Reiko Sadamoto, Kenichi Niikura, Kenji Monde, Hirosato Kondo, and Shin-Ichiro Nishimura**

Insulin, a peptide hormone, is virtually the only effective drug for the treatment of diabetes.[1] The quantity of insulin required is increasing along with the number of patients, and the demand for insulin can be as much as 0.5–1.0 g per year per patient.[2] However, native insulin is broken down in the blood stream within a few hours of administration, mainly as a result of decomposition in the liver.[3] Several insulin derivatives have been investigated as possible long-acting insulins for decreasing the required frequency of administration.[3,4] Most of these insulin derivatives are less soluble than native insulin and are gradually released from the injection site into the blood stream.[4] The disadvantages of these long-acting insulins are in the complex administration procedure and the difficulty in controlling the circulating glucose levels.[5] It is expected that water-soluble long-acting insulin derivatives would overcome these problems.[5,6]

We focused on the increased solubility[7] and decreased degradation offered by the addition of a sugar chain, especially one containing a sialic acid moiety, for the development of new long-acting insulin derivatives. In the case of erythropoietin,[8] another widely used protein-based drug, it is known that the introduction of sugar chains containing sialic acid moieties through genetically controlled protein expression, for example, in mammalian cells, enhances the duration of the drug's effectiveness because the proteins with sialic acid on the surface are not easily decomposed in the liver. However, optimization of the glycosylation is difficult because of the inevitable variation, arising from the expression, in the number of sugar moieties per protein and in the site of introduction. Several chemical methods have been reported for making a uniformly glycosylated protein;[9] however, these chemical methods usually require relatively harsh conditions that may abolish protein functions.

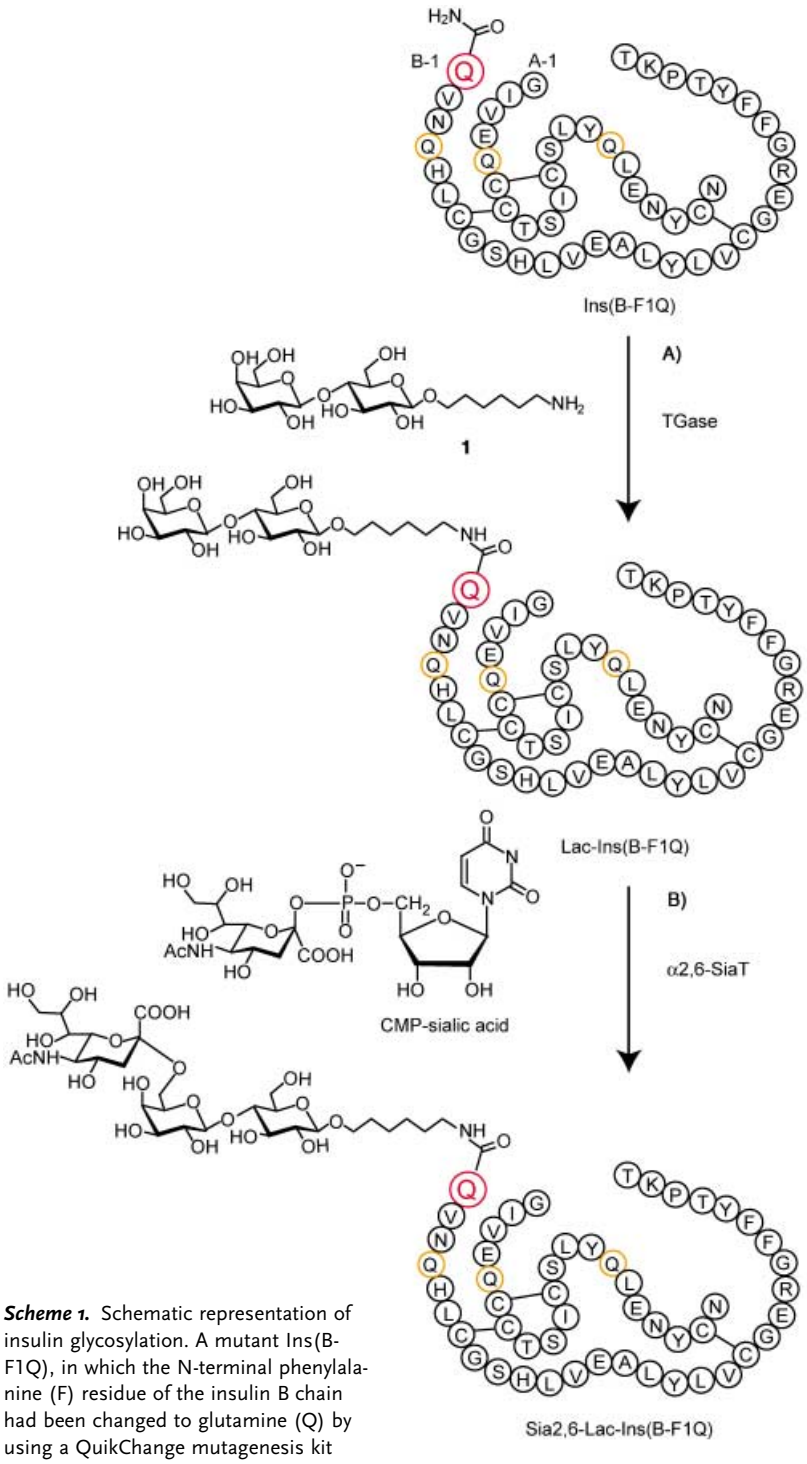
Herein, we describe a method for introducing a defined sugar moiety onto mutant insulins by enzymatic reactions (Scheme 1). After insertion of the lactose derivative through a transglutaminase reaction,[10] sialic acid can be introduced through a recombinant α2,6-sialyltransferase reaction.[11,12] These enzymatic reactions can proceed under physiological conditions. The insulin with the sialyl oligosaccharide is expected to demonstrate prolonged activity as a result of its resistance to decomposition in the liver.

Introduction of a sugar moiety to the glutamine residues by using TGase was reported to be a very effective method for obtaining sugar-modified peptides and proteins.[13] Recently, it was also demonstrated that an efficient synthesis of potent blockers of the influenza virus based on cyclic glycopeptides can be performed by using a TGase reaction.[12] Initially, we attempted to use the three glutamine residues in wild-type human insulin (Ins(WT)) as the reaction anchor to conjugate lactose derivative **1** through a TGase reaction. However, the sugar moiety was not effectively attached to the insulin (Figure 1 A), a result suggesting that all three glutamine residues are inaccessible to TGase. Therefore, we thought that some reactive glutamine residues should be added to the native insulin by point mutations. Human insulin consists of two separate chains, the A chain (21 amino acids) and the B chain (30 amino acids), joined together by a characteristic pattern of disulfide bridges[14] (Scheme 1). While the amino acid sequence of insulin is highly conserved in most animal species, species variations do exist, mainly in the amino acid sequence of the N-terminal region of the B chain. It has also been shown that modification of the N terminus of either the B or the A chain with a poly(ethylene glycol) chain or a simple monosaccharide residue has almost no negative effect on insulin bioactivity.[9b,15] Therefore, we created two mutagenetic insulins, Ins(B-F1Q) which has a glutamine residue instead of phenylalanine at the N terminus of the B chain and

[*] M. Sato, Dr. K. Niikura, Dr. K. Monde, Prof. S.-I. Nishimura
Division of Biological Sciences, Graduate School of Science
Frontier Research Center for Post-Genomic Science and Technology
Hokkaido University, Kita 21 Nishi 11, Sapporo 001-0021 (Japan)
Fax: (+81) 11-706-9042
E-mail: shin@glyco.sci.hokudai.ac.jp
Prof. S.-I. Nishimura
Japan Bioindustry Association
Sapporo 060-0810 (Japan)
Prof. S.-I. Nishimura
National Insutitute of Advanced Industrial Science and Technology (AIST)
Sapporo 062-8517 (Japan)
Dr. R. Sadamoto
Shionogi Laboratory of Biomolecular Chemistry
Division of Biological Science, Graduate School of Science
Frontier Research Center for Post-Genomic Science and Technology
Hokkaido University, Kita 21 Nishi 11, Sapporo 001-0021 (Japan)
Dr. H. Kondo
Discovery Research Laboratories
Shionogi & Co., Ltd., Osaka 553-0002 (Japan)

[**] This work was supported partly by a grant for "Research and Development on Glycocluster Controlling Biomolecules" from the New Energy and Industrial Technology Development Organization (NEDO). M.S. would like to thank Mr. A. Toda, Dr. K. Yamada, Dr. K. Iwata, Mr. T. Ohta, and Dr. H. Nakagawa for helpful discussions.

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/ange.200353058 




***Scheme 1.*** Schematic representation of insulin glycosylation. A mutant Ins(B-F1Q), in which the N-terminal phenylalanine (F) residue of the insulin B chain had been changed to glutamine (Q) by using a QuikChange mutagenesis kit (Stratagene), was overexpressed and purified from *Escherichia coli*. A) Chemoenzymatic glycosylation of Ins(B-F1Q) was carried out in 20 mM Tris buffer (pH 7.5) containing Ins(B-F1Q) (104 nmol), compound **1** (0.9 mmol), 5 mM $CaCl_2$, and TGase (4 U; Oriental Yeast Co.). The reaction mixture was incubated for 30 min at 37 °C. B) The sugar-elongation reaction of Lac-Ins(B-F1Q) was carried out in 50 mM sodium cacodylate buffer (pH 7.5, 1 mL) containing Lac-Ins(B-F1Q) (0.5 mg, 80.5 nmol), 0.6 mM CMP-sialic acid, 0.2% BSA, 1.6 mM $MnCl_2$, 0.2% Triton CF54, calf-intestine alkaline phosphatase (20 U), and recombinant α2,6-SiaT (150 mU, TOYOBO Co.). The reaction mixture was incubated for 2 h at 37 °C. TGase = transglutaminase, α2,6-SiaT = α2,6-sialyltransferase, Lac-Ins(B-F1Q) = lactose-displaying Ins(B-F1Q), Sia2,6-Lac-Ins(B-F1Q) = α2,6-sialylated Lac-Ins(B-F1Q), CMP = cytidine monophosphate, Tris = tris(hydroxymethyl)aminomethane, U = international unit, BSA = bovine serum albumin.

Ins(A-G1Q) which has a glutamine residue instead of glycine at the N terminus of the A chain.

First, we tested whether the mutagenetic insulins Ins(B-F1Q) and Ins(A-G1Q) retained their bioactivity by using streptozotocin-treated diabetic mice (STZ-treated mice). STZ-treated mice were separated into three groups at random, treated with a single subcutaneous dose of Ins(WT), Ins(B-F1Q), or Ins(A-G1Q), and then monitored for changes in blood-glucose level every hour for five hours (Figure 2). The dosage of all insulin preparations was adjusted to 1 U. The Ins(B-F1Q) action profile was very similar to that of wild-type insulin, a result showing that the biological potency of Ins(B-F1Q) was equal to that of Ins(WT). However, the glucose-lowering action of Ins(A-G1Q) was much weaker than that of the wild-type insulin. Therefore, we chose the active Ins(B-F1Q) for further modification with sugars.

The enzymatic reaction of Ins(B-F1Q) with a large excess of lactose derivative **1** was carried out in a Tris buffer (pH 7.5; Scheme 1). Mass spectrometric stidies showed that the modified insulin mostly had a single lactose moiety (Figure 1B). Purification of glycosylated insulin was performed by HPLC to afford the desired insulin Lac-Ins(B-F1Q) in around 50% yield. To determine the glycosylation site, Lac-Ins(B-F1Q) was treated with 2-thioethanol to cleave the disulfide bonds and divide it into two chains (the A and B chains); the molecular weights of the two peptide chains were then determined (Figure 1C). Mass spectrometry confirmed the glycosylation of the B chain occurred as expected. After purification of the B chain by HPLC, amino acid sequence analysis by using aminopeptidase M degradation was performed to determine the glycosylation site (Figure 1D). Treatment of the glycosylated B chain with aminopeptidase M, an exo-type protease that hydrolyzes peptides one by one from the N terminus, showed that not only the terminal glutamine residue (Q1) of the B chain but also the fourth glutamine residue (Q4) from the N terminus had been glycosylated by TGase. From the intensity of the fragment peaks corresponding to the peptides lacking only the N-terminal residues it seems that the glycosylation ratio between Q1 and Q4 is almost identical. Since the Q4 residue of the native insulin was not glycosylated by TGase, the phenylalanine → glutamine substitution at the N terminus is

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

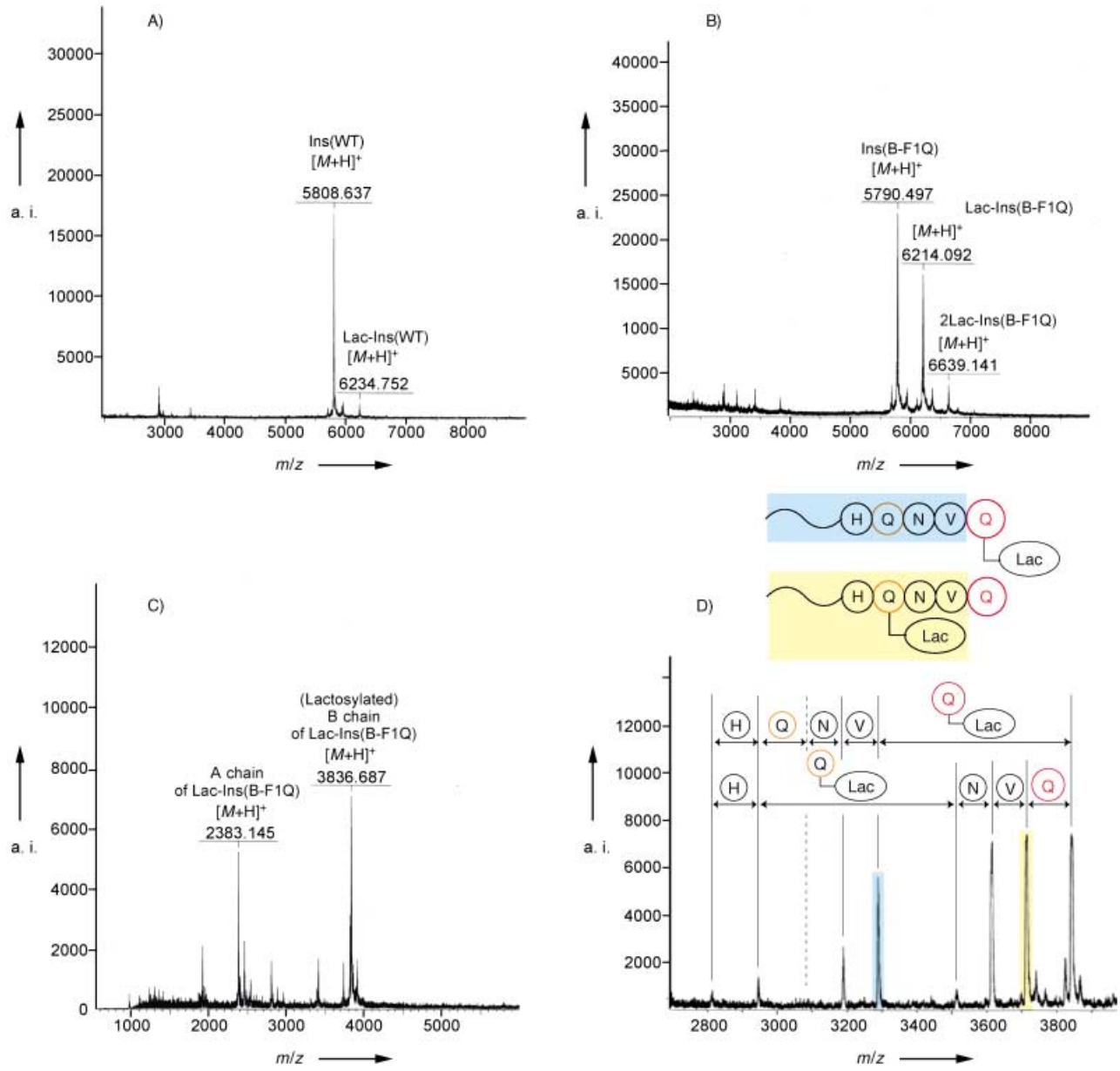



***Figure 1.*** MALDI-TOF mass spectra of insulin and glycosylated insulin derivatives. A) Wild-type insulin after the TGase reaction. B) Ins(B-F1Q) after the TGase reaction. C) Lac-Ins(B-F1Q) after cleavage of the chains by 2-thioethanol (final concentration = 0.1 mmol mL$^{-1}$). D) Amino acid sequence analysis of the N terminus of the lactosylated B chain with aminopeptidase M. The lactosylated B chain was incubated with aminopeptidase M (final concentration = 1 U mL$^{-1}$; Roche) for 10 min at 37 °C. The reaction mixture was analyzed by MALDI-TOF (Bruker REFLEXIII) mass spectrometry. a.i. = arbitrary intensity.

thought to affect the microenvironment around Q4, thereby resulting in enhanced accessibility for TGase. The lactose-displaying insulin, which showed a single peak in the HPLC analysis, was found to be a mixture of two isomers, and we employed the mixture of Q1- and Q4-glycosylated insulin for further experiments.

The sialylation of Lac-Ins(B-F1Q) was performed by means of recombinant α2,6-SiaT, which transfers the sialic acid of CMP-sialic acid to position 6 of galactose (Scheme 1). The reaction mixtures were purified with reversed-phase HPLC to yield Sia2,6-Lac-Ins(B-F1Q).

To investigate if changes in the secondary/tertiary structure of the insulin molecule had occurred because of the mutation and glycosylation, far-UV/CD spectra of the insulin derivatives were compared to that of Ins(WT) (Figure 3). The spectrum of Ins(A-G1Q) was different from that of Ins(WT), a fact suggesting that the α-helix structure of the N terminus of the insulin A chain was changed by the point mutation. The spectra of Ins(B-F1Q) and Sia2,6-Lac-Ins(B-F1Q) were similar to that of Ins(WT). Therefore, the F1Q mutation and glycosylation at the N terminus of the insulin B chain had no effect on the overall structure, as evident from the CD spectra.

The biological activity of the two glycosylated insulin derivatives was also evaluated by using STZ-treated mice (Figure 2). The glycosylated insulins showed similar activity

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

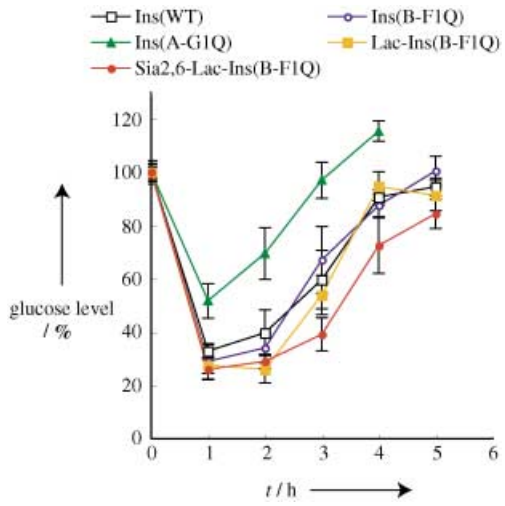



***Figure 2.*** Changes in blood-glucose levels after the administration of insulin and glycosylated insulin derivatives. Mice (C57BL/6J, male, 7 weeks of age) were fasted for 18–20 h and then injected intraperitoneally with a single dose of streptozotocin (STZ; 200 mg kg$^{-1}$ body weight; Wako). Mice with initial blood-glucose levels of above 400 mg dL$^{-1}$ were selected for the monitoring of blood-glucose levels. Five mice were used for each insulin derivative. Blood samples were obtained from the tail vein at 0, 1, 2, 3, 4, and 5 h after subcutaneous administration of Ins(WT), Ins(B-F1Q), Ins(A-G1Q), Lac-Ins(B-F1Q), or Sia2,6-Lac-Ins(B-F1Q) (1 U of each). Blood-glucose levels were measured by using a Glucocard (Kyoto Daiichi Kagaku Co.). Data are presented as a percentage of the initial values ± standard error.

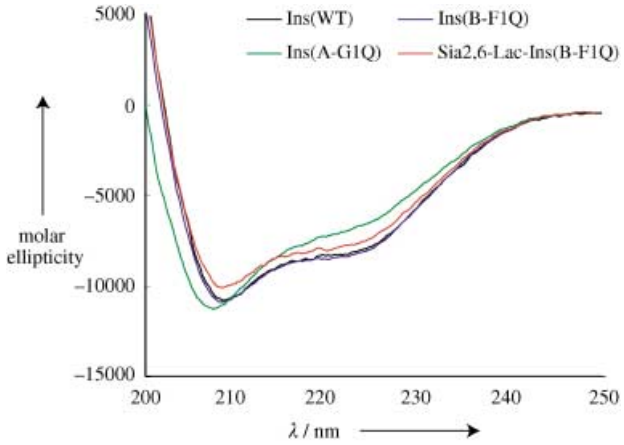



***Figure 3.*** Circular dichroism spectra of insulin derivatives. All the insulin samples were dissolved in phosphate-buffered saline (pH 7.4). The protein concentrations were all adjusted to 40 μM by measurement of the UV absorption at 280 nm. CD measurements were performed on a Jasco-820 CD spectropolarimeter at 25 °C. For the far-UV/CD spectra, samples were scanned from 200–250 nm and accumulated 10 times at a resolution of 0.2 nm with a scanning speed of 100 nm min$^{-1}$. The cell length was 0.1 cm. All the CD data are expressed as the mean residue ellipticity.

shortly after subcutaneous injection, but the activity levels varied significantly with time. Lac-Ins(B-F1Q) had a slightly extended glucose-lowering effect compared to that of the wild type. However, Sia2,6-Lac-Ins(B-F1Q) displayed clearly prolonged glucose-lowering action, especially after three and four hours. We speculate that the long-acting effect of Sia2,6-Lac-Ins(B-F1Q) might be because the introduced sialic acid moiety increases the resistance against degradation of insulin in the liver.

In conclusion, we demonstrated that the point mutation of the N terminus of the insulin B chain enabled chemoenzymatic glycosylation to be achieved by combined use of a transglutaminase and a glycosyltransferase, thereby providing enzymatic access to link insulin and sugars. Sialyllactose-displaying insulin was found to show a prolonged duration of glucose-lowering action in the blood in vivo compared to the action of wild-type or lactose-displaying insulin, a result implying that the introduction of sialic acid to the protein is essential for prolonged activity in vivo. Our chemoenzymatic method generates artificial glycoproteins with a homogeneous glycoform and is suitable for large-scale production. This strategy is widely applicable for the synthesis of novel artificial glycopeptides and glycoproteins with improved and potent biological activity as protein-based drugs. Further study on the applications of this method is currently underway in our laboratory.

## Experimental Section

Wild-type and mutant proinsulins were expressed and purified according to the procedure in ref. [16], except that arginine was used instead of methionine as a linker between the polyhistidine tag and the proinsulins for easy single cleavage.[17] The proinsulins were properly folded according to the protocol described in ref. [18]. The histidine tag and the C chain of the proinsulins were removed by trypsin cleavage.[17] Compound **1** was chemically synthesized according to the procedure described in ref. [12].

Received: October 10, 2003 [Z53058]

**Keywords:** drug design · glycoproteins · glycosylation · insulin · sialic acids

[1] a) O. Pillai, R. Panchagnula, *Drug Discovery Today* **2001**, *6*, 1056–1061; b) D. E. DeWitt, I. B. Hirsch, *JAMA J. Am. Med. Assoc.* **2003**, *289*, 2254–2264; c) D. E. DeWitt, D. C. Dugdale, *JAMA J. Am. Med. Assoc.* **2003**, *289*, 2265–2269.

[2] M. Schmidt, K. R. Babu, N. Khanna, S. Marten, U. Rinas, *J. Biotechnol.* **1999**, *68*, 71–83.

[3] E. Gershonov, Y. Shechter, M. Fridkin, *Diabetes* **1999**, *48*, 1437–1442.

[4] a) J. Markussen, P. Hougaard, U. Ribel, A. R. Sorensen, E. Sorensen, *Protein Eng.* **1987**, *1*, 205–213; b) J. Markussen, I. Diers, A. Engesgaard, M. T. Hansen, P. Hougaard, L. Langkjaer, K. Norris, U. Ribel, A. R. Sorensen, E. Sorensen, H. O. Voigt, *Protein Eng.* **1987**, *1*, 215–223; c) J. Markussen, I. Diers, P. Hougaard, L. Langkjaer, K. Norris, L. Snel, A. R. Sorensen, E. Sorensen, H. O. Voigt, *Protein Eng.* **1988**, *2*, 157–166.

[5] J. Markussen, S. Havelund, P. Kurtzhals, A. S. Andersen, J. Halstrom, E. Hasselager, U. D. Larsen, U. Ribel, L. Schaffer, K. Vad, I. Jonassen, *Diabetologia* **1996**, *39*, 281–288.

[6] a) E. Gershonov, I. Goldwaser, M. Fridkin, Y. Shechter, *J. Med. Chem.* **2000**, *43*, 2530–2537; b) P. Kurtzhals, B. Kiehr, A. R. Sorensen, *J. Pharm. Sci.* **1995**, *84*, 1164–1168.

[7] G. Walsh, *Protein Biochemistry and Biotechnology*, Wiley, Chichester, **2002**, p. 30.

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

[8] a) J. C. Egrie, E. Dwyer, J. K. Browne, A. Hitz, M. A. Lykos, *Exp. Hematol.* **2003**, *31*, 290 – 299; b) J. C. Egrie, J. K. Browne, *Br. J. Cancer* **2001**, *84* (Suppl. 1), 3 – 10.
[9] a) M. Brownlee, A. Cerami, *Science* **1979**, *206*, 1190 – 1191; b) T. Uchio, M. Baudys, F. Liu, S. C. Song, S. W. Kim, *Adv. Drug Delivery Rev.* **1999**, *35*, 289 – 306; c) E. J. Bilsky, R. D. Egleton, S. A. Mitchell, M. M. Palian, P. Davis, J. D. Huber, H. Jones, H. I. Yamamura, J. Janders, T. P. Davis, F. Porreca, V. J. Hruby, R. Polt, *J. Med. Chem.* **2000**, *43*, 2586 – 2590; d) D. Macmillan, R. M. Bill, K. A. Sage, D. Fern, S. L. Flitsch, *Chem. Biol.* **2001**, *8*, 133 – 145.
[10] a) J. E. Folk, P. W. Cole, *Biochim. Biophys. Acta* **1966**, *122*, 244 – 264; b) T. Ohtsuka, M. Ota, N. Nio, M. Motoki, *Biosci. Biotechnol. Biochem.* **2000**, *64*, 2608 – 2613.
[11] a) H. J. Gross, R. Brossmer, *Glycoconjugate J.* **1995**, *12*, 739 – 746; b) A. Nagpurkar, D. Hunt, S. Mookerjea, *Int. J. Biochem. Cell Biol.* **1996**, *28*, 1337 – 1348c) E. R. Sjoberg, H. Kitagawa, J. Glushka, H. van Halbeek, J. C. Paulson, *J. Biol. Chem.* **1996**, *271*, 7450 – 7459.
[12] T. Ohta, N. Miura, N. Fujitani, F. Nakajima, K. Niikura, R. Sadamoto, C.-T. Guo, T. Suzuki, Y. Suzuki, K. Monde, S.-I. Nishimura, *Angew. Chem.* **2003**, *115*, 5344 – 5347; *Angew. Chem. Int. Ed.* **2003**, *42*, 5186 – 5189.
[13] D. Ramos, P. Rollin, W. Klaffke, *J. Org. Chem.* **2001**, *66*, 2948 – 2956.
[14] G. I. Bell, R. L. Pictet, W. J. Rutter, B. Cordell, E. Tischer, H. M. Goodman, *Nature* **1980**, *284*, 26 – 32.
[15] K. D. Hinds, S. W. Kim, *Adv. Drug Delivery Rev.* **2002**, *54*, 505 – 530.
[16] R. B. Mackin, *Protein Expression Purif.* **1999**, *15*, 308 – 313.
[17] P. Jonasson, J. Nilsson, E. Samuelsson, T. Moks, S. Stahl, M. Uhlen, *Eur. J. Biochem.* **1996**, *236*, 656 – 661.
[18] D. J. Cowley, R. B. Mackin, *FEBS Lett.* **1997**, *402*, 124 – 130.

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 