Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters 46 (2005) 4197–4200
Tetrahedron Letters

First chemical synthesis of triglucosylated tetradecasaccharide ($Glc_3Man_9GlcNAc_2$), a common precursor of asparagine-linked oligosaccharides

Ichiro Matsuo, Toshinori Kashiwagi, Kiichiro Totani and Yukishige Ito*

RIKEN (The Institute of Physical and Chemical Research), 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan
CREST, JST Kawaguchi 322-1102, Japan

Received 7 March 2005; revised 12 April 2005; accepted 13 April 2005
Available online 29 April 2005

Abstract—Triglucosylated high-mannose-type tetradecasaccharide ($Glc_3Man_9GlcNAc_2$), the oligosaccharide part of the donor substrate of oligosaccharyl transferase (OST) complex, and diglucosylated tridecasaccharide ($Glc_2Man_9GlcNAc_2$) were synthesized. These oligosaccharides were assembled in a convergent and stereoselective manner. Undecasaccharide **5** was employed as the common intermediate, and coupling with trisaccharide (**4**) and disaccharide (**3**) donor afforded fully protected tetradeca-(**17**) and tridecasaccharide (**16**), respectively. These oligosaccharides were deprotected to give $Glc_3Man_9GlcNAc_2$ and $Glc_2Man_9GlcNAc_2$, respectively.

Addition of asparagine (Asn)-linked oligosaccharide (*N*-glycosylation) is an important and highly conserved modification of secretory and membrane-bound eukaryotic proteins.[1] The key step along this pathway is the transfer of tetradecasaccharide ($Glc_3Man_9GlcNAc_2$), which occurs in the lumen of the endoplasmic reticulum (ER).[2] This oligosaccharide is assembled on dolichyl pyrophosphate ($Glc_3Man_9GlcNAc_2$-pp-Dol, **1b**) and transferred en bloc to the Asn-X-Ser/Thr sequence of nascent polypeptide chains (Fig. 1). This transformation is catalyzed by oligosaccharyltransferase (OST), a hetero-oligomeric membrane-bound protein complex.[3] After being transferred, protein-bound $Glc_3Man_9GlcNAc_2$ (**1a**) is trimmed by glucosidase-I (to $Glc_2Man_9GlcNAc_2$, **2a**) and then by glucosidase-II (to $Glc_1Man_9GlcNAc_2$ and $Man_9GlcNAc_2$).

Since $Glc_3Man_9GlcNAc_2$-pp-Dol is the preferred substrate of OST, the presence of glucotriose (Glcα1→2-Glcα1→3Glcα; Glc_3) has been inferred to be important for OST recognition.[4] It has been suggested that this trisaccharide part has a well-defined conformation and, possibly, functions as the recognition epitope.[5] However, due to the poor availability of the full length substrate such as $Glc_3Man_9GlcNAc_2$-pp-Dol,[6,7] most in vitro studies have been carried out with truncated derivatives.[8] Although $Glc_3Man_9GlcNAc_2$ is the key

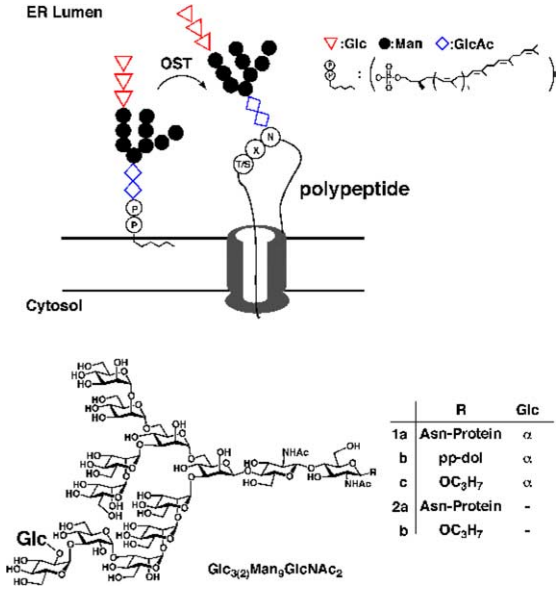


Figure 1. Transfer of tetradecasaccharide from dolichyl pyrophosphate to protein.

Keywords: Oligosaccharyl transferase; Synthesis; N-Glycan.
* Corresponding author. Tel.: +81 48 4467 9430; fax: +81 48 462 4680; e-mail: yukito@postman.riken.go.jp

0040-4039/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2005.04.056

intermediate in glycoprotein biosynthesis, its chemical synthesis has yet to be achieved.

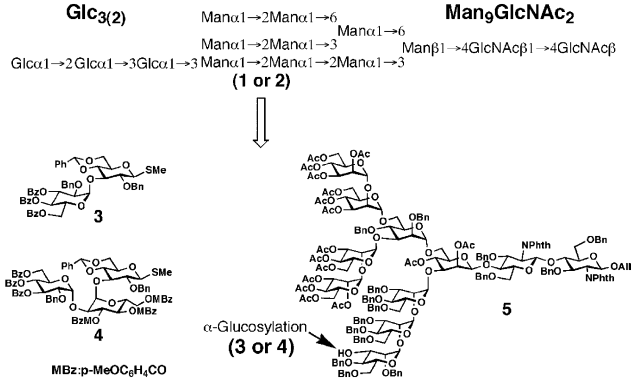
As part of our effort toward comprehensive synthesis of glycan chains[9] related to glycoprotein processing,[10] we wish to report herein the first chemical synthesis of $Glc_3Man_9GlcNAc_2$ (**1c**) as well as its glucosidase-I product $Glc_2Man_9GlcNAc_2$ (**2b**).

The target sugar chains were divided into di/trisaccharide and undecasaccharide fragments as depicted in Scheme 1. Because, we previously observed that, when thioglycoside was used as a donor, the use of 4,6-*O*-benzylidene protected donor was important for α-selective glucosylation at 3-position of mannose,[9a] Glc_2 and Glc_3 donors were designed as **3** and **4**, respectively. Synthesis of the undecasaccharide fragment **5** was described previously,[9a] which can now be performed in multigram scale.

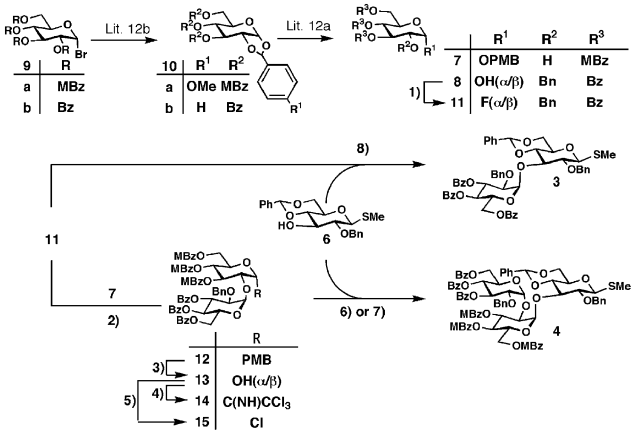
Preparation of **3** and **4** was conducted as shown in Scheme 2. Previously reported synthesis of the Glc_3 sequence either required a large excess of glucosyl bromide or resulted in incomplete selectivity.[7] We planned to use 3,4,6-*O*-acylated glucosyl donors (**11**, **14/15**) for stepwise elongation. It was envisaged that these donors were sufficiently deactivated to exhibit the desired α-selectivity[11] and to minimize premature decomposition during activation. These components were prepared from 2-OH derivative **7** and hemiacetal **8**, which in turn were prepared as described by Suzuki et al. from glucosyl bromides **9a/b** via cyclic acetals **10a/b**.[12] Hemiacetal **8** was then converted to the fluoride **11** using diethylaminosulfur trifluoride (DAST) in 92% yield. Glycosylation of **7** with **11** (2 equiv) using Cp_2HfCl_2 and AgOTf[13] in CH_2Cl_2 afforded disaccharide **12** as a single isomer in 90% yield.[14] Removal of the *p*-methoxybenzyl (PMB) group using DDQ afforded hemiacetal **13** (73%) that was converted to imidate **14** (quant.) and chloride **15** (quant.). When imidate **14** (1.3 equiv) was used as the glycosyl donor[15] to react with **6**, under activation with TMSOTf in toluene, trisaccharide **4** was obtained in 45% yield as a single stereoisomer.[14] The same product was obtained with chloride **15** (1.5 equiv) in slightly higher yield, however, with reduced selectivity (57%, α:β = 8.5:1). On the other hand, coupling of fluoride **11** (1.5 equiv) with **6** provided 63% yield of disaccharide **3** as a pure α-siomer [δ_H 5.73 (d, J 3.60 Hz); δ_C 95.56 ($^1J_{C-H}$ 172 Hz, C-1)].[14]

Although yields of glycosylation steps toward **3** and **4** were modest, these di/triglucosyl fragments were obtained with high stereochemical purity and can be used directly for the next coupling without further manipulation of the anomeric position.

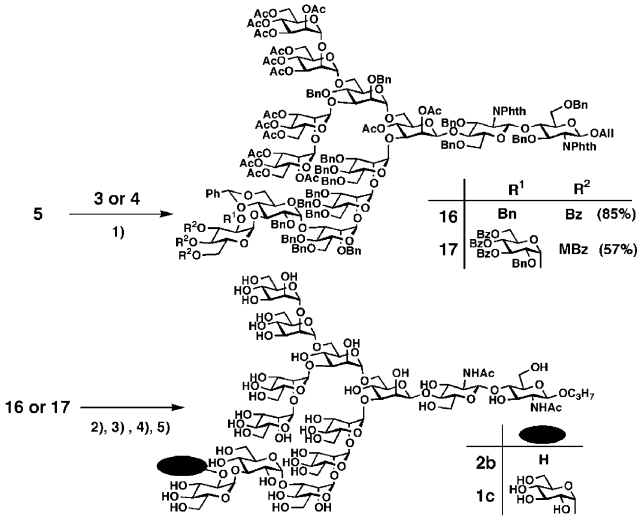
With all oligosaccharide components in hand, fragment condensation was undertaken as shown in Scheme 3. Thus, tridecasaccharide skeleton was constructed using **3** (1.7 equiv) and **5**, through thioglycoside activation



Scheme 1. Oligosaccharide fragments **3** and **4** for the synthesis of $Glc_3Man_9GlcNAc_2$ **1** and $Glc_2Man_9GlcNAc_2$ **2**.



Scheme 2. Reagents and conditions: (1) DAST, CH_2Cl_2, −40 °C, 3 h, 92%; (2) Cp_2HfCl_2, AgOTf, MS 4A, CH_2Cl_2, −40 °C, 7 h, 90%; (3) DDQ, CH_2Cl_2/H_2O, rt, 12 h, 73%; (4) CCl_3CN, DBU, 0 °C, 1 h, quant.; (5) $SOCl_2$, DMF (cat.), CH_2Cl_2, 0 °C, 4 h, quant.; (6) **14**, TMSOTf, MS-AW300, toluene, −40 °C, 2 h, 45%; (7) **15**, AgOTf, di-*tert*-butyl-4-methylpyridine (DTBMP), MS 4A, toluene/$ClCH_2CH_2Cl$ = 2:1, −20 °C–rt, 12 h, 51% (β-isomer 6%); (8) Cp_2HfCl_2, AgOTf, MS 4A, ether/toluene = 2:1, −40 °C, 7 h, 63%.



Scheme 3. Reagents and conditions: (1) MeOTf, MS 4A, DTBMP, $ClCH_2CH_2Cl$/cyclohexane = 1:4, 50 °C, 12 h, 85% (**14**), 57% (**15**; α:β = 8.5:1); (2) ethylenediamine, *n*-BuOH, 90 °C; (3) Ac_2O, pyridine, DMAP, rt; (4) NaOMe/MeOH, 40 °C; (5) $Pd(OH)_2$, H_2, MeOH/$AcOH/H_2O$ = 5:1:1, rt.

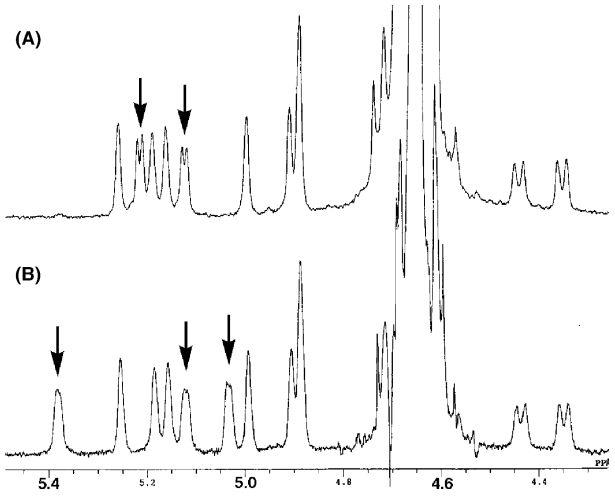


Figure 2. ^{1}H NMR spectra (400 MHz, D_2O, 23 °C, referenced to HOD adjusted to 4.65 ppm) of tridecasaccharide **2b** (A) and tetradecasaccharide **1c** (B). Anomeric signals derived from α-Glc are indicated by arrows.

by MeOTf[16] to afford **16** as a single isomer in 85% yield. On the other hand, glycosylation of **5** with trisaccharide donor **4** (1.6 equiv) gave 57% yield of tetradecasaccharide **17**, which, unfortunately, contaminated with a detectable amount of the β-isomer (α/β = 19:1). Since no chromatographic conditions were found to separate these isomers in a preparative scale, the mixture was carried on further. Complete deprotection of **16** was performed as follows: (1) conversion of phthalimide groups to acetamide, (2) $Pd(OH)_2$ catalyzed debenzylation, and (3) deacetylation, to give tridecasaccharide **2b**[17] in 48% overall yield. The presence of 2 α-linked Glc was confirmed by ^{1}H NMR [δ_H 5.212 (d, J 3.66 Hz), 5.121 (d, J 3.66 Hz)]. Tetradecasaccharide **17** was deprotected in the same manner, and final purification by HPLC using Fluorex column (conditions; 20ϕ × 250 mm, eluent; H_2O, flow rate: 10 ml/min, detection UV 214 nm) allowed the isolation of anomerically pure **1c** [δ_H 5.381 (d, J 2.2 Hz), 5.122 (d, J = 3.17 Hz, 1H), 5.034 (d, J = 2.9 Hz, 1H)] in 56% yield.[18] The ^{1}H NMR spectrum of **1c** and **2b** were in excellent agreement with those reported for closely related compounds (Fig. 2).[19]

In conclusion, convergent and stereoselective synthetic routes to $Glc_3Man_9GlcNAc_2$ (**1c**) and $Glc_2Man_9GlcNAc_2$ (**2b**) were established. These oligosaccharides will be valuable standards to reveal protein–oligosaccharide interactions involved in glycoprotein biosynthesis. Future studies are going to be directed to the development of practical methodology to convert them to oligosaccharide–dolichyl pyrophosphate conjugates.

Acknowledgements

We thank Ms. Akemi Takahashi for technical assistance. Financial support from Ministry of Education, Culture, Sports, Science and Technology [Grant-in-Aid for Scientific Research (B) No. 16390011] is acknowledged.

References and notes

1. Trombetta, E. S. *Glycobiology* **2003**, *13*, 77R–91R.
2. Burda, P.; Aebi, M. *Biochim. Biophys. Acta* **1999**, *1426*, 239–257.
3. Knauer, R.; Lehle, L. *Biochim. Biophys. Acta* **1999**, *1426*, 259–273.
4. Breuer, W.; Bause, E. *Eur. J. Biochem.* **1995**, *228*, 689–696.
5. (a) Petrescu, A. J.; Butters, T. D.; Reinkensmeier, G.; Petrescu, S.; Platt, F. M.; Dwek, R. A.; Wormald, M. R. *EMBO J.* **1997**, *16*, 4302–4310; (b) Alvarado, E.; Nukada, T.; Ogawa, T.; Ballou, C. E. *Biochemistry* **1991**, *30*, 881–886.
6. Preparation of G3M9GN2-pp-dol from yeast: Kelleher, D. J.; Karaoglu, D.; Gilmore, R. *Glycobiology* **2001**, *11*, 321–333.
7. Synthesis of partial structures: (a) Ennis, S. C.; Cumpstey, I.; Fairbanks, A. J.; Butters, T. D.; Mackeen, M.; Wormald, M. R. *Tetrahedron* **2002**, *58*, 9403–9411; (b) Scaman, C. H.; Hindsgaul, O.; Palcic, M. M.; Srivastava, O. P. *Carbohydr. Res.* **1996**, *296*, 203–213; (c) Ogawa, T.; Nukada, T.; Kitajima, T. *Carbohydr. Res.* **1983**, *123*, C12–C15.
8. (a) Imperiali, B.; Shannon, K. L.; Rickert, K. W. *J. Am. Chem. Soc.* **1992**, *114*, 7942–7944; Imperiali, B.; Shannon, K. L.; Unno, M.; Rickert, K. W. *J. Am. Chem. Soc.* **1992**, *114*, 7944–7945; (b) Lee, J.; Coward, J. K. *J. Org. Chem.* **1992**, *57*, 4126–4135.
9. (a) Matsuo, I.; Wada, M.; Manabe, S.; Yamaguchi, Y.; Otake, K.; Kato, K.; Ito, Y. *J. Am. Chem. Soc.* **2003**, *125*, 3402–3403; (b) Matsuo, I.; Ito, Y. *Carbohydr. Res.* **2003**, *338*, 2163–2168; (c) Takatani, M.; Matsuo, I.; Ito, Y. *Carbohydr. Res.* **2003**, *338*, 1073–1081; (d) Totani, K.; Matsuo, I.; Takatani, M.; Arai, M. A.; Hagihara, S.; Ito, Y. *Glycocojugate J.* **2004**, *21*, 67–74.
10. (a) Spiro, R. G. *Cell. Mol. Life Sci.* **2004**, *61*, 1025–1041; (b) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364–2369; (c) Ellgaard, L.; Helenius, A. *Nature Rev. Mol. Cell Biol.* **2003**, *4*, 181–191.
11. Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 155–173; Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B. *J. Am. Chem. Soc.* **1988**, *110*, 5583–5584.
12. (a) Suzuki, K.; Mizuta, T.; Yamaura, M. *J. Carbohydr. Chem.* **2003**, *22*, 143–147; (b) Suzuki, K.; Nonaka, H.; Yamaura, M. *Tetrahedron Lett.* **2003**, *44*, 1975–1977.
13. Suzuki, K.; Maeta, H.; Matsumoto, T. *Tetrahedron Lett.* **1989**, *30*, 4853–4854.
14. **12**: ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.03 (t, J 9.6 Hz, 1H), 5.82 (t, J 9.6 Hz, 1H), 5.41 (t, J 9.6 Hz, 1H,), 5.29 (t, J 9.6 Hz, 1H), 5.14 (d, J 3.6 Hz, 1H, H-1$^{\alpha\text{-Glc}}$), 4.93 (d, J 3.2 Hz, 1H, H-1$^{\alpha\text{-Glc}}$); ^{13}C NMR ($CDCl_3$) δ 96.91 ($^{1}J_{C-H}$ 168 Hz, C-1$^{\alpha\text{-Glc}}$), 94.98 ($^{1}J_{C-H}$ 170 Hz, C-1$^{\alpha\text{-Glc}}$); MS (MALDI-TOF) $C_{72}H_{66}O_{21}Na$: Calcd 1289.1. Found: 1290.3 $(M+Na)^+$. **4**: ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.00–6.63 (m, Ar), 6.14 (t, J 10 Hz, 1H) 5.86 (t, J 10 Hz, 1H), 5.82 (d, J 3.6 Hz, 1H, H-1$^{\alpha\text{-Glc}}$), 5.45, (t, J 10 Hz, 1H), 5.40 (s, 1H, PhC*H*), 5.27, (t, J 10 Hz, 1H), 4.93 (d, J 3.2 Hz, 1H, H-1$^{\alpha\text{-Glc}}$), 3.83, 3.81, 3.50 (s × 3, OMe), 2.28 (s, 3H, SMe); ^{13}C NMR (100 MHz, $CDCl_3$) δ 95.95 ($^{1}J_{C-H}$ 173 Hz, C-1$^{\alpha\text{-Glc}}$), 95.03 ($^{1}J_{C-H}$ 175 Hz, C-1$^{\alpha\text{-Glc}}$), 86.05 ($^{1}J_{C-H}$ 154 Hz, C-1$^{\beta\text{-Glc}}$); MS (MALDI-TOF) $C_{85}H_{80}O_{24}Na$: Calcd 1540.5. Found: 1540.2 $(M+Na)^+$. **3**: ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02–6.69 (m, Ar), 6.00 (t, J 9.6 Hz, 1H), 5.73 (d, J 3.60 Hz, 1H, H-1$^{\alpha\text{-Glc}}$), 5.47 (s, 1H, PhC*H*), 5.41 (t, J 9.9 Hz, 1H), 4.51 (d, J 9.6 Hz, 1H, H-1$^{\beta\text{-Glc}}$), 2.29 (s, 3H, SMe); ^{13}C NMR ($CDCl_3$) δ 102.21, 95.56 ($^{1}J_{C-H}$ 172 Hz, C-1$^{\alpha\text{-Glc}}$), 86.09 ($^{1}J_{C-H}$ 158 Hz, C-1$^{\beta\text{-Glc}}$); MS (MALDI-TOF) $C_{55}H_{52}O_{13}NaS$: Calcd 975.3. Found: 975.7 $(M+Na)^+$.
15. Schmidt, R. R.; Kinzy, W. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 21–123.

16. Lönn, H. *J. Carbohydr. Chem.* **1987**, *6*, 301–306.
17. **2b**: ^{1}H NMR (400 MHz, D_2O) δ 5.256 (s, 1H, H-1^{Man}), 5.212 (d, *J* 3.66 Hz, 1H, H-1^{Glc}), 5.186 (s, 1H, H-1^{Man}), 5.159 (s, 1H, H-1^{Man}), 5.121 (d, *J* 3.66 Hz, 1H, H-1^{Glc}), 4.994 (s, 1H, H-1^{Man}), 4.906 (s, 1H, H-1^{Man}), 4.888 (s, 2H, H-1^{Man}), 4.439 (d, *J* 7.0 Hz, 1H, H-1^{GlcNAc}), 4.350 (d, *J* 7.0 Hz, 1H, H-1^{GlcNAc}), 1.917 (s, 3H, Ac), 1.876 (s, 3H, Ac), 1.391 (m, 2H), 0.709 (t, 3H); MS (MALDI-TOF) $C_{91}H_{154}N_2O_{71}Na$: Calcd 2433.7. Found: 2432.8 $(M+Na)^+$.
18. **1c**: ^{1}H NMR (400 MHz, D_2O) δ 5.381 (d, *J* 2.2 Hz, 1H, H-1^{Glc}), 5.254 (s, 1H, H-1^{Man}) 5.183 (s, 1H, H-1^{Man}), 5.156 (s, 1H, H-1^{Man}), 5.122 (d, *J* 3.2 Hz, 1H, H-1^{Glc}), 5.034 (d, *J* 2.9 Hz, 1H, H-1^{Glc}), 4.992 (s, 1H, H-1^{Man}), 4.904 (s, 1H, H-1^{Man}), 4.886 (s, 2H, H-1^{Man}), 4.436 (d, *J* 7.3 Hz, 1H, H-1^{GlcNAc}), 4.350 (d, *J* 7.3 Hz, 1H, H-1^{GlcNAc}), 1.919 (s, 3H, Ac), 1.878 (s, 3H, Ac), 1.394 (m, 2H), 0.711 (t, 7.07 Hz, 3H); MS (MALDI-TOF) $C_{85}H_{144}N_2O_{66}Na$: Calcd 2272.9. Found: 2273.6 $(M+Na)^+$.
19. (a) Wormald, M. R.; Wooten, E. W.; Bazzo, R.; Edge, C. J.; Feinstein, A.; Rademacher, T. W.; Dwek, R. A. *Eur. J. Biochem.* **1991**, *198*, 131–139; (b) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Eur. J. Biochem.* **1998**, *258*, 372–386.