Enzymatic and chemo-enzymatic synthesis of carbohydrates

Chi-Huey Wong

Department of Chemistry, The Scripps Research Institute
10666 North Torrey Pines Road, La Jolla, CA 92037

Abstract: Carbohydrate-mediated biochemical recognition processes are potential targets for drug development. The synthesis of complex carbohydrates and related structures, however, has been a difficult task, especially from practical point of view. New solution- and solid-phase methods based on enzymatic and a combined enzymatic and chemical procedure have recently been developed for the practical synthesis of various carbohydrate-related structures. Monosaccharides, aza- and thiosugars, and their analogs can be effectively prepared based on aldolase-catalyzed reactions. Many oligosaccharides can now be synthesized on multi-gram or kilogram scales based on recombinant glycosyltransferases with *in situ* regeneration of sugar nucleotides. Serine proteases can be engineered to peptide ligases for glycopeptide coupling in aqueous solution, and further enzymatic incorporation of sugars to form complex glycopeptides can be achieved. Phospholipids containing sugars and azasugars can be prepared using phospholipase D-catalyzed exchange reaction followed by self-assembly to form liposome-like structures with sugars displayed on the surface. These newly developed chemo-enzymatic methods have been applied to the synthesis of specific inhibitors of glycosidases and glycosyltransferases, and of receptors of carbohydrate-mediated cell adhesion.
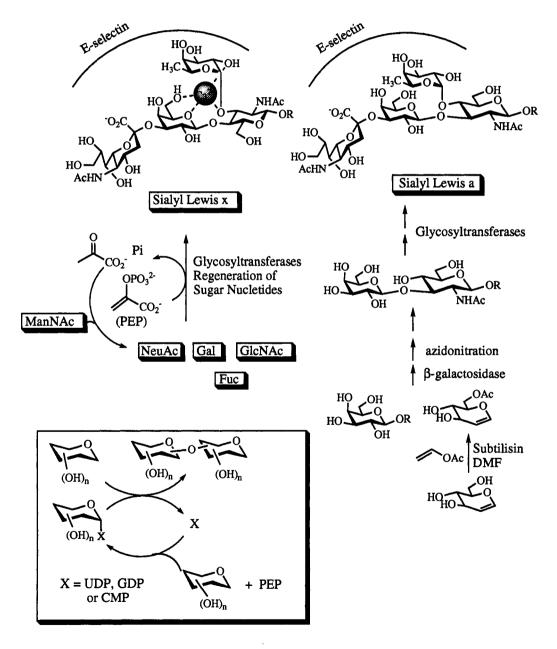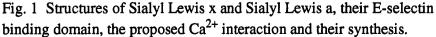
Introduction

Carbohydrates on cell surfaces are effective information molecules (1). Although only seven to eight different monosaccharides are commonly used in mammalian systems as building blocks (2,3), the multifunctionality of these monomers and the number of possibilities for their linkage can lead to the assembly of an immense variety of complex structures. Multimillion different tetrasaccharide structures, for example, can be constructed from this small number of building blocks, when considering the branching, stereochemistry of glycosidic linkages, and modification of hydroxyl and amino groups. Oligosaccharides therefore represent a class of effective biomolecules coding for a vast amount of information required in various biological recognition processes, such as bacterial and viral infection, cell adhesion, signal transduction, differentiation, development, regulation and many other intercellular communications (1). It seems that carbohydrate-mediated recognition is a fascinating area for drug discovery. The pace of development of carbohydrate-based pharmaceuticals has, however, been much slower than that of other classes of biomolecules. Part of the reason is due to the lack of significant glycobiology information as well as the difficulty encountered in the synthesis, especially from a practical point of view, of complex carbohydrates. There is no method available for the amplification of oligosaccharides to facilitate sequence analysis. There is no machine available for a rapid synthesis of oligosaccharides for the study of structure-function relationship. Furthermore, there are also concerns about the poor oral activity of carbohydrates and the weak binding of carbohydrates to their receptors.

New discoveries in glycobiology

In the past five years, however, several new exciting discoveries in glycobiology have renewed our interest in carbohydrate research, and the one of particular interest to us is sialyl Lewis x (SLe^x) mediated adhesion of leukocytes to E-selectin containing endothelial cells in response to inflammatory reactions (4). Since the tetrasaccharide SLe^x was determined to be the ligand for E-selectin (5), it has been considered to be potentially useful as an inhibitor of E-selectin and therefore as a new antiinflammatory agent. In addition to E-selectin, P- and L-selectins and their ligands have

also been discovered (4,6), and they are also involved in inflammatory reactions. The solution conformation of SLe^X has been determined based on the 2-D NMR analysis of a ^{13}C-labeled derivative (7), and the location of Ca^{++} bound to SLe^X has been proposed to be in the Le^X moiety based on the results of collision induced decomposition study of SLe^X and derivatives using ion-spray mass spectrometry (Fig. 1) (8). Although the x-ray crystal structure of E-selectin is now available (9), the structure of the receptor-ligand-Ca^{++} complex remains to be determined. Extensive studies on the structure-function relationship of SLe^X (10,11) have led to the assumption that the active binding domain of SLe^X resides in the topostructure determined by the fucose, galactose and the carboxylate of sialic acid. Current efforts are directed toward the design of non-carbohydrate organic molecules which resemble the topostructure of this active domain and hopefully have better and desirable biological activities.

Fig. 1 Structures of Sialyl Lewis x and Sialyl Lewis a, their E-selectin binding domain, the proposed Ca^{2+} interaction and their synthesis.

Solution- and solid-phase synthesis of oligosaccharides and glycopeptides

Our synthesis of SLe^x depends on the use of glycosyltransferases coupled with *in situ* regeneration of sugar nucleotides (7) (Fig. 1). Two and perhaps three different glycosyltransferases can be used in one pot to construct the tetrasaccharide from the four monosaccharides, and NeuAc can be generated *in situ* from pyruvate and N-acetylmannosamine catalyzed by sialic acid aldolase (12). This procedure not only reduces the cost of sugar nucleotides but also eliminates the problem of product inhibition caused by nucleoside phosphates, and is applicable to a large-scale process. Methods are now available for the regeneration of each of the seven sugar nucleotides commonly found in mammalian systems (13). Glycosidases can also be used in the synthesis of a glycoside (e.g. N-acetyllactosamine) which is then used *in situ* as a substrate for a glycosyltransferase (14,15). As recombinant soluble glycosyltransferase are becoming more readily available (based on the baculous virus or fungal expression system) (7, and personal communication with J.C. Paulson), enzymatic synthesis of oligosaccharides on large scales will be no longer difficult. A major drawback is, however, the relatively narrow substrate specificity of glycosyltransferase (7, 16-19). Whether the regio- and substrate specificity of glycosyltransferases can be easily altered by site-directed mutagenesis remains to be an open question. In any case, cloning of glycosyltransferase genes, especially into *E. coli*, for overproduction of the enzymes may facilitate the study; however, it has been very difficult to express glycosyltransferases in *E. coli*. The recombinant *E. coli* strain recently developed for the expression of an α-1,2-mannosyl-transferase provides a reasonable amount of the enzyme (~ one unit/liter) for study and the mannosylation reactions can be performed with the use of whole cells (20). No product inhibition caused by GDP was observed. The system is being used as a model for engineering the enzyme specificity.

The study of enzymatic synthesis of SLe^x has also led to the development of a new solid-phase method for glycopeptide synthesis (21) (Fig. 2). Using aminopropyl silica gel as a solid support, a polyglycine spacer and an enzyme-sensitive cleavable bond are incorporated, followed by addition of amino acids (one of them is glycosylamino acid) using conventional solid-phase method. Additional sugars are then incorporated catalyzed by glycosyltransferases. This method allows a rapid synthesis of various glycopeptides on a solid support which may be useful for screening and for affinity isolation of receptors. For large-scale processes, a new solution-phase method has been developed for glycopeptide synthesis using engineered thio-subtilisins for ligation of glycopeptides and glycosyltransferases for incorporation of additional sugars (22). These chemo-enzymatic methods for glycopeptide synthesis have the advantage that minimal protection and deprotection steps are required and that product isolation is straightforward, as minimal deleted oligosaccharide

Fig. 2 Strategies for the solid-phase and solution-phase chemo-enzymatic synthesis of glycopeptides (A$_1$-----A$_n$, amino acid residues; (S) , solid support; P, R, protection groups)

sequences are generated. It is expected that when glycosyltransferases become readily available, solution- or solid-phase synthesis of complex oligosaccharides and glycopeptides will be straightforward and certain unnatural structures will also be accessible.

As mentioned, despite their great utility in the synthesis of natural and related oligosaccharides, glycosyltransferases generally do not accept unnatural sugars as good substrates. Synthesis of oligosaccharide analogs and derivatives containing modified monosaccharides therefore often depends on conventional chemical methods. Our efforts in this regard have resulted in the development of a new glycosylation method based on glycosylphosphite (23), and the synthesis of monosaccharide-related structures based on aldolase-catalyzed addition reactions (19) (Fig. 3, 4). The glycosylphosphite chemistry has proven to be applicable to all monosaccharides and particularly useful for sialylation and fucosylation (24). When TMSOTf is used as a catalyst for coupling, the real catalyst was determined to be the triflic acid released from the initial reaction of TMSOTf and the sugar acceptor (24).

Synthesis of monosaccharides and related structures based on enzymatic aldol reactions
For the synthesis of modified monosaccharides, various aldolases can be used (19). Approximately twenty aldolases have been discovered and several of them have been cloned and overexpressed. These enzymes generally accept a wide range of unnatural aldehydes as substrates and form a new carbon-carbon bond stereospecifically in aqueous solution, providing a useful new route to various monosaccharides and related structures. Amino-, deoxy-, thio-, fluoro-O-acyl or alkyl, epimeric sugars and carbocycles, for example, can be easily prepared. The aldehydes can be used as racemate and often only one of the enantiomers is accepted as substrate in either a kinetic or thermodynamic mode of reaction. Enantiomerically pure aldehydes can also be prepared easily via lipase catalyzed resolution, or catalytic asymmetric dihydroxylation or epoxidation (Fig. 3).

Inhibition of glycosidases and glycosyltransferases: Synthesis of azasugars based on aldolases
The two major classes of enzymes involved in carbohydrate processing and biosynthesis, i.e. glycosidases and glycosyltransferases, are interesting targets for inhibition. Development of specific inhibitors of these enzymes could lead to the discovery of new agents for the treatment of diseases and metabolic disorders associated with carbohydrate recognition. For the inhibition of
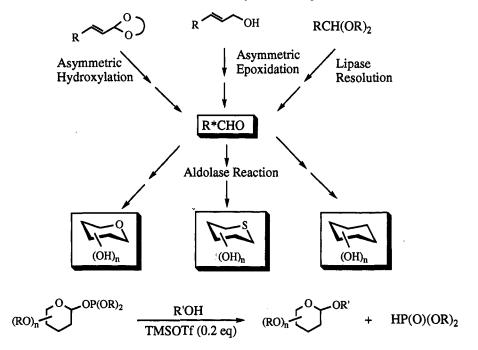

Fig. 3 Strategy for the synthesis of monosaccharides, thiosugars and carbocycles based on aldolase-catalyzed reactions, and for glycosylation using glycosylphosphite

glycosidases, azasugars have proven to be effective as they are considered to mimic the transition state or high-energy intermediate of the glycosidic bond cleavage, a process believed to proceed through a half-chair transition state with a substantial positive charge developed at the anomeric center, and the cleavage is assisted by a carboxylate and a carboxylic acid in the enzyme active site (25). Both five- and six-membered azasugars and their homo-analogs can be prepared based on aldolase-catalyzed reaction with azidoaldehydes, followed by Pd-catalyzed reductive amination (26) (Fig. 4). The reactions are mild and stereoselective, and are carried out in aqueous solution with minimal protection of the substrates. The azidoaldehydes used in the synthesis may be racemic, because often only one enantiomer is accepted. If enantiomerically pure azidoaldehydes are preferred, they are accessible by lipase-catalyzed enantioselective hydrolysis or acylation.
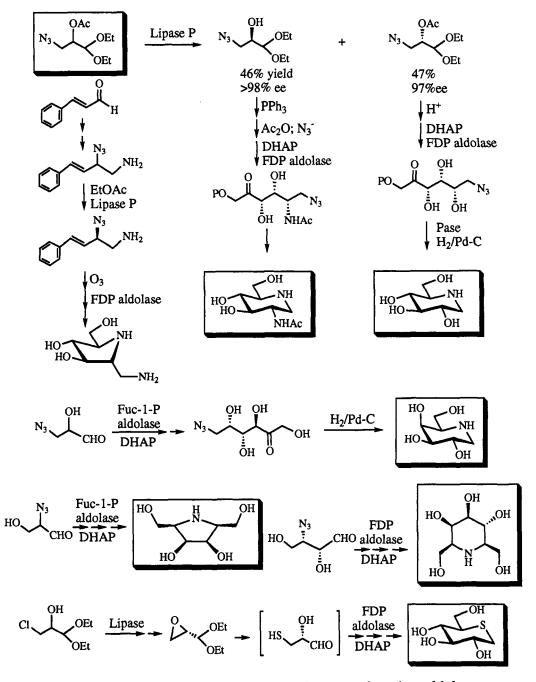

Fig. 4 Examples for the synthesis of aza- and thiosugars based on aldolases
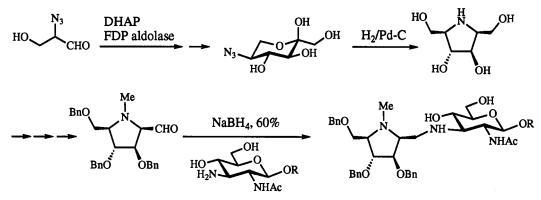
Fig. 5 Synthesis of sequence-specific glycosidase inhibitor

Azasugars can be used in the development of specific inhibitors of endoglycosidases or glycosyltransferases. For example, incorporation of an appropriate carbohydrate group to a homoazasugar at the position equivalent to the anomeric carbon gives a sequence specific glycosidase inhibitor (Fig. 5). Mixing an azasugar that inhibits α-fucosidase and GDP would form a complex in the active site of α-1,3-fucosyltransferase and becomes an effective inhibitor of the enzyme (7). Attachment of an acceptor moiety of α1,3-fucosyltransferase (i.e. N-acetyllactosamine or the sialyl derivative) to the ring nitrogen of the azasugar and a GDP equivalent to the carbon one position provides a sequence-specific inhibitor of α-1,3-fucosyltransferase (Fig. 6).
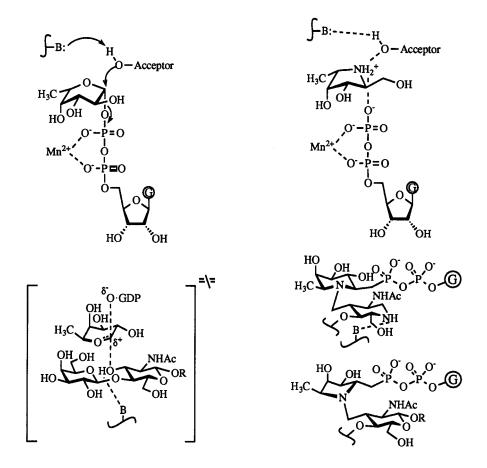
Fig. 6 Synergistic Inhibition of α-1,3-fucosyltransferase by an azasugar and GDP, and transition-state analog inhibitors of the enzyme

Issues of multivalency

It appears that inhibition of cell-surface receptors or enzymes can be effectively enhanced if the momeric inhibitor is oligomerized or clustered, as illustrated in the inhibition of influenza hemagglutinin (27) and E-selectin (14). The increase of binding is mainly due to the decrease of entropy associated with the second and subsequent steps of inhibitor binding. Although multivalent inhibitors can be easily prepared by polymerization of the inhibitor or by attachment of the inhibitor to a template such as proteins or organic molecules (28), an alternative approach is to incorporate the inhibitor into a phospholipid via exchange with the choline group catalyzed by phospholipase D. The lipase reaction is very versatile and the enzyme is selective for primary hydroxyl groups in the presence of secondary hydroxyls, thiols and amines (28). The newly formed phospholipids are then self-assembled into a liposome with many of the inhibitors displayed on the surface (Fig. 7). The size of liposomes can be controlled by the type of fatty acids used, and the distance between the inhibitors can be manipulated by mixing with an appropriate portion of natural phospholipids. This approach has been proven effective in the inhibition of glucosylceremide β-glycosidase and hemagglutinin, and may be applicable to the inhibition of selectins.
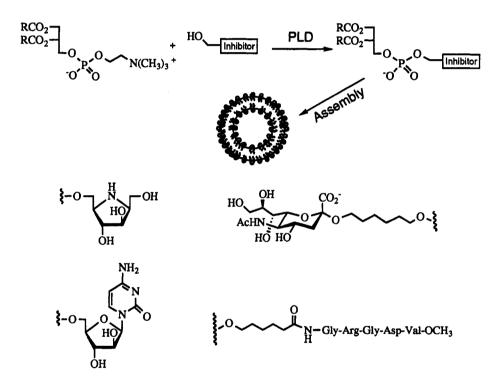

Fig. 7 Phospholipase D catalyzed synthesis of phospholipid-inhibitor conjugates and self assembly into liposome as multivalent inhibitors

References
1. A. Varki. *Glycobiology* **3**, 97 (1993).
2. L.F. Leloir. *Science* **172**, 1299 (1971).
3. A.T. Beyer, J.E. Sadler, J.I. Rearick, J.C. Paulson, R.L. Hill. *Adv. Enzymol.* **52**, 24 (1981).
4. L.A. Lasky. *Science* **258**, 964 (1992).

5. M.L. Phillips, E. Nudelman, F.C.A. Gaeta, M. Perez, A.K. Singhal, S. Hakomori, J.C. Paulson. *Science* **250**, 1130 (1990).

6. S. Hemmerich, S.D. Rosen. *Biochemistry* **33**, 4830 (1994).

7. Y. Ichikawa, Y.-C. Lin, D.P. Dumas, G.-J. Shen, E. Garcia-Junceda, M.A. Williams, R. Bayer, C. Ketcham, L.E. Walker, J.C. Paulson, C.-H. Wong. *J. Am. Chem. Soc.* **114**, 9283 (1992).

8. G. Siuzdak, J.R. Ramphal, S. DeFrees, Y. Ichikawa, K.C. Nicolaou, C.-H. Wong. *Bioorg. & Med. Chem. Lett.*, submitted.

9. B.J. Graves, R.L. Crowther, C. Chandran, J.M. Rumberger, S. Li, K.-S. Huang, D.M. Presky, P.C. Familetti, B.A. Wolitzky, D.K. Burns. *Nature* **367**, 532 (1994).

10. S.A. DeFrees, F.C.A. Gaeta, Y.-C. Lin, Y. Ichikawa, C.-H. Wong. *J. Am. Chem. Soc.* **115**, 7549 (1993).

11. B.K. Brandley, M. Kiso, S. Abbas, P. Nikrad, O. Srivasatava, C. Foxall, Y. Oda, A. Hasegawa. *Glycobiology* **3**, 633 (1993).

12. Y. Ichikawa, J.L.-C. Liu, G.-J. Shen, C.-H. Wong. *J. Am. Chem. Soc.* **113**, 6300 (1991).

13. Y. Ichikawa, G.-J. Shen, R. Wang, C.-H. Wong. *Method. Enzymol.* 1994, in press. For the first sugar nucleotide regeneration (UDP-Gal), see C.-H. Wong, S.L. Haynie, G.M. Whitesides. *J. Org. Chem.* **47**, 5416 (1982).

14. G.F. Herrmann, Y. Ichikawa, C. Wandrey, F.C.A. Gaeta, J.C. Paulson, C.-H. Wong. *Tetrahedron Lett.* **19**, 3091 (1993).

15. G.C. Look, Y. Ichikawa, G.-J. Shen, P.-W. Cheng, C.-H. Wong. *J. Org. Chem.* **58**, 4326 (1993).

16. O. Hindsgaul. *Seminars in Cell Biology* **2**, 319 (1991).

17. C.-H. Wong, Y. Ichikawa, T. Krach, C. Gautheron, D. Dumas, G. Look. *J. Am. Chem. Soc.* **113**, 8137 (1991).

18. Y. Nishida, T. Wiemann, V. Sinwell, J. Thiem. *J. Am. Chem. Soc.* **115**, 2536 (1993).

19. For a recent view, see C.-H. Wong, R.L. Halcomb, Y. Ichikawa, T. Kajimoto. *Angew. Chem.*, in press.

20. G.F. Herrmann, P. Wang, G.-J. Shen, C.-H. Wong. *Angew. Chem. Int. Ed. Engl.* **33**, 1241 (1994).

21. M. Schuster, P. Wang, J.C. Paulson, C.-H. Wong. *J. Am. Chem. Soc.* **116**, 1135 (1994).

22. C.-H. Wong, M. Schuster, P. Wang, P. Sears. *J. Am. Chem. Soc.* **115**, 5893 (1993); P. Sears, P. Wang, K. Witte, M. Schuster, C.-H. Wong. *J. Am. Chem. Soc.* **116**, 6521 (1994).

23. H. Kondo, Y. Ichikawa, C.-H. Wong. *J. Am. Chem. Soc.* **114**, 8748 (1992). Schmidt and Watanabe have also reported the use of glycosylphosphite in glycosylation: T.J. Martin, R.R. Schmidt. *Tetrahedron Lett.* **33**, 6123 (1992); Y. Watanabe, C. Nakamoto, S. Ozaki. *Synlett* **115** (1993).

24. H. Kondo, S. Aoki, Y. Ichikawa, R.L. Halcomb, H. Ritzen, C.-H. Wong. *J. Org. Chem.* **59**, 864 (1994).

25. M.L. Sinnott. *Chem. Rev.* **90**, 1171 (1990).

26. G.C. Look, C.-H. Fotsch, C.-H. Wong. *Acc. Chem. Res.* **26**, 182 (1993). For most recent work in this area, see Y.-F. Wang, Y. Takaoka, C.-H. Wong. *Angew. Chem. Int. Ed. Engl.* **33**, 1242 (1994).

27. A. Spaltensin, G.M. Whitesides. *J. Am. Chem. Soc.* **113**, 686 (1991); G.D. Glick, P.L. Toogood, D.C. Wiley, J.J. Skehel, J.R. Knowles. *J. Biol. Chem.* **266**, 23660 (1991); S. Sabesan, J.O. Duus, S. Neira, P. Domaille, S. Kelm, J.C. Paulson. K. Bock. *J. Am. Chem. Soc.* **114**, 8363 (1992); R. Roy, C.A. Laferriere. *J. Chem. Soc. Chem. Commun.* 1709 (1990).

28. P. Wang, M. Schuster, Y.-F. Wang, C.-H. Wong. *J. Am. Chem. Soc.* **115**, 10487 (1993).