Glycoprotein Synthesis

DOI: 10.1002/anie.200904376

Design and Synthesis of a Homogeneous Erythropoietin Analogue with Two Human Complex-Type Sialyloligosaccharides: Combined Use of Chemical and Bacterial Protein Expression Methods**

*Kiriko Hirano, Derek Macmillan, Katsunari Tezuka, Takashi Tsuji, and Yasuhiro Kajihara**

Glycoproteins, which comprise oligosaccharides and a protein, have been prepared by using mammalian cells such as Chinese hamster ovary cells (CHO)[1] and have been used as pharmaceuticals.[2] However, the resulting glycosylation pattern is highly heterogeneous,[3] although mutagenesis has been powerful in regulation of the glycosylation position and structure.[4] Several chemical and enzymatic synthetic methods have been reported for the production of homogeneous glycoproteins that contain complex-type oligosaccharides, in order to investigate the functions of the oligosaccharides.[5] Chemical approaches for the synthesis of glycoproteins share the advantage that the oligosaccharide structure and glycosylation position can be predetermined through synthetic design. Therefore, it is expected that tailor-made glycoproteins can be synthesized for both the understanding of oligosaccharide function[2] and their application to the preparation of glycoprotein drug candidates. The clear protein function and bioactivity of erythropoietin (EPO) make it the best glycoprotein with which to elucidate oligosaccharide protein functions. Herein, we report the design and robust synthesis of an EPO analogue that has two homogeneous complex-type sialyloligosaccharides.

EPO is known to act as a drug for the treatment of renal anemia. EPO consists of 166 amino acid residues and has three N-linked complex-type oligosaccharides attached to the asparagine side chain at the 24, 38, and 83 positions, and a short O-linked oligosaccharide at serine 126.[6] X-ray analysis of EPO and its receptor complex have been reported, though the data is for nonglycosylated EPO.[7] Details of the complex between glycosylated EPO and its receptor have been discussed,[8] and we have also made a simple model of the complex (Figure 1 a).[9] It is known that the three complex type sialyloligosaccharides are mainly tri- or tetra-antennary type structures and this model, which is based on X-ray data, suggests that all three oligosaccharides face toward one side of the molecule when EPO binds to the receptor (Figure 1 a). These sialyloligosaccharides increase the lifetime of EPO in blood[10] because of interference from glomerular filtration or interaction of galactose-binding lectin. We have demonstrated the synthesis of homogeneous glycopeptides and glycoproteins by using dibranched complex-type sialyloligosaccharides isolated from egg yolk,[11] and therefore we aimed to establish an efficient synthesis of a glycosylated EPO

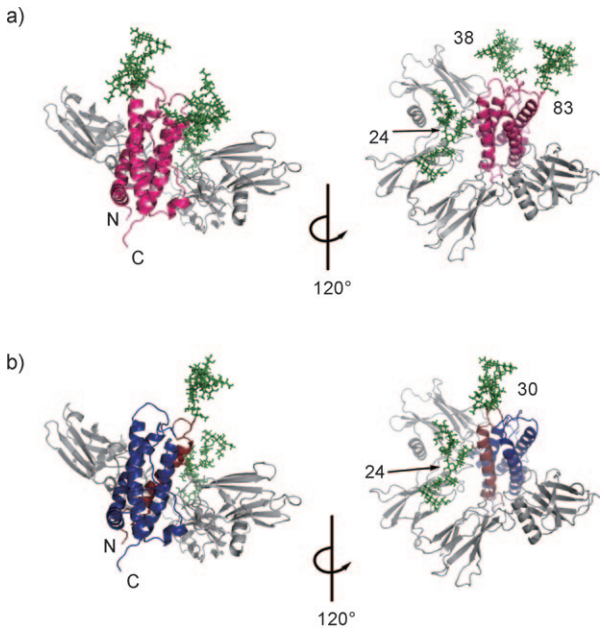


Figure 1. Binding model of a) natural EPO and the b) EPO analogue with the receptor. The oligosaccharides, EPO, and the receptor are shown in green, blue and red, and gray, respectively.

[*] K. Hirano, Prof. Dr. Y. Kajihara[+]
International Graduate School of Arts and Sciences
Yokohama City University
22-2, Seto, Kanazawa-ku, Yokohama, 236-0027 (Japan)
E-mail: kajihara@chem.sci.osaka-u.ac.jp

Prof. Dr. D. Macmillan
Department of Chemistry, University College London
20 Gordon Street, London WC1H 0AJ (UK)

Dr. K. Tezuka
Research Project for Glycoprotein Engineering funded by Otsuka Chemical Co., Ltd., Tokyo University of Science
2641 Yamazaki, Noda, Chiba 278-8510 (Japan)

Prof. Dr. T. Tsuji
Department of Biological Science and Technology
Faculty of Industrial Science and Technology
Tokyo University of Science
2641 Yamazaki, Noda, Chiba 278-8510 (Japan)

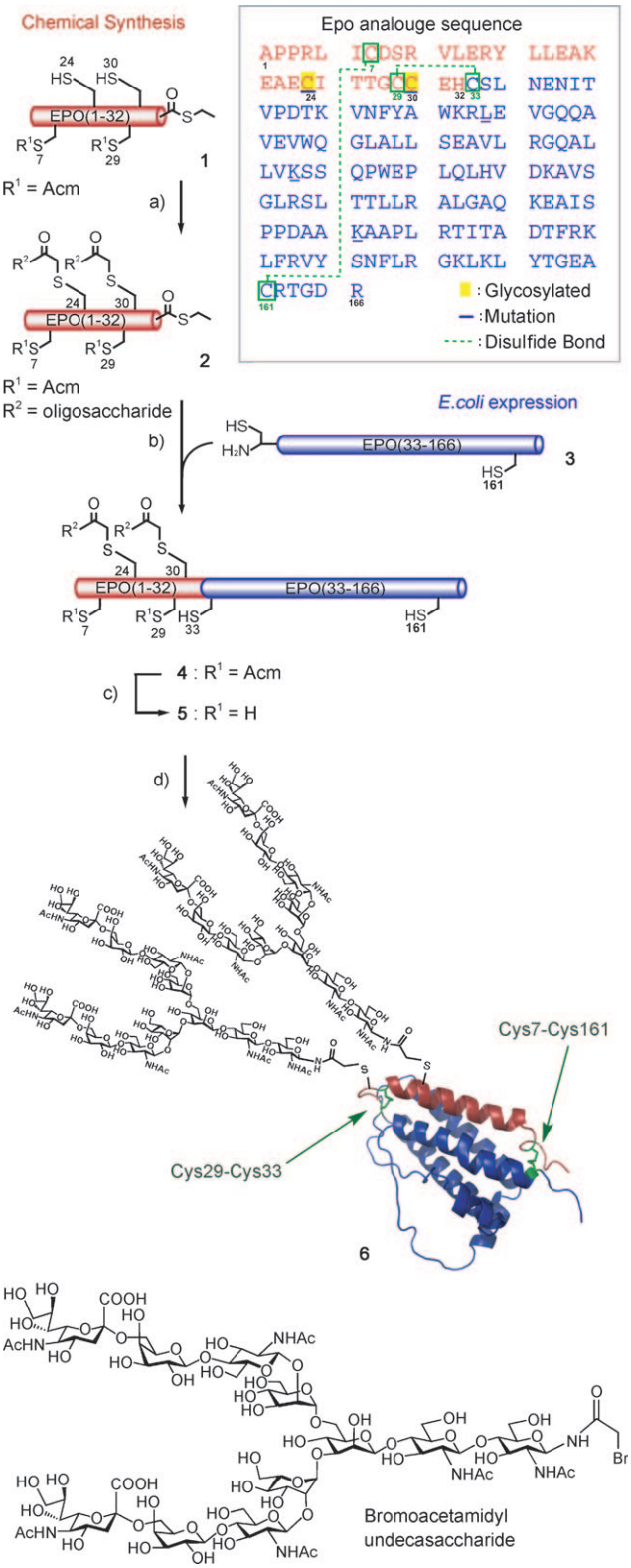
[+] Current address:
Department of Chemistry, Graduate School of Science, Osaka University
1-1 Machikaneyama, Toyonaka, Osaka 560-0043 (Japan)

[**] Financial support from the Japan Society for the promotion of Science (Grant-in-Aid for Creative Scientific Research no.17GS0420 and 18550155) is acknowledged. We thank Dr. Yukishige Ito (Riken), Dr. Yukio Nishina (Yokohama City University), Yuri Hatakeyama, Kyoko Isumi, Ai Himeno (Otsuka Chemical Co.), and Ayako Kajihara for technical support and encouragement.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200904376.

analogue by using this methodology. For the efficient synthesis of such a compound, we employed a strategy to prepare the whole polypeptide chain by use of native chemical ligation (NCL)[12] between a synthetic glycopeptide-α-thioester that contains complex-type sialyloligosaccharides, and the N-terminal cysteine of a large polypeptide chain, prepared through expression in *E. coli*.[13] Fortunately, EPO has a cysteine residue at position 33 and thus we thought that we could utilize this position as a ligation site for NCL.[13a] Because *E. coli* do not perform mammalian-type glycosylation on proteins, incorporation of an oligosaccharide at the natural position Asn83 (Asn = asparginine) was not possible and therefore we opted to incorporate additional sialyloligosaccharides at alternative suitable positions. When Cys33 was used as the ligation site, introduction of an additional glycosylation site was required.

Consequently, we designed a glycosylated EPO analogue, which has two glycosylation sites at position 24 (natural) and position 30 (unnatural; Figure 1b). Glycosylation at position 30, though unnatural, is found in the modified EPO therapeutic Aranesp (Darbepoetin alfa) within a relatively unstructured loop between helix A and helix B of the four α-helical bundle structure.[14] It is anticipated that these oligosaccharides may face toward one side as in natural EPO, thus indicating that sialyloligosaccharides may not interfere with the binding of the glycosylated EPO analogue to a receptor. Based on this concept, we set out to complete the semisynthesis of the EPO analogue by the combined use of chemical glycopeptide synthesis and protein expression in *E. coli*. The first aim was to conduct the synthesis of a glycopeptide-α-thioester (32 amino acid residues) that contain two complex-type sialyloligosaccharides at positions 24 and 30 by the haloacetamide method, which afforded a non-native linkage between protein and oligosaccharides. We employed the haloacetamide method in order to synthesize the glycopeptide-α-thioester that contains two complex-type sialyloligosaccharides **2**.[15] As has been demonstrated,[15a,b] this haloacetamide coupling is a convenient method and its peptide substrates can be easily prepared by a conventional solid phase peptide synthesis (SPPS). As shown in Scheme 1, a peptide comprising the 32 N-terminal residues of EPO (shown in red) was synthesized by SPPS and the subsequent haloacetamide reaction was examined. EPO contains two disulfide bonds: one between Cys7 and Cys161, and the other between Cys29 and Cys33. In order to employ the haloacetamide reaction between bromoacetamidyl undecasaccharides[15b] and the thiol groups of Cys24 and Cys30, the thiol groups of Cys7 and Cys29 were protected with acetamidomethyl (Acm) groups. The peptide **1** was prepared by general 9-fluorenylmethyloxycarbonyl (Fmoc) SPPS. For SPPS, HMPB–PEGA resin (HMPB = 4-(4-hydroxymethyl - 3 - methoxyphenoxy)-butyric acid, PEGA = poly(ethylene glycol)–poly(dimethylacrylamide) copolymer) was used and peptide elongation was conducted by using 2-(1*H*-benzotriazole-1-yl)-



Scheme 1. Synthesis of glycosylated EPO analogue **6**. a) Bromoacetamidyl undecasaccharide, 30 mM urea, 100 mM sodium phosphate/DMF (2:1), pH 7.2, RT, 7 h, 38%; b) EPO(33–166) segment **3**, 6 M Gn-HCl, 0.2 M sodium phosphate, 40 mM TCEP, 0.2 M MPAA, pH 7.3, RT, 15 h, ca. 75%; c) 42 mM AgOAc in 90% AcOH, RT, 4 h, ca. 90% (based on HPLC, mass, and protein estimation by OD280); d) 1) 3 M Gn-HCl, 100 mM Tris-HCl, 4 μM cysteine, 0.5 μM cystine, pH 8.5, 4°C, 12 h; 2) 1 M Gn-HCl, 100 mM Tris-HCl, pH 8.0, 4°C, 8 h; 3) 10 mM Tris-HCl, pH 7.0, 4°C, 12 h, ca. 90% (based on HPLC, mass, and protein estimation by OD280). MPAA = 4-(carboxymethyl)thiophenol, Tris = tris(hydroxymethyl)aminomethane, TCEP = tris(2-carboxyethyl)phosphine hydrochloride.

1,1,3,3-tetramethylaminium hexafluorophosphate (HBTU), 1-hydroxy-1*H*-benzotriazole (HOBt), and *N*,*N*-diisopropylethylamine (DIPEA). After construction of the 32-residue peptide, the fully protected peptide was released from the HMPB linker after exposure to AcOH/trifluoroethanol. The selectively deprotected C-terminal α-carboxylic acid was converted to the peptide-α-thioester using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), DIPEA, and excess EtSH at −20 °C. Under these conditions, even a C-terminal histidine residue did not cause epimerization during the thioesterification and the thioester was obtained in good yield.[16] The haloacetamide reaction was next performed between peptide thioester **1** and the bromoacetamidyl undecasaccharide to obtain the homogeneous glycopeptide analogue **2**. This coupling reaction occurred smoothly and afforded the glycopeptide-α-thioester **2** within 7 hours. HPLC analysis and the mass spectrum of this product confirmed the desired structure (see the Supporting Information: ESI-MS *m*/*z* 8332.6 [*M*+H$^+$] (deconvoluted) calcd: 8333.5). The preparation of the polypeptide chain **3**, which comprises residues 33 to 166, was performed as previously described.[13b] This process afforded homogeneous polypeptide chain **3**, which contain a cysteine residue at the N-terminus.

After the successful synthesis of the glycopeptide-α-thioester **2** and the polypeptide chain **3**, which has a cysteine residue at the N-terminus, we next investigated the NCL reaction between glycopeptide **2** and polypeptide segment **3**. The reaction was performed in a solution of 0.2 M sodium phosphate containing 6 M guanidine hydrochloride (Gn-HCl) and employed MPAA as catalyst.[17] After 15 hours, this NCL reaction afforded the complete EPO–polypeptide chain **4**. Monitoring this reaction by SDS electrophoresis and HPLC indicated that NCL afforded the desired product in moderate yield. After purification, the glycosylated polypeptide **4** was analyzed by mass analysis. This analysis clearly verified that the NCL reaction afforded the glycosylated polypeptide chain **4**. In order to remove the Acm groups, the glycosylated polypeptide **4** was treated with silver acetate to afford the EPO analogue **5**. This deprotection step was monitored by ESI mass spectrometry. Within 4 hours, substrate **4** was converted into the desired product **5** and ESI mass spectrometry clearly showed the desired molecular weight of **5**. Because we could prepare the glycosylated EPO polypeptide **5**, we finally carried out protein folding experiments in order for this glycosylated polypeptide to attain its ideal three-dimensional structure.

The folding process was performed by using dialysis.[18] This protocol initially employed 3 M Gn-HCl to denature the glycosylated polypeptide chain **5** followed by a three-step dilution process of Gn-HCl in the presence of a cysteine–cystine redox system.[18] This folding experiment successfully afforded the glycosylated EPO **6**. The structure of folded EPO **6** in the solution was then evaluated after purification. The analytical data for **6** is shown in Figure 2; HPLC analysis showed only one major peak (Figure 2c.). SDS-PAGE (Figure 2c, inset) showed that glycosylated EPO appears as a single band of approximately 23 kDa in the SDS-PAGE gel. The ESI mass spectrum (Figure 2a) showed multiple charge states, and the observed peak was identical to that of the theoretical value of the target molecule. The CD spectrum of EPO analogue **6** (Figure 2b.) was also consistent with the helical secondary structure of a correctly folded EPO.[19] To confirm that the correct disulfide bonds had formed in the folded **6**, we examined peptidase digestion and subsequent mass analysis of resultant peptide fragments (see the Supporting Information). The resulting disulfide-bonded fragments were found to be consistent with the anticipated peptide fragments and the data clearly confirmed that the folding process afforded two correctly formed disulfide bonds. Therefore we concluded that the glycosylated EPO in the final folding buffer solution was correctly folded and the yield of folding experiment was approximately 90%.

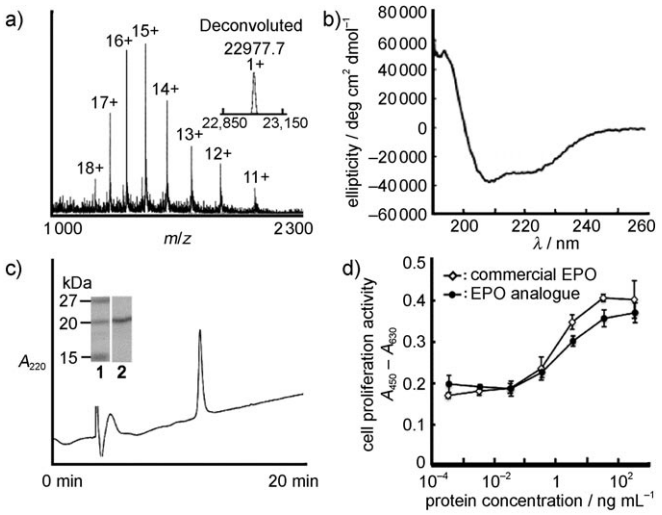


Figure 2. a) ESI mass spectrum (calcd *m*/*z* 22977.2; found m/z 22977.7), b) CD spectrum, c) HPLC and SDS electrophoresis: lane 1: marker; and lane 2: folded EPO **6**, d) estimation of cell (TF-1 cell) proliferation.

The biological activity of the folded EPO analogue **6** was also evaluated. We examined a cell proliferation assay using TF-1 cells, which were established from bone marrow cells.[20] Assays using synthetic EPO successfully showed the induction of TF-1 cell proliferation. The degree of cell proliferation relative to the concentration of **6** was estimated (Figure 2d). Both the native and synthetic EPO analogues exhibited cell proliferation activity from 50 pg mL^{-1} of **6**. These results suggest that the assembly of the EPO analogue **6** successfully exhibited suitable biological activity in vitro. In terms of an in vivo assay with **6**, we could not observe clear hematocrit activity with 20 μg kg^{-1} of **6**, although native EPO, which contains several tri- or tetrabranched sialyloligosaccharides, and polymer-modified EPO analogues can exhibit a definite activity with doses of 22.5 μg kg^{-1}.[18] EPO produced from yeast also has three dibranched sialyloligosaccharides, and doses of 45 μg kg^{-1} might be also needed to observe hematocrit activity.[21] Because **6** exhibited potential stability in the buffer solution and serum solution (see the Supporting information), research is under way to incorporate additional di- and tribranched complex type sialyloligosaccharides into this EPO analogue in order to enhance bioactivity in vivo.

We have demonstrated the design and synthesis of a homogenous erythropoietin analogue that has two human complex-type sialyloligosaccharides by the combined use of chemical synthesis and protein expression in *E. coli*. Although bioactive glycosylated cytokines have been prepared by bioengineering, our demonstration shows that the synthesis of bioactive homogeneous glycoproteins is now within the reach of organic chemists. This strategy will also be useful for the elucidation of the oligosaccharide function of such proteins.

Received: August 5, 2009
Published online: October 1, 2009

Keywords: erythropoietin · glycopeptides · glycoproteins · native chemical ligation · protein expression

[1] E. Grabenhorst, P. Schlenke, S. Pohl, M. Nimtz, H. S. Conradt, *Glycoconjugate J.* **1999**, *16*, 81 – 97.

[2] a) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131 – 163; b) H. Ledford, *Nature* **2007**, *449*, 259; H. Ledford, *Nature* **2007**, *449*, 274 – 276; c) T. U. Gerngross, *Nat. Biotechnol.* **2004**, *22*, 1409 – 1414; d) B. G. Davis, *Chem. Rev.* **2002**, *102*, 579 – 601.

[3] M. Gawlitzek, U. Valley, M. Nimtz, R. Wagner, H. S. Conradt, *J. Biotechnol.* **1995**, *42*, 117 – 131.

[4] S. R. Hamilton, P. Bobrowicz, B. Bobrowicz, R. C. Davidson, H. Li, T. Mitchell, J. H. Nett, S. Rausch, T. A. Stadheim, H. Wischnewski, S. Wildt, T. U. Gerngross, *Science* **2003**, *301*, 1244 – 1246.

[5] a) P. Nagorny, B. Fasching, X. Li, G. Chen, B. Aussedat, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5792 – 5799; b) C. Kan, J. D. Trzupek, B. Wu, Q. Wan, G. Chen, Z. Tan, Y. Yuan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5438 – 5443; c) Y. Yuan, J. Chen, Q. Wan, Z. Tan, G. Chen, C. Kan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5432 – 5437; d) Z. Tan, S. Shang, T. Halkina, Y. Yuan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5424 – 5431; e) W. Huang, C. Li, B. Li, M. Umekawa, K. Yamamoto, X. Zhang, Lai-Xi Wang, *J. Am. Chem. Soc.* **2009**, *131*, 2214 – 2223; f) C. Piontek, D. V. Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem.* **2009**, *121*, 1974 – 1978; *Angew. Chem. Int. Ed.* **2009**, *48*, 1941 – 1945; g) N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson, Y. Kajihara, *J. Am. Chem. Soc.* **2008**, *130*, 501 – 510; h) Y.-Y. Yang, S. Ficht, A. Brik, C. H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 7690 – 7701; i) C. P. R. Hackenberger, C. T. Friel, S. E. Radlford, B. Imperiali, *J. Am. Chem. Soc.* **2005**, *127*, 12882 – 12889; j) L. A. Marcaurelle, L. S. Mizoue, J. Wilken, L. Oldham, S. B. H. Kent, T. M. Handel, C. R. Bertozzi, *Chem. Eur. J.* **2001**, *7*, 1129 – 1132.

[6] a) M. Takeuchi, S. Takasaki, H. Miyazaki, T. Kato, S. Hoshi, N. Kochibe, A. Kobata, *J. Biol. Chem.* **1988**, *263*, 3657 – 3663; b) H. Sasaki, B. Bothner, A. Dell, M. Fukud, *J. Biol. Chem.* **1987**, *262*, 12059 – 12076; c) P.-H. Lai, R. Everett, F.-F. Wang, T. Arakawa, E. Goldwasser, *J. Biol. Chem.* **1986**, *261*, 3116 – 3121.

[7] R. S. Syed, S. W. Reid, C. Li, J. C. Cheetham, K. H. Aoki, B. Liu, H. Zhan, T. D. Osslund, A. J. Chirino, J. Zhang, J. Finer-Moore, S. Elliott, K. Sitney, B. A. Katz, D. J. Matthews, J. J. Wendoloski, J. Egrie, R. M. Stroud, *Nature* **1998**, *395*, 511 – 516.

[8] S. Elliott, T. Lorenzini, S. Asher, K. Aoki, D. Brankow, L. Buck, L. Busse, D. Chang, J. Fuller, J. Grant, N. Hernday, M. Hokum, S. Hu, A. Knudten, N. Levin, R. Komorowski, F. Martin, R. Navarro, T. Osslund, G. Rogers, N. Rogers, G. Trail, J. Egrie, *Nat. Biotechnol.* **2003**, *21*, 414 – 421.

[9] The model was made by use of free software (http://www.glycosciences.de/modeling/glyprot/php/main.php) and Protein data bank (ID: 1BUY).

[10] a) E. Tsuda, G. Kawanishi, M. Ueda, S. Masuda, R. Sasaki, *Eur. J. Biochem.* **1990**, *188*, 405 – 411; b) M. N. Fukuda, H. Sasaki, L. Lopez, M. Fukuda, *Blood* **1989**, *73*, 84 – 89.

[11] Y. Kajihara, Y. Suzuki, N. Yamamoto, K. Sasaki, T. Sakakibara, L. R. Juneja, *Chem. Eur. J.* **2004**, *10*, 971 – 985.

[12] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779.

[13] a) J. P. Richardson, D. Macmillan, *Org. Biomol. Chem.* **2008**, *6*, 3977 – 3982; b) D. Macmillan, L. Arham, *J. Am. Chem. Soc.* **2004**, *126*, 9530 – 9531.

[14] G. Walsh, R. Jefferis, *Nat. Biotechnol.* **2006**, *24*, 1241 – 1252.

[15] a) T. Murase, T. Tsuji, Y. Kajihara, *Carbohydr. Res.* **2009**, *344*, 762 – 770; b) N. Yamamoto, T. Sakakibara, Y. Kajihara, *Tetrahedron Lett.* **2004**, *45*, 3287 – 3290; c) D. Macmillan, R. M. Bill, K. A. Sage, D. Fern, S. L. Flitsch, *Chem. Biol.* **2001**, *8*, 133 – 145; d) S. Y. C. Wong, G. R. Guile, R. A. Dwek, G. Arsequell, *Biochem. J.* **1994**, *300*, 843 – 850.

[16] Y. Kajihara, A. Yoshihara, K. Hirano, N. Yamamoto, *Carbohydr. Res.* **2006**, *341*, 1333 – 1340.

[17] E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640 – 6646.

[18] G. G. Kochendoerfer, S.-Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. H. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884 – 887.

[19] J. M. Davis, T. Arakawa, *Biochemistry* **1987**, *26*, 2633 – 2638.

[20] T. Kitamura, A. Tojo, T. Kuwaki, S. Chiba, K. Miyazono, A. Urabe, F. Takaku, *Blood* **1989**, *73*, 375 – 380.

[21] S. R. Hamilton, R. C. Davidson, N. Sethuraman, J. H. Nett, Y. Jiang, S. Rios, P. Bobrowicz, T. A. Stadheim, H. Li, B-K. Choi, D. Hopkins, H. Wischnewski, J. Roser, T. Mitchell, R. R. Strawbridge, J. Hoopes, S. Wildt, T. U. Gerngross, *Science* **2006**, *313*, 1441 – 1443.

