Chem. Soc. Rev., 1989, **18**, 347–374

RHÔNE-POULENC LECTURE*

The Origin of the Specificity in the Recognition of Oligosaccharides by Proteins

By R. U. Lemieux, F.R.S.
DEPARTMENT OF CHEMISTRY, UNIVERSITY OF ALBERTA, EDMONTON, ALBERTA, CANADA T6G 2G2

The rich mosaic of carbohydrate structures which occur as complex oligosaccharides in covalent attachment to lipids and proteins both at cell surfaces and in biological fluids play key roles both in the maintenance of health and the establishment of disease. The length of these oligosaccharides is normally less than 20 sugar residues. Nevertheless, the possible changes in configuration, in points of attachment, in states of oxidation or reduction, and in substitution provide a wide range of topographies which serve as an extremely large and sophisticated vocabulary to assist in the expression of the exquisite poetry we experience as life. Thus, the correct tailoring of oligosaccharides serves as the source of the intelligence required for the control and guidance of key events for healthy cellular development. These include embryogenesis, fertilization, neuronal development, hormonal activities, cell proliferation, and their organization into specific tissues. Remarkable changes in the cell-surface carbohydrates occur with tumour progression and these changes appear to be intimately involved with the dreaded state of metastasis. Normally, more than 70% of the antibodies obtained on immunization with whole cells are directed against oligosaccharide structures. Recognition of bacterial cell-wall polysaccharides in immune response to infection is a major contributor to survival of the host. Some of these oligosaccharides provide adhesion sites for the attachment of bacteria, viruses, and viral toxins. This appreciation of the ubiquitous importance of complex oligosaccharide structures to cell biology has largely developed over the past two decades and now represents perhaps the major frontier of molecular biology. Consequently, an understanding of how and why these structures are recognized by the receptor sites of antibodies, enzymes, and lectins has become a subject of major concern. This lecture will outline contributions that my laboratory has made to this intriguing field for inquiry. I firmly believe that important progress was made and that our findings are of significance to the general subject of molecular recognition for the triggering of physiological responses.

We became actively involved in this field about 15 years ago as the result of successes in the chemical synthesis of so-called human blood group antigenic determinants. This involvement arose because it had become known that the

*Delivered at the Annual Chemical Congress of the Royal Society for Chemistry, University of Hull, on 6 April, 1989.

structural features responsible for these antigenic properties are present in complex oligosaccharides which occur as building units of glycolipids and glycoproteins found in large abundance at the surface of red blood cells. The chemical synthesis of these extremely rare structures represented an important challenge and my first review of our progress was published in the *Chemical Society Reviews* in 1978.[1] By then, we had established methodologies for the provision of the terminal tri- and tetra-saccharide units of oligosaccharides responsible for the human ABO and Lewis blood groups. Thus, amounts adequate for the study of their conformational and immunochemical properties became available. Composite structures for two types of these blood group determinants are presented in Figure 1.

An important reason for presenting Figure 1 is to introduce the so-called Lewis b (Le^b) and Y tetrasaccharides since much of my presentation will be concerned with the binding of both these structures by a glycoprotein which was the fourth lectin to be isolated by Baker and associates[2] from the seeds of a West-Central African plant called *Griffonia simplicifolia* and now commonly referred to as GS4. It is to be noted that the structures differ only in the interchange of the linkage of the β-D-Gal and one of the α-L-Fuc units to the β-D-GlcNAc residue.

Figure 2 illustrates our general strategy for entry into the field of immunochemistry by way of the human blood determinants using the Lewis b tetrasaccharide as an example.[1,3] It is seen that oligosaccharides were synthesized as simple methyl glycosides to serve as inhibitors in the immunological assays used to examine the binding affinities of antibodies and lectins. The syntheses were also devised to provide artificial antigens and immunoadsorbents by synthesizing the oligosaccharide with the 8-methoxycarbonyloctyl group as the aglycon. Standard procedures served to convert the carboxymethyl group into acyl azide for reaction with amino groups of suitable carrier molecules. These matters were outlined in detail in previous lectures[1,4,5] and are mentioned again to emphasize that these syntheses provided the oligosaccharides in amounts not only adequate for the immunochemical studies but also for detailed conformational analyses based in ^{1}H-nuclear magnetic resonance spectroscopy,[6,7] The kind of data relevant to the conformational analyses is illustrated for the Lewis b determinant in Figure 3.[8] This Figure is presented to stress that we very early realized that these oligosaccharides possess well defined conformational preferences which likely are maintained in the complexes with proteins. Although nuclear Overhauser experiments clearly inferred a strong conformational prefer-

[1] R. U. Lemieux, *Chem. Soc. Rev.*, 1978, **7**, 423.
[2] S. Shibata, I. J. Goldstein, and D. A. Baker, *J. Biol. Chem.*, 1982, **257**, 9324.
[3] R. U. Lemieux, D. A. Baker, W. M. Weinstein, and C. M. Switzer, *Biochemistry*, 1981, **20**, 199.
[4] R. U. Lemieux, Proceedings of the 28th IUPAC Congress, 'Frontiers in Chemistry', ed. K. J. Laidler, Pergamon Press, Oxford, 1982, p. 3.
[5] R. U. Lemieux, VIIIth International Symposium on Medicinal Chemistry, Proceedings, Vol. I, Swedish Pharmaceutical Press, Stockholm, 1985, p. 329.
[6] R. U. Lemieux, K. Bock, L. T. J. Delbaere, S. Koto, and V. S. Rao, *Can. J. Chem.*, 1980, **58**, 631.
[7] H. Thøgersen, R. U. Lemieux, K. Bock, and B. Meyer, *Can. J. Chem.*, 1982, **60**, 44.
[8] U. Spohr and R. U. Lemieux, *Carbohydr. Res.*, 1988, **174**, 211.

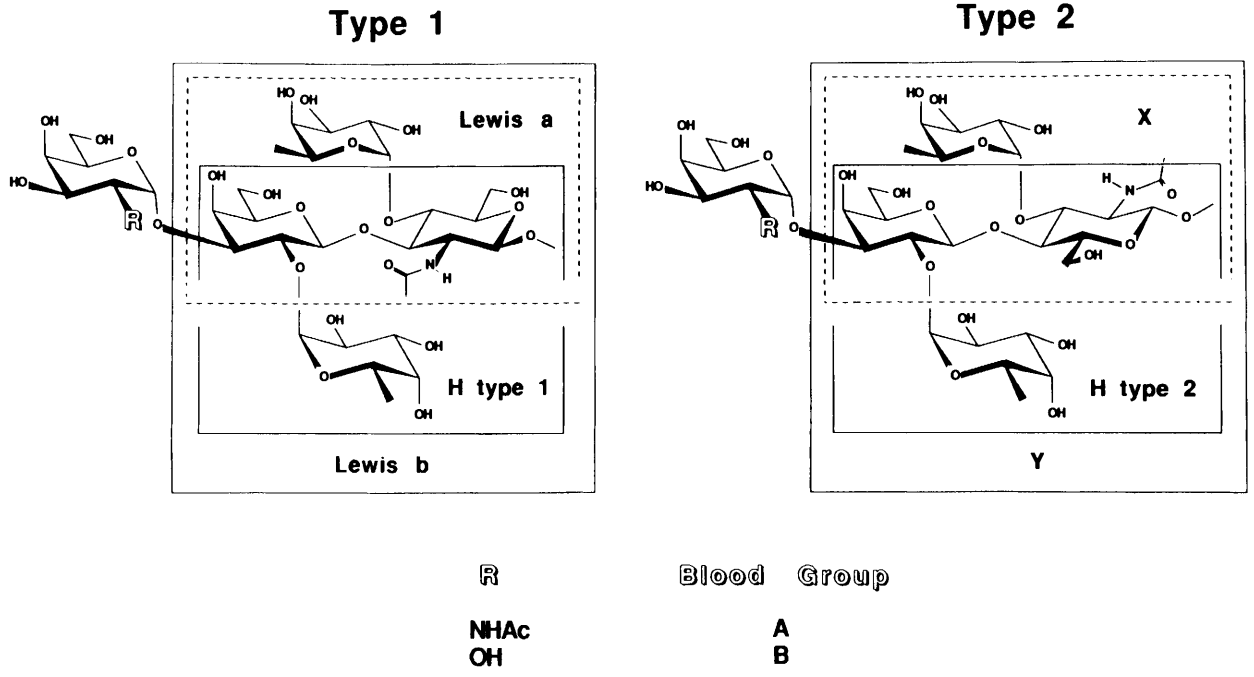
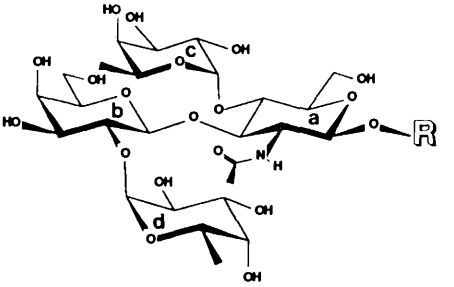


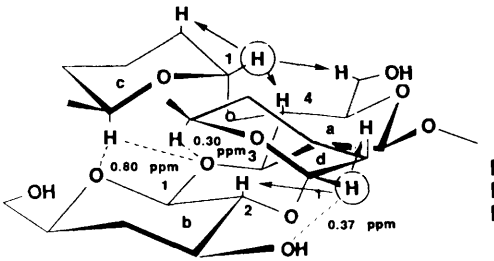
Figure 1 *Composite presentation of ABH and Lewis human blood group antigenic determinants. Fucosylation of the* β-D-Gal *unit of the core disaccharides provides the H type* 1 *and H type* 2 *trisaccharides which are the biological precursors to the A, B, Lewis b, and Y determinants. The Lewis a and X trisaccharides arise from fucosylation of the* β-D-GlcNAc *unit of the type* 1 *and type* 2 *core disaccharides, respectively*



Lewis b

ℝ	
-Me	Inhibitor
$-(CH_2)_8CONH$-Protein	Artificial antigen
$-(CH_2)_8CONH(CH_2)_3$-$Si(OEt)_2$-Silicate	Immunoadsorbent

Figure 2 *Illustration, using the Lewis b tetrasaccharide, of the products achieved by chemical synthesis and necessary to the investigation*[1]



Torsion Angles

ϕ^H	ψ^H
H-1^b/C-3^a = 55°	C-1^b/H-3^a = 10°
H-1^c/C-4^a = 55°	C-1^c/H-4^a = 25°
H-1^d/C-2^b = 45°	C-1^d/H-2^b = 15°

..... **Specific interunit deshielding**

○ ⟶ **Nuclear Overhauser enhancement**

Figure 3 *Outline, using a partial formula for the Lewis b determinant as an example, of the procedures used to establish the conformational preferences of oligosaccharides by* ^{1}H-*n.m.r. spectroscopy*[7]

ence, the strong interunit deshieldings of specific hydrogen atoms was most convincing especially since the nOe conformer has only these hydrogen atoms in van der Waals' contact with oxygen atoms on different sugar units. It is for these reasons that we have consistently attempted to rationalize our oligosaccharide-to-protein binding data on the basis of the ^{1}H-n.m.r. conformer.

Primarily because we had at our disposal artificial antigens that could serve as

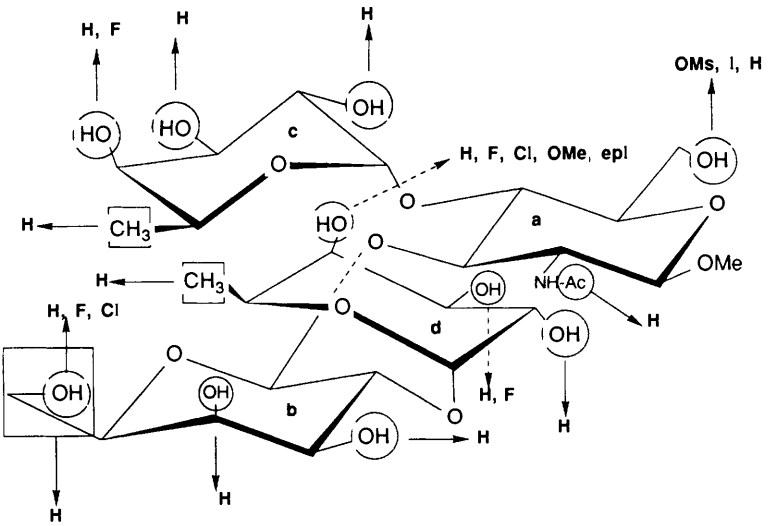


Figure 4 *Some of the modified Lewis b determinants synthesized in order to probe the combining sites of the lectin 4 of* Griffonia simplicifolia *and a hybridoma monoclonal anti-Le*^{b} *antibody*[10]

$$[\text{Oligosaccharide}\cdot x\text{H}_2\text{O}] + [\text{Protein}\cdot y\text{H}_2\text{O}]$$

$$\Big\downarrow \uparrow K_{\text{Assoc}}$$

$$[\text{Oligosaccharide}\cdot\text{Protein}\cdot z\text{H}_2\text{O}] + n\text{H}_2\text{O}$$

where $x + y = z + n$ and the solvent is water

Figure 5 *The problem is to achieve an appreciation of the origin of the specific interactions of oligosaccharides with proteins; i.e. antibodies, lectins, and enzymes. The solution requires an understanding of the role played by water in addition to the interactions between the protein and the oligosaccharide*

monospecific reagents for the detection of immunological activities and monospecific immunoadsorbents for the affinity purification of antibodies and lectins, we were in an excellent position by 1978[1] to become involved in strategies that could lead to an understanding of why certain protein structures display high affinities for the noncovalent binding of specific oligosaccharides. Our synthetic abilities provided the key to these investigations since, as displayed in Figure 4 and using the Lewis b tetrasaccharide as example, synthesis would make available not only the natural structure but a wide range of slightly chemically modified structures which could be used to probe the combining sites of antibodies and lectins that recognize the natural determinant.[8–10] Our main effort in this regard was to prepare all the monodeoxy derivatives of a given oligosaccharide epitope as is illustrated for the Lewis b tetrasaccharide in Figure

[9] U. Spohr, O. Hindsgaul, and R. U. Lemieux, *Can. J. Chem.*, 1985, **63**, 2644.
[10] R. U. Lemieux, R. Cromer, and U. Spohr, *Can. J. Chem.*, 1988, **66**, 3083.

4. *Nor*-derivatives were also prepared by substituting hydrogen for the terminal carbon of hexose units. Some of the results obtained with the deoxy compounds suggested that certain hydroxyl groups were incorporated into the complex intramolecularly hydrogen bonded for nonpolar interaction with the protein. In order to better examine such possibilities, a number of deoxy-halogeno derivatives were also prepared.

As indicated in Figure 5, the binding reactions are reversible. The objective in our studies was to determine the association equilibrium constants (K_{Assoc}) and, by determining this constant for a given reaction at different temperatures in the range 10—40 °C, achieve useful estimates of the changes in enthalpy and entropy which occur for equilibrium to be achieved. The method based on the change in ultraviolet absorption with complex formation[9] proved useful in this regard.

We also determined the amount of an inhibitor which provides 50% inhibition of the binding of an artificial antigen by an antibody or lectin using a solid-phase competitive radioimmunoassay (RIA) by labelling the artificial antigen with iodine-125.[9] This procedure provided a very convenient method to determine the relative potencies of two inhibitors. As pointed out by Pauling and Pressman,[11] the differential change in free energy for the binding reaction can be calculated from the 50% inhibition data. It was very comforting to find that the $\Delta\Delta G°$ values obtained in this way corresponded well to those determined from ultraviolet absorption data.[12]

The inhibition studies led to the following general conclusions.

1 Key Polar Interactions

Invariably, a number of the hydroxyl groups of an oligosaccharide proved outstandingly essential to complex formation. Most often these hydroxyl groups occur in clusters of two or three hydroxyl groups and I have termed these clusters the key polar groupings.[5] However, it is now established that the binding of methyl *N*-acetyl-β-lactosaminide by the monoclonal myeloma antibody known as anti-I Ma involves only one key hydroxyl group[13,14] and the binding of the H-type 2 trisaccharide to the newly discovered lectin of *Galactia tenuiflora*[15] involves recognition of a key polar grouping provided by four hydroxyl groups.[16] In all our studies, the results require the binding to occur in clefts at the surface of the proton and the key polar interaction to be within the cleft. Quiocho[17] has shown that the binding of L-arabinose by the L-arabinose-binding protein almost completely encloses the sugar and leads to involvement of all the hydroxyl groups in polar interaction with the protein. These findings are not surprising since the energy required to isolate an hydroxyl group within an hydrophobic environment would normally be prohibitively high.

[11] L. Pauling and D. Pressman, *J. Am. Chem. Soc.*, 1945, **67**, 1003.
[12] U. Spohr, M. Bach, and R. U. Lemieux, Abstracts of the XIIIth International Carbohydrate Symposium, Ithaca, New York, 1986, p. 353.
[13] R. U. Lemieux, T. C. Wong, J. Liao, and E. A. Kabat, *Molecular Immunology*, 1984, **21**, 751.
[14] R. U. Lemieux, P. Nikrad, U. Spohr, E. Petrakova, J. Liao, and E. A. Kabat, to be published.
[15] LePendu, G. Gérard, F. Lambert, R. Mollicone, and R. Oriol, *Glycoconjugate J.*, 1986, **3**, 203.
[16] R. Cromer, U. Spohr, J. LePendu, R. Oriol, and R. U. Lemieux, to be published.
[17] F. A. Quiocho, *Ann. Rev. Biochem.*, 1986, **55**, 287.

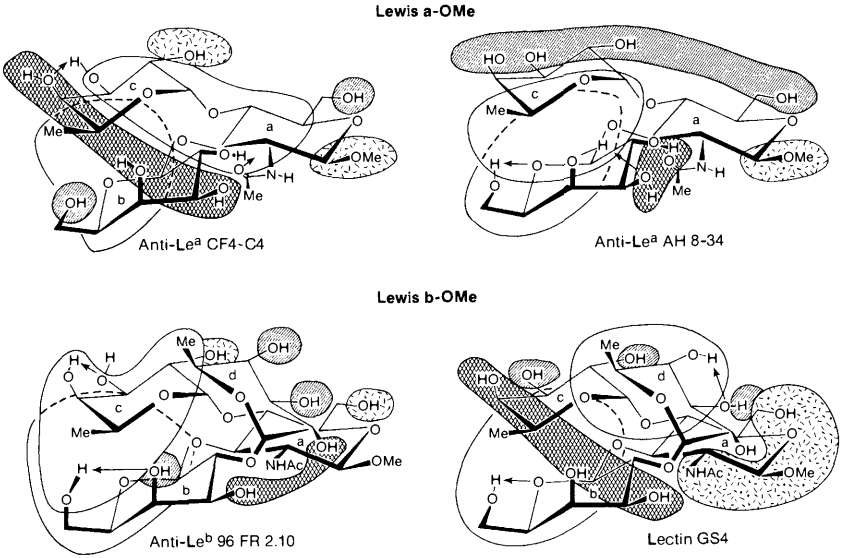


Figure 6 *The conclusions reached in studies of the binding properties of two different monoclonal anti-Lewis a antibodies,*[19] *a monoclonal anti-Lewis b antibody,*[18] *and the lectin* 4 *of* Griffonia simplicifolia *(GS4).*[8–10] *The cross-hatched regions identify the key polar interactions, the hatched regions indicate hydroxyl groups involved in relatively unessential polar interactions near the peripheries of the combining sites, and the speckled regions those portions of the oligosaccharides that remain in contact with the aqueous phase. The regions circled, above and below the oligosaccharide, indicate the surfaces that appear to become involved in nonpolar interactions with the protein (see Figure 7)*

As displayed in Figure 6, the key polar grouping appears most often presented by hydroxyl groups from different sugar units.[9,18,19] This finding is of major significance to efforts designed to establish the specificity of a protein in the sense that strong binding requires recognition of the full key polar grouping which most often will not be provided by a simple monosaccharide unit. In view of the paucity of complex oligosaccharides that are available for such studies, the establishment of the full specificity possessed by a given protein remains a fortuitous experience. From Nature's point of view, the distribution of key hydroxyl groups over more than one sugar unit provides a fine means to establish a high level of fidelity in the recognitions offered by oligosaccharides for the guidance of living processes. Bock and co-workers[20] have discovered that the key polar grouping required for the hydrolysis of maltosides by the enzyme glucoamylase involves hydroxyl groups of both glucose units (OH-4, OH-4′, and OH-6′). It is to be expected, I believe, that multipoint key polar interactions will prove to be of importance throughout

[18] U. Spohr, N. Morishima, O. Hindsgaul, and R. U. Lemieux, *Can. J. Chem.*, 1985, **63**, 2659.
[19] R. U. Lemieux, O. Hindsgaul, P. Bird, S. Narasimhan, and W. W. Young, Jr., *Carbohydr. Res.*, 1988, **178**, 293.
[20] K. Adelhorst, K. Bock, H. Pedersen, and S. Refn, *Acta Chem. Scand.*, 1988, **B42**, 196.

molecular biology for ensuring high selectivity in noncovalent molecular associations.

2 Peripheral Polar Interactions

As mentioned above, the substitution of only one of the key hydroxyl groups of an oligosaccharide by an hydrophobic atom such as hydrogen or chlorine can result in virtually complete loss of the affinity by the protein. On the other hand, we have also observed that similar substitutions of other hydroxyl groups can have relatively much smaller effects on the affinity, the differential changes in free energy being less than one kcal $mole^{-1}$. In fact, in many instances, the potency for binding is actually increased. As illustrated in Figure 6, these hydroxyl groups appear to occur near the periphery of the combining site. Consequently, it could be imagined[19] that the interaction with a polar group offered by an amino acid unit at near the periphery of the combining site may not be of critical importance to the stability of the complex. That is, an only slight distortion of the combining site may be required for the polar group of the amino acid to remain in contact with water. The energy required for the distortion could be, in part, compensated by the 'micro' hydrophobic effect introduced as a consequence of the deoxygenation.

3 Lipophilic Interactions

Definitive evidence for the participation of large surfaces which are lipophilic in character was first obtained in studies of the binding of methyl *N*-acetyl-β-lactosaminide by the monoclonal antibody anti-I Ma.[13] The lipophilic regions arise because of the presence in each sugar unit of an oligosaccharide of hydrogens bonded to carbon and the ethereal oxygens which interact relatively weakly with water. Also, it appears that intramolecular hydrogen bonding can be invoked in order to render a polar region more lipophilic in character.[9,21] This idea is discussed in more detail below.

The conclusions reached as to the regions of oligosaccharides which come into interaction with nonpolar parts of the combining site are depicted in Figure 6 for proteins that bind the Lewis a and Lewis b human blood group determinants. The results obtained in studies of the affinities of the anti-H type 2 lectin I of *Ulex europaeus*[22] and the anti-I Ma[13] and anti-B[23] antibodies were very similar in kind. It is to be noted that in each case, the proposal was made that certain hydroxyl groups are accepted intramolecularly hydrogen bonded. The experimental reason for making this proposal is that, in each case, the replacement of such an hydroxyl group by hydrogen enhanced the strength of the bonding reaction. The theoretical basis for the proposal is that, whenever this enhancement was observed, an atom which can serve as proton acceptor was sterically well disposed relative to the hydroxyl group for the formation of an intramolecular

[21] R. U. Lemieux, P. H. Boullanger, D. R. Bundle, D. A. Baker, A. Nagpurkar, and A. Venot, *Nouv. J. Chim.*, 1978, **2**, 321.
[22] O. Hindsgaul, D. P. Khare, M. Bach, and R. U. Lemieux, *Can. J. Chem.*, 1985, **63**, 2653.
[23] R. U. Lemieux, A. P. Venot, U. Spohr, P. Bird, G. Mandal, N. Morishima, O. Hindsgaul, and D. R. Bundle, *Can. J. Chem.*, 1985, **63**, 2664.

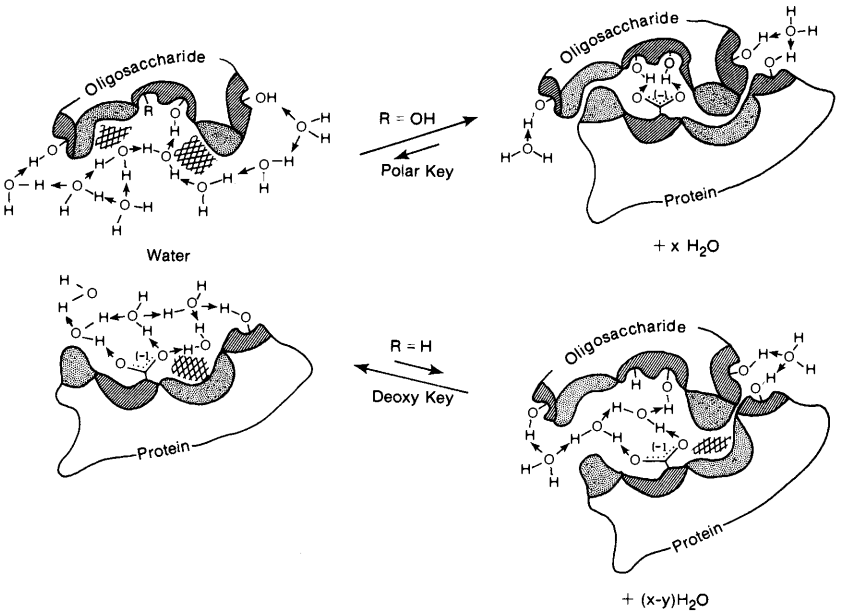


Figure 7 *The speckled and hatched regions at the surface of the oligosaccharide and the combining site of the protein represent lipophilic and polar regions, respectively. The cross-hatched regions indicate the possible existence of void space that becomes filled on complex formation*

hydrogen bond. Experimental support for this postulation is provided by the observation that carbohydrate structures form such energetically favourable intramolecular hydrogen bonds for acceptance into aprotic weakly polar solvents.[24,25] Of course, the charge delocalization that results from intramolecular hydrogen bonding does not provide a nonpolar grouping. Both the proton donating and accepting oxygen atoms remain capable, although more weakly, of interacting with water. The question then arose, 'What is the origin of the force that would cause polar groupings to leave water and become intramolecularly hydrogen bonded in order to enter a nonpolar region of the combining site?' Part of the answer may be found in that the hydrogen-bonded structures always appear to be near the periphery of the combining site and, therefore, may remain partially bonded to water. Nevertheless, it was the notion of intramolecular hydrogen bonding for acceptance into nonpolar regions of a combining site that led to the question, 'How do water molecules interact with amphiphilic surfaces?' My postulations[5] in this regard are schematically illustrated by Figure 7. An examination of a molecular model for an oligosaccharide will show that the surfaces of these structures are made up of nonpolar and polar regions. Thus, the

[24] R. U. Lemieux and A. A. Pavia, *Can. J. Chem.*, 1969, **46**, 1453.
[25] R. U. Lemieux and J. T. Brewer, 'Carbohydrate in Solution', Advances in Chemistry Series No. 117, ed. R. G. Gould, American Chemical Society, Washington, D.C., 1973, p. 121.

oligosaccharide can be considered to present amphiphilic surfaces for solvation by water. Of course, the polar regions of the oligosaccharides will interact much more strongly with the water and, being Coulombic in nature, these polar interactions will tend to orientate the water molecules in certain specific directions depending on the orientation of the hydroxyl groups. Rotation about the C–OH bond of an isolated hydroxyl group must be a very low energy process and, except for restrictions imposed by bonding with water, be essentially free. Therefore, important interaction of a water molecule with the hydroxyl group in all three staggered orientations must normally be expected and, in the case of an hydroxyl group adjacent to a nonpolar grouping, the interaction may favour, on the average, orientation of the water molecule in directions away from the adjacent nonpolar groupings for which it has relatively much weaker attraction. Therefore, I anticipated that the hydration of the amphiphilic surface of an oligosaccharide may well effectively lead to the presence of void spaces over the nonpolar region.[5] These voids would, of course, be continuously fluctuating in size but could serve to provide a driving force, in addition to dispersion forces of attraction and dipole–dipole interactions, including hydrogen bonding, for complex formation. As will be mentioned below, this idea found support in Monte Carlo simulations for the hydration of the Le^b-OMe tetrasaccharide performed by Helmut Beierbeck.[26]

Figure 7 schematically presents what I have termed the 'hydrated polar gate effect'.[5] The polar gate is represented by a carboxylate group because this negatively charged group can provide a strong polar interaction with both water molecules and hydroxyl groups of the substrate as was actually demonstrated by Fersht *et al.*[27] using an enzyme for which individual amino acids were altered. We have now found, as we shall see later, that an aspartate unit provides part of the polar gate for the binding of the Lewis b tetrasaccharide by the lectin 4 of *Griffonia simplicifolia*.

Both the oligosaccharide and the combining site of the protein are presented hydrogen-bonded to water molecules and, as discussed above, it is inferred that the directionality required for strong hydrogen bond formation will cause water molecules, on the average, not to be in van der Waals' contact with the nonpolar regions of the oligosaccharide. Since the oligosaccharide is conformationally rather rigid, this unstable condition is not expected to be corrected by conformational change. Similar consideration applies to the combining site and 'voids' are expected to exist along the amphiphilic surfaces presented by the protein. On this basis, should the two amphiphilic surfaces be so complementary that their interaction effectively reduces the void spaces, an important contribution to the driving force for complexation would be established. A further consideration in this regard is that the association of the complementary surfaces may lead to the establishment of stronger intermolecular

[26] H. Beierbeck and R. U. Lemieux, *Can. J. Chem.*, submitted.
[27] A. R. Fersht, J.-P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, and G. Winter, *Nature (London)*, 1985, **314**, 235.

Metals —Ca^{2+} and Mn^{2+} [30]
Carbohydrates—Three asparagine-linked side chains [29]
Protein —Molecular weight of protein 53 394 [29]
—Crystalline, 52% water [30]
—Two noncovalently linked subunits [2]
Subunits —Differ in molecular weight [2]
—Each as 243 amino acid residues of which five are different [29]
—One has the carbohydrate at positions 5 and 18, the other only at 18 [29]
Affinity —Le^b and Y tetrasaccharides [9]

	K_{Assoc}	ΔG (kcal mole $^{-1}$, 298 °C)
Le^b-OMe	4.4×10^4	−6.3
Y-OMe	2.0×10^4	−5.9

Figure 8 *Some properties of the fourth lectin isolated from the seeds of the West-Central African plant* Griffonia simplicifolia [2]*—a dimeric metalloglycoprotein*

dispersion forces of attraction [28] than existed when the interactions were with water molecules.

On the basis of the scheme presented in Figure 7, the replacement of a key hydroxyl group by hydrogen leads to loss of affinity for complex formation because, as illustrated, due to the loss of an intermolecular hydrogen bond there no longer exists an energetically acceptable mechanism for dehydration of the gate. From this point of view, the contribution of *intermolecular* hydrogen bonding to the stabiliity of the complex is largely derived from the release of water molecules from the amphiphilic surfaces for the formation of stronger hydrogen bonds in the bulk water. In other words, I am not aware of a reason to believe that the *intermolecular* hydrogen bonds within the complex, taken together, are stronger than the hydrogen bonds that existed, prior to complex formation, between both the oligosaccharide and the protein with the surrounding water molecules.

Some of the properties of the above-mentioned lectin known as GS4 are reported in Figure 8. It is seen to be a metalloglycoprotein. As reported by Shibata and co-workers [2] the protein exists as a dimer and each subunit has now been found [29] to have 243 amino acids. The dimer can be dissociated by sodium dodecyl sulphate for separation of the subunits by polyacrylamide gel electrophoresis.[2] Not surprisingly, the strongly bound metals are calcium and manganese.[30] The amino acid sequences of the two units have been established [29] in Larry Smillie's group at the University of Alberta and found to differ at five positions in each of the subunits. The crystal structure of the lectin was recently established to near 2.3 Å resolution and that of its complex with the Lewis b tetrasaccharide (Le^b-OMe) to 2.8 Å resolution by Louis Delbaere's group [30,31] at the University of Saskatchewan. Contrary to the first surmise,[30] the two subunits have proven to be crystallographically different.[31] This is not surprising since in addition to the difference in amino acid composition, one of the subunits proved

[28] R. U. Lemieux and K. Bock, *Arch. Biochem. Biophys.*, 1983, **221**, 125.
[29] P. Nikrad, R. U. Lemieux, J. Pearlstone, M. R. Carpenter, and L. B. Smillie, unpublished results.
[30] M. Vandonselaar, L. T. J. Delbaere, U. Spohr, and R. U. Lemieux, *J. Biol. Chem.*, 1987, **262**, 10848.
[31] L. T. J. Delbaere, M. Vandonselaar, L. Prasad, J. W. Quail, U. Spohr, and R. U. Lemieux, *American Crystallographic Association, Abstracts of Papers*, 1989, **17**, V5.

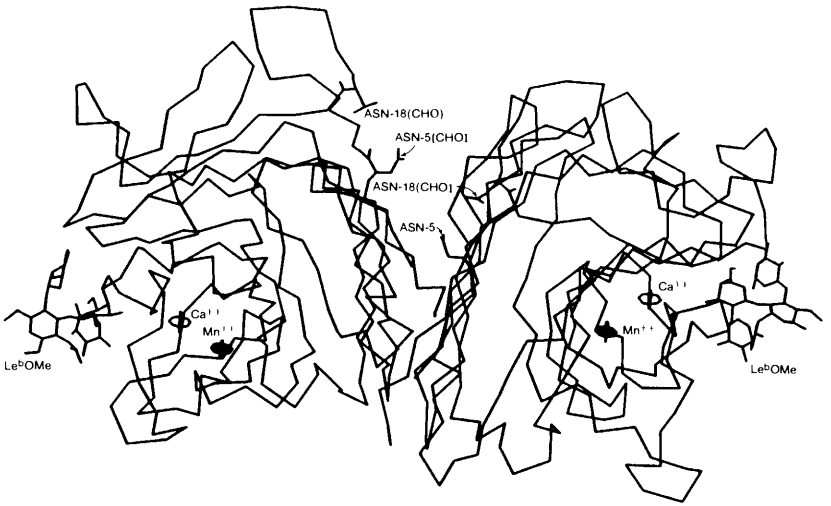


Figure 9 *Abbreviated X-ray crystallographic structure for the complex formed by the metalloglycoprotein* GS4 *with the Lewis b tetrasaccharide* (Le^b-OMe)[31]

to have oligosaccharides at asparagines 5 and 18 whereas the other subunit has oligosaccharide at only asparagine 18. The structures of these oligosaccharide units are not as yet known but together appear to be of the composition $[(Man)_3(GlcNAc)_2(Xyl)]_3(Fuc)_2$.[29] It is seen that at 25 °C, the lectin binds Le^b-OMe about twice as strongly as Y-OMe but this amounts to a differential change in free energy of only 0.4 kcal $mole^{-1}$.[9]

These findings[29–31] are presented in Figure 9. The specific asparagine unit (ASN-5) to which an oligosaccharide side-chain is bound, is not yet established. Nevertheless, it appears that it will reside at least partially in a groove between the two subunits. Thus, it seems probable that the recognition of this oligosaccharide ligand by the other subunit may provide the force of adhesion required for the dimerization.

Of major interest is that the complex of GS4 with Le^b-OMe also crystallized[31] and the crystals proved isomorphous to those of the native protein.[30] Moreover, isomorphous crystals were obtained with a wide variety of derivatives of Le^b-OMe.[30] Indeed, the positions of the combining sites were first located by examining a crystal of the complex with 6a-deoxy-6a-iodo-Le^b-OMe where it was found that the iodine atoms reside in water channels that separate the protein structure.[30] Later, the positions identified in Figure 9 were established through the differences in the diffraction patterns for the native GS4 and its complex with Le^b-OMe.

The interpretation of the *X*-ray diffraction data in terms of electron density maps for the GS4•Le^b-OMe complex was not completed prior to presentation of the lecture. Nevertheless, I decided to present our attempt to achieve a correlation between the structure of the native protein and the results we had

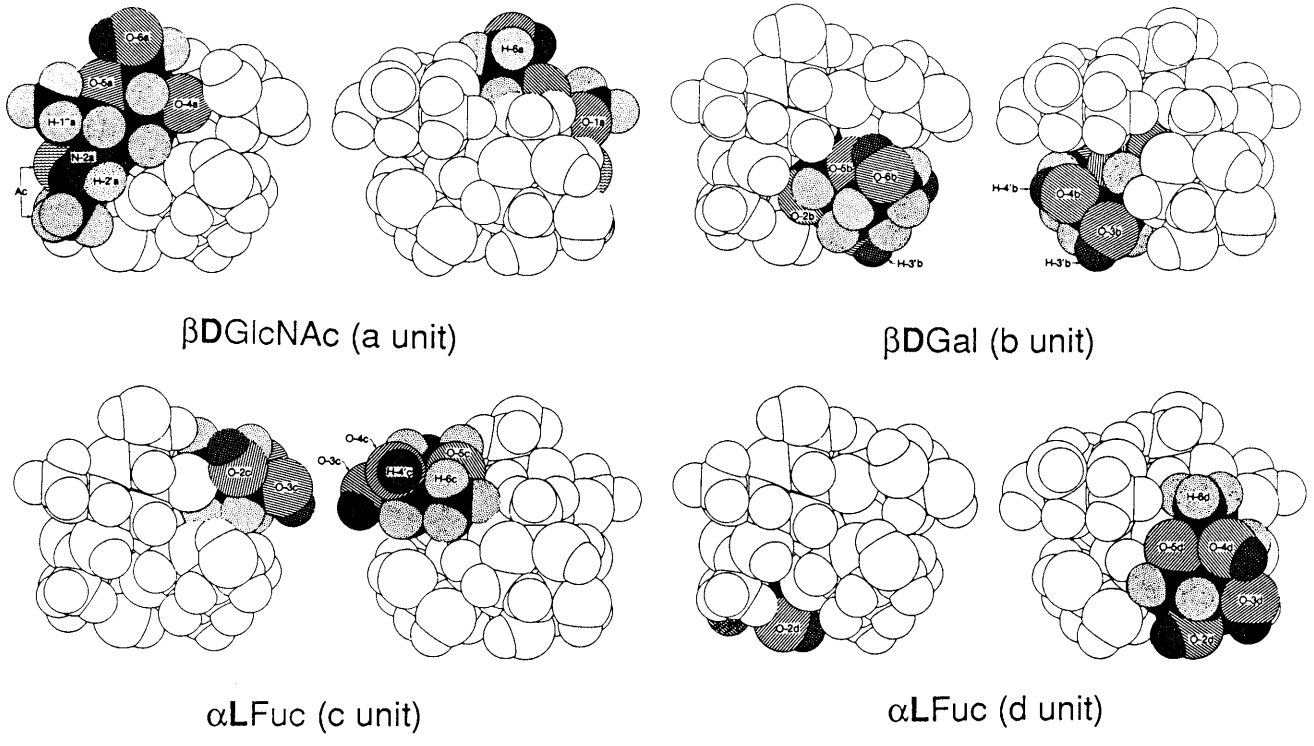


Figure 10 *Molecular model of the methyl glycoside of the Lewis b tetrasaccharide* (α-L-Fuc(1 ⟶ 2)β-D-Gal(1 ⟶ 3)[α-L-Fuc(1 ⟶ 4)]β-D-GlcNAc-OMe). *The cross-hatched atoms are hydrogens of hydroxyl groups. The oxygen atoms are hatched and hydrogen atoms bonded to carbon are speckled. The* NHC=O *portion of the acetamido group has the hydrogen marked with small bars and the carbonyl oxygen is identified by wavy lines. All carbon atoms are black. These designations will be used throughout the remainder of this presentation*

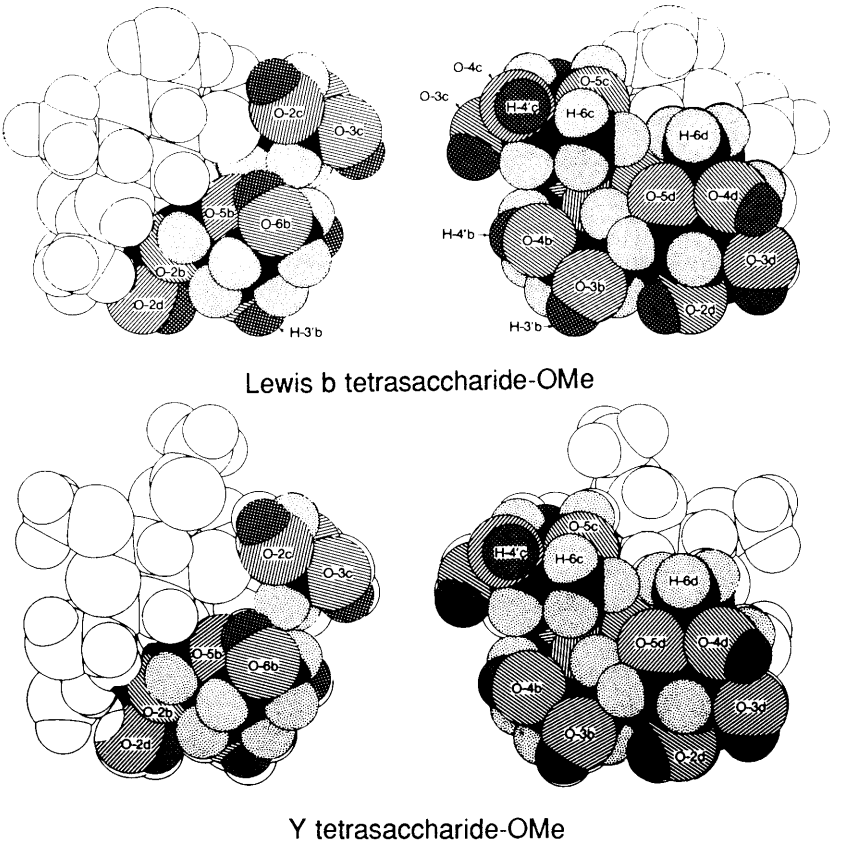
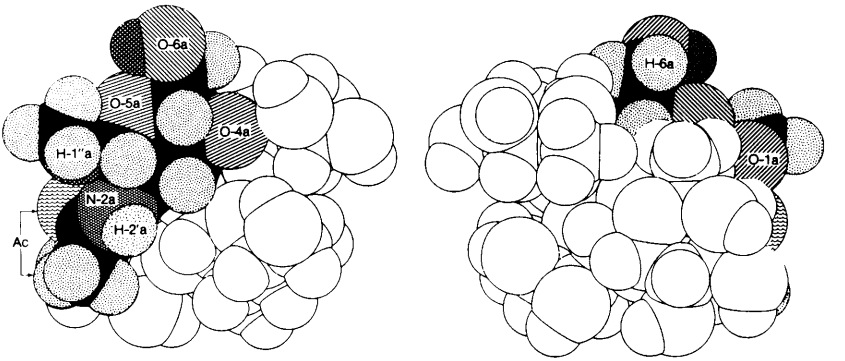


Figure 11 *The topographies held in common by the Lewis b and Y tetrasaccharides which are both bound by the lectin* GS4[9]

obtained[8–10] on probing the combining site with modified Le^b-OMe structures. For this purpose, it was assumed that the conformation of the backbone peptide chain does not change with complex formation and that the combining site recognizes the tetrasaccharide only by rotations about the C_α to C_β and C_β to C_α bonds of the amino acid residues that are involved. The conclusions reached in fact proved to be in good general agreement with the probing data. Nevertheless, errors and omissions were made which became apparent once the structure of the complex became known in May, 1989.[32]

However, prior to going into these matters in more detail, I wish to present hard evidence, based in probing experiments, for non-involvement of the β-D-

[32] L. T. J. Delbaere, M. Vandonselaar, L. Prasad, J. W. Quail, U. Spohr, and R. U. Lemieux, *Proceedings of the Canadian Federation of Biological Societies, Abstracts of Papers*, 1989, **32**. 305.



kcal mole^{-1}, 25 °C

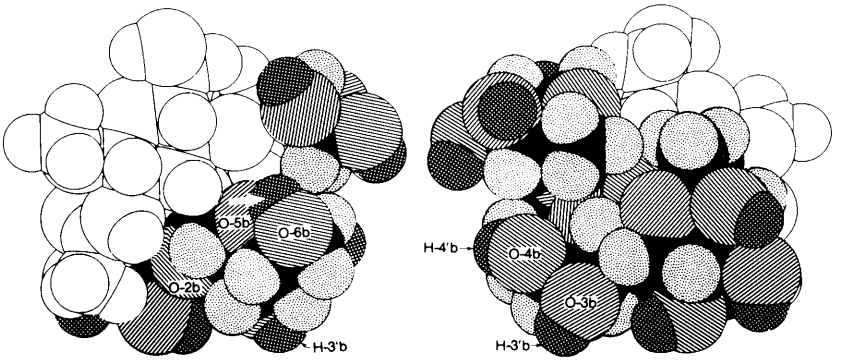
	$\Delta G°$	$\Delta\Delta G°$	$\Delta H°$	$\Delta\Delta H°$	$T\Delta S°$	$\Delta(T\Delta S°)$
Le^b-OMe	−6.3	0	−13.3	0	−7.0	0
Derivatives						
6a-Deoxy	−6.3	0	−13.1	0.2	−6.8	0.2
6a-*O*-Methanesulphonyl	−6.3	0	−12.3	1.0	−6.0	1.0
6a-Deoxy-6a-iodo	—	0	—	—	—	—
1a-*O*-8-Methoxycarbonyloctyl	(−6.5)	−0.2	—	—	—	—
2a-Amino-2a-deoxy	(−6.1)	0.2	—	—	—	—

Figure 12 *Thermodynamic data for the binding by* GS4 *of* Le^b-OMe *and of derivatives with changes in the* β-D-GlcNAc *a-unit*[8]

GlcNAc residue in the binding reaction. Although the *X*-ray structure appears to have rendered such data obsolete, it is to be recognized that *X*-ray structures can become available for only a small fraction of the receptor sites of interest to the health sciences. Consequently, it will be important to learn how to interpret properly the results obtained on probing combining sites with appropriately chosen substrates.

In order to present correlations between the structure of the lectin and the inhibition data, the computer-drawn CPK model of the Le^b-OMe tetrasaccharide, shown in Figure 10, is used since this allows the molecule to be viewed one side at a time and, thereby, allows a ready appreciation of the structural features under consideration. Furthermore, the shading of atoms in the space-filling models allows a convenient presentation of the kinds of topographical features that are recognized. The tetrasaccharide is presented in the conformation that is known to be preferred in aqueous solution.[7,8] The partial labelling of atoms is considered sufficient for the identification of all of the atoms of the four sugar units. The orientations of the hydroxyl-group hydrogens correspond to orientations that appear to conform with the interpretations placed on the results obtained in the probing experiments and are strictly tentative.

As mentioned above with reference to Figure 8, GS4 binds both the Le^b and Y determinants. Examination of the HSEA models[9] indicated that substantial topography is held in common and corresponds to the darkened areas shown in Figure 11. It was expected, therefore, that the β-D-GlcNAc unit would not be



kcal mole $^{-1}$, 25 °C

		$\Delta G°$	$\Delta\Delta G°$	$\Delta H°$	$\Delta\Delta H°$	$T\Delta S°$	$\Delta(T\Delta S°)$
Le^b-OMe		−6.3	0	−13.3	0	−7.0	0
Modification							
3b-Deoxy-	changes in	all					
4b-Deoxy-	the key	highly					
4c-Deoxy	polar	inactive					
4c-Deoxy-4c-fluoro	group						
6b-Deoxy-		−6.2	0.1	−8.7	4.6	−2.5	4.5
6b-Deoxy-6b-chloro		(−6.7)	−0.4	—	—	—	—
6b-Deoxy-6b-fluoro		(−6.4)	−0.1	—	—	—	—
6b-Deoxy-6b-*nor*-		inactive					

Figure 13 *Thermodynamic data for the binding by* GS4 *of* Le^b-OMe *and of derivatives with changes in the* β-D-Gal *b-unit.*[9,10] *Changes at the fourth position of the* α-L-Fuc *c-unit are included to complete the description of the key polar interaction*

involved in the binding and this expectation was confirmed both by the data presented in Figure 12[8] and the position of the iodine atom in the above-mentioned complex formed with 6a-deoxy-6a-iodo-Le^b-OMe.[30] As seen in Figure 9, this conclusion was confirmed by the *X*-ray structure of the crystalline GS4•Le^b-OMe complex.

As indicated in Figure 12, the replacement of the 6a-hydroxyl by hydrogen had no significant effect on the thermodynamic parameters. Similarly, mesylation of OH-6a did not significantly affect the binding. As expected, the molar amount of the 6a-deoxy-6a-iodo-Le^b-OMe required to reach 50% inhibition in the RIA was the same as that for the parent compound. The fact that the aglycon could be changed from CH_3O– to $CH_3CO_2(CH_2)_8O$– with little change in activity supports the conclusion, based on Figure 11, that the aglycon of the β-D-GlcNAc unit is also not involved in the binding. De-*N*-acetylation to form the 2a-amino compound also had little effect on binding and, clearly, the binding was not dependent on recognition of the acetamido group. The fact that the thermodynamic data for Le^b-OMe, 6a-deoxy-Le^b-OMe, and 6a-*O*-mesyl-Le^b-OMe compounds all agreed within experimental error was of major importance since it provided confidence in our ability to measure the thermodynamic parameters for the binding reactions to useful approximations.[8,12]

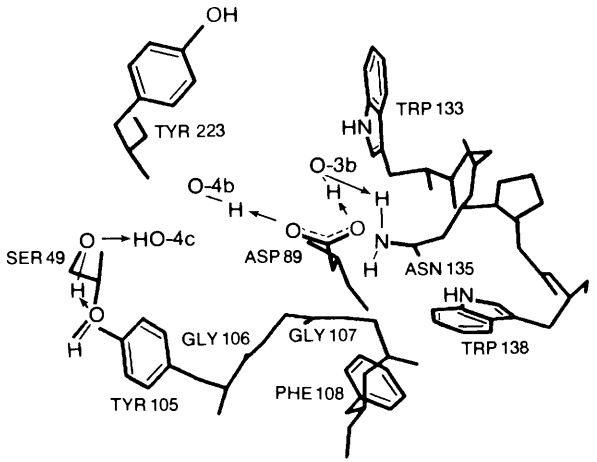


Figure 14 *The interaction of the polar gate provided by serine* 49, *aspartate* 89, *and asparagine* 135 *with the polar key provided by* OH-4b, OH-3b, *and* OH-4c.[31] *It is expected that immobilized water molecules will be found involved in the networks of hydrogen bonds*

In view of its non-involvement in the binding, the β-D-GlcNAc unit of Le^b-OMe will be left unshaded in the presentations of the CPK molecular model that follow.

The data in Figure 13 show that the β-D-Gal unit contributes two hydroxyl groups (OH-3b and OH-4b) to the key polar interaction. As was indicated earlier with reference to Figure 6, this interaction also includes OH-4c. Prior to further consideration of the probing data for the binding of Le^b-OMe by GS4, it is of interest to examine the nature of the combining site and the key polar interaction as was revealed by the *X*-ray crystal structure of the complex. The amino acid residues that surround the Le^b-OMe molecule in Figure 14 are viewed as if looking down through the β-D-GlcNAc residue. Thus, when viewed in stereo, the negatively charged carboxylate group of aspartate 89 is at the bottom of the shallow depression which forms the combining site. This charged group is undoubtedly the dominating portion of the site's polar gate and acts as proton acceptor in hydrogen bonding networks involving both OH-3 and OH-4 of the β-D-Gal b-unit. As indicated in Figure 14, OH-3b also serves as a proton acceptor from asparagine 135. The ease for the donation of the proton by ASN 135 is undoubtedly enhanced by the complexing of the carbonyl oxygen of its amide grouping with the Ca^{2+} ion. Also, the charge delocalization from the carboxylate of ASP 89 to O-3b as a result of the hydrogen bond will help strengthen the O-3 to NH hydrogen bond. At the time that we reported direct experimental evidence for this phenomenon, I termed it 'hydrogen-bond conjugation.'[33] However, the effect is currently referred to as 'hydrogen-bond cooperativity,'[34] a term which I have also adopted. Figure 14 shows OH-4b hydrogen-

[33] R. U. Lemieux and A. A. Pavia, *Can. J. Chem.*, 1969, **47**, 4441.
[34] J. E. Del Bene and J. A. Pople, *J. Chem. Phys.*, 1973, **58**, 3605.

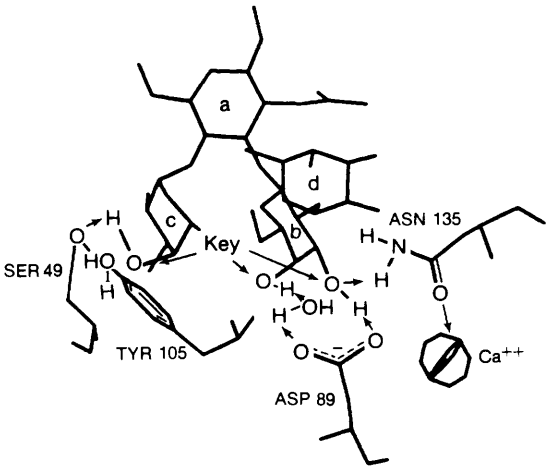


Figure 15 *The interaction of the polar gate with the polar key provided by* Le^b-OMe [31]

bonded to the aspartate carboxylate group. However, as will be discussed below, it seems most probable that once higher resolution is achieved, a water molecule will be found to bridge the two groups (see Figure 15). It appears definite that OH-4c becomes hydrogen bonded as a proton donor to the hydroxyl group of serine 49. The inactivity of the 4c-deoxy-4c-fluoro derivative (Figure 14) supports this conclusion that OH-4c donates a proton. The serine hydroxyl group, in turn, appears to donate a proton to the hydroxyl of tyrosine 105. It is likely that further water molecules will be found to be involved in the network of hydrogen bonds that form the key to gate polar interaction for the establishment of a specific complex. It is to be noted that these polar interactions become surrounded by hydrophobic groupings. Thus, the low dielectric will favour strong polar interactions in a region protected from exchanges with water. The indications are that the polar gate is more exposed to water prior to complex formation and that substantial reorientations, especially of the two tyrosine and the serine residues, occur in the course of the reaction. Tyrosine 105 should be especially prone to such a displacement since it is bridged to phenylalanine 108 by way of a diglycine unit. Glycine units are well known to provide 'hinges' for facile changes of the conformations of peptide chains. The engagement of the key polar group of the Le^b-OMe tetrasaccharide in the complex to the extent known to date is presented in Figure 15.

Figure 13 also presents the effects on binding of modifications at C-5 of the β-D-Gal b-unit. It is seen that replacement of the 6b-hydroxyl by hydrogen had no significant effect on the difference in binding free energy. Consequently, it was at first considered that this hydroxyl group may not be involved in complex formation.[9] However, once the thermodynamic parameters became available, it was evident that the hydroxyl group is in fact involved since the differential changes in enthalpy and entropy were very substantial; that is, each about 4.5

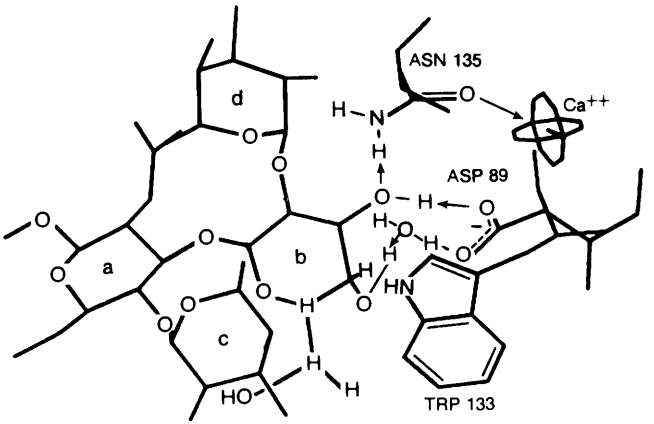
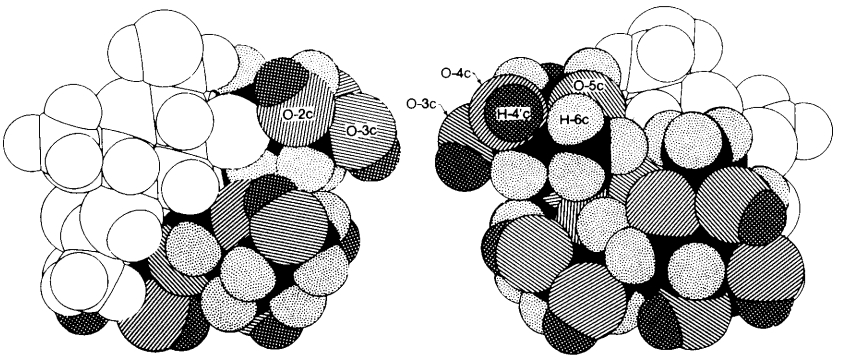


Figure 16 *The involvement of the* β-D-Gal *b-unit in the binding reaction*[31]

kcal mole^{-1}.[12] These observations suggested that the 6b-hydroxyl is accepted into a nonpolar region of the combining site by becoming intramolecularly hydrogen bonded to O-5b. In order to test this hypothesis, the 6b-chloro and 6b-fluoro derivatives were prepared and, as seen in Figure 13 and in accord with the intramolecular H-bonding theory, these compounds are both bound somewhat more strongly than either Le^b-OMe or its 6b-deoxy derivative.[10] In view of this evidence that the 6b-hydroxymethyl group comes into nonpolar interaction with the protein, it was to be expected that its replacement by hydrogen to form an α-L-arabinopyranosyl unit would have an important deleterious effect on the binding. Indeed, as seen in Figure 13, the loss in complementarity results in a virtually inactive compound.

In addition to the involvement of OH-3b and OH-4b in the key polar interaction, the β-D-Gal b-unit comes into van der Waals' contact with the indolyl group of tryptophan 133 (see Figure 16). In view of the data presented in Figure 13, interaction with an hydrophobic group of the protein was to be expected. Surprisingly, however, the 2.8 Å structure appears to place the plane of the indolyl group approximately in the plane of the atoms that form the pyranose ring and near the hydrocarbon H-4 and H-5 atoms of the β-D-Gal unit. The resolution does not allow the drawing of firm conclusions as to the complementarity that is involved. Quite evidently the edge of the indolyl group may dock within the groove presented with H-3 and H-5 on one side and H-4 and an H-6 on the other. Both the aromatic and NH group of the indole unit appear to be too removed from polar groups of the β-D-Gal unit for direct intermolecular hydrogen bonding.[35] Intramolecular hydrogen bonding of OH-6b with O-5b was previously invoked[9,10] to rationalize the data presented in Figure 13. It now appears that the orientation of OH-6b which would allow such a bond is

[35] M. Levitt and M. F. Peretz, *J. Mol. Biol.*, 1988, **201**, 751.



kcal mole^{-1}, 25 °C

	$\Delta G°$	$\Delta\Delta G°$	$\Delta H°$	$\Delta\Delta H°$	$T\Delta S°$	$\Delta(T\Delta S°)$
Le^b-OMe	6.3	0	−13.3	0	−7.0	0
Modification						
2c-Deoxy-	−5.7	0.6	−10.0	3.3	−4.3	2.7
3c-Deoxy-	−5.5	0.8	−6.6	6.7	−1.1	5.9
4c-Deoxy-Key	inactive					
6c-*Nor*	−4.9	1.4	−11.7	1.6	−6.8	0.2

Figure 17 *Thermodynamic data for the binding by* GS4 *of* Le^b-OMe *and of derivatives with changes in the* α-L-Fuc *c-unit* [12]

required to avoid conflict with TRP 133. Certainly, it appears that much remains to be learned as to what can comprise complementarity. It will be seen below that geometries of the nonpolar interactions of the α-L-Fuc d-unit with tryptophan 138 and phenylalanine 108 are similar to that of the interaction of the β-D-Gal unit and tryptophan 133.

As already mentioned, the α-L-Fuc(1 ⟶ 4) c-unit contributes OH-4c to the key polar interaction (Figure 15). Judging from the data in Figure 17, OH-2c and OH-3c are also involved in interactions with the lectin since the replacement of either one of these hydroxyls by hydrogen has quite profound effects on the thermodynamic parameters. In both cases, the entropy barrier to binding is strongly reduced but especially so for the 3c-deoxy compound. It is apparent that this latter deoxygenation, as was observed for 6b-deoxy-Le^b-OMe (Figure 13), introduces a particularly strong contribution to binding that is akin to the hydrophobic effect.[36] Apparently, water molecules become more highly immobilized about the deoxy positions and their increased degrees of freedom on being liberated to bulk water make the positive contribution to the change in entropy.[36,37] It was considered[9] (Figure 6) that the interactions involving OH-2c and OH-3c are polar and near the periphery of the combining site since their roles are not of crucial importance. In other words, by being near the interface

[36] C. Tanford, 'The Hydrophobic Effect: Formation of Micelles and Biological Membranes', 2nd Edition, John Wiley & Sons, New York, 1980.

[37] N. Kauzmann, in 'Advances in Protein Chemistry', ed. C. B. Anfinsen, Jr., K. Bailey, M. L. Anson, and J. T. Edsall, Vol. 14, Academic Press Inc., New York, 1959, 1.

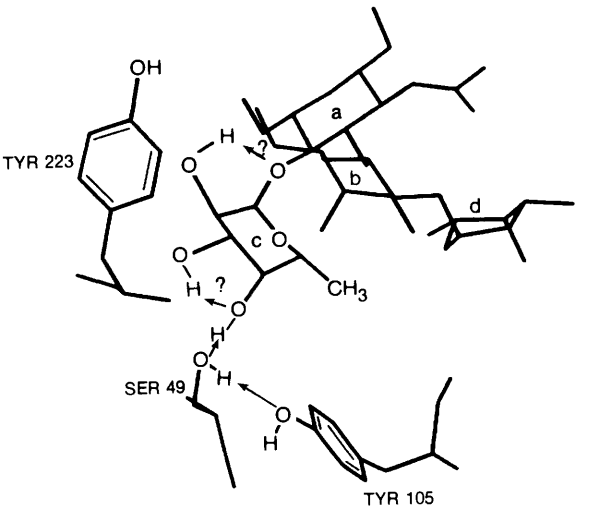
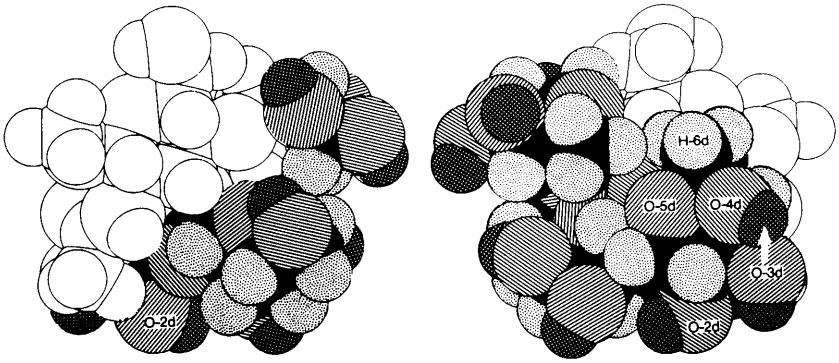


Figure 18 *The environment of the* α-L-Fuc *c-unit in the complex*[31]

with water where the protein chains are presumably more flexible, the necessary adjustments to the combining site can be made for the replacement of an hydrophilic interaction by one that is hydrophobic.

As illustrated by Figure 18, the *X*-ray crystallographic data at 2.8 Å resolution indicate that the phenyl group of tyrosine 223 comes into van der Waals' contact with both OH-2c and OH-3c. The placing of these hydroxyl groups in a nonpolar environment can conceivably lead to the formation of intramolecular hydrogen bonds as is indicated in Figure 18. If so, presumably the hydroxyl groups could also serve as proton acceptors for hydrogen-bond formation with water. In all likelihood, the hydroxyl group of Tyr 223 remains in the aqueous phase since it appears to be about 5.4 Å removed from the Le^b-OMe. It is noteworthy that this amino acid is bonded to glycine 222, a structural feature that can act as a hinge for conformational mobility. It is conceivable, therefore, that the position of Tyr 223 is different in the complexes with the 2c- and 3c-deoxy derivatives of Le^b-OMe (Figure 17). It is noteworthy in this regard that several of the complexes for GS4 and the modified Le^b-OMe structures[12] have crystallized and all are isomorphous to that for the GS4•Le^b-OMe complex.[31] Therefore, the effects of deoxygenation on the orientation of the tyrosine unit and changes in hydration about the c-unit will become known. This same exciting prospect exists for many other derivatives of Le^b-OMe and provides unique opportunities to learn about the role played by water in molecular recognition (Figure 5) and how the flexibility of the combining site can become involved.

Involvement of the α-L-Fuc(1 ⟶ 2) d-unit in the complexation proved particularly interesting.[9] The thermodynamic parameters for the bonding of the



kcal mole^{-1}, 25 °C

	$\Delta G°$	$\Delta\Delta G°$	$\Delta H°$	$\Delta\Delta H°$	$T\Delta S°$	$\Delta(T\Delta S°)$
Le^b-OMe	−6.3	0	−13.3	0	−7.0	0
Modification						
2d-Deoxy	−5.7	0.6	−12.1	1.2	−6.4	0.6
3d-Deoxy	−5.6	0.4	−8.6	4.7	−3.1	3.9
3d-Deoxy-3d-fluoro	−6.2	0.1	−7.5	5.8	−1.3	5.7
4d-Deoxy	−6.4	−0.1	−7.4	5.9	−1.0	6.0
4d-Chloro-4d-deoxy	−6.5	−0.2	−13.6	−0.3	−7.1	−0.1
4d-Deoxy-4d-fluoro	−6.5	−0.2	−11.1	2.2	−4.6	2.4
6d-*Nor*	−6.1	0.2	−10.8	2.5	−4.7	2.3

Figure 19 *Thermodynamic data for the binding by the lectin* GS4 *of* Le^b-OMe *and derivatives with modifications in the* α-L-Fuc(1 ⟶ 2) d-unit [12]

three monodeoxy derivatives are presented in Figure 19. It is seen that replacement of OH-2d by hydrogen had a relatively minor effect. This hydroxyl is in close proximity to the acetamido group which, as seen above, remains in the aqueous phase. Therefore, the small differential change in entropy [$\Delta(T\Delta S°) = 0.6$ kcal mole^{-1}] suggests that OH-2d remains bonded to bulk water. In this connection, it is to be remembered that the increased hydrophobicity about CH_3-6a on replacing OH-6a by hydrogen (Figure 12) did not affect the thermodynamic parameters for complex formation.

Surely, the data reported for the 3d-deoxy compound (Figure 19) requires OH-3d of Le^b-OMe to be involved in the binding reaction. The changes in the thermodynamic parameters that occur on replacing OH-3d by either hydrogen or fluorine are remarkable in that the differential changes in free energy are small in spite of very much lower decreases in enthalpy. The data also appear to require OH-4d to become involved in a nonpolar interaction with the combining site. This conjecture follows from the thermodynamic data presented in Figure 19, and particularly from the observation that the binding of the 4d-chloro-4d-deoxy-Le^b-OMe (Figure 20) occurs with very near the same changes in thermodynamic parameters as does the binding of the parent compound. For these reasons, it was considered that OH-4d is accepted into the combining site intramolecularly hydrogen bonded to either O-3d or O-5d. On this basis, the

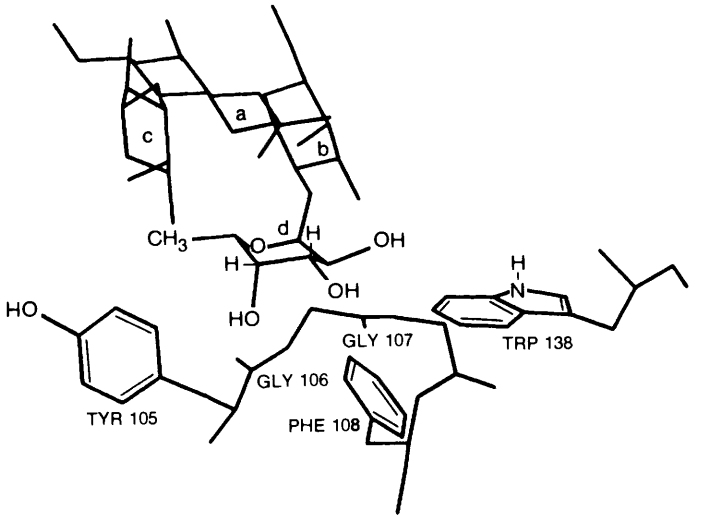


Figure 20 *The environment of the* α-L-Fuc *d-unit in the complex*[31]

high affinity exhibited for the binding of the 4d-deoxy derivative (Figure 19) would be attributed to the introduction of an hydrophobic effect as evidenced by the substantial lowering of the entropy barrier to binding.[12]

The *X*-ray crystallographic data have now shown that, as seen in Figure 20, the α-L-Fuc d-unit is in fact surrounded by hydrophobic groups. The low resolution (2.8 Å) available at this time does not allow a detailed interpretation. Nevertheless, it is of interest to note that the edges of the aromatic residues of both Phe 108 and Trp 138 appear to be in van der Waals' contact with the H-4d and H-3d atoms, respectively, of the α-L-Fuc d-unit. The diglycine unit that separates Phe 108 and Tyr 105 is also nearby. It seems probable that the crystal structures for the complexes of the derivatives described in Figure 19 will show some rearrangement of the phenyl group of Phe 108 from the position it occupied in the complex with Le^b-OMe. The same possibility was suggested above for the interaction of Tyr 223 with the α-L-Fuc c-unit.

The inhibition data in Figure 21 can be appreciated on this basis, especially the strong affinity for the derivative which has the α-L-Fuc unit substituted by the relatively much less polar 4R-methyltetrahydropyranyl group (c).[38] The importance of complementarity is demonstrated by the greater potency of the 2b-methoxymethyl (f) derivative as compared to that of the 2b-n-propyl (e) compound since, for reasons of the anomeric effect,[39] the methoxymethyl compound can readily accept the *syn*-clinal conformation shown whereas the n-propyl group must be expected strongly to prefer the *anti*-periplanar conformation.[9,38] It could not be expected that the introduction of methyl group in the *R*-

[38] U. Spohr and R. U. Lemieux, to be published.
[39] J.-P. Praly and R. U. Lemieux, *Can. J. Chem.*, 1987, **65**, 212.

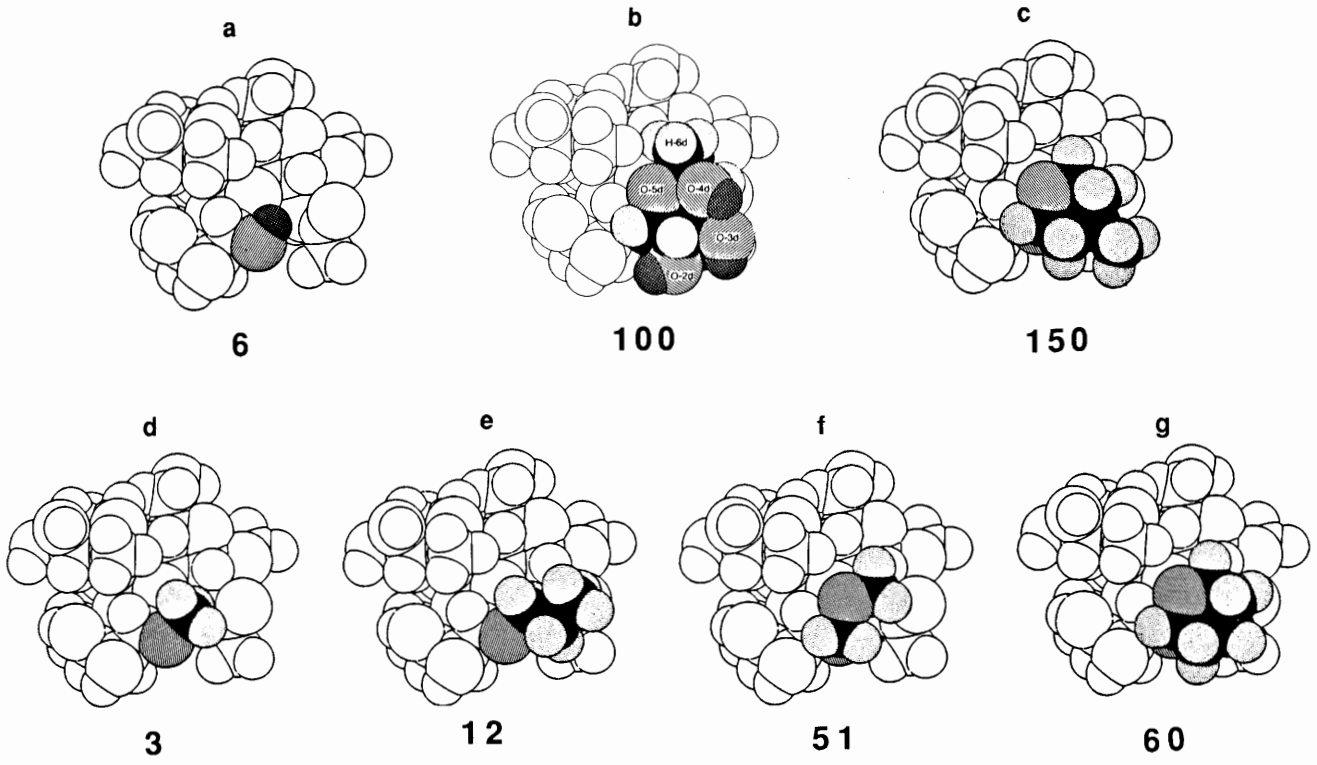


Figure 21 *Effects, on relative potency as inhibitors, of changing the* α-L-Fuc *substituent at the 2b-position of the* β-D-Gal *unit of* Le^b-OMe: (*a*) Le^a-OMe (*no substituent*); (*b*) Le^b-OMe; (*c*) 4-*R-methyltetrahydropyranyl*; (*d*) *methyl*; (*e*) *n-propyl*; (*f*) *methoxymethyl*; (*g*) *tetrahydropyranyl*[38]

configuration at the 4-position of the tetrahydropyranyl compound would yield a substance that is 50% more potent than the reference Le^b-OMe. It will be of decided interest to learn why. An understanding of the rather bizarre data reported in Figure 21 will surely be of important assistance toward the achievement of an appreciation of the role of water in these binding reactions, and especially how it participates in the recognition for binding by the complementary hydrophobic groups. The crystal structures of the various complexes are expected to provide much of the required information.

As was reported in Figure 17, the replacement of the CH_3-6c group by hydrogen resulted in a substantial loss in the strength of the binding. This effect arises from a lowering of the decrease in enthalpy and, therefore, appears to result from the loss in complementarity but without significant change in the hydrophobic contribution. As is illustrated by Figure 20, the nonpolar interaction of both the 6c- and 6d-methyl groups within the combining site can be provided by tyrosine 105. Being attached to the diglycine unit formed by Gly 106 and Gly 107, the hydroxyphenyl group, like the phenyl group of Phe 108, should have substantial mobility and, therefore, an ability to adjust to differing demands.

In conclusion, it seems fair to say that the results of this investigation of the binding of Le^b-OMe by the lectin GS4 leave little doubt that the establishment of nonpolar interactions plays an important role in the setting of the stability of the complex. These nonpolar interactions, as evidenced by the data presented in Figure 21, can be reduced substantially without resulting in loss of observable complex formation. The importance of complementarity for complex formation is expected to be greatest when the nonpolar interaction is at the bottom of the combining site where the protein structure is most rigid (for example, the 6b-deoxy-6b-*nor* derivative of Figure 13). The situation appears most forgiving when the interaction is near the periphery of the combining site and this seems reasonable since flexibility in the protein structure in the region of the periphery of the combining site must be a prerequisite for acceptance of the oligosaccharide at the beginning of complex formation. Some ability to reach a compromise also applies to polar interactions about the periphery. However, polar interactions within the combining site are the key to complex formation and these, together with the demands for complementarity, lie at the core of the origin of the specificity for the recognition of oligosaccharides by proteins. It is noteworthy in this regard that the key polar group is still intact in Le^a-OMe (Figure 21a) and, therefore, its activity as an inhibitor, although low, is not surprising. What is surprising is that the substituent at the 2b position can offer either an hydroxylated surface, as in Le^b-OMe, or a hydrocarbon surface (Figure 21c) to the lectin for the formation of complexes of near equal stability. How these reactions are driven will remain a subject for conjecture, at least until the *X*-ray structures for the various complexes are known.

As was mentioned with regard to Figure 7, water plays an important role in all of these complexation reactions. We have seen that high entropy barriers to complexation occur, but that these can be significantly reduced by the substitution of only one hydroxyl group by a hydrogen atom. Evidently, therefore, the

observed changes in entropy reflect the result of a miscellany of contributions on both sides of the equation. Sorting these out promises to be quite a problem!

To date, all the binding reactions we have examined proceed with a decrease in entropy. This must arise for the following kinds of reasons. One reason would be that a greater number of water molecules are immobilized, along with the oligosaccharide, within the complex. Also, although the main framework of an oligosaccharide such as Le^b-OMe is expected to remain conformationally well defined, the movement of appendant substituents (hydroxyl, hydroxymethyl, *etc.*) must become much more restricted once the complex has been formed. Furthermore, as discussed above, the combining site is not well formed in the native protein and the complexation therefore will require the fixing of R groups of the relevant amino acids into the proper orientations. In this regard, as was already pointed out, three of the amino acids that become engaged in the combining site are linked to glycine units. Thus, there can be no surprise that entropy barriers to complexation exist and that, in the absence of cancelling contributions by attendant hydrophobic effects, these barriers would be considerably higher than those observed.

The changes in enthalpy arise from exchanges in forces of attraction—both polar, including hydrogen bonds, and nonpolar dispersion forces. As already mentioned, and as long recognized, water cannot be excluded from these considerations and in this connection I wish to bring to your attention the results we have obtained in Monte Carlo simulations of the hydration of Le^b-OMe. The scope of this lecture does not allow the presentation of more than a very brief outline of the theoretical calculations by Dr. Helmut Beierbeck.[26] Nevertheless, I was tempted to present an insight to what we are finding with help of Figure 22.

It was decided to approach the calculations by building a shell of water about the oligosaccharide as thick as the available computer time would allow about the oligosaccharide. It was soon evident that the achievement of water bridges over the nonpolar regions represented a difficult challenge. Dr. Beierbeck eventually succeeded in building uniformly spherical solvation shells, 3 to 4 water molecules deep, by always adding water molecules, one at a time, at those potential energy wells on the developing cluster surface which are closest to the molecule's centre of gravity. The procedure will be described in detail elsewhere.[26] Suffice it to say at this time that it involves, at each step, the calculation of the interactions between all of the molecules with each other, using Clementi's additive intermolecular pairwise potentials.[40] The electric charges for the various atoms in the assembly depend on the electronegativities of the atoms to which they are attached. These charges are averages of atomic charges for amino acids, derived quantum mechanically.[41] Some adjustments were made using an expression developed by Fraga[42] to achieve the correct zero molecular charge for Le^b-OMe.

[40] E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, 1977, **99**, 5531, 5545.
[41] O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, 1976, **64**, 1351.
[42] S. Fraga, *J. Comput. Chem.*, 1982, **3**, 329.

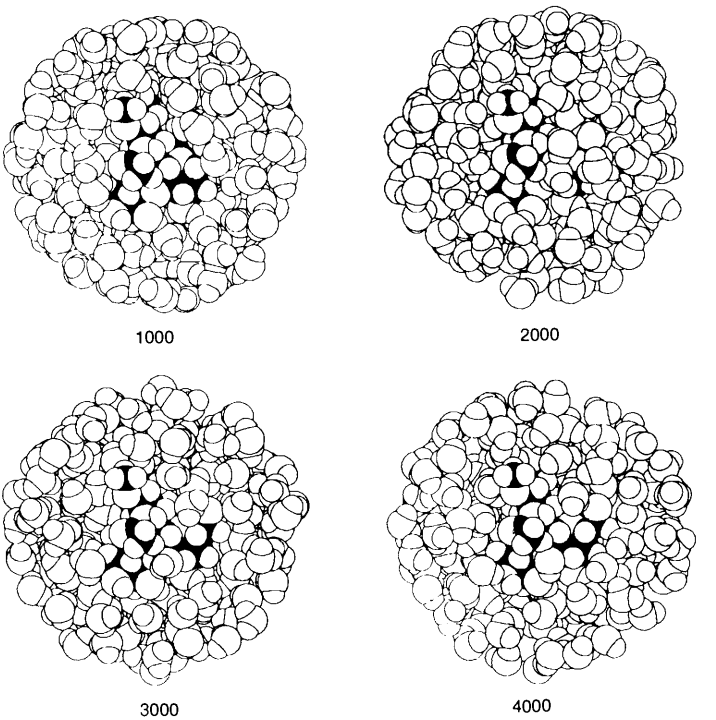


Figure 22 *Monte Carlo simulations for* 250 *water molecules solvating the methyl glycoside of the Lewis b tetrasaccharide* (Le^b-OMe) *at* 300 K *with the number of sampling cycles shown.*[26] *A* 2.3 Å *thick layer of water molecules was removed from the front of each assembly in order to reveal part of the* Le^b-OMe *molecule (carbon atoms in black). The hydrophobic* α*-side of the* β-D-Gal *b-unit is seen at* 4 *o'clock except after* 2 000 *cycles*

Once the solvent shell was established, Monte Carlo simulations[43] were made with fixed coordinates for the Le^b-OMe 'solute', a density of 1 for the assembly of Le^b-OMe and the 250 water molecules and at 300 K. A typical simulation involved 2000 equilibration and 4000 sampling cycles and required about three hours of Cyber 205 supercomputer time. A typical result is presented in Figure 22 for Le^b-OMe with its hydroxyl groups oriented as shown in Figure 10. The total shell is not presented. Instead, all the water molecules for which the nucleus of the oxygen atoms were more than 2.3 Å outwards from the plane defined by H-1b, H-3b, and H-5b of the β-D-Gal unit are removed in order to display part of the Le^b-OMe molecule. When examined in stereo, the presence of void space over the hydrophobic α-side of the β-D-Gal was clearly evident, and especially so after 1000 and 4000 cycles. This can be readily appreciated from the projections shown in Figure 22 since, after 2000 cycles, water molecules are seen to occupy space

[43] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087.

that is void after cycles 1000, 3000, and 4000 of the simulation. Indeed, the calculation of interatomic distances between the oxygen nuclei of water molecules and the closest hydrogens bonded to carbon was found to leave, *on the average*, over 1 Å between the van der Waals' surfaces. It is expected that this distance is minimum in view of the restrictions placed on radial distances of water molecules from the cluster centre during the Monte Carlo simulation, in order to control the 'solution' density. Water molecules hydrogen-bonded to oxygen atoms of Le^b-OMe were found to be even shorter than those expected from the van der Waals' radii and in the range expected for hydrogen-bonded oxygen atoms (2.7–3.0 Å). The calculations therefore make it apparent that, as displayed in Figure 16, a departure of the indolyl group of tryptophan 133 from its interaction with water molecules to come into van der Waals' contact with the α-side of the β-D-Gal b-unit may result in an important contribution to the driving force for complex formation simply because of better packing being achieved. Our investigations along this line are continuing and we plan to include, in due course, simulations for the hydration of the combining site. We are encouraged in this regard by the attractive hydrogen-bonded structures for water that are provided by the four 'snap-shots' in time shown in Figure 22.

In conclusion, I wish to express my deep gratitude and thanks to the many individuals and agencies that provided me with the help, guidance, facilities, and finance necessary for the accomplishing of the results outlined in this lecture. I am particularly indebted to my research associations with Drs. Ulrike Spohr and Ole Hindsgaul, who both participated and directed much of the work that I have described. Dr. Beierbeck's entry into the Monte Carlo calcuations resulted from an earlier collaboration with Professor Serafin Fraga to whom we are deeply grateful. The research could not have been accomplished in the absence of generous grants from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Council of Canada. Successful execution of the project was dependent on the collaborations of Drs L. T. J. Delbaere and L. B. Smillie, and I will always be indebted to them for their willingness to become involved and the provision of high priority to the engagement.