Acc. Chem. Res. **1996,** *29,* 373−380

How Water Provides the Impetus for Molecular Recognition in Aqueous Solution

Raymond U. Lemieux

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Received January 19, 1996

Introduction

This Account deals with my current views on matters related to molecular recognition in general and to carbohydrate recognition by proteins in particular. The urge to summarize my thoughts on the subject arose because of important but not widely appreciated developments, especially on the role of water, since the publication[1] in 1990 of an autobiography for the American Chemical Society. The book ends with a brief coda on the importance of hydration to the stability of complexes and the postulation of perturbed layers of water[2] over what we later termed polyamphiphilic surfaces.[3]

The popular notion that the dominating factor was the intermolecular attraction between the complementary surfaces did not seem plausible in view of the large excess of water. Hydrophobic effects exercised in the presence of complementarity seemed to offer the best answer—but how? At the beginning, an attempt was made to force interpretations of our experimental findings in terms of the natural entropy-driven tendency of nonpolar molecules to leave water. However, the enigma existed that associations of oligosaccharides with lectins that had been examined were enthalpic[4] and, moreover, involved a decrease, rather than an increase, in entropy. Why? In fact, but unbeknown to me, the notion of perturbed water at molecular surfaces was incubating in the minds of biophysicists. This can be appreciated from the review in 1989 of the fluctuation theory of hydration forces by Kornyshev and Leikin[5] and the appearance in 1993 of the paper entitled “Measured Change in Protein Solvation with Substrate Binding and Turnover” by Rand and co-workers.[6] Quite evidently, the subject was in need of precise structural information, at the atomic level, about the interacting surfaces.

The Strategy

It was about 20 years ago that I decided to contribute to an understanding of the selectivity and driving forces for the binding of carbohydrates by proteins as exhibited by antibodies, enzymes, and lectins. This seemed of first rate importance since it had become amply evident that the mosaics of complex oligosaccharides at cell surfaces likely play important roles in the maintenance of biological processes.

The research required methodologies drawn from the neighboring disciplines of immunochemistry and glycobiology. In this regard, it is to be appreciated that the immunization of a test animal using either a natural or artificial (normally semisynthetic) antigen stimulates the production of a polyclonal population of antibodies because the immune response results in genetically different antibody-producing cells. Cloning procedures exist for the production of monoclonal antibodies (identical molecules produced by the progeny of one antibody-producing cell) that all have an affinity for a specific area at the surface of the antigen known as the epitope. A small fragment of the antigen that contains the epitope is termed the hapten. Our artificial antigens were prepared by attaching an oligosaccharide hapten by way of a nine-carbon linking arm to bovine serum albumin (BSA). Congeners are structural analogs of a hapten, useful as affinity probes for the mapping of a combining (receptor) site. Competitive inhibition studies were used to establish the relative affinities (activities) of haptens and their congeners. The ratio (K_{rel}) of the molar amounts of two inhibitors that cause 50% inhibition of the binding of a reference ligand (n our work an appropriate artificial antigen) is a measure of their relative potencies and can be used to calculate the differential change in free energy ($\Delta\Delta G = RT \ln K_{rel}$). Lectins are proteins that possess specific affinities for the binding of carbohydrate structures and are neither enzymic nor products of an immune system. They serve in plant and animal tissues as “biological glues”. The epitopes recognized by antibodies and lectins can vary from a topography presented by a simple monosaccharide to one presented by parts of the several sugar units of an oligosaccharide. Enzymes that act on carbohydrate structures recognize their substrates for similar reasons. The glycosyl groups α-L-fucopyrano-

Raymond U. Lemieux was born in 1920 in Lac la Biche, Alberta, Canada. He received his B.Sc. (honors) in chemistry from the University of Alberta in 1943 and his Ph.D. in organic chemistry from McGill University in 1946. He was Professor and Vice-Dean at the University of Ottawa from 1954 to 1961 and Professor at the University of Alberta from 1961 to 1985. He became University Professor Emeritus at the University of Alberta in 1985 and still holds that position. Other interests include reaction mechanisms, stereochemistry, synthesis, conformational analysis, physical methods, and immunochemistry as related to carbohydrate structures. Major distinctions and awards include Companion of the Order of Canada, 1994; “Albert Einstein” World Award of Science, 1992; Canada Gold Medal for Science and Engineering, 1991; King Faisal International Prize in Science, 1990; Gairdner Foundation International Award, 1985; Haworth Award, The Royal Society of Chemistry, 1978; Officer of the Order of Canada, 1968; Fellow, Royal Society of London, 1967; C.S. Hudson Award, American Chemical Society, 1966; Palladium Medal, Chemical Institute of Canada, 1964; and Fellow of the Royal Society of Canada, 1955. He is the author of over 200 publications, mainly in the *Canadian Journal of Chemistry.*

(1) Lemieux, R. U. In *Explorations with Sugars. How Sweet it Was. Autobiographies of Eminent Chemists*; Seeman, J. I. Ser. Ed.; American Chemical Society: Washington, DC, 1990; 185 pp.
(2) Lemieux, R. U.; Delbaere, L. T. J.; Beierbeck, H.; Spohr, U. In *Ciba Foundation Symposium No. 158, Host-Guest Molecular Interactions: From Chemistry to Biology, July 3−5, 1990, London, U.K.,* Wiley: Chichester, 1991; pp 231−248.
(3) Spohr, U.; Paszkiewicz-Hnatiw, E.; Morishima, N.; Lemieux, R. U. *Can. J. Chem.* **1992**, *70*, 254−271.
(4) Hindsgaul, O.; Khare, D. P.; Bach, M.; Lemieux, R. U. *Can. J. Chem.* **1985**, *63*, 2653−2658.
(5) Kornyshev, A. A.; Leikin, S. *Am. Phys. Soc.* **1989**, *40*, 6431−6437.
(6) Rand, R. P.; Fuller, N. L.; Butko, P.; Francis, G.; Nicholls, P. *Biochemistry* **1993**, *32*, 5925−5929.

S0001-4842(96)00008-8 CCC: $12.00 © 1996 American Chemical Society

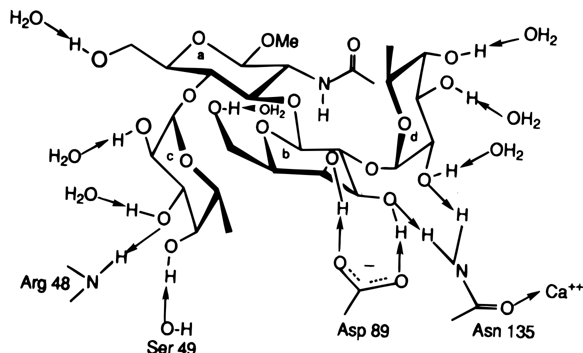


Figure 1. Hydrogen bonds formed between the epitope of Le^b-OMe and the receptor site of GS-IV to illustrate[27–29] (1) the three key interactions (OH-4c to Ser 49, OH-3b to both Asp 89 and Asn 135, and OH-4b to Asp 89), (2) the bonding at the periphery of OH-3c to Arg 48 and water and OH-2d to Asn 135 and water, and (3) that the other five hydroxyl groups remain entirely bonded to water.

syl, β-D-galactopyranosyl, 2-acetamido-2-deoxy-β-D-glucopyranosyl, and β-D-xylopyranosyl are abbreviated α-L-Fuc, β-D-Gal, β-D-GlcNAc, and β-D-Xyl, respectively. The linkage of two or more of these groups provides an oligosaccharide that may be branched, as shown for the β-methyl glycoside of the Lewis-b tetrasaccharide in Figure 1 which can be represented as α-L-Fuc(1→2)-β-D-Gal(1→3)[α-L-Fuc(1→4)]-β-D-GlcNAc-OMe. The H type 2-OMe trisaccharide of Tables 2 and 3 is α-L-Fuc(1→2)-β-D-Gal(1→4)-β-D-GlcNAc-OMe.

The first stage was to develop methodologies for the synthesis of both naturally occurring and structurally modified complex oligosaccharides (congeners) to serve as probes in studies of the recognition sites of the receptor proteins. Once these became available, the synthetic oligosaccharides allowed the preparation of monospecific artificial antigens and immunoabsorbents. The immunoabsorbents were required for the affinity purification of lectin and antibody reagents. These goals were essentially reached in the mid-1970s.[7]

Our first publication on molecular recognition[8] reviewed the early work on the probing of the combining sites of antibodies and presented the results obtained in the inhibition of a population of refined anti-β-D-Gal polyclonal antibodies generated in rabbits using the artificial antigen (β-D-Gal-O$(CH_2)_8$-CO)$_{24}$-BSA. The results, using a variety of congeners, were interpreted as the β-D-Gal group being accepted by combining sites that focused on the hydrophobic patch that is present on its α-side. In other words, the main driving force arose from a sequestering from water of the hydrophobic patch and a complementary hydrophobic cavity at the surface of the protein. This view was in accord with the proposal by Kauzmann[9] that water molecules use the opportunity to leave highly ordered, "frozen" arrays over nonpolar surfaces in a largely entropy driven process. Exchanges of polar interactions were not expected to contribute importantly to the stability of the complex.[10,11] Salem[12] had suggested that, in addition to hydrophobic hydration forces, multipoint attractive dispersion forces could contribute importantly to the association of complementary nonpolar surfaces.

In due course of time, more complex synthetic oligosaccharides and their congeners were used to probe the combining sites of monoclonal antibodies and lectins.[1] On occasion, a deoxygenation provided a more strongly bonded congener, and such observations appeared to reinforce the hypothesis that the bindings were importantly the results of hydrophobic effects. For example, quite convincing evidence[8] for the formation of intramolecular hydrogen bonds in the course of an association was obtained for the binding of congeners of the Lewis-b tetrasaccharide with a monoclonal anti-Lewis-b antibody.[13] The basis for the postulation of the formation of intramolecular hydrogen bonds was their occurrence in similar structures when dissolved in the weakly polar solvents.[14] It was stressed throughout these early investigations that these postulations were strictly provisional.

Hydrophobic forces remained the basic tenet of our working hypothesis until about 1985.[15] The further studies were mainly with lectins since these were more readily available and generally useful than monoclonal antibodies. It was considered that an understanding of ground-state equilibria was a necessary prelude to the study of enzyme–oligosaccharide interactions leading to high-energy transition states.

Key Polar Interactions and Amphiphilicity

It was discovered that, in all cases, the deoxygenation of certain hydroxyl groups essentially abolished binding, and this brought attention to the fact that all the epitopes were amphiphilic in character.[16] In the case of complex oligosaccharides, the epitopes normally involved a cluster of two to four hydroxyl groups, and these were designated the key hydroxyl groups. Since, prior to complex formation, the polar groups of both the epitope and the receptor site were surely extensively hydrated, water molecules would have to be displaced for the complex to form.[17,18] In view of the directional demands for the formation of hydrogen bonds,[19] a high degree of complementarity was necessary; otherwise,[18] dehydration would be energetically difficult and complex formation strongly discouraged. For this reason, the water molecules of the hydration shell were considered a hindrance to access to the receptor site by noncomplementary structures, and the hydrated polar gate concept was proposed.[13] Paraphrasing Emil Fischer,[20] the hy-

(7) Lemieux, R. U. *Chem. Soc. Rev.* **1978**, *7*, 423–452.
(8) Lemieux, R. U.; Boullanger, P. H.; Bundle, D. R.; Baker, D. A.; Nagpurkar, A.; Venot, A. *Nouv. J. Chim.* **1978**, *2*, 321–329.
(9) Kauzmann, W. In *Advances in Protein Chemistry*, Afinsen, C. B., Jr., Bailey, K., Anson, M. L., Edsall, J. T., Eds.; Academic Press Inc.: New York, 1959; Vol. 14, pp 1–63.
(10) Karush, F. *J. Am. Chem. Soc.* **1957**, *79*, 3380–3384.
(11) Hindsgaul, O.; Norberg, T.; LePendu, J.; Lemieux, R. U. *Carbohydr. Res.* **1982**, *109*, 109–142.
(12) Salem, L. In *Electronic Aspects of Biochemistry*, Pullman, B., Ed.; Academic Press Inc.: New York, 1964; pp 293–294.
(13) Lemieux, R. U.; Cromer, R.; Spohr, U. *Can. J. Chem.* **1988**, *66*, 3083–3098.
(14) Lemieux, R. U. *Pure Appl. Chem.* **1971**, *27*, 527–547.
(15) Lemieux, R. U. In *Proceedings of the VIIIth International Symposium on Medicinal Chemistry*, Uppsala, Sweden, Aug 27–31, 1984; Swedish Pharmaceutical Press: Stockholm, 1985; Vol. 1, pp 329–351.
(16) Lemieux, R. U. Proceedings of the 28th IUPAC Congress, Vancouver, Canada, Aug 16–22, 1981. In *Frontiers in Chemistry*; Laidler, K. J., Ed.; Pergamon Press: Oxford, 1982; pp 3–26.
(17) Spohr, U.; Hindsgaul, O.; Lemieux, R. U. *Can. J. Chem.* **1985**, *63*, 2644–2652.
(18) Lemieux, R. U. *Chem. Soc. Rev.* **1989**, *18*, 347–374.
(19) Wade, R. C.; Goodford, P. J. *J. Med. Chem.* **1993**, *36*, 148–156.
(20) Fischer, E. *Ber.* **1894**, *27*, 2895. See also Lemieux, R. U.; Spohr, U. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 1–20.

Table 1. Thermodynamic Parameters[a] for the Binding of Le[b]-OMe Tetrasaccharide and Monodeoxy Congeners by the Lectin GS-IV at 298 °C

position deoxygenated[b]	$\Delta G°$ (kcal/mol)	$\Delta\Delta G°$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta\Delta H°$ (kcal/mol)	$T\Delta S°$ (kcal/mol)	internuclear distance[c] (Å)
none	−6.3 (−6.5)	0	−13.3 (−11.9)	0	−7.0	
6a	−6.3 (−7.0)	0	−13.1 (−10.7)	0.2	−6.8	6.44 (O-Tyr 105)
6b	−6.2 (−6.4)	0.1	−8.7 (−7.4)	4.6	−2.5	3.62 (C-Tyr 223)
2c	−5.7	0.6	−10.0	3.3	−4.3	3.32 (N-Arg 48)
3c	−5.5 (*d*)	0.8	−6.6 (*d*)	6.7	−1.1	2.97 (N-Arg 48)
2d	−5.7 (*d*)	0.5	−12.1 (*d*)	1.2	−6.4	2.96 (N-Asn 135)
3d	−5.6 (−6.5)	0.7	−8.6 (−12.4)	4.7	−3.1	3.01 (N-Trp 138)
4d	−6.4	−0.1	−7.4	5.9	−1.0	2.88 (N-His 114)

[a] The values in parentheses were obtained by microcalorimetry and provided by Dr. Eric Toone, Duke University. [b] The 3b-, 4b-, and 4c-monodeoxy congeners were too inactive for significant measurements. [c] The distance between the oxygen of the hydroxyl group that was replaced by hydrogen and the nearest non-hydrogen atom in GS-IV (identified in parentheses). [d] Because of the weak binding and paucity of materials, reproduced results were not obtained. Definitely, however, these reactions were also exothermic ($\Delta H \simeq -9$ kcal/mol) with a compensating decrease in entropy ($\Delta G \simeq -7$ kcal/mol).

drated polar groups within the combining site were viewed as a locked gate that could be opened only by the key polar groups of the epitope. In this context, water molecules are intimately connected to the specificity of binding. With regard to structural requirements for effective binding, it is noteworthy, as recently reviewed by Lee and Lee,[21] that simple monosaccharides often display low but detectable activities. It was exciting to learn, through the use of deoxy congeners, that, in fact, a strong recognition of an oligosaccharide normally involved key hydroxyl groups on more than one sugar unit.[22] It is also noteworthy that, in order to effect complementarity, stereoelectronically well stabilized water molecules are often occluded within the complex.[23,24] It was recently estimated that the cost in entropy for the imprisonment of a single water molecule can be as high as 2 kcal/mol.[25]

As already mentioned, the hypothesis for binding based on a dominant hydrophobic effect involving nonpolar molecular surfaces seemed reinforced by finding that often the replacement of a hydroxyl group by hydrogen produced a more strongly bound congener.[4,14] Furthermore, in every instance wherein such reduction in the polarity of the compound strengthened the binding, the hydroxyl group was sterically well disposed for intramolecular hydrogen bonding. However, this working hypothesis had to be dramatically altered in the course of studies of the complex formed by the Lewis-b tetrasaccharide as its methyl glycoside (Le[b]-OMe; see Figure 1 for its structure), with the fourth lectin of the plant *Griffonia simplicifolia* found to have a human blood group activity[26] and which we refer to as GS-IV.[18]

As seen in Table 1, the deoxygenation of any one of the three key hydroxyl groups of Le[b]-OMe had a major effect on binding, and it was suggested that the epitope was a rather large nonpolar region guarded by the three key hydroxyl groups and did not involve the 6b-hydroxyl group of the β-D-Gal unit.[17] This possibility was rejected when the thermodynamic parameters for the complexation by GS-IV of both Le[b]-Me and its 6b-deoxy congener became known.[14] As had been previously observed,[17] the structural change caused an only slight (0.1 kcal/mol) change in free energy. However, it was now found that there existed a large 4.5 kcal/mol differential change in enthalpy.[14] Consequently, the original postulation that the 6b-hydroxyl group does not interact with the receptor site was reversed, and with enhanced confidence, it was proposed that this hydroxyl group was in fact accepted into the receptor site intramolecularly hydrogen bonded to the ring oxygen.[14] As will be seen below, both of these postulations proved to be in error.

About two years earlier, we had undertaken a collaboration with Louis Delbaere and associates to determine the X-ray crystal structures of GS-IV and its complex with Le[b]-OMe.[27−29] The structure of the complex at 2.8 Å resolution became known in 1989,[27] and it could be concluded soon thereafter that, in fact, intramolecular hydrogen bonds were not present in the complex.[28] As illustrated by Figures 1 and 2, of the ten hydroxyl groups in Le[b]-OMe, in accord with the results of affinity probing, the three at positions 3b, 4b, and 4c form the key polar grouping. Of the other seven, two at positions 3c and 2d become hydrogen bonded to the lectin as proton acceptors at the periphery of the combining site. The five others stay fully in contact with water. These conclusions were later unequivocally confirmed to exist in homogeneous aqueous solution by a comparison of the effects on binding of monodeoxygenations with those resulting from mono-*O*-methylations.[30] As was to be expected from molecular modeling of the complex, only the congeners from *O*-methylation of any one of the five hydroxyl groups that remain fully in contact with water remained highly active.[30] For this reason, 6b-OH and 4b-OH were definitely not part of the epitope and intramolecular hydrogen bonding was not involved in the formation of the complex. The concept for intramolecular hydrogen bonding in the course of complex formation should not be altogether discarded

(21) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 323−327.
(22) Lemieux, R. U.; Venot, A. P.; Spohr, U.; Bird, P.; Mandal, G.; Morishima, N.; Hindsgaul, O.; Bundle, D. R. *Can. J. Chem.* **1985**, *63*, 2664−2668.
(23) Quiocho, F. A.; Wilson, D. K.; Vyas, N. K. *Nature* **1989**, *340*, 404−407.
(24) Bundle, D. R.; Eichler, E.; Gidney, M. A. J.; Meldal, M.; Ragauskas, A.; Sigurskjold, B. W.; Sinnott, B.; Watson, D. C.; Yaguchi, M.; Young, N. M. *Biochemistry* **1994**, *33*, 5172−5182.
(25) Dunitz, J. D. *Science* **1994**, *264*, 670.
(26) Shibata, S.; Goldstein, I. J.; Baker, D. A. *J. Biol. Chem.* **1982**, *257*, 9324−9325.
(27) Delbaere, L. T. J.; Vandonselaar, M.; Prasad, L.; Quail, J. W.; Nikrad, P. V.; Pearlstone, J. R.; Carpenter, M. R.; Smillie, L. B.; Spohr, U.; Lemieux, R. U. *Trans. Am. Crystallogr. Assoc.* **1989**, *25*, 65−76.
(28) Delbaere, L. T. J.; Vandonselaar, M.; Prasad, L.; Quail, J. W.; Nikrad, P. V.; Pearlstone, J. R.; Carpenter, M. R.; Smillie, L. B.; Spohr, U.; Lemieux, R. U. *Can. J. Chem.* **1990**, *68*, 1116−1121.
(29) Delbaere, L. T. J.; Vandonselaar, M.; Prasad, L.; Quail, J. W.; Wilson, K. S.; Dauter, Z. *J. Mol. Biol.* **1993**, *230*, 950−965.
(30) Nikrad, P. V.; Beierbeck, H.; Lemieux, R. U. *Can. J. Chem.* **1992**, *70*, 241−253.

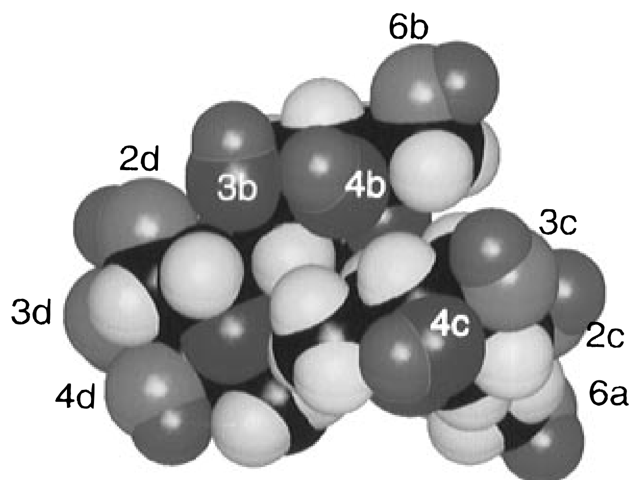


Figure 2. Polyamphiphilic topography of the epitope of Le^b-OMe that is recognized by the lectin GS-IV.[2,28] Note that hydration of this surface must include the six hydroxyl groups (in orange) that, in the complex, are at or very near the periphery of the combining site (see Table 1) in a network of hydrogen-bonded water molecules that also includes hydrogen bonding to the three key hydroxyl groups (in red) at positions 3b, 4b, and 4c. Hydroxyl group hydrogens are green, and those attached to carbon are white.

since Bundle and co-workers[31] found such a bond in the crystal structure of the complex of a trisaccharide with an antibody. As was to be expected in view of the high molar excess of water, this bond was not detected in aqueous solution.

At the point in the research when the importance of key polar interactions had become recognized, it seemed necessary to conclude that the exchanges of interactions between polar groups, including those of water, favor complex formation even in the presence of a 54000-fold molar excess of water molecules. Why was not obvious. It was hoped that a clue would appear as the result of the increasing number of structures of oligosaccharide epitopes that we were obtaining by chemical mapping. The early studies had only provided differential changes in free energy. It was now obvious that the thermodynamic parameters must also be determined.[17]

The tedious task of determining the thermodynamic parameters for the binding by GS-IV of the seven active monodeoxy derivatives of Le^b-OMe was then undertaken. The effects on binding, as measured by changes in ultraviolet absorption, caused by increasing amounts of the ligand up to saturation at temperatures in the range 15−45 °C, were used to establish van't Hoff plots from which the changes in enthalpy (ΔH) and entropy (ΔS) were estimated.[2,17] It was found, and in part confirmed by microcalorimetry (see Table 1), that, although the deoxygenations of hydroxyl groups near the periphery of the combining site (see Figure 1) caused little change in equilibrium ($\Delta G = \Delta H - T\Delta S$), rather large but virtually compensating changes occurred for the values of ΔH and ΔS. The changes found were not significant in the case of the 6a-position of the GlcNAc residue, which is distant (6.4 Å) from the protein. It is seen (Table 1) that the association of GS-IV with all of the monodeoxy congeners of Le^b-OMe caused a decrease in entropy, whereas an increase was to be expected had the structural change created a hydrophobic effect conducive to so-called "hydrophobic bonding". Clearly, the idea of a dominant hydrophobic effect had to be abandoned.

The high level of enthalpy−entropy compensation (Table 1) was of considerable theoretical interest since, as was surmised on the basis of NMR data,[2] and now confirmed by X-ray crystal structures,[32] the various deoxy compounds are all bound in the same conformation as is Le^b-OMe and, therefore, the enthalpy−entropy compensation had to arise from changes in the dynamics of the water structure.[2] A compensation of changes in thermodynamic parameters had been observed earlier in a rather surprising way in that I noticed good compensations in the thermodynamic parameters for the binding of a miscellaneous group of oligosaccharides by a miscellaneous group of lectins.[4] Since the only component that the various systems had in common was the aqeuous solvent, it appeared that the phenomenon was somehow related to water. The compensation phenomenon[33] had been observed for the binding by a wide variety of proteins including first an enzyme[34] and then antibodies[35] and several lectins[4,36] and may be related to the manifestation of a unique property of water. Perhaps the release of a water molecule to form stronger hydrogen bonds leads to a loss in mobility (decrease in entropy) that virtually cancels the decrease in enthalpy.[2] It is important to note that the binding of congeners of Le^b-OMe by GS-IV does not necessarily result in enthalpy−entropy compensation.[37] The largest deviations resulted from the introduction of polar groups at positions in the Le^b-OMe tetrasaccharide that, in the complex, reside near the periphery of the protein's combining site. Certainly, it is reasonable to expect that structure-building hydration effects involving ionic groups can have an important disruptive influence on complementarity.

It is seen in Table 1 that all of the deoxygenations led to a lesser decrease in enthalpy. Since prior to deoxygenation these hydroxyl groups were hydrogen bonded to water, there can be no doubt, and of major importance to note, that these changes must mainly result from the formation of less stable hydration shells about the positions affected. It is noteworthy that *O-n*-propylation at the 6b-position produced stronger binding.[37] In fact, the structure of the complex formed by GS-IV with Le^b-OMe has the 6b-position of Le^b-OMe bordering a hydrophobic cavity. This result exemplifies the fact that hydrophobic effects that favor the sequestering of complementary nonpolar surfaces remain relevant to the binding of oligosaccharides by proteins.[36,38] For this reason, I propose that discussions of these surface properties would be facilitated by reserving the term *hydrophobic* for associations that arise from water leaving nonpolar surfaces and to introduce the term *hydraphobic* for

(31) Bundle, D. R.; Baumann, H.; Brisson, J.-R.; Gagné, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183−5192.

(32) Delbaere, L. T. J.; Vandonselaar, M. Private communication.

(33) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709−712.

(34) Belleau, B. *Ann. N. Y. Acad. Sci.* **1967**, *144*, 705−719.

(35) Szewczuk, M. R.; Mukkur, T. K. S. *Immunology* **1977**, *32*, 111−119.

(36) Lemieux, R. U.; Du, M.-H.; Spohr, U.; Acharya, S.; Surolia, A. *Can. J. Chem.* **1994**, *72*, 158−163.

(37) Lemieux, R. U.; Du, M.-H.; Spohr, U. *J. Am. Chem. Soc.* **1994**, *116*, 9803−9804.

(38) Petrakova, E.; Spohr, U.; Lemieux, R. U. *Can. J. Chem.* **1992**, *70*, 233−240.

"hydrophobicity" related to the hydration of polyamphiphilic surfaces.

Polyamphiphilicity and Perturbed Water Molecules

In 1984, I was still puzzling over the fact that the binding of oligosaccharides by lectins appeared most often driven by decreases in both enthalpy and entropy. Recourse had been taken to the multipoint electrodynamic (dispersion) forces that had been proposed by Salem.[12] Certainly, I had become impressed with how simple hard-sphere calculations, based on the Kitaigorodsky expressions, well produced conformational models in good agreement with NMR data for oligosaccharides in aqueous solution.[39,40] For this reason, it was conceivable that it was the establishment of stronger dispersion forces of attraction between the complementary surfaces than with water that caused the decrease in enthalpy. The problem was how to gain supporting experimental evidence.

At that time, one of my colleagues, Serafin Fraga, was publishing papers on the hydration of amino acids. This caught my attention, and I asked if he would be interested in simulating the hydration of methyl α-L-fucopyranoside. I chose this compound since the fucosyl unit is present in many oligosaccharidic antigenic determinants and one side of the pyranose ring is highly hydrophobic as compared to the other. The results of his calculations are reported in a lecture I presented in 1984.[13] As I had come to anticipate, the hydration shell showed a cavity over the nonpolar α-side of the molecule. I remarked "...should the void spaces...be essentially real, the calculation of hydration shells about compounds that interact with receptor sites will become a useful exercise..." This idea sparked the beginning of a line of thought that has culminated with this present Account.

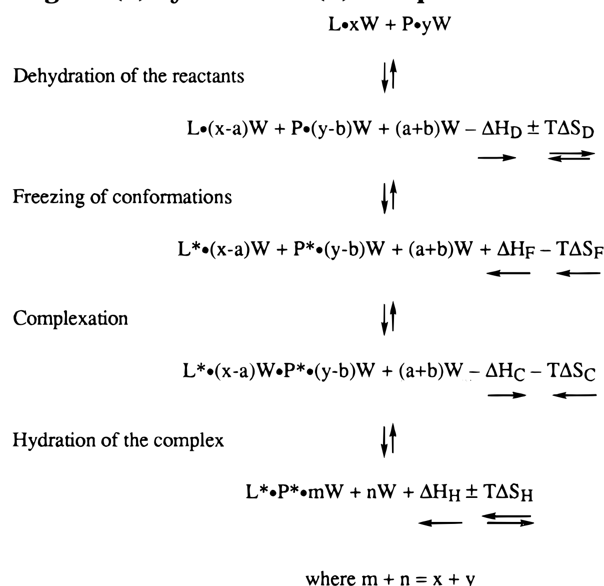
I had to cope with a prolonged bout of illness throughout 1985. Fortunately, Ulrike Spohr[1] had joined by research group, and under her guidance the probing of receptor sites and the thermodynamic studies continued without interruption. It became evident that all the epitopes that we had mapped were arrays of small amphiphilic structural units, as displayed in Figure 2 for the epitope of Le^b-OMe that is recognized by GS-IV. On examining these surfaces and having learned that hydration forces are important to the stability of oligosaccharide−protein complexes,[2] the thought came to mind that these polyamphiphilic surfaces were so structured that their hydration can result in a perturbation of water molecules. This was a most exciting moment; finally, I might have found the trail to why the binding reactions were exothermic. In view of the polyamphiphilicity, it had to be expected that, prior to association, the polar groups of both the epitope and the receptor site are all extensively hydrogen bonded to molecules of water. These water molecules, in turn, were then necessarily hydrogen bonded to other water molecules. However, it seemed unlikely that the monohydrated polar groups of the surfaces could always present their water molecules for hydrogen bonding to other water molecules in as energetically favorable orientations as is in bulk, where the water molecules are relatively free to adopt whatever orientations are most favorable to hydrogen bonding with their neighbors. Higher energy perturbed water molecules must result and heat liberated should these be returned to bulk.

To my good fortune, an already well-experienced computational chemist, Helmut Beierbeck,[1] joined my research group in order to write and execute programs for Monte Carlo simulations. It was decided to begin by simulating the hydration of Le^b-OMe at 300 °C using 100 molecules of water at a density of 1.0, an effort compatible with the supercomputer time we could expect to have at our disposal. Our first publication[41] outlines the difficulties that had to be overcome to achieve uniform hydration. As expected, at thermal equilibrium, the simulations displayed a density of water over the nonpolar regions that was always substantially less than that in bulk. Helmut then undertook examining the hydration of the receptor site of the lectin GS-IV, the X-ray structure of which was known to 2.15 Å resolution.[42] The procedure was to first set a rectangular block of 250 water molecules at density 1.0 to cover and fill the receptor site. As anticipated, the calculated water-to-water interaction energies in fact strengthen with increasing distance from the receptor surface. The main significance we attach to this result is that it supports our contention that water over polyamphiphilic surfaces can be perturbed and the perturbation is attenuated over a distance of about three water molecules. Considering that these layers cover the areas of both the receptor site and the complementary epitope, very many water molecules are involved and the coalescence of these higher energy water molecules to bulk water can make a strongly exothermic contribution to a decrease in the free energy of the system. Therefore, in the absence of a compensatory decrease in entropy, complexation should result. The extent to which the decrease in enthalpy is compensated by a decrease in entropy is surely variable and cannot, at this time, be anticipated. Although in many cases the compensation observed for congeners was nearly exact, as already mentioned, exceptions appear to arise from the structuring of water about ionic substituents.[37]

As already mentioned, the binding of an oligosaccharide by a lectin is normally exothermic, with the decrease in enthalpy compensated in part by a decrease in entropy. A main source of compensation surely originates in the dehydration of polyamphiphilic surfaces (the first stage in Scheme 1). However, it is to be expected that prior to binding some of the surface water molecules are highly ordered over nonpolar patches of the complementary surfaces and that their release to bulk will, in contrast to the release of perturbed water from the polyamphiphilic surfaces, contribute an increase in entropy. Furthermore, the water molecules directly bonded to the polar groups of these surfaces, as is already widely recognized, likely have restricted motion as compared to those in bulk and, on release, contribute an increase in entropy that will further counteract the decrease

(39) Thøgersen, H.; Lemieux, R. U.; Bock, K.; Meyer, B. *Can. J. Chem.* **1982**, *60*, 44−57.
(40) Sabesan, S.; Bock, K.; Lemieux, R. U. *Can. J. Chem.* **1984**, *62*, 1034−1045.
(41) Beierbeck, H.; Lemieux, R. U. *Can. J. Chem.* **1990**, *68*, 820−827.
(42) Beierbeck, H.; Delbaere, L. T. J.; Vandonselaar, M.; Lemieux, R. U. *Can. J. Chem.* **1994**, *72*, 463−470.

Scheme 1. Artificial Expressions for the Binding of a Ligand (L) by a Protein (P) in Aqueous Solution[a]



[a] The hydration of the reactants and the product is expected to involve *x*, *y*, and *m* water (W) molecules that have thermodynamic parameters significantly different from those in bulk. The asterisks are to represent L and P in specific conformations. The arrows under the thermodynamic parameters are to indicate the direction that their change is expected to have on the various hypothetical equilibria. The $\pm T\Delta S_D$ and $\pm T\Delta S_H$ are to indicate both increases and decreases may contribute to the net $T\Delta S$.

in entropy expected to occur on the release of the less ordered perturbed water molecules. It appears that the escape to bulk of the perturbed water molecules from hydraphobic surfaces is Nature's way of providing the excess of the energy required to meet the various entropy barriers to complexation that are present in the overall process (Scheme 1). I stress at this point that Scheme 1 is strictly crude "bookkeeping" to keep track of contributions to the overall change in the thermodynamic parameters.

That the restructuring of perturbed surface water provides an important force for molecular association was the central feature of a lecture that I presented to the American Chemical Society in 1991 entitled "How proteins recognize and bind oligosaccharides".[43] Our conclusions soon thereafter received compelling experimental support by way of osmotic stress measurements. Thus, in 1992, Colombo and co-workers[44] arrived at the conclusion that "the work of solvation in allosteric regulation can no longer go unrecognized." In the same issue, Rand[45] in an article entitled "Raising water to new heights" wrote "The nature of the surface perturbation of water remains an enigma." In 1993, Leikin, Parsegian, Rau, and Rand[46] published a noteworthy review entitled "Hydration forces".

In the case of the enzyme hexokinase,[6] the variation in the activity of the water over the range of osmotic pressures studied indicated that, on binding glucose, at least 65 ± 10 water molecules are liberated. It was concluded that most of these water molecules had come from the cleft in the protein that accepts the glucose and thereby made a significant contribution to the energetics of the conformational changes. In this regard, they stated "...glucose induces the conformational changes required for catalysis, the classical example of an 'induced fit' mechanism of enzyme specificity." [6]

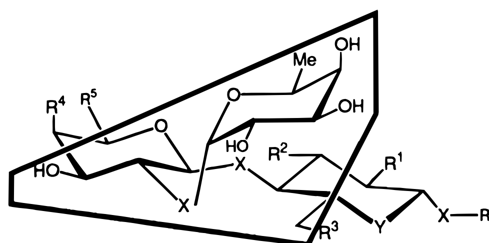
Further experimental evidence in support of the thesis[43] that large numbers of water molecules are engaged in the overall dynamics of molecular associations in aqueous solution was independently provided by Chervenak and Toone.[47] They discovered that the amount of heat liberated on the binding of ligands with a variety of proteins was significantly less (0.4–1.8 kcal/mol) when heavy water was the solvent. Since it can be assumed that the conformations of the solutes and the complex were unaffected by the change in solvent, it was concluded that the negative differential changes in enthalpy arose from dynamic reorganization of the solvent molecules. Analyses of the microcalorimetric data for associations with proteins of carbohydrates, a peptide, and a nucleic acid indicated that 25–100% of the observed enthalpies of binding was the result of solvent reorganization. Since the decreases in enthalpy were in the range −7 to −17 kcal/mol, and must represent the sum of many small changes in energy, the number of water molecules involved in the reorganization had to be large. An excellent but not completely comprehensive critical review of the relevant literature was presented.[47]

In summary, it seems unequivocally established that the interaction of water molecules with either a nonpolar or a polyamphiphilic surface leads to molecules that, on the average, have higher energy contents than those in bulk. At nonpolar surfaces, as was proposed by Kauzmann,[9] an organized layer of molecules is formed that, on being released to bulk, provides an increase in entropy. In contrast, the liberation of water molecules from polyamphiphilic surfaces causes important decreases in both ΔH and $T\Delta S$ (Scheme 1, expression 1). A contribution of significance to the change in free energy must also arise from the change in the electrodynamic (dispersion) forces of attraction between water and the reacting species and their hydrated complex.[12,48] To my knowledge, the significance of this contribution to the change in free energy remains unknown.

The data in Table 2 are presented to illustrate that the effect of a structural change in the ligand that does not involve the epitope becomes stronger the nearer it comes to reside at the periphery of the receptor site. It is seen that, whereas deoxygenation at the 4-position of the β-D-Gal residue decreased the stability of the complex by 0.60 kcal/mol, deoxygenation at the 6-position of the β-D-GlcNAc unit (R^3 = H) stabilized the binding by 0.55 kcal/mol. Evidently, the latter deoxygenation revealed a *hydrophobic* effect whereas the former identified a *hydraphobic* effect. No comparable effects on strength of binding resulted from deoxygenations at more remote positions. Participation of the methylene group of the β-D-GlcNAc unit was further established by finding α-L-Fuc(1→2)-B-D-

(43) Lemieux, R. U. In *Carbohydrate Antigens*; Garegg, P. J., Lindberg, A. A., Eds.; ACS Symposium Series 519; American Chemical Society: Washington, DC, 1993; pp 5–18.
(44) Colombo, M. F.; Rau, D. C.; Parsegian, V. A. *Science* **1992**, *256*, 655–659.
(45) Rand, R. P. *Science* **1992**, *256*, 618.
(46) Leikin, S.; Parsegian, V. A.; Rau, D. C. *Annu. Rev. Phys. Chem.* **1993**, *44*, 369–395.
(47) Chervenak, M. C.; Toone, E. J. *J. Am. Chem. Soc.* **1994**, *116*, 10533–10539.
(48) Privalov, P. L.; Gill, S. J. *Pure Appl. Chem.* **1989**, *61*, 1097–1104.

Table 2. Effects on Binding by *Ulex europaeus* of Defunctionalization of H-Type 2-OMe Both near and Remote from the Epitope



R	R1	R2	R3	R4	R5	X	Y	ΔΔG° (kcal/mol)
a	NHAc	OH	OH	OH	CH_2OH	O	O	0.0
b	NHAc	OH	OH	OH	CH_2OH	O	O	−0.24
c	NHAc	OH	OH	OH	CH_2OH	CH_2	O	0.2
d	NHAc	OH	H	OH	CH_2OH	O	O	−0.55
d	NHAc	OH	OH	H	CH_2OH	O	O	0.60
d	H	H	H	OH	CH_2OH	O	CH_2	0.36
d	H	H	H	OH	H	O	O	0.25

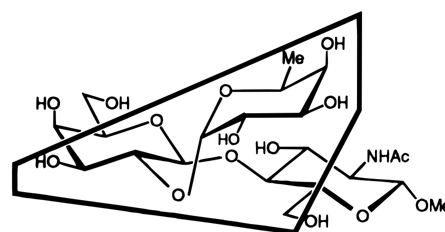
[a] Me.[3,4] [b] $(CH_2)_8COOMe$.[3,4] [c] $(CH_2)_8COOH$.[51] [d] Me.[3]

Gal(1→4)-*β*-D-Xyl-OMe to be highly inactive.[49] It is this participation that provided the clue as to why the de-*N*-acetylated form of H-type 1-OMe is much more active than the parent compound.[50] For similar reasons, it was to be expected that, as recently reported by Kishi and co-workers,[51] the replacement of the hydroxymethyl group of the *β*-D-GlcNAc unit of H-type 2-OMe by hydrogen would lead to an important loss of activity.

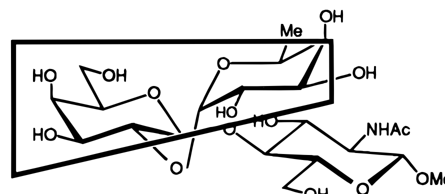
The data in Table 3 support the contention that the formation of intermolecular hydrogen bonds does not dominate in the setting of the stability of a complex. The chemical mapping procedure[30,52] was used to reveal the epitopes for the binding of H-type 2-OMe by two different lectins.[49] It is seen that, in both cases, four hydroxyl groups are required to become hydrogen bonded to the receptor site. However, the decreases in enthalpy differ by 23.6 kcal/mol, that is, nearly 5 kcal/mol per hydroxyl group. This is not plausible in view of the large excess of water. The decreases in both enthalpy and entropy observed for the binding by the *Ulex* lectin signify a dominating hydraphobic effect. In contrast, the increase in entropy found for the binding by the *Psophocarpus* lectin indicates the participation of an important hydrophobic effect. Obviously, the partitioning of the various parameters inferred in Scheme 1 presents a difficult challenge. Microcalorimetry, the mutagenesis of proteins, and the always increasing rapidity for the achievement of high-resolution X-ray structures and increasing power of computers offer encouragement in this regard.

The 2.0 ± 0.2 kcal/mol decrease in the stability of the complex of H-type 2-$O(CH_2)_8COOH$ that occurred on deoxygenation of the 3-hydroxyl group of the *β*-D-Gal unit recently reported by Kishi and co-workers[51] was in agreement with our finding of a 2.2 kcal/mol deactivation for the same congener of H-type 2-OMe.[4] It appears, therefore, that our data are compatible.

Table 3. Thermodynamic Parameters for the Binding of H-Type 2-OMe by Two Different Lectins



Ulex europaeus I



Psophocarpus tetragonolobus II

	UE-1	PT-II[a]
$\Delta H°$ (kcal/mol)	−29	−5.4
$T\Delta S°$ (kcal/mol)	−20.5	+0.8

[a] Determined by Acharya and Surolia.[36]

They had shown[53] that their substitution of the three glycosidic oxygens by methylene groups did not significantly alter the conformational preference. Therefore, it is not surprising that this congener was very nearly as active as the tri-*O*-glycoside because, as seen in Table 2, none of the three interunit methylene groups are positioned to interact with the receptor site. The thermodynamic parameters may, however, be quite different but compensating.

Water as a Universal Catalytic Chaperon

Water behaves as a catalyst in the sense that it forms activated complexes that are necessary to the molecular association but is regenerated in the course of the reaction. The source of activation is the mass effect exerted by a large excess of water that forces the water molecules near the molecular surfaces (*x* and *y* of Scheme 1) to adopt states of greater enthalpy and/or entropy than prevails in the bulk. The values of *x* and *y* can never be discrete since there can be no discrete boundary between the perturbed and bulk water molecules. It may be noteworthy in this regard that the receptor sites of lectins and antibodies commonly contain several aromatic amino acid residues that in the complex are stacked against the sugar units.[24,28] This structural feature can provide a zone of low dielectric that favors attractive polar interactions, such as hydrogen bonds. Indeed, the X-ray crystal structure of the native GS-IV shows the combining site to be strongly hydrated by seven water molecules.[42] It is the further hydration of these well-"anchored" water molecules that leads to the formation of zones of perturbed water over and about the reacting surfaces.

(49) Du, M.-H.; Spohr, U.; Lemieux, R. U. *Glycoconjugate J.* **1994**, *11*, 443−446.
(50) Cromer, R.; Spohr, U.; Khare, D. P.; LePendu, J.; Lemieux, R. U. *Can. J. Chem.* **1992**, *70*, 1511−1530.
(51) Wei, A.; Boy, K. M.; Kishi, Y. *J. Am. Chem. Soc.* **1995**, *117*, 9432−9436.
(52) Lemieux, R. U. *The Alfred Benzon Symposium No. 36, Complex Carbohydrates in Drug Research;* Munskgaard: Copenhagen, 1994; pp 188−201.
(53) Wei, A.; Haudrecky, A.; Audin, C.; Jun, H.-S.; Haudrecky-Bretel, N.; Kishi, Y. *J. Org. Chem.* **1995**, *60*, 2160−2169.

It must be that, whenever hydrated polyamphiphilic surfaces come into sufficiently close proximity for the zones of perturbed water molecules to interact, these water molecules will tend to escape to bulk. The resulting decrease in perturbation must cause a weak attraction whether or not the surfaces are complementary. That this is so appears supported by the phenomenon of leucocytes rolling at the surface of endothelial tissue as these are pushed along by vascular flow. This phenomenon has received much attention because of its importance to disease. Springer[54] has recently reviewed the subject and discussed how white cell migrations appear to be under the control of traffic signals. In this regard, it is reasonable to imagine that the rates of migration and residence times at specific sites are the result of varying degrees of complementarity.

As suggested for the rolling of leucocytes, some migration of a ligand along the protein surface may occur prior to encountering the receptor site. The recognition of complementarity can be expected to result in an extremely fast leakage to bulk of the intervening perturbed water molecules and a highly rapid association. For example, Otting and co-workers[55] have demonstrated that the residence times of water molecules at the surface of proteins are in the subnanosecond range. Thus, the dehydration, freezing, complexation, and rehydration stages presented in Scheme 1 are surely part and parcel of a very fast and continuous process. Indeed, it has recently been shown that the half-times for the binding of oligosaccharides by antibodies at millimolar concentrations are near oneone-hundredth of a second.[56]

(54) Springer, T. A. *Cell* **1994**, *76*, 301−314.
(55) Otting, G.; Liepinsh, E.; Wüthrich, K. *Science* **1991**, *254*, 974−980.
(56) MacKenzie, C. R., Hirama, T., Deng, S., Bundle, D. R., Narang, S. A., Young, N. M. *J. Biol. Chem.* **1996**, *271*, 1527−1533.

The realization that a large force for association can be generated from the persistent presence of perturbed water appears highly relevant to the ease and speed of enzyme-catalyzed reactions. Water is normally in a high molecular excess in functioning biological systems and, because of its high heat capacity, can serve as a huge heat reservoir to maintain perturbed zones of water over the various biological surfaces that serve as donor and receptor sites. In terms of Scheme 1, the amount of energy liberated will depend in part on the number (n) of perturbed water molecules that are released to bulk. On this basis, it can be expected that the activation sites of enzymes have larger surface areas than those concerned with simple associations in the ground state. Of course, the amount of energy liberated from a given area will vary depending on the geometry and composition of the surface involved.

Summary

It seems unequivocal that water plays a central role in all biological associations because its interactions with molecular surfaces lead to perturbation by either ordering (nonpolar surface) or strain (polyamphiphilic surface). This “bottled” energy is always on tap to help drive specific molecular events. Like a chaperon, water accompanies the reactants in their search for each other. Since the epitope of the substrate is complementary to the active site of the enzyme, it can displace the always rapidly moving perturbed water molecules with ease to form a properly bonded complex with coordinates close to those of the transition state.

I owe thanks to the Department of Chemistry, University of Alberta, for allowing me to continue, after retirement, the research that led to this Account.

AR9600087