

Semisynthesis of Biologically Active Glycoforms of the Human Cytokine Interleukin 6**

Andreas Reif, Sebastian Siebenhaar, Andrea Tröster, Marina Schmälzlein, Carolin Lechner, Phanindra Velisetty, Karen Gottwald, Claudia Pöhner, Irene Boos, Volker Schubert, Stefan Rose-John, and Carlo Unverzagt*

In memory of Gilbert Ashwell

Abstract: Human interleukin 6 (IL-6) is a potent cytokine with immunomodulatory properties. As the influence of N-glycosylation on the *in vivo* activities of IL-6 could not be elucidated so far, a semisynthesis of homogeneous glycoforms of IL-6 was established by sequential native chemical ligation. The four cysteines of IL-6 are convenient for ligations and require only the short synthetic glycopeptide 43–48. The Cys-peptide 49–183 could be obtained recombinantly by cleavage of a SUMO tag. The fragment 1–42 was accessible by the simultaneous cleavage of two inteins, leading to the 1–42 thioester with the native N-terminus. Ligation and refolding studies showed that the inherently labile Asp–Pro bond 139–140 was detrimental for the sequential C- to N-terminal ligation. A reversed ligation sequence using glycopeptide hydrazides gave full-length IL-6 glycoproteins, which showed full bioactivity after efficient refolding and purification.

The majority of recombinant human therapeutic proteins is glycosylated. Typically their stability, biological activity, and serum half-life is influenced by the sugar part. As cellular biosynthesis renders the carbohydrate moieties of glycoproteins heterogeneous, the isolation of pure glycoforms from natural sources is rarely possible. However, with the recent advances in chemoselective ligation methods, homogeneous glycoproteins have become accessible by synthesis.^[1] Native chemical ligation^[2] has enabled chemists to assemble complex N-glycoproteins^[3] with more than one hundred residues by means of chemical and semisynthetic^[4] methods. As a model for synthetic therapeutic glycoproteins, we chose the pleiotropic cytokine human interleukin 6 (IL-6), which can activate the immune system and regenerative processes.^[5] Here we describe a semisynthetic approach to homogeneous glyco-

forms of biologically active human IL-6 by sequential native chemical ligations. Key features of this synthesis are the optimization of a two-intein system for thioesters with a native N-terminus and a reversed ligation sequence using latent thioester glycopeptides.

Processed human IL-6 (183 amino acids) contains four cysteines and an N-glycan at Asn44. Glycosylation of IL-6 is highly cell-type specific.^[6] Analysis of the N-glycans of human IL-6 showed a mixture of common and truncated structures.^[7] The bioactivity of glycosylated hIL-6 was reported to be equal to nonglycosylated variants in assays under *in vitro* conditions.^[8] For *in vivo* applications^[9] of recombinant glycosylated hIL-6, structure–activity relationships could not be established, mainly because of the lack of well-defined IL-6 glycoforms.

For the synthesis of homogeneous IL-6 glycoproteins, we identified the cysteines 43 and 49 as most suitable for sequential native chemical ligation using thiazolidines^[10] (Scheme 1a). This disconnection strategy led to the short synthetic glycopeptide thioester **B** (43–48), flanked by thioester **A** (1–42) and Cys-peptide **C** (49–183), which we envisioned to produce by recombinant methods using inteins. Despite the accessibility of thioester **A** by existing solid-phase methods, we were attempting a two-intein strategy by the expression of fusion protein **D**. The use of these selectively cleavable expression and purification tags was one of the key methods in the semisynthesis of glycosylated ribonuclease C.^[11] Controlled removal of both inteins from **D** should directly provide IL-6 thioester **A** with its native N-terminus. To the best of our knowledge the N- and C-terminal cleavage of inteins from fusion proteins was only reported in the context of cyclic peptides and proteins.^[12]

The synthetic gene for fragment **A** was cloned into the pTWIN1 vector (NEB) bearing an N-terminal *Ssp* DnaB and a C-terminal *Mxe* GyrA intein fused to chitin-binding domains (CBD). Overexpression of the protein in *E. coli* at 37 °C gave inclusion bodies exclusively, whereas at 15 °C some soluble fusion protein **D** was obtained. The soluble protein was loaded onto a chitin affinity column and both inteins were cleaved individually (1. *Ssp* DnaB **F** pH 7.0; 2. *Mxe* GyrA **G** pH 8.5, 50 mM MESNA). LC-MS analysis of the eluate from the column showed MESNA-derived thioester **A1**, but also equal amounts of hydrolyzed peptide **A^{OH}** (data not shown). A similar mixture of **A1** and **A^{OH}** was found when the purified inclusion bodies of **D** were refolded using 0.5 M arginine^[13]

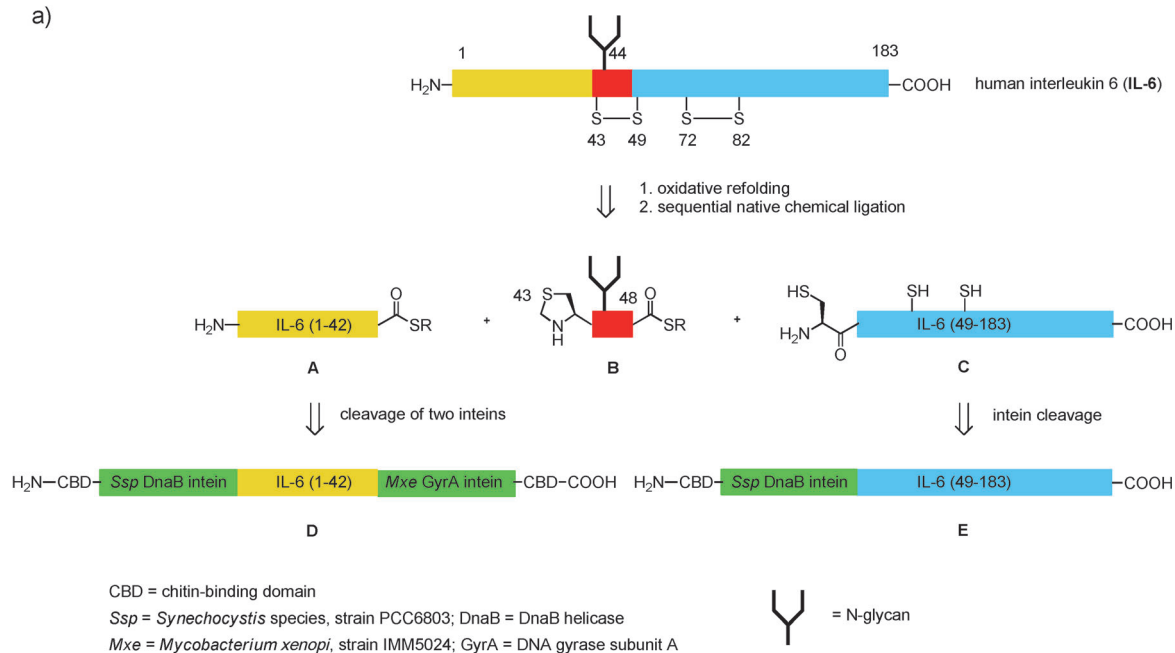
[*] A. Reif, S. Siebenhaar, A. Tröster, M. Schmälzlein, C. Lechner, P. Velisetty, K. Gottwald, C. Pöhner, I. Boos, V. Schubert, Prof. C. Unverzagt
Bioorganic Chemistry, Gebaeude NW1, University of Bayreuth
95440 Bayreuth (Germany)
E-mail: carlo.unverzagt@uni-bayreuth.de

Prof. S. Rose-John
Department of Biochemistry, Christian-Albrechts-University Kiel
24098 Kiel (Germany)

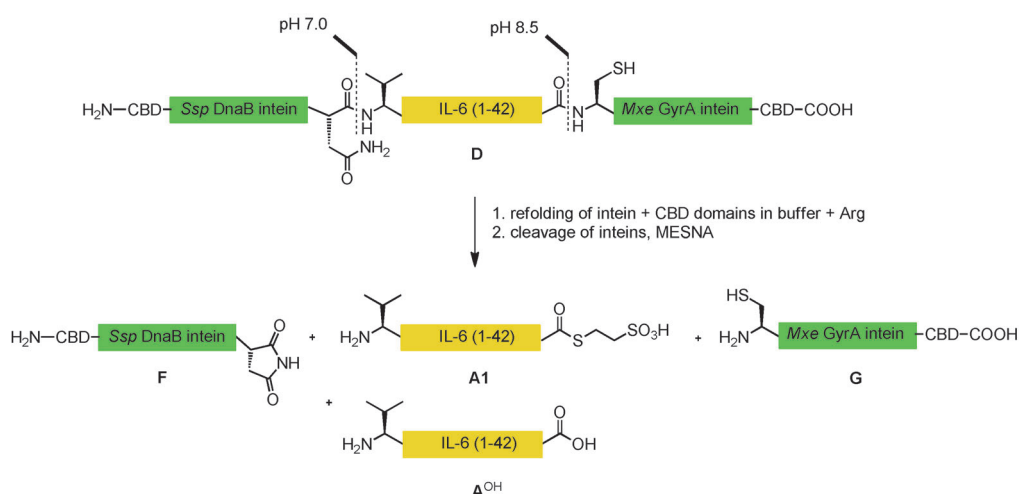
[**] This work was supported by the Deutsche Forschungsgemeinschaft (DFG).

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201407160>.

a)



b)

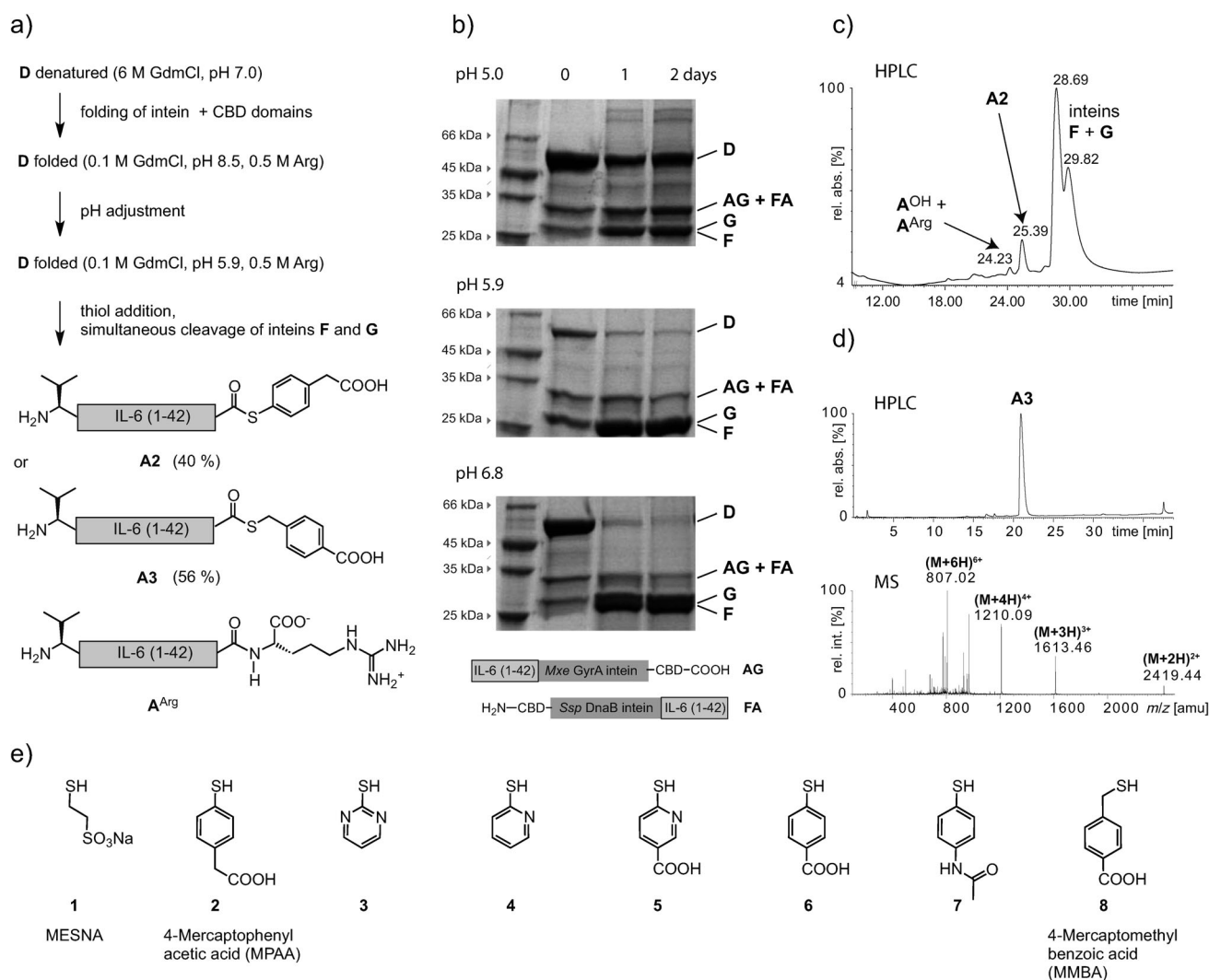


Scheme 1. a) Retrosynthesis of N-glycosylated human IL-6. b) Intein cleavages of **D** generating IL-6 (1–42) MES thioester **A1**.

followed by selective cleavage conditions for each intein (Scheme 1b). Unexpectedly, after cleaving only the N-terminal intein **F** from **D** at pH 7, LC-MS analysis showed some free peptide **A^{OH}**, thus indicating that the C-terminal intein **G** was labile below the recommended pH value of 8.5 (Figure S5).^[4,12a] Thus, we explored the possibility to cleave both inteins simultaneously. Gratifyingly, when adding thiol (50 mM MESNA) under the conditions for the cleavage of intein **F** (pH 7), simultaneous liberation of both inteins **G** and **F** was found. LC-MS showed the desired thioester **A1** co-eluting with the hydrolysis product **A^{OH}** (Figure S6).

By cleaving **D** in the presence of the aromatic thiol MPAA^[14] (**2**, 50 mM) at pH 6.8, the corresponding MPAA thioester **A2** was obtained, which could be separated from **A^{OH}** and the aminolysis^[15] product **A^{Arg}**. The refolding of **D**

was further optimized and simultaneous intein cleavage was compared at different pH values (Scheme 2b). Refolding of **D** from 6 M guanidine hydrochloride (GdmCl) by rapid 60-fold dilution with a 0.5 M Arg buffer at pH 8.5 gave only soluble protein. This solution was divided, the pH value of the batches was adjusted to 5.0, 5.9, and 6.8, and MPAA **2** was added. SDS-PAGE analysis showed only partial cleavage of the inteins at pH 5.0 and accumulation of the intermediates **AG** and **FA**. Nearly complete cleavages were observed at pH 5.9 and 6.8. The highest conversion to MPAA thioester **A2** and less side reactions were found at pH 5.9 (Scheme 2c). On a preparative scale, thioester **A2** was isolated by solid-phase extraction and HPLC. However, lyophilization of the purified MPAA thioester **A2** led to 5–10% of cyclic and oligomeric

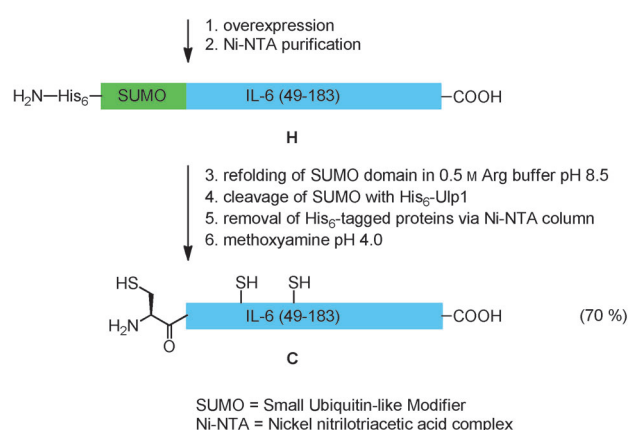


Scheme 2. a) Optimized simultaneous intein cleavages of **D** giving IL-6 thioesters **A2** and **A3**. b) SDS-PAGE analysis of simultaneous intein cleavages of **D** at pH 5.0, 5.9, and 6.8 + MPAA. c) HPLC of simultaneous intein cleavage of **D** at pH 5.9. d) HPLC-MS of purified **A3**. e) Different thiols tested for simultaneous intein cleavage.

side products, which were presumably formed by aminolysis (Figure S13).^[15]

In order to maintain the intein-cleavage efficiency observed for the aromatic thiol MPAA **2** and to increase the stability of the corresponding thioester, a series of seven thiols with aromatic moieties (**2–8**) was screened (Scheme 2e). Addition of the nonheterocyclic aryl thiols **2**, **6**, or **7** to the cleavage solutions gave the desired thioesters, however, accompanied by significant amounts of hydrolysis and aminolysis (Arg) side products. In contrast, for the benzylic thiol MMBA **8**, only traces of side products were found, reflecting an increased stability of the aliphatic thioester **A3** (compare Figures S9 and S10). After further optimization of cleavage and purification, **A3** was obtained in a yield of 56 % after HPLC (70 mg of **A3** from 10 L of bacterial culture, Scheme 2a).

Subsequently, we investigated the preparation of IL-6 49–183 (**C**; Scheme 3). Expression of the intein fusion protein **E** in *E. coli* at 37 °C gave inclusion bodies. Despite the cleavability of the intein moiety after refolding, the isolation of the



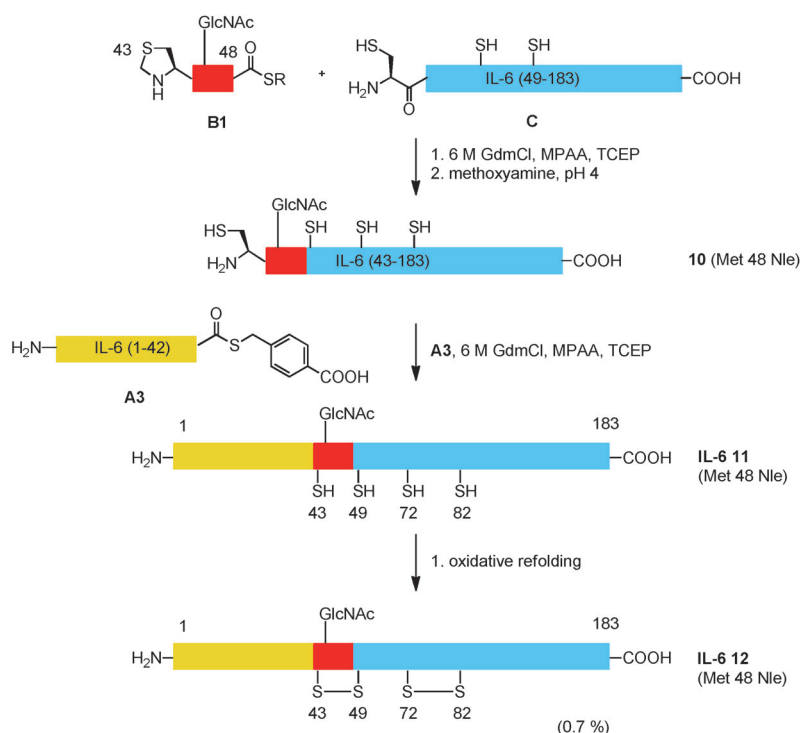
Scheme 3. Overexpression of **H**, cleavage of the SUMO domain, and purification of IL-6 (49–183) **C**.

Cys-peptide **C** by HPLC gave very low yields (data not shown). Thus, an alternative expression system^[16] bearing the

enzymatically cleavable SUMO tag was tested. Overexpression of the SUMO fusion protein **H** in *E. coli* at 40 °C gave inclusion bodies, which were purified by Ni-NTA chromatography under denaturing conditions. The SUMO domain of **H** was refolded by dialysis and cleaved by the SUMO specific protease His₆-Ulp1 (Figures S17 and S18). As HPLC purification of **C** was not feasible, the precipitated and resolubilized protease digest was passed over a Ni-NTA column, thus eliminating all His₆-tagged proteins. A highly enriched Cys-peptide **C** was obtained, which was partly blocked by an N-terminal thiazolidine derived from formaldehyde. This modification was removed by incubation with methoxyamine^[10,17] at pH 4 and Cys-peptide **C** was precipitated by dialysis against H₂O (70 % yield; 75 mg of **C** from 5 L of bacterial culture).

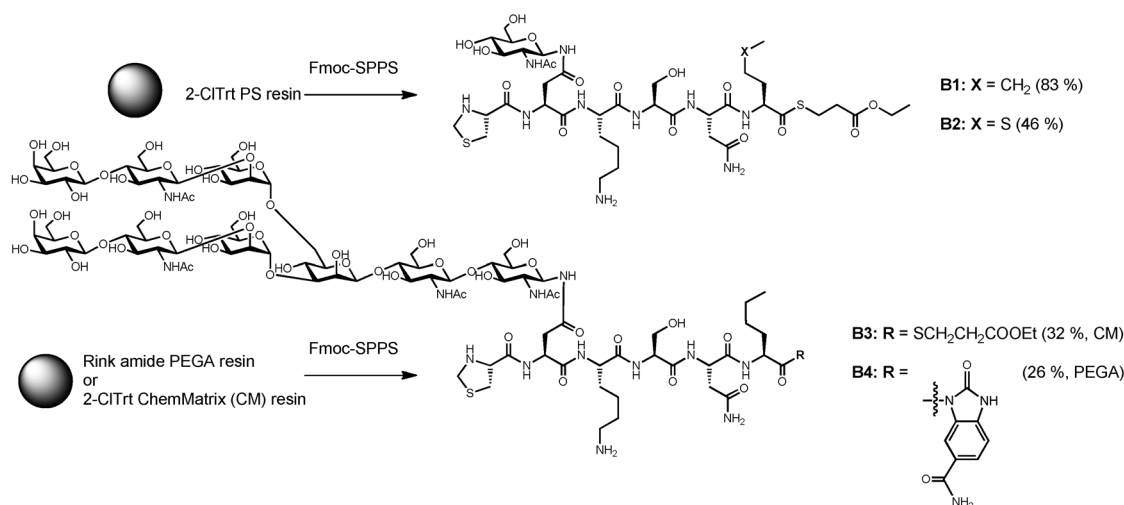
Unexpectedly, the solid-phase synthesis of the short glycopeptide thioesters **B** was not straightforward. Yields were variable, depending on the resin and the linker system (Scheme 4). Best results were obtained for the GlcNAc thioesters **B1** (83 %) and **B2** (46 %) on a polystyrene resin with a trityl linker, followed by in situ thioesterification.^[18] In the case of thioester **B3**, polar resins were required for efficient incorporation of the oligosaccharide-Asn building block.^[19] **B3** was synthesized on PEGA resin/safety catch double linker^[20] (18 %) and on ChemMatrix resin/2Cl-trityl linker^[21] with an optimized in situ thioesterification (32 %). The glycopeptide **B4**, activated as a benzimidazolone,^[22] was obtained in 26 % yield on a PEGA resin. In all cases, the polar resins were retaining the cleaved glycopeptides and led to low yields. In compounds **B1**, **B3**, and **B4**, Met 48 was replaced by norleucine in order to prevent unwanted oxidation.

With the required fragments **A–C** in hand, we investigated the assembly of the glycosylated IL-6 1–183 polypeptide chain



Scheme 5. Native chemical ligations and refolding of IL-6 GlcNAc **11**.

by sequential native chemical ligation^[10] (Scheme 5). Fragment **C** was ligated with the glycopeptide thioester **B1** (2 equiv) in the presence of MPAA and TCEP. After two days, the ligation did not proceed further, and the thiazolidine of ligation product **9** was converted to a cysteine by methoxyamine/HCl at pH 4. All attempts to isolate intermediate **10** by preparative HPLC were not successful, but the polypeptide could be isolated by TCA precipitation. The ligation of **10** with the thioester **A3** was carried out as above and gave the full-length IL-6 (GlcNAc) **11**, albeit, accompanied by a series of side products. These arose from a missing or doubly inserted fragment **B1** or from a capping reaction that blocked



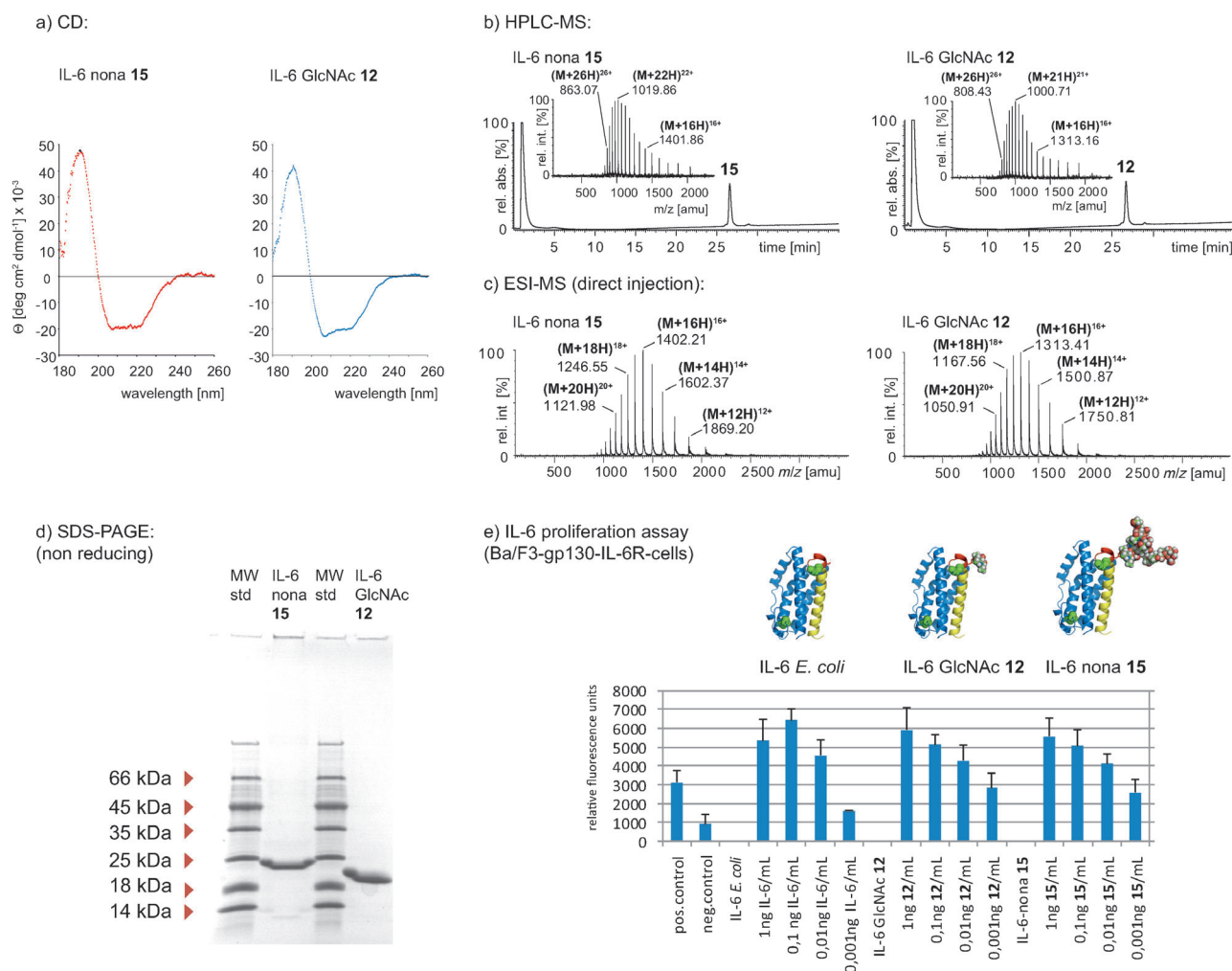
Scheme 4. Solid-phase peptide synthesis (SPPS) of glycopeptides **B1–B4**.

with methoxyamine (74 % yield). The glycopeptide hydrazide **B5** was ligated with the IL-6 thioester **A3** for seven days and the 1–48 hydrazide **13** was obtained in good yield (66 %). According to Liu et al., the C-terminal hydrazide was subsequently converted to an azide by diazotation at -10°C and pH 3 in 6 M GdmCl followed by thiolysis with MPAA **2** at pH 7.0.^[26] HPLC-MS analysis showed the desired MPAA thioester **14**, however, in conjunction with the internal thiolactone **14a** (Figure S37). Purification of the thioesters **14** and **14a** by gel filtration (92 % yield) was essential for the following conversions. The mixture of **14** and **14a** was ligated with segment **C**. After seven days, the crude ligation mixture was refolded by dialysis under reducing conditions followed by oxidative disulfide formation. After gel filtration, the refolded IL-6 **12** was obtained as a homogeneous material in a yield of 2 %.

The yield of refolding was improved to 36 % by a rapid dilution of the crude ligation mixture from 6 M to 0.4 M GdmCl under reducing conditions, followed by gel filtration (Figure S41). The freshly purified IL-6 glycoprotein showed a typical CD spectrum, however, when analyzing this sample by HPLC-MS after one day, an additional isobaric

peak appeared. By reacting a mixture that showed both peaks with the thiol-modifying reagent MMTS,^[27] we found that both species initially contained IL-6 with free thiols. After air oxidation for ten days, disulfide formation was complete. This process also generated small amounts of disulfide-linked IL-6 dimers, which were removed by a final gel filtration (Figure S43). This optimized refolding protocol rendered full-length IL-6 GlcNAc glycoprotein **12** in high purity according to HPLC-MS and nonreducing SDS-PAGE analysis (Scheme 7). The CD spectrum (Scheme 7a) of our synthetic IL-6 GlcNAc was nearly identical to a reference spectrum^[28] from recombinant IL-6 and we thus assumed that the formation of the native four helix bundle fold was achieved. For the final ligation followed by refolding, oxidation, and two gel filtrations, an overall yield of 26 % was obtained, providing 4.3 mg of **12**.

We then applied this approach to another glycoform of IL-6 that carries a complex type biantennary nonasaccharide (Scheme 6b). In analogy to fragment **B5**, the glycopeptide hydrazide **B6** was synthesized from **B4** (90 % yield) and submitted to sequential ligation with fragments **A3** (86 %) and **C**. The refolding and formation of the oxidative disulfide



Scheme 7. Characterization of semisynthetic IL-6: a) CD spectroscopy, b) HPLC-MS, c) ESI-MS direct injection, d) SDS-PAGE analysis, and e) biological activity of semisynthetic IL-6 glycoproteins **12** and **15**.

were carried out as above and gave IL-6 **15** with a biantennary nonasaccharide in 22 % yield (5 mg).

The two refolded IL-6 glycoforms **12** and **15** were characterized by CD, HPLC, MS, and SDS-PAGE analysis. In contrast to HPLC-MS, the folded state of **12** and **15** was observable by direct injection ESI-MS (compare charge distribution in Scheme 7b and c). Additionally, the bioactivity of both cytokines was compared with recombinant IL-6 from *E. coli* in an assay using Ba/F3 cells.^[29] (Scheme 7e). The growth of these cells is dependent on external IL-6 and was assayed in triplicates. By analysis of a series of dilutions, the activity of the two semisynthetic IL-6 glycoforms **12** and **15** was found to be equal to the fully recombinant reference IL-6. The high in vitro activity of the two semisynthetic IL-6 glycoforms further indicated that the correct folding was obtained, as IL-6, lacking one of the two disulfides, already displays low activity.^[30] A key feature of the optimized ligation and refolding scheme is that purification by gel filtration could be applied to all final steps. This led to a high recovery and purity of the semisynthetic IL-6 glycoforms. By varying the short glycopeptide building block **B**, this highly efficient approach should be amenable to the rapid generation of libraries of glycoforms of IL-6.

In summary, we have established a robust semisynthesis for glycoforms of human IL-6. The longest fragments **A** and **C** were obtained recombinantly in high yields. For the stable thioester **A3** bearing a native N-terminus, a two-intein approach was established. By use of glycopeptide hydrazides **B5** and **B6** as latent thioesters, the ligation of the smallest fragments **A** and **B** could be carried out prior to ligation with **C**. This permitted a simplified work-up of each ligation mixture and eliminated side reactions. An efficient refolding was developed that provided rapid access to libraries of glycoforms of human IL-6 in high purity and quantities sufficient for extensive biochemical and biophysical experiments.

Received: July 13, 2014

Published online: September 22, 2014

Keywords: cytokines · glycopeptides · glycoproteins · native chemical ligation · solid-phase synthesis

- [1] C. P. Hackenberger, D. Schwarzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 10030–10074; *Angew. Chem.* **2008**, *120*, 10182–10228.
- [2] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
- [3] a) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131–163; b) L. Liu, C. S. Bennett, C. H. Wong, *Chem. Commun.* **2006**, 21–33; c) Y. Yuan, J. Chen, Q. Wan, R. M. Wilson, S. J. Danishefsky, *Biopolymers* **2010**, *94*, 373–384; d) C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408–4420.
- [4] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705–6710.
- [5] J. Scheller, A. Chalaris, D. Schmidt-Arras, S. Rose-John, *Biochim. Biophys. Acta Mol. Cell Res.* **2011**, *1813*, 878–888.
- [6] L. T. May, J. E. Shaw, A. K. Khanna, J. B. Zabriskie, P. B. Sehgal, *Cytokine* **1991**, *3*, 204–211.
- [7] R. B. Parekh, R. A. Dwek, T. W. Rademacher, G. Opdenakker, J. Van Damme, *Eur. J. Biochem.* **1992**, *203*, 135–141.
- [8] C. S. Dela Cruz, Y. Lee, S. R. Viswanathan, A. S. El-Guindy, J. Gerlach, S. Nikiforow, D. Shedd, L. Gradoville, G. Miller, *J. Exp. Med.* **2004**, *199*, 503–514.
- [9] a) F. Herodin, J. C. Mestries, D. Janodet, S. Martin, J. Mathieu, M. P. Gascon, M. O. Pernin, A. Ythier, *Blood* **1992**, *80*, 688–695; b) J. C. Mestries, E. K. Kruithof, M. P. Gascon, F. Herodin, D. Agay, A. Ythier, *Eur. Cytokine Network* **1994**, *5*, 275–281.
- [10] D. Bang, S. B. Kent, *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538; *Angew. Chem.* **2004**, *116*, 2588–2592.
- [11] a) C. Piontek, D. Varón Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem. Int. Ed.* **2009**, *48*, 1941–1945; *Angew. Chem.* **2009**, *121*, 1974–1978; b) C. Piontek, P. Ring, O. Harjes, C. Heinlein, S. Mezzato, N. Lombana, C. Pöhner, M. Püttner, D. Varón Silva, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem. Int. Ed.* **2009**, *48*, 1936–1940; *Angew. Chem.* **2009**, *121*, 1968–1973.
- [12] a) T. C. Evans, Jr., J. Benner, M. Q. Xu, *J. Biol. Chem.* **1999**, *274*, 18359–18363; b) M. Q. Xu, T. C. Evans, Jr., *Methods* **2001**, *24*, 257–277; c) C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnou, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638–13643.
- [13] R. Rudolph, G. Böhm, H. Lilie, R. Jaenicke in *Protein Function* (Ed.: T. E. Creighton), Oxford University Press, Oxford, **1997**, p. 57.
- [14] E. C. Johnson, S. B. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
- [15] R. J. Payne, S. Ficht, W. A. Greenberg, C. H. Wong, *Angew. Chem. Int. Ed.* **2008**, *47*, 4411–4415; *Angew. Chem.* **2008**, *120*, 4483–4487.
- [16] C. D. Lee, H. C. Sun, S. M. Hu, C. F. Chiu, A. Homhuan, S. M. Liang, C. H. Leng, T. F. Wang, *Protein Sci.* **2008**, *17*, 1241–1248.
- [17] M. Villain, J. Vizzavona, K. Rose, *Environ. Chem. Chem. Biol.* **2001**, *8*, 673–679.
- [18] S. Flemer, Jr., *J. Pept. Sci.* **2009**, *15*, 693–696.
- [19] Y. Kajihara, Y. Suzuki, N. Yamamoto, K. Sasaki, T. Sakakibara, L. R. Juneja, *Chem. Eur. J.* **2004**, *10*, 971–985.
- [20] S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chem. Int. Ed.* **2005**, *44*, 1650–1654; *Angew. Chem.* **2005**, *117*, 1677–1681.
- [21] C. Heinlein, D. Varon Silva, A. Tröster, J. Schmidt, A. Gross, C. Unverzagt, *Angew. Chem. Int. Ed.* **2011**, *50*, 6406–6410; *Angew. Chem.* **2011**, *123*, 6530–6534.
- [22] J. B. Blanco-Canosa, P. E. Dawson, *Angew. Chem. Int. Ed.* **2008**, *47*, 6851–6855; *Angew. Chem.* **2008**, *120*, 6957–6961.
- [23] D. Piszkiwicz, M. Landon, E. L. Smith, *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1173–1178.
- [24] D. Ejima, M. Watanabe, Y. Sato, M. Date, N. Yamada, Y. Takahara, *Biotechnol. Bioeng.* **1999**, *62*, 301–310.
- [25] D. Bang, B. L. Pentelute, S. B. Kent, *Angew. Chem. Int. Ed.* **2006**, *45*, 3985–3988; *Angew. Chem.* **2006**, *118*, 4089–4092.
- [26] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chem. Int. Ed.* **2011**, *50*, 7645–7649; *Angew. Chem.* **2011**, *123*, 7787–7791.
- [27] R. Wynn, F. M. Richards, *Methods Enzymol.* **1995**, *251*, 351–356.
- [28] A. Krüttgen, S. Rose-John, C. Moller, B. Wroblewski, A. Wollmer, J. Mullberg, T. Hirano, T. Kishimoto, P. C. Heinrich, *FEBS Lett.* **1990**, *262*, 323–326.
- [29] K. J. Kallen, J. Grötzinger, E. Lelievre, P. Vollmer, D. Aasland, C. Renne, J. Müllberg, K. H. Meyer zum Büschenfelde, H. Gascan, S. Rose-John, *J. Biol. Chem.* **1999**, *274*, 11859–11867.
- [30] F. L. Rock, X. Li, P. Chong, N. Ida, M. Klein, *Biochemistry* **1994**, *33*, 5146–5154.