standard cleavage and polyadenylation reaction (*15*) except that reactions were incubated for 60 min at 30°C.
20. S. Bienroth, W. Keller, E. Wahle, *EMBO J.* **12**, 585 (1993).
21. A. Moreira, M. Wollerton, J. Monks, N. J. Proudfoot, *EMBO J.* **14**, 3809 (1995).
22. G. M. Gilmartin, E. S. Fleming, J. Oetjen, B. R. Graveley, *Genes Dev.* **9**, 72 (1995).
23. A. Bilger, C. A. Fox, E. Wahle, M. Wickens, *ibid.* **8**, 1106 (1994).
24. F. W. Studier, A. H. Rosenberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol.* **185**, 60 (1990).
25. S. Heidmann, B. Obermaier, K. Vogel, H. Domdey, *Mol. Cell. Biol.* **12**, 4215 (1992).
26. S. Heidmann, C. Schindewolf, G. Stumpf, H. Domdey, *ibid.* **14**, 4633 (1994).
27. J. Hani, G. Stumpf, H. Domdey, *FEBS Lett.* **365**, 198 (1995).
28. For the reconstitution of specific polyadenylation, immunodepleted extract (2 μl) was mixed with 1 μl of partially purified PAP (contained in fraction 11) and 1 μl of Mono Q fractions 19 to 27. The PAP-containing fraction was identified by an unspecific polyadenylation assay as described (*6*).
29. We thank J. Hani for advice with the computer search, U. Sterzer for technical assistance, H. Würl for materials, and A. Goppelt and K. Wiebauer for discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft.

21 June 1996; accepted 21 October 1996

Parallel Synthesis and Screening of a Solid Phase Carbohydrate Library

Rui Liang, Lin Yan, Jennifer Loebach, Min Ge, Yasuhiro Uozumi, Klara Sekanina, Nina Horan, Jeff Gildersleeve, Chris Thompson, Andri Smith, Kaustav Biswas, W. Clark Still, Daniel Kahne*

A solid phase carbohydrate library was synthesized and screened against *Bauhinia purpurea* lectin. The library, which contains approximately 1300 di- and trisaccharides, was synthesized with chemical encoding on TentaGel resin so that each bead contained a single carbohydrate. Two ligands that bind more tightly to the lectin than Gal-β-1,3-GalNAc (the known ligand) have been identified. The strategy outlined can be used to identify carbohydrate-based ligands for any receptor; however, because the derivatized beads mimic the polyvalent presentation of cell surface carbohydrates, the screen may prove especially valuable for discovering new compounds that bind to proteins participating in cell adhesion.

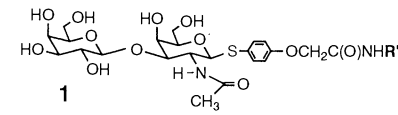
Cell surface carbohydrates play central roles in many normal and pathological biological recognition processes (*1*). For example, cell surface carbohydrates have been implicated in chronic inflammation, in viral and bacterial infection, and in tumorigenesis and metastasis (*2*). Considerable effort has been directed toward (i) understanding how carbohydrates function as recognition signals, and (ii) developing strategies to block undesirable interactions between cell surface carbohydrates and their protein targets. Ligands whose binding to the protein targets is better than that of the natural cell surface carbohydrates could provide an effective means of preventing or treating various diseases. But, progress in understanding how structure and function are related in biologically active carbohydrates has been slow because obtaining synthetic carbohydrate derivatives for biochemical studies is extremely difficult. In contrast, excellent chemical and biological methods to obtain large quantities of peptides and nucleic acids have been available for decades. As a consequence, much more is known about these other biopolymers.

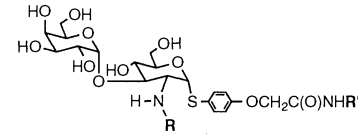
Screening combinatorial libraries of compounds is a very fast way to identify promising leads and elucidate structure-activity relationships (*3*). The first combinatorial libraries were built around peptides and nucleic acids because the chemistry to make them already existed (*4*). The successful use of peptide and nucleic acid libraries has stimulated efforts to develop combinatorial approaches to make other classes of molecules (*5*). One obvious area where a combinatorial approach could have tremendous impact is in carbohydrate chemistry and biochemistry.

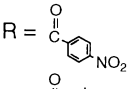
For the time being, the most satisfactory chemical approaches to the construction of compound libraries involve synthesis of the molecules on a solid support (*6*). Despite the efforts made, over more than 20 years, there has been little progress in the solid phase synthesis of carbohydrates. Carbohydrates present special difficulties for solid phase synthesis because the bonds between monomers must be formed stereospecifically and in high yield. Most glycosylation methods are extremely sensitive to structural variations in the glycosyl donor-acceptor pairs (*7*). Reaction conditions that provide excellent yields with one donor-acceptor pair may give virtually no product for another donor-acceptor pair. Furthermore, the stereochemical outcome is often difficult to predict. The unreliability of most glycosylation methods has precluded the construction of a solid phase carbohydrate library, which requires the ability to make a wide range of different glycosidic linkages both stereoselectively and in high yield.

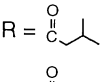
Several years ago we discovered a glycosylation method that makes use of anomeric sulfoxides as glycosyl donors. Anomeric sulfoxides can be activated almost instantaneously at low temperature regardless of the protecting groups on the sugar hydroxyl groups. The low temperature of reaction leads to excellent stereochemical control for a wide range of glycosyl donor-acceptor pairs while preventing side reactions so that it is possible to get nearly quantitative yields on the solid phase (*8*, *9*). We now report the synthesis of a solid phase carbohydrate library using the sulfoxide glycosylation reaction. The library was screened against a carbohydrate binding protein, and two ligands that bind more tightly than the natural ligand were identified. The strategy described can be used to identify carbohydrate-based ligands for any receptor; however, the on-bead screen may prove especially useful for discovering carbohydrate ligands that bind to their receptors in a polyvalent fashion.

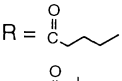
We designed the carbohydrate library to contain a diverse array of glycosidic linkages. The library consists of approximately





2 R = 

3 R = 

4 R = 

5 R = 

A series: R' = TentaGel
B series: R' = $CH_2CH_2OCH_3$

Fig. 1. Relevant structures contained in the library.

R. Liang, L. Yan, J. Loebach, M. Ge, N. Horan, J. Gildersleeve, C. Thompson, A. Smith K. Biswas, D. Kahne, Department of Chemistry, Princeton University, Princeton, NJ 08544, USA.
Y. Uozumi, K. Sekanina, W. C. Still, Department of Chemistry, Columbia University, New York, NY 10027, USA.
*To whom correspondence should be addressed.

1300 di- and trisaccharides, including both the α- and β-thiophenyl derivatives of the known *B. purpurea* ligand Gal-β-1,3-GalNAc(1A) (Fig. 1) (*10*). *Bauhinia purpurea* lectin is a protein that contains many carbohydrate binding sites. This lectin binds to carbohydrates on the surfaces of erythrocytes, causing them to agglutinate. The lectin is thus a good model system for cell adhesion proteins and other carbohydrate binding proteins that recognize cell surface carbohydrates.

The library was synthesized with the use of a split and mix strategy from the monomers shown in Fig. 2 (*11*). Six different carbohydrate monomers were attached separately to TentaGel resin. Then 12 different glycosyl sulfoxide donors were coupled separately to mixtures of beads containing all six monomers. The beads were recombined, the sugar azides were reduced to amines, and the beads were split again. The separate pools of beads were then *N*-acylated with different reagents. Finally, all the beads were recombined and deprotected. To facilitate identification of the products on each bead, the beads were encoded with chemical tags at each combinatorial step to record the reaction history of each bead (*12*).

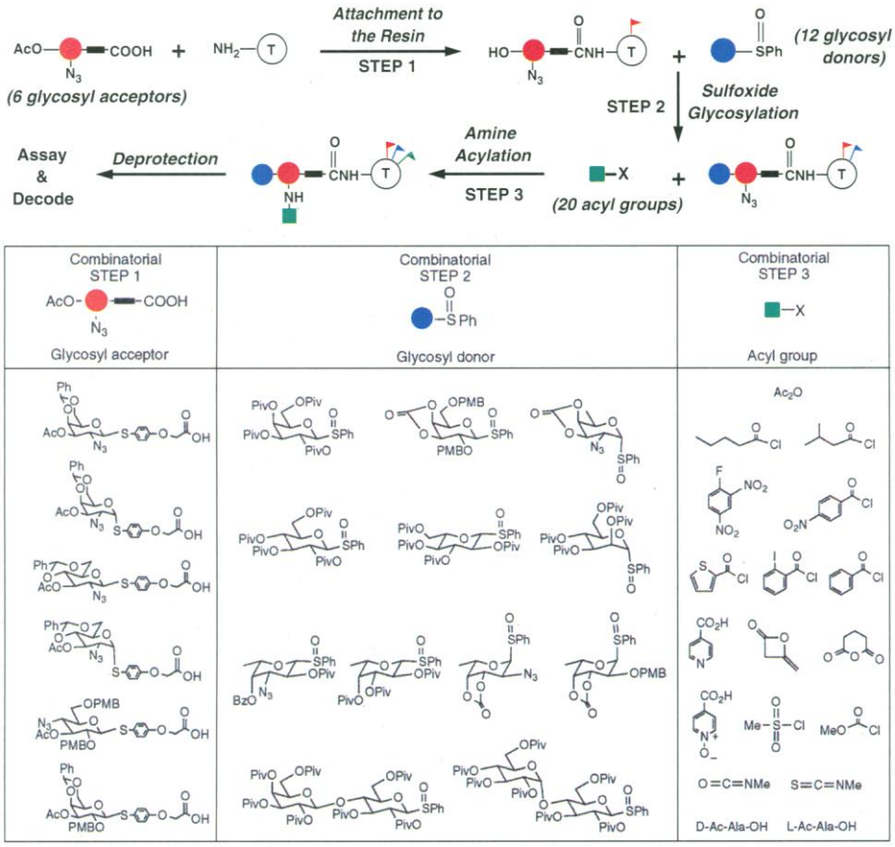


Fig. 2. First, six glycosyl acceptors were coupled to 500 mg of TentaGel amine resin (Rapp Polymere, 0.3 mmol/g) with 1-hydroxybenzotriazole (HOBT) and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), *N,N*-diisopropylethylamine, 1-methyl-2-pyrrolidinone (NMP) at room temperature for 2 to 5 hours. The resins were washed with dichloromethane (CH_2Cl_2), NMP, and *N,N*-dimethylformamide (DMF), treated with hydrazine in DMF (1:7) at room temperature for 9 hours, and washed with portions of DMF, water, methanol (CH_3OH), and CH_2Cl_2. The six resin portions were encoded (*12*). In step 2, the resins were mixed, divided into 12 glycosylation vessels, and suspended in CH_2Cl_2. The glycosyl donor (4 equivalents) and 2,6-di-*tert*-butyl-4-methylpyridine (8 equivalents) were dissolved in CH_2Cl_2, added to the resins, and cooled at −65°C. Trifluoromethanesulfonic anhydride (2 equivalents) in 1 ml of CH_2Cl_2 was added over 10 min, and the reactions were warmed (1 to 2 hours) to 0°C. The resins were washed with aqueous sodium bicarbonate, water, CH_3OH, diethyl ether, CH_2Cl_2, and toluene. The resins were lyophilized for 12 hours and subjected to the reaction conditions again. The 12 resins were encoded. In step 3, the resins were mixed and divided into 20 portions, and 19 were treated with trimethylphosphine in tetrahydrofuran (THF) at room temperature for 4 hours and then with THF and H_2O (16:1) at 70 to 75°C for 24 hours. The resins were washed with THF and CH_2Cl_2, and dried. Portions of resin (*18*) were acylated by one of two procedures: 1,3-dicyclohexylcarbodiimide, and a mixture of DMF, and CH_2Cl_2 (1:1) at room temperature for 12 to 24 hours or catalytic 4-dimethylaminopyridine, triethylamine, CH_2Cl_2 at 4°C for 12 hours. The original 20 portions of resin were encoded. The resins were combined and deprotected with 20% trifluoroacetic acid in CH_2Cl_2 at room temperature for 30 min followed by lithium hydroxide monohydrate in 20% THF in CH_3OH at room temperature for 12 hours.

The carbohydrates were synthesized from the reducing to the nonreducing end because this procedure permits the use of excess glycosyl donor, which drives the glycosylation reaction to completion. Complete reaction at each chemical step is critical in the construction of a solid phase library because mixtures of products on the beads can compromise both screening and identification of hits. In addition, carbohydrates constructed from the reducing to the nonreducing end are presented on the beads in a way that mimics their presentation on cell (or protein) surfaces. Hence, on deprotection the resin-bound library could be screened immediately against carbohydrate binding proteins that recognize cell surface carbohydrates.

A colorimetric assay was used to screen the resin-bound carbohydrate library against *Bauhinia purpurea* lectin (Fig. 3). Briefly, 10 mg of derivatized resin containing approximately 9000 beads, or six copies of the 1300-member library, was incubated with biotin-labeled lectin. The beads were then exposed to streptavidin-linked alkaline phosphatase and stained. Remarkably, only a small percentage of the 10-mg sample of beads (<0.3%) stained over a period of 20 minutes (Fig. 4). During this time, a total of 25 dark purple beads were picked out of the library and decoded (*12*). Four of the five beads that stained heavily within the first 5 minutes after addition of the enzyme substrate were identified as Gal-α-1,3-GlcNR-α-thiophenyl glycoside acylated with either 4-nitrobenzoyl (**2A**) or isovaleroyl (**3A**) (Fig. 1). After 20 minutes of staining, five copies of **2A** and five copies of **3A** had been pulled out of the library along with three other *N*-acyl derivatives of the same disaccharide (two copies of **4A** and one of **5A**). Thus, of 25 stained beads, 13 contained the same core disaccharide acylated with a hydrophobic group. In the remaining 12 stained beads, no carbohydrate structure appeared more than once and no pattern was evident. We consider the structures that appear only once to be noise. It is remarkable that essentially all the copies of two specific structures were selected from a pool of beads that contained approximately six copies each of ~1300 other related carbohydrate ligands.

For verification of the results, the known ligand **1A**, as well as the hit ligands **2A** and **3A**, were independently resynthesized on TentaGel, mixed, and stained. Beads containing **2A** and **3A** stained rapidly, but beads containing the known ligand did not. However, beads derivatized with the known ligand stained preferentially to beads derivatized with several other carbohydrates. The results confirmed that **1A** is a good ligand for the lectin, but **2A** and **3A** are better ligands.

To evaluate the relative solution affinities of the "hit" ligands relative to the natural ligand, we synthesized compounds **1B**, **2B**, and **3B** and tested for their ability to inhibit binding of the lectin to beads derivatized with the known ligand (**1A**). **1B**, **2B**, and **3B** all inhibit binding of the lectin to beads (derivatized with the known ligand) at concentrations of 20 to 50 μg/ml, indicating that **2B** and **3B** can block the pocket of the known ligand.

It is interesting that the lectin discriminates so well between beads containing the hit ligands and beads containing the known ligand given that the inhibitory concentrations are so similar. One of the paradoxes of carbohydrate binding is that typical carbohydrate binding proteins bind a wide range of related saccharides with similar affinities in solution, and yet appear to function with remarkable specificity in cell-cell recognition (*13*). The polyvalent presentation of carbohydrates on cell surfaces is believed to amplify the affinity and specificity of their interactions with carbohydrate binding proteins containing multiple carbohydrate binding sites (*14*, *15*). Our work definitively demonstrates that a carbohydrate binding protein can recognize a particular polyvalent carbohydrate ligand with exquisite specificity in the presence of a large number of related structures.

The work also shows that presenting the carbohydrate ligands on the beads affects their interactions with the lectin. Otherwise, the known ligand should stain as well as the hit ligands. Although presentation effects complicate on-bead screening of some libraries, we believe that on-bead screens are more biologically relevant than off-bead screens for studying binding interactions such as these that normally occur at surfaces. From the standpoint of design, one critical implication of our results is that the solution affinity of a monovalent carbohydrate ligand is not a reliable indicator of how well it will function in a polyvalent context.

Our work shows that it is possible to synthesize and screen a large carbohydrate library in a parallel fashion. The selection of two specific ligands from a collection of 1300 polyvalent carbohydrate structures validates the fidelity of the chemistry used to make the library as well as the utility of the assay. This combinatorial approach to studying carbohydrate binding could accelerate research on many carbohydrate recognition events dramatically. In addition, this approach should be tremendously valuable for discovering polyvalent ligands that block interactions at cell surfaces.

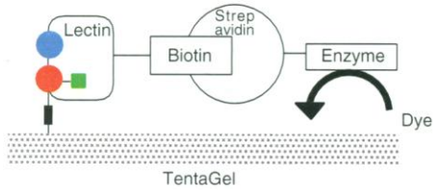


Fig. 3. Derivatized TentaGel beads (10 mg) were washed three times with 1 ml of PBST buffer [10 mM sodium phosphate, pH 7.2; 150 mM sodium chloride (NaCl); 0.05% Tween-20] and then suspended in 1 ml of PBST containing 3% bovine serum albumin (BSA). After being shaken for 30 minutes at room temperature, the beads were washed three times with 1 ml of PBST containing 1% BSA. The beads were incubated at room temperature for 3 hours on a rotary shaker in 1 ml of a biotin-labeled lectin solution (10 μg/ml in PBST containing 1% BSA) and then washed three times with 1 ml of TBST buffer [20 mM tri(hydroxymethyl)aminomethane hydrochloride (tris·HCl), pH 7.5; 500 mM NaCl; 0.05% Tween-20) containing 1% BSA. The beads were incubated on a rotary shaker for 20 min at room temperature in 1 ml of alkaline phosphatase-coupled streptavidin (10 μg/ml in TBST containing 1% BSA). The beads were washed three times with 1 ml of alkaline phosphatase buffer (100 mM tris·HCl, pH 9.2; 100 mM NaCl and 5 mM $MgCl_2$) and kept in the alkaline phosphatase buffer prior to staining. A portion of the beads was transferred to a petri dish and the alkaline phosphatase buffer was replaced with 200 μl of a solution containing 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro blue tetrazolium (NBT). Color development was observed under a low-power microscope. The staining was terminated by washing the beads twice with 200 μl of sodium EDTA solution (20 mM, pH 7.4). The colored beads were picked out manually under the microscope for decoding.

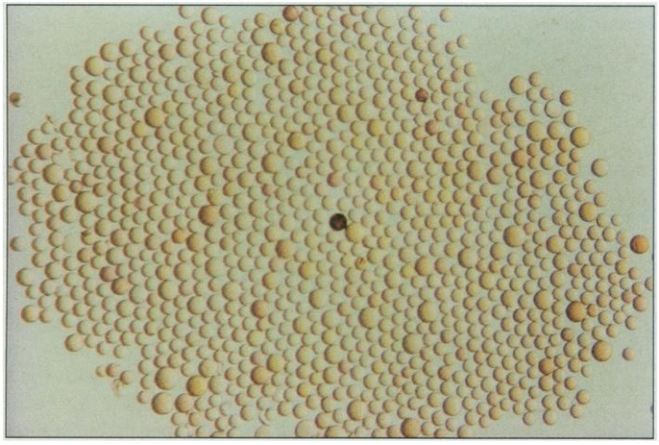


Fig. 4. A portion of the beads in the library after 5 min of staining. The dark bead in the center was identified as a hit. This level of contrast shown in the photograph was representative [60X].

REFERENCES AND NOTES

1. A. Varki, *Glycobiology* **3**, 97 (1993).
2. N. Sharon and H. Lis, *Sci. Am.* **268**, 82 (January 1993).
3. K. D. Janda, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10779 (1994).
4. A. D. Ellington and J. W. Szostak, *Nature* **346**, 818 (1990); C. Tuerk and L. Gold, *Science* **249**, 505 (1990); K. S. Lam *et al.*, *Nature* **354**, 82 (1991); R. A. Houghten *et al.*, *ibid.*, p. 84; L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, J. J. Toole, *Nature* **355**, 564 (1992).
5. E. M. Gordon, R. W. Barrett, W. J. Dower, S. P. A. Fodor, M. A. Gallop, *J. Med. Chem.* **37**, 1385 (1994); L. A. Thompson and J. A. Ellman, *Chem. Rev.* **96**, 555 (1996).
6. One attempt to form glycosidic linkages in the construction of a library has been reported. The glycosylations were carried out in solution on unprotected sugars to produce a mixture of carbohydrates. O. Kanie *et al.*, *Angew. Chem. Int. Ed. Engl.* **34**, 2720 (1995).
7. K. Toshima and K. Tatsuta, *Chem. Rev.* **93**, 1503 (1993).
8. L. Yan, C. M. Taylor, R. Goodnow Jr, D. Kahne, *J. Am. Chem. Soc.* **116**, 6953 (1994).
9. For other work on polymer-supported carbohydrate synthesis; see: J. M. J. Frechet, in *Polymer-supported Reactions in Organic Synthesis*, P. Hodge and D. C. Sherrington, Eds. (Wiley, New York, 1980), pp. 407–434; R. Verduyn, P. A. M. van der Klein, M. Douwes, G. A. van der Marel, J. H. van Boom, *Tecl. Trav. Chim. Pays-Bas* **112**, 464 (1993); S. J. Danishefsky, K. F. McClure, J. T. Randolph, R. B. Ruggeri, *Science* **260**, 1307 (1993); M. Schuster, P. Wang, J. C. Paulson, C. H. Wong, *J. Am. Chem. Soc.* **116**, 1135 (1994); S. P. Douglas, D. M. Whitfield, J. J. Krepinsky, *ibid.* **117**, 2116 (1995); J. Rademann and R. R. Schmidt, *Tetrahedron Lett.* **37**, 3989 (1996). J. Hunt and W. Roush, *J. Am. Chem. Soc.* **118**, 9998 (1996).
10. T. Osawa, T. Irimura, T. Kawaguchi, *Methods Enzymol.* **50**, 367 (1978); A. M. Wu, E. A. Kabat, F. G. Gruezo, H. J. Allen, *Arch. Biochem. Biophys.* **204**, 622 (1980).
11. A. Furka, F. Sebestyen, M. Asgedom, G. Dibo, *Int. J. Peptide Protein Res.* **37**, 487 (1991).
12. Encoding and decoding were essentially as described in H. P. Nestler, P. A. Bartlett, W. C. Still, *J. Org. Chem.* **59**, 4723 (1994); M. H. J. Ohlmeyer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10922 (1993).
13. Y. C. Lee and R. T. Lee, *Acc. Chem. Res.* **28**, 321 (1995); L. L. Kiessling and N. L. Pohl, *Chem. Biol.* **3**, 71 (1996).
14. R. L. Schnaar, P. H. Weigel, M. S. Kuhlenschmidt, Y. C. Lee, S. Roseman, *J. Biol. Chem.* **253**, 7940 (1978); R. Roy and C. A. Laferriere, *Carbohydr. Res.* **177**, c1 (1988); A. Spaltenstein and G. M. Whitesides, *J. Am. Chem. Soc.* **113**, 686 (1991); G. D. Glick and J. R. Knowles, *ibid.*, p. 4701; J. E. Kingery-Wood, K. W. Williams, G. B. Sigal, G. M. Whitesides, *ibid.* **114**, 7303 (1992); W. Spevak *et al.*, *ibid.* **115**, 1146 (1993); R. Roy, W. K. C. Park, Q. Wu, S. Wang, *Tetrahedron Lett.* **36**, 4377 (1995).
15. K. H. Mortell, R. V. Weatherman, L. L. Kiessling, *J. Am. Chem. Soc.* **118**, 2297 (1996).

16. Supported by the Office of Naval Research and a James A. Shannon Director's Award from the National Institutes of Health (to D.K.) and the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation (DRG-1365) (to J.L.).

9 July 1996; accepted 1 October 1996