Generation of oligosaccharide and glycoconjugate libraries for drug discovery

Michael J. Sofia

Oligosaccharides play many important roles in biological processes and in pathology. Combinatorial library technologies for the rapid identification of potential oligosaccharide and glycoconjugate drugs could provide valuable tools for drug discovery, although rapid and efficient access to library constituents would be required. Developments in the solid-phase synthesis of oligosaccharides and glycoconjugates will provide the tools for medicinal chemists to generate combinatorial libraries of these complex systems in the future.

Oligosaccharides play many important roles in biological processes[1–3], including involvement in the processes of cellular trafficking, cellular differentiation, cell–cell adhesion, hormone–cell recognition, and viral–host cell[4] and bacterial–host cell[5] attachment. Consequently, their involvement is implicated in many diseases, such as cancer, inflammatory disorders, cardiovascular diseases, microbial infections and lysosomal storage diseases[1]. Saccharides have long been known to be important constituents of medicinally important agents, such as the antitumor antibiotics, cardiac glycosides, and gangliosides. For example, the oligosaccharide moieties of the antitumor antibiotics ciclamycin and calicheamicin are important in the DNA recognition processes critical for the activity of the compounds. Cardiac glycosides, such as digoxin, are steroids conjugated to a trisaccharide. Gangliosides, lipid–oligosaccharide conjugates, have been used to treat spinal cord injury. Attachment of carbohydrates to peptides to form neoglycoconjugates helps to stabilize the peptide against degradation and occasionally facilitates transport of the peptide across biological barriers[6–8]. The use of saccharides as rigid scaffolding for presenting key functional elements in a defined arrangement for the identification of new receptor ligands is another area of great promise[9–11]. Access to the molecular diversity of carbohydrates would, therefore, be valuable for drug discovery in the areas of biologically significant oligosaccharides and analogs, glycoconjugates and analogs, neoglycoconjugates and carbohydrate molecular scaffolding.

Combinatorial library technology as a new paradigm for accelerating the identification of novel lead structures by rapidly generating vast chemical diversity has gained wide acceptance in the drug discovery community[12,13]. The rapid synthesis of many compounds, either as mixtures or in a spatially arrayed format, has depended on the application of reliable chemistry, either on solid phase or in solution[14–16]. Coupled with novel deconvolution strategies and microsequencing techniques that enable the identification of the biologically active members of a mixture or an array, these developments in chemical library generation have significantly expanded the capabilities of medicinal chemists.

The new techniques have, however, been primarily applied to the generation of peptide[17,18] and oligonucleotide[19] libraries and a few libraries of small molecules[20–25]. Meanwhile, the application of library technology to the synthesis of oligosaccharides and glycoconjugates has largely been ignored. Primary reasons include the complexity of the monosaccharide building block involving multiple reactive sites, the lack

Michael J. Sofia, Transcell Technologies Inc., 2000 Cornwall Rd, Monmouth JCT, NJ 08852, USA. tel: +1 908 940 6909, fax: +1 908 821 4470

of a single universal reaction for forming glycosidic bonds that can be applied reliably to a wide variety of substrates, either in solution or on the solid phase, and the absence of well-developed solid-phase technology that is applicable to the synthesis of oligosaccharides or glycoconjugates using combinatorial approaches. Other issues to be addressed include methods and strategies for deconvoluting complex mixtures of oligosaccharides or glycoconjugates. Microsequencing techniques and amplification technologies similar to those used for peptides and oligonucleotides are not available, and library-encoding methods being developed for some small molecule libraries have not yet been applied to such complex molecules as oligosaccharides or glycoconjugates[26–32]. Of all of these issues, however, the most significant is the ability to prepare the library constituents rapidly, efficiently and with sufficient purity.

Solid-phase synthesis of oligosaccharides

Much effort has been directed towards the development of technologies such as solid-phase oligosaccharide synthesis that are suitable for rapid and efficient generation of the constituents of a carbohydrate-based library. The attractive aspects of solid-phase synthesis include the use of excess reagents to drive the reactions to completion and therefore obtain high chemical yields, and the elimination of tedious work-up and purification steps. Compared with the solid-phase synthesis of peptides and oligonucleotides, however, the solid-phase synthesis of oligosaccharides is hampered by both stereochemical and concentrated functional group complexity. The stereochemistry of a glycosidic linkage can be either α (axial) or β (equatorial), and the presence of multiple reactive sites on a monosaccharide unit complicates molecular construction because of the need to ensure site-selective glycosylation. Oligosaccharide construction is different from that of peptides and oligonucleotides in the need to access both linear and branched systems and to construct glycosidic linkages to sites with widely different reactivities.

Solid-phase synthesis of oligosaccharides using existing glycosylation chemistries was first attempted approximately 25 years ago (Figure 1)[33–37]. These attempts demonstrated that it is possible to attach a carbohydrate to a polymer support, then glycosylate the polymer-bound monosaccharide and

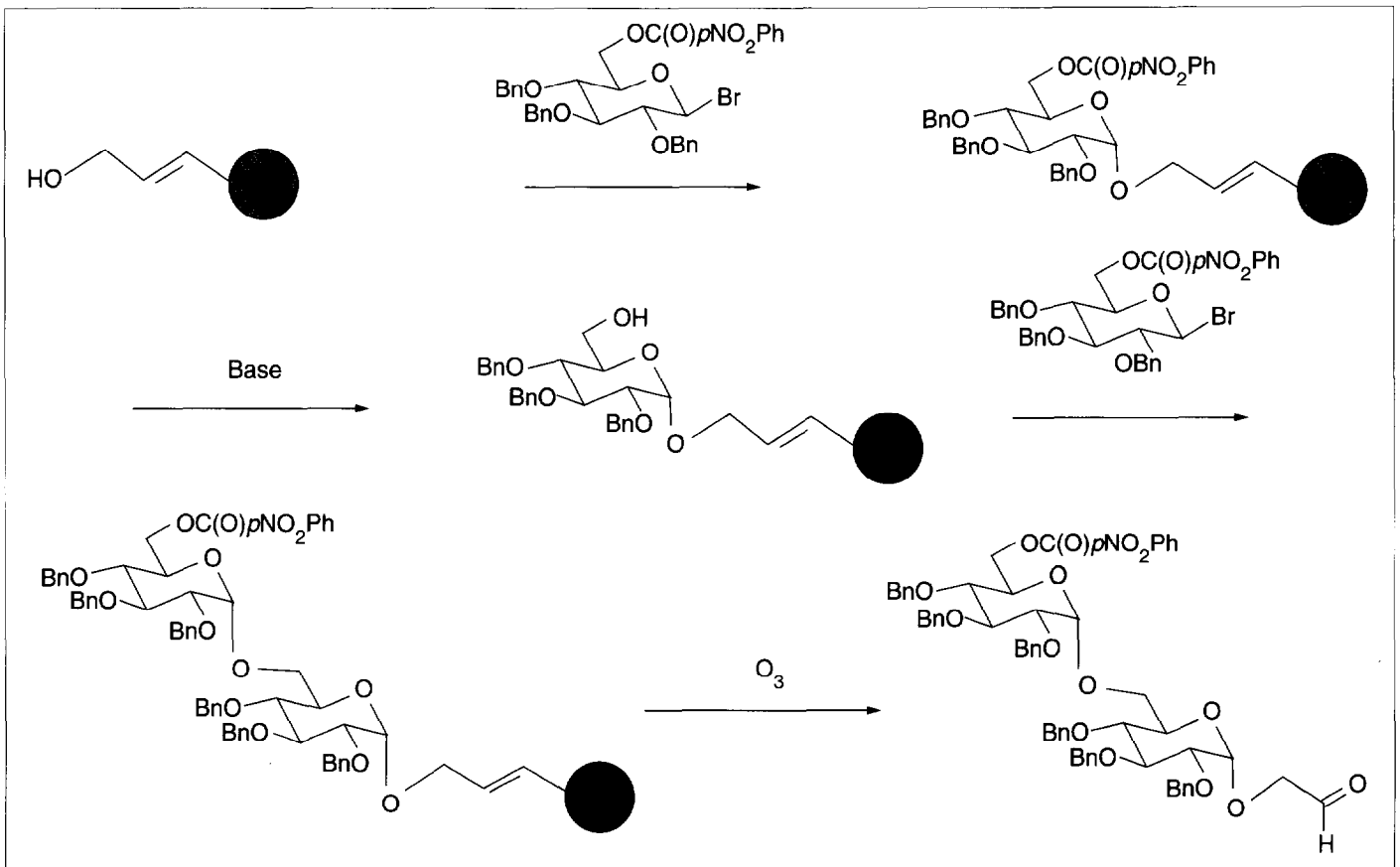

Figure 1. *The first reported example of solid-phase synthesis of an oligosaccharide. The glycosidic bond is formed by alcoholysis of an anomeric bromide glycosyl donor. The polymeric support is polystyrene divinylbenzene copolymer.*

remove the product from the solid support. Solid-supported synthesis of oligosaccharide-like systems, in which the monosaccharides were linked by phosphate units, was also reported by different authors, but in these cases the difficult-to-form glycosidic linkage was replaced by the well-developed phosphodiester bond construction (Figure 2)[38–40].

Recently, new efforts directed toward solid-phase synthesis of oligosaccharides are providing solutions. Application not only of traditional solid supports, such as polystyrene resins and controlled-pore glass, but also of soluble polymeric systems based on polyethyleneglycol are now being used for polymer-supported oligosaccharide synthesis (Figure 3)[41–44]. Although the application of more traditional glycosylation chemistries is being investigated for polymer-supported synthesis[45,46], it is the development of new and efficient glycosylation technologies that increasingly appears to provide a solution to the problem of solid-phase carbohydrate chemistry.

Solid-phase technologies that show promise for potential practical application include those based on glycosylsulfoxide (Figure 4)[47,48], glycal (Figure 5)[49–51], trichloromethylimidate (Figure 3)[41,42,44,45] and enzymatic (Figure 6)[52,53] glycosylation technologies. These technologies for solid-phase oligosaccharide synthesis hold great promise for application to the generation of oligosaccharide and glycoconjugate libraries, but many issues remain unresolved[54]. The ideal strategy for solid-phase synthesis of oligosaccharides would, for example, use a single glycosyl donor technology to construct each glycosidic linkage, irrespective of whether it is to a primary or secondary alcohol or of α or β stereochemistry at the anomeric center. The single donor approach reduces the number of glycosyl donors and simplifies the optimization of the chemistries required to synthesize an oligosaccharide library on the solid phase. This ability has only been demonstated for the glycosylsulfoxide-based solid-phase method[47]. In addition, even under optimized solution conditions, the formation of α linkages to sialic acid and β linkages to mannose is very difficult and still remains a serious challenge for current solid-phase methodologies. Also, because saccharides are polyfunctional molecules with multiple potential glycosylation sites and so a unified strategy for protecting functional groups is needed for use with any nonenzymatic solid-phase approach. Even with present limitations, however, it is possible to generate oligosaccharide libraries with considerable molecular diversity.

Solid-phase synthesis of glycoconjugates

Solid-phase preparation of glycoconjugates has focused primarily on glycopeptides, because of their immense biological

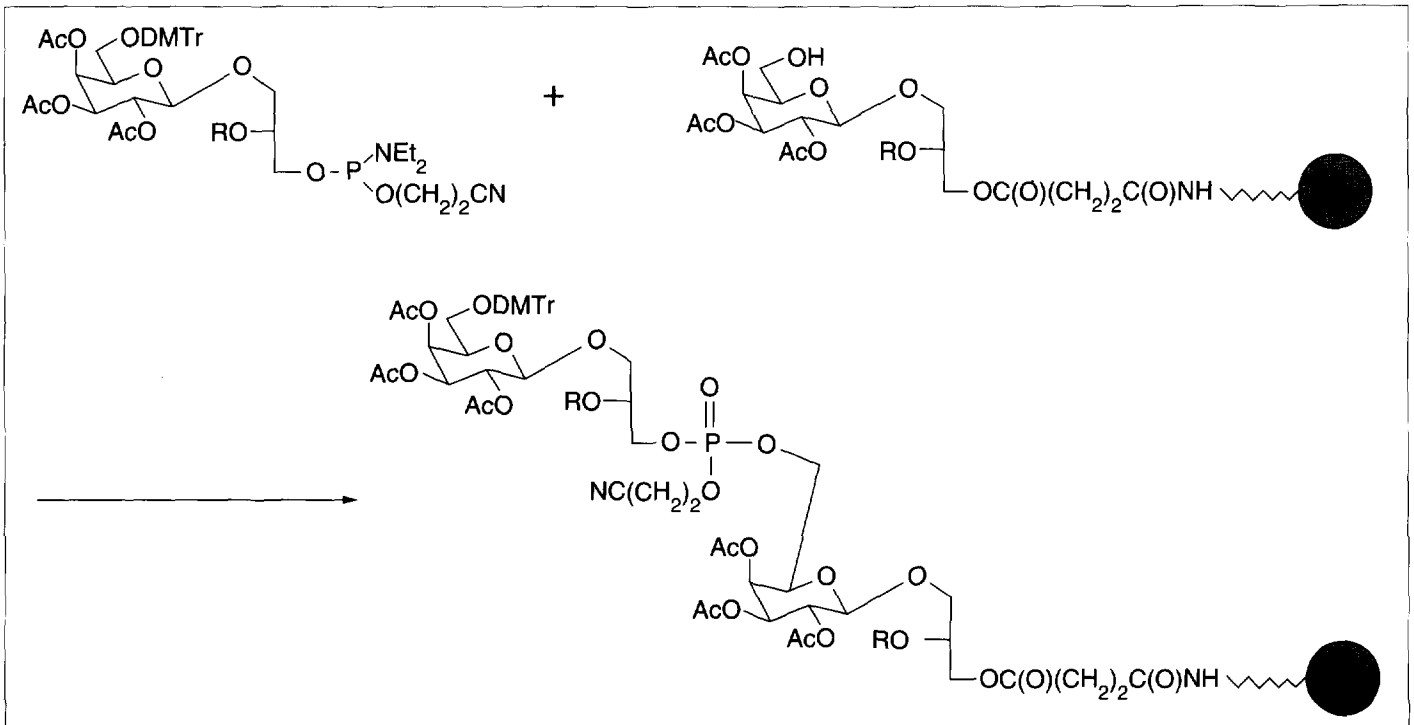

Figure 2. *Synthesis of teichoic acid substructure using solid-phase synthesis – an example of solid-phase synthesis in which the building blocks consist of monosaccharide substructures linked by phosphodiester bonds. R, BOM.*

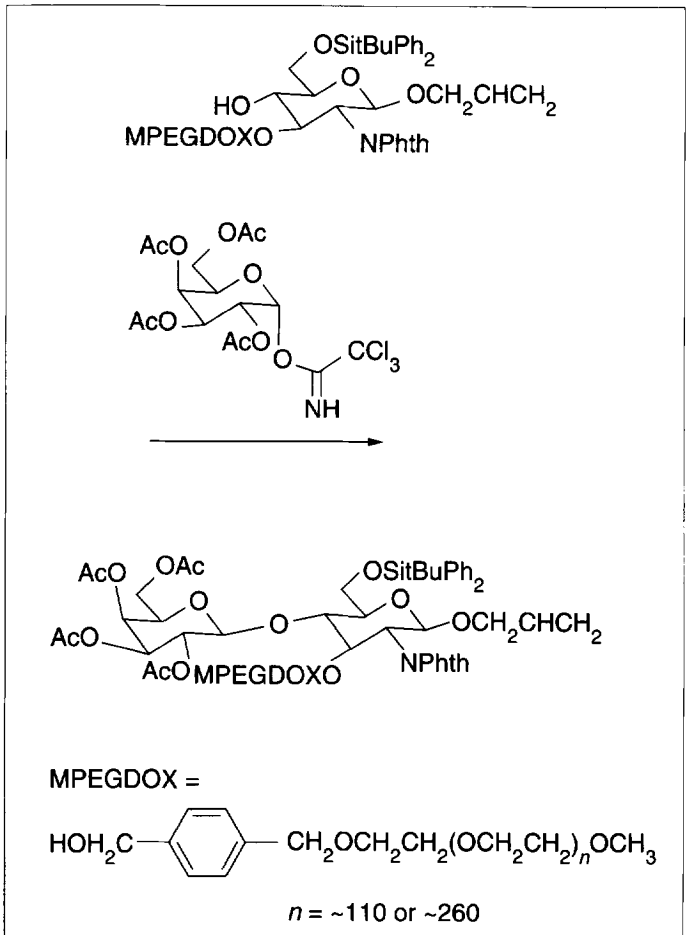

Figure 3. *Glycosylation using a soluble polymer-bound glycosyl acceptor and trichloromethylimidate glycosyl donor technology. The reactants are completely soluble in the reaction mixture and the polymer-bound product is precipitated when the reaction is complete, thus facilitating product purification.*

significance[55–59]. Solid-phase *O*- or *N*-linked glycopeptide synthesis has relied on two approaches, the 'convergent' or 'building blocks' methods (Figures 7 and 8)[60–62]. Typically, however, in neither approach is glycosylation performed on the solid phase. Each approach takes advantage of the well-established amide-bond-forming reaction to build the glycopeptide. The 'convergent' approach uses a C-1 amino-derivatized saccharide and constructs an amide bond to an activated carboxylate group of a resin-bound peptide[63]. It has been demonstrated that the use of the convergent approach for the formation of an amide linkage between a C-1 amino-functionalized saccharide and an active ester of the peptide does not require the use of hydroxyl-protecting groups on the saccharide unit (Figure 9)[64,65]. In another example of the use of the convergent strategy, the oligosaccharide was constructed on the solid phase and coupled to a peptide via a carboxylate group on the peptide and the C-1 amino group of the polymer-bound oligosaccharide[66].

The 'building blocks' approach requires the synthesis of the desired *N*- or *O*-linked glycosylated and appropriately protected amino acids, which are then used in standard solid-phase peptide synthesis (Figure 10)[67–78]. This approach has been used extensively to build glycopeptides in which the position of the saccharide unit varies within the peptide sequence.

Enzyme technology has also been applied to the solid-phase synthesis of glycopeptides. The use of enzymes to construct glycopeptides has focused primarily on the use of glycosyltransferases to extend the saccharide unit from a preformed glycopeptide fragment (Figure 11)[52,79,80]. Advantages associated with enzyme technology include the ability to

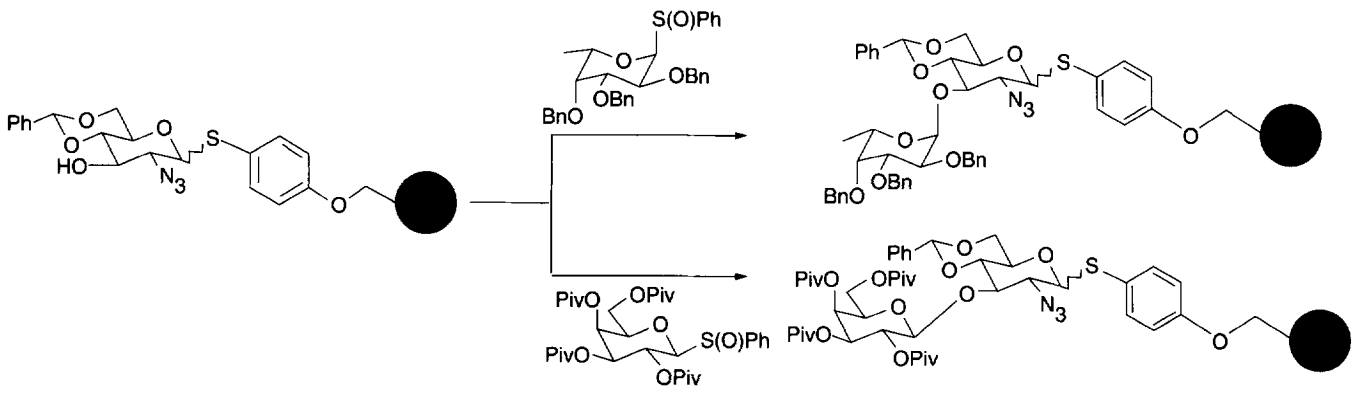

Figure 4. *Solid-phase synthesis of oligosaccharides using glycosylsulfoxide donor technology. The polymer support is Merrifield resin, a polystyrene divinylbenzene copolymer.*

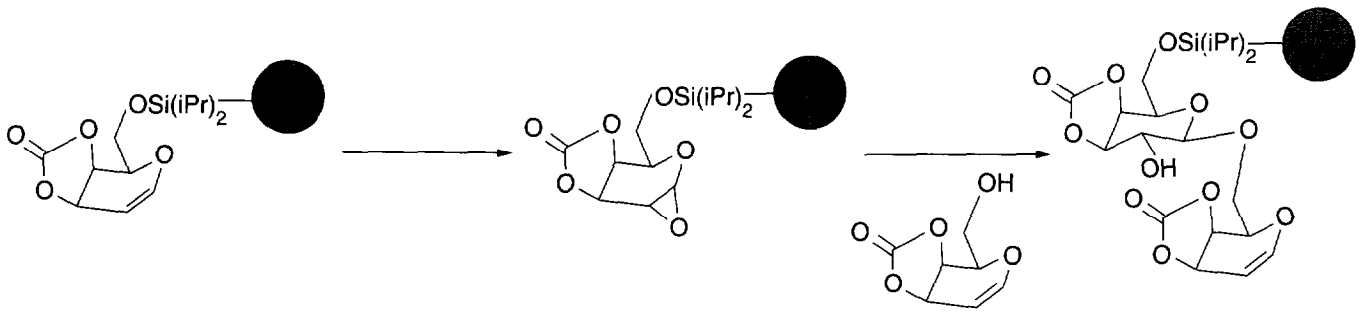

Figure 5. *Solid-phase synthesis of oligosaccharides using glycal-based donor technology. The polymer support is polystyrene divinylbenzene copolymer.*

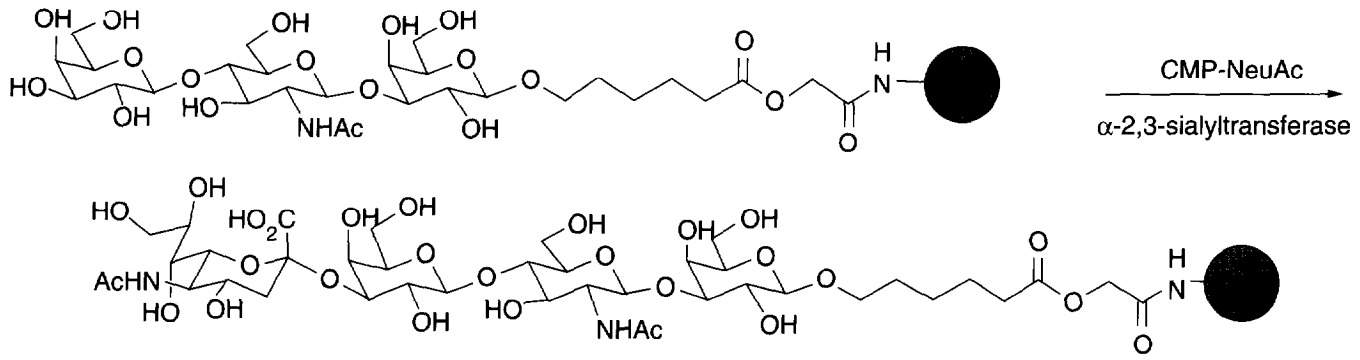

Figure 6. *Solid-phase synthesis of oligosaccharides using enzyme-based glycosylation technology. The polymer support is controlled-pore glass.*

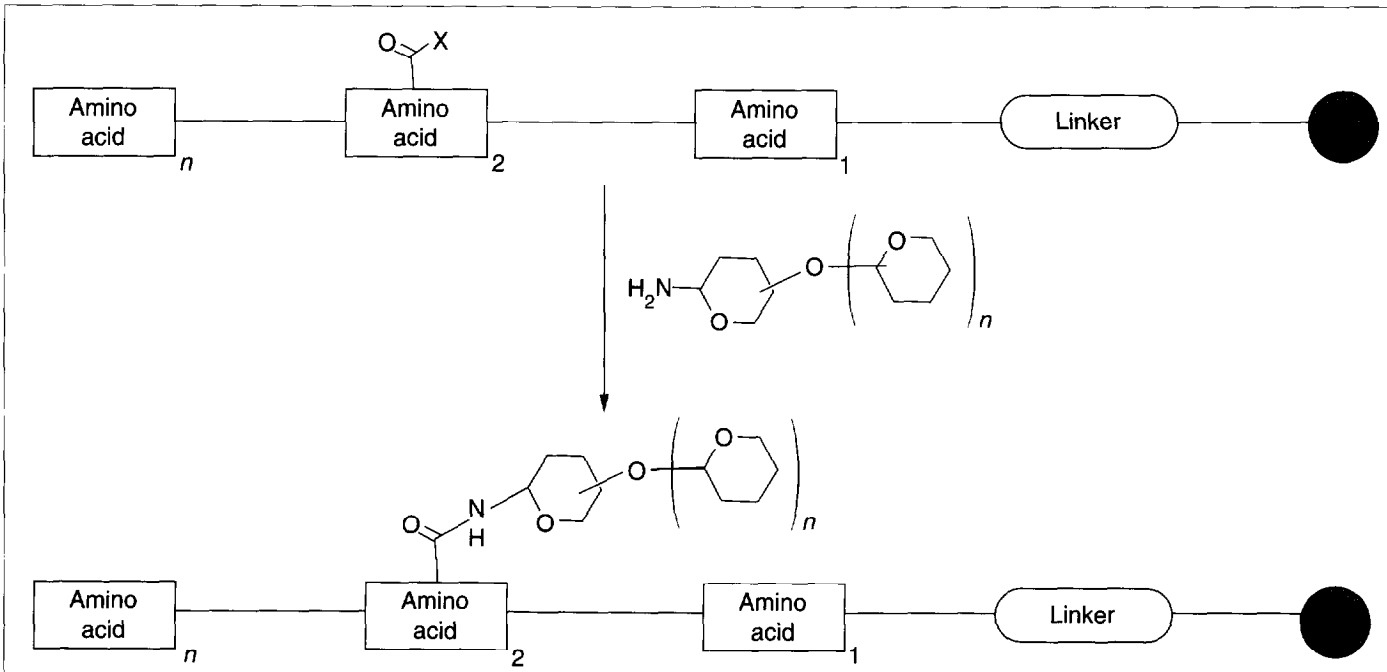

Figure 7. *The convergent approach for the construction of glycopeptides on the solid phase. Typically, the amino acid chain is built on the polymer support, and the presynthesized saccharide unit is then coupled to the resin-bound peptide by an amide bond employing the C-1 amino derivative of the saccharide. X, Acid activating group.*

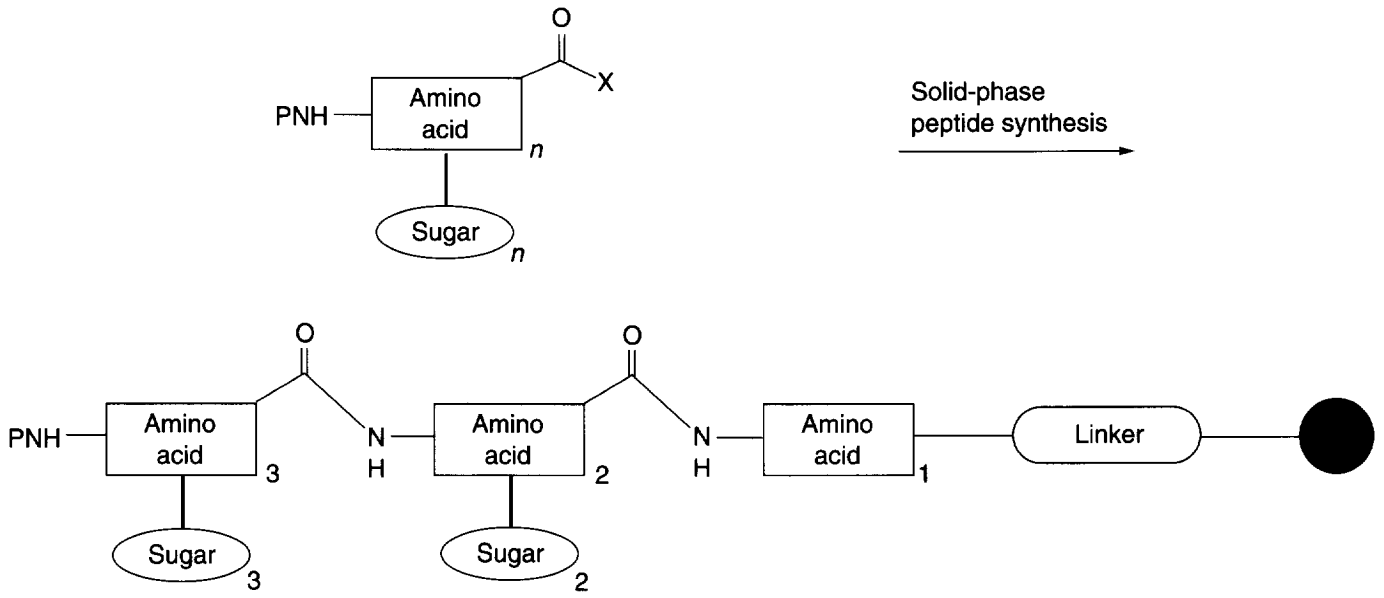

Figure 8. *The building blocks approach for the construction of a glycopeptide on the solid phase. Typically, a monosaccharide or oligosaccharide is attached to a single amino acid in which the α-amino group is protected. The carboxylic acid functionality is then converted to an active ester followed by use as a building block in solid-phase peptide synthesis. X, acid activating group; P, amine protecting group.*

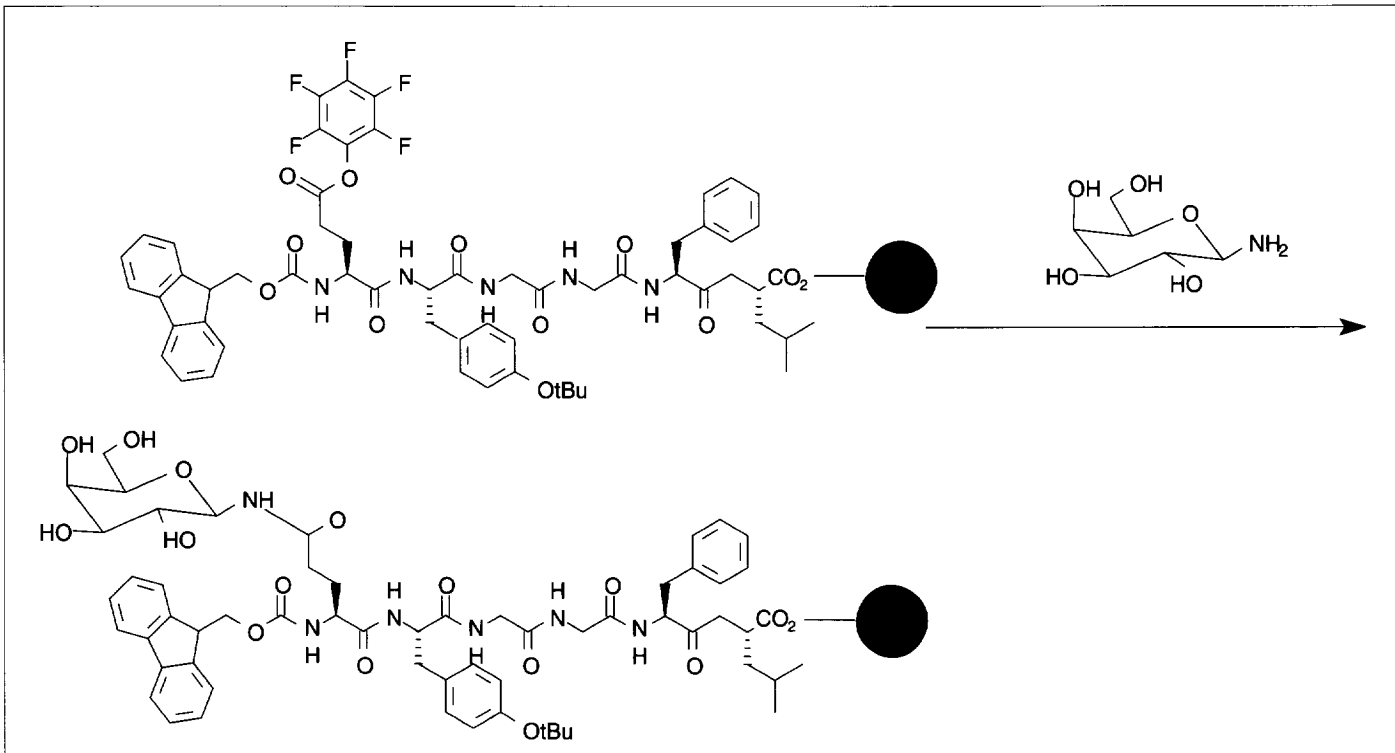

Figure 9. *Application of the convergent strategy for solid-phase glycopeptide synthesis. An unprotected C-1 amino-functionalized monosaccharide or oligosaccharide is coupled to a glutamic acid residue. The polymer support is a polystyrene resin with a highly acid-labile linker (super acid-sensitive resin; SASRIN).*

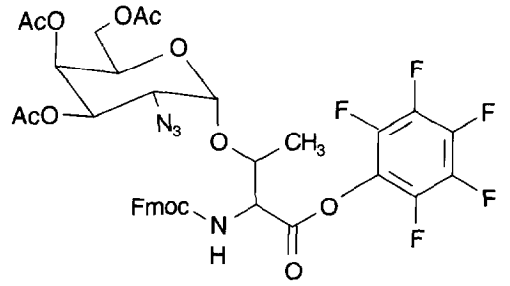

Figure 10. *Typical example of a 'building block' for glycopeptide synthesis. Use of pentafluorophenyl esters as activated forms of the acid terminus is a key feature of this type of building block.*

provide efficient and stereospecific glycosidic bond formation and the elimination of the need to use protecting groups on the carbohydrate moieties. Although enzyme-compatible resins have been developed, reducing the general problem of the accessibility of the interior of polymeric resin beads to the enzyme[79,81,82], enzyme-based solid-phase chemistry is still limited by low reaction rates and the availability of glycosyltransferases and their sugar-phosphate-nucleotide cosubstrates[83].

Access to the carbohydrate diversity of a glycopeptide is currently only limited by the availability of the necessary saccharide units, irrespective of strategy. Application of the 'building blocks' strategy requires that each *N*- or *O*-linked amino acid saccharide unit be synthesized separately. Thus, with the 'building blocks' approach although it is easy to vary the position of the carbohydrate-containing amino acid, it is not as easy to vary the nature of the carbohydrate itself without constructing a new building block. This approach also requires glycosidic linkages to survive repeated exposure to various types of chemistry during the construction of a full glycopeptide. In spite of these limitations, by application of the split resin technique, the 'building blocks' strategy has yielded glycopeptide libraries of modest size[84–86]. The generation of a small library of glycopeptides by the convergent strategy has been accomplished by amino-derivatization of the reducing end of several commercially available oligosaccharides then formation of amide bonds to a series of carboxyl-containing resin-bound peptides[65]. Currently, only *N*-linked oligosaccharides can be effectively accessed by the convergent approach. If access to other than the limited number of commercially available saccharides is required, it will be necessary to prepare the saccharide units.

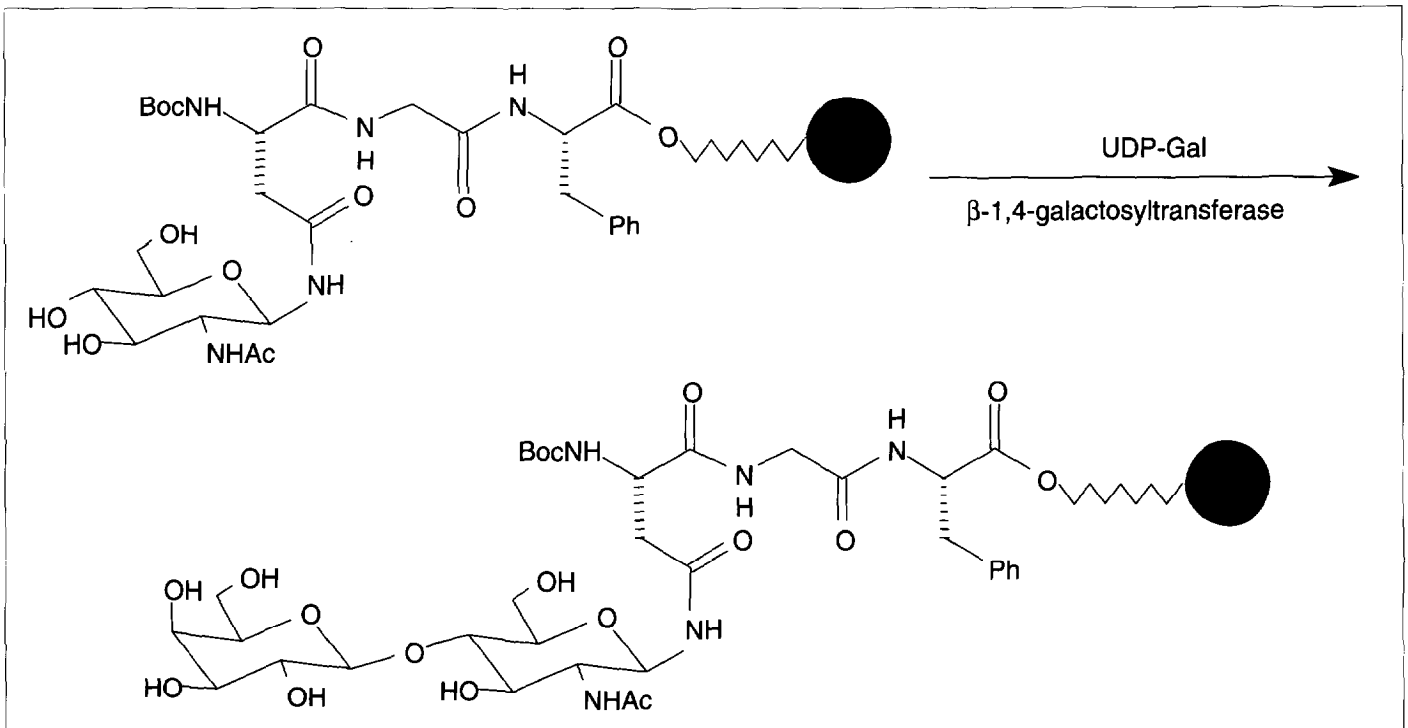

Figure 11. *Solid-phase synthesis of glycopeptides using enzyme-based glycosylation technology. The solid support is an amino-functionalized silica.*

Conclusion

The key element needed for generating carbohydrate-based chemical libraries for screening against biological targets is a chemical technology that enables rapid and efficient generation of the constituents of the library. Associated technologies, such as sequencing methods and library deconvolution methods, would be very valuable but are currently not necessary to begin the preparation of carbohydrate-based libraries. Through use of a spatially arrayed reaction format and information handling systems, access to oligosaccharide and glycoconjugate diversity is possible with currently available solid-phase synthesis technologies.

REFERENCES

1 Varki, A. (1993) *Glycobiology* 3, 97–130
2 Rademacher, T.W., Parekh, R.B. and Dwek, R.A. (1988) *Annu. Rev. Biochem.* 57, 785–838
3 Sairam, M.R. (1989) *FASEB J.* 3, 1915–1926
4 Gambaryan, A.S. *et al.* (1995) *FEBS Lett.* 366, 57–60
5 Boren, T. *et al.* (1993) *Science* 262, 1892–1895
6 Lee, Y.C. (1989) in *Carbohydrate Recognition in Cellular Function* (Bock, G. and Harnett, S., eds), pp. 80–95, John Wiley & Sons
7 Fisher, J.F. *et al.* (1991) *J. Med. Chem.* 34, 3140–3143
8 Rodriguez, R.E. *et al.* (1989) *Neurosci. Lett.* 101, 89–94
9 Hirschmann, R. *et al.* (1993) *J. Am. Chem. Soc.* 115, 12550–12568
10 Allanson, N.M., Davidson, A.H. and Martin, F.H. (1993) *Tetrahedron Lett.* 34, 3945–3948
11 Defrees, S.A. *et al.* (1995) *J. Am. Chem. Soc.* 117, 66–79
12 Gallop, M.A. *et al.* (1994) *J. Med. Chem.* 37, 1233–1251
13 Gordon, E.M. *et al.* (1994) *J. Med. Chem.* 37, 1385–1400
14 DeWitt, S.H. *et al.* (1994) *Drug Dev. Res.* 33, 116–124
15 Desai, M.C., Zuckermann, R.N. and Moos, W.H. (1994) *Drug Dev. Res.* 33, 174–188
16 Han, H. *et al.* (1995) *Proc. Natl Acad. Sci. USA* 92, 6419–6423
17 Lam, K.S. *et al.* (1993) *Bioorg. Med. Chem. Lett.* 3, 419–424
18 Pinilla, C. *et al.* (1995) *Biopolymers* 37, 221–240
19 Sherman, M.I., Bertelsen, A.H. and Cook, A.F. (1993) *Bioorg. Med. Chem. Lett.* 3, 469–475
20 Simon, R.J. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 9367–9371
21 Bunin, B.A., Plunkett, M.J. and Ellman, J.A. (1994) *Proc. Natl Acad. Sci. USA* 91, 4708–4712
22 Campbell, D.A. *et al.* (1995) *J. Am. Chem. Soc.* 117, 5381–5382
23 Murphy, M.M. *et al.* (1995) *J. Am. Chem. Soc.* 117, 7029–7030
24 Burbaum, J.J. *et al.* (1995) *Proc. Natl Acad. Sci. USA* 92, 6027–6031
25 Baldwin, J.J. *et al.* (1995) *J. Am. Chem. Soc.* 117, 5588–5589
26 Janda, K.D. (1994) *Proc. Natl Acad. Sci. USA* 91, 10779–10785
27 Needels, M.C. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 10700–10704
28 Ohlmeyer, M.H.J. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 10922–10926
29 Nikolaiev, V. *et al.* (1993) *Peptide Res.* 6, 161–170
30 Nielsen, J., Brenner, S. and Janda, K.D. (1993) *J. Am. Chem. Soc.* 115, 9812–9813
31 Deprez, B. *et al.* (1995) *J. Am. Chem. Soc.* 117, 5405–5406
32 Pinilla, C. *et al.* (1994) *Drug Dev. Res.* 33, 133–145
33 Frechet, J.M. and Schuerch, C. (1971) *J. Am. Chem. Soc.* 93, 492–496
34 Frechet, J.M. and Schuerch, C. (1972) *Carbohydr. Res.* 22, 399–412
35 Frechet, J.M. and Schuerch, C. (1972) *J. Am. Chem. Soc.* 94, 604–609
36 Guthrie, R.D., Jenkins, A.D. and Stehlicek, J. (1971) *J. Chem. Soc. (C)* 2690–2696
37 Guthrie, R.D., Jenkins, A.D. and Roberts, G.A.F. (1973) *J. Chem. Soc. Perkin Trans. 1*, 2414–2417
38 Veeneman, G.H. *et al.* (1990) *Carbohydr. Res.* 195, C1–C4
39 Kandil, A.A. *et al.* (1992) *Synlett* 555–557
40 Westerduin, P. *et al.* (1987) *Tetrahedron Lett.* 28, 1557–1560
41 Douglas, S.P., Whitfield, D.M. and Krepinsky, J.J. (1991) *J. Am. Chem. Soc.* 113, 5095–5097
42 Douglas, S.P., Whitfield, D.M. and Krepinsky, J.J. (1995) *J. Am. Chem. Soc.* 117, 2116–2117
43 Leung, O-T. *et al.* (1994) *New J. Chem.* 18, 349–363
44 Krepinsky, J.J., Douglas, S.P. and Whitfield, D.M. (1994) *Methods Enzymol.* 242, 280–293
45 Verduyn, R. *et al.* (1993) *Recl. Trav. Chim. Pays-Bas* 112, 464–466
46 Veeneman, G.H. *et al.* (1987) *Tetrahedron Lett.* 28, 6695–6698
47 Yan, L. *et al.* (1994) *J. Am. Chem. Soc.* 116, 6953–6954
48 Wang, Y. *et al.* (1995) *Chem. Lett.* 273–274
49 Danishefsky, S.J. *et al.* (1993) *Science* 260, 1307–1309
50 Randolph, J.T. and Danishefsky, S.J. (1994) *Angew. Chem. Int. Ed. Engl.* 33, 1470–1473
51 Randolph, J.T., McClure, K.F. and Danishefsky, S.J. (1995) *J. Am. Chem. Soc.* 117, 5712–5719
52 Schuster, M. *et al.* (1994) *J. Am. Chem. Soc.* 116, 1135–1136
53 Halcomb, R.L., Huang, J. and Wong, C-H. (1994) *J. Am. Chem. Soc.* 116, 11315–11322
54 Vetter, D., Tate, E.M. and Gallop, M.A. (1995) *Bioconj. Chem.* 6, 319–322
55 Kunz, H. (1987) *Angew. Chem. Int. Ed. Engl.* 26, 294–308
56 Paulsen, H. (1990) *Angew. Chem. Int. Ed. Engl.* 29, 823–839
57 Meldal, M. (1994) *Curr. Opin. Struct. Biol.* 4, 710–718
58 Christensen, M.K. *et al.* (1994) *J. Chem. Soc. Perkin Trans. 1*, 1299–1310
59 Christensen, M.K., Medal, M. and Bock, K. (1993) *J. Chem. Soc. Perkin Trans. 1*, 1453–1460
60 Norberg, T., Luning, B. and Tejbrant, J. (1994) *Methods Enzymol.* 242, 87–106
61 Andrews, D.M. and Seale, P.W. (1993) *Int. J. Peptide Res.* 42, 165–170
62 Hollosi, M. *et al.* (1991) *Tetrahedron Lett.* 32, 1531–1534
63 Cohen-Anisfeld, S.T. and Lansbury, P.T., Jr (1993) *J. Am. Chem. Soc.* 115, 10531–10537
64 Vetter, D. and Gallop, M.A. (1995) *Bioconj. Chem.* 6, 316–318
65 Vetter, D. *et al.* (1995) *Angew. Chem. Int. Ed. Engl.* 34, 60–63
66 Roberge, J.Y., Beebe, X. and Danishefsky, S.J. (1995) *Science* 269, 202–204
67 Meldal, M. and Bock, K. (1990) *Tetrahedron Lett.* 31, 6987–6990
68 Christiansen-Brams, I., Meldal, M. and Bock, K. (1993) *J. Chem. Soc. Perkin Trans. 1*, 1461–1471
69 Meldal, M., Mouritsen, S. and Bock, K. (1993) in *Carbohydrate Antigens, American Chemical Society Symposium Series 519* (Garegg, P.J. and Lindberg, A.A., eds), pp. 19–33, Academic Press
70 Jensen, K.J., Meldal, M. and Bock, K. (1993) *J. Chem. Soc. Perkin Trans. 1*, 2119–2129
71 Meldal, M. and Bock, K. (1994) *Glycoconjugate J.* 11, 59–63
72 Bielfeldt, T. *et al.* (1992) *Angew. Chem. Int. Ed. Engl.* 31, 857–859
73 Jansson, A.M., Meldal, M. and Bock, K. (1992) *J. Chem. Soc. Perkin Trans. 1*, 1699–1707
74 Reimer, K.B. *et al.* (1993) *J. Chem. Soc. Perkin Trans. 1*, 925–932
75 Seitz, O. and Kunz, H. (1995) *Angew. Chem. Int. Ed. Engl.* 34, 803–805
76 Rio-Anneheim, S. *et al.* (1995) *J. Chem. Soc. Perkin Trans. 1*, 1071–1080
77 Szabo, L., Li, Y. and Polt, R. (1991) *Tetrahedron Lett.* 32, 585–588
78 Polt, R. *et al.* (1992) *J. Am. Chem. Soc.* 114, 10249–10258
79 Wong, C-H. *et al.* (1993) *J. Am. Chem. Soc.* 115, 5893–5901
80 Meldal, M. *et al.* (1994) *J. Chem. Soc. Chem. Commun.* 1849–1850
81 Auzanneau, F.I., Meldal, M. and Bock, K. (1994) *J. Peptide Sci.* 1, 31–44
82 Meldal, M., Auzanneau, F.I. and Bock, K. (1994) in *Innovation and Perspectives in Solid-phase Synthesis* (Epton, R., ed.), pp. 259–266, Mayflower Worldwide
83 Wong, C.-H. *et al.* (1995) *Angew. Chem. Int. Ed. Engl.* 34, 521–546
84 Meldal, M. *et al.* (1993) *Int. J. Peptide Protein Res.* 41, 250–260
85 Elofsson, M. *et al.* (1993) *Carbohydr. Res.* 246, 89–103
86 Peters, S. *et al.* (1992) *J. Chem. Soc. Perkin Trans. 1*, 1163–1171