

Conformational Analysis of the Milk Oligosaccharides

MARGARET BISWAS and V. S. R. RAO, *Molecular Biophysics Unit,
Indian Institute of Science, Bangalore-560 012, India*

Synopsis

Possible conformations of lacto-*N*-tetraose, lacto-*N*-neotetraose, related disaccharides, and other milk oligosaccharides have been studied by an energy-minimization procedure using empirical potential functions. Lacto-*N*-tetraose favors a "curved" conformation, while lacto-*N*-neotetraose favors an approximately "straight" conformation. These two conformations differ mainly in the position of the terminal galactose residue with respect to the rest of the molecule. This difference explains the greater strength of lacto-*N*-neotetraose compared with lacto-*N*-tetraose in its ability to inhibit the cross-reaction of blood group P1 fractions with Type XIV pneumococcal antipolysaccharide. Although the favored conformation of lacto-*N*-tetraose (inactive) agrees with the model proposed by the earlier workers, that for lacto-*N*-neotetraose (active) differs. The favored conformations for the disaccharides galactose- β (1-4)-*N*-acetylglucosamine, galactose- β (1-3)-*N*-acetylglucosamine, and lactose are similar in overall shape, differing only in the nature and orientation of the side groups. This explains their nearly equal inhibitory activity. These theoretical models also explain the increased activity of lacto-*N*-fucopentaose I over that of lacto-*N*-tetraose and the relative activities of the substituted lactoses. The present studies suggest that it is the overall shape of the molecule which is important for activity, rather than the terminal β (1-4)-linked galactose residue alone.

INTRODUCTION

A set of criteria is being used to determine the specificity of combining sites on macromolecules, including antibodies and lectins.¹ Once the combining site size and specificities of antibodies and of lectins are known, they become of great value in establishing structures of unknown oligosaccharides.²

A great deal of work has been done on antibody-antigen reactions in solution.³ Kabat and coworkers,^{4,5} in a comparative study of various oligosaccharides from human milk, found that lacto-*N*-neotetraose was better than lacto-*N*-tetraose in inhibiting the cross-reaction of blood group P1 fractions with Type XIV pneumococcal antipolysaccharide. The difference in the activity of these two tetrasaccharides has been attributed to the important role of a terminal galactocyl- β (1-4) linkage in SXIV specificity. This reasoning could not, however, explain the equal inhibiting potency of the disaccharides galactose- β (1-3)-*N*-acetylglucosamine and lactose. The increase in the activity of lacto-*N*-tetraose and the decrease in the

activity of lactose on addition of an L-fucose residue to the C-2 atom of the terminal galactose residues also could not be understood.

It is reasonable to expect that antibodies will be specific for an antigen in its minimum-energy conformation or close to it, since these are the most likely conformations for the molecules in solution. From a study of the favored conformations of the antigens it is possible to obtain information about the specificity of binding of antibodies and lectins. We have attempted to compute theoretically the most favored conformations of lacto-*N*-tetraose, lacto-*N*-neotetraose, and the related milk oligosaccharides. This would throw further light on the nature of antigen-antibody interactions.

METHOD OF CALCULATION

The various milk oligosaccharides described in the present work are shown in Fig. 1. Steric maps were constructed using the contact criteria.⁶ The monomeric sugar residues were taken as rigid, and all possible conformations of di- and oligosaccharides were generated by making rotations about the interunit glycosidic bonds, C1-O (ϕ rotation) and O-CX' (ψ rotation), through -180° to $+180^\circ$ at intervals of 10° . All sugar residues were assumed to be in the 4C_1 (D) form, except for L-fucose, which was assumed to be in the 1C_4 (L) form. The atomic coordinates of each residue were based on the standard residue of Arnott and Scott,⁷ compiled from crystal structure data. The acetamido group was fixed using Pauling-Corey geometry⁸ so that the C2-H2 and N-H bonds were *trans*. The bond angle at the glycosidic oxygen atom⁹ was fixed at the average value of 117.5° . The initial conformation corresponding to $(\phi, \psi) = 0^\circ, 0^\circ$ is the same as that described earlier.¹⁰ A clockwise rotation was taken as positive.

The effects of the L-fucose residue on the allowed conformations of the two terminal disaccharide fragments of the tetrasaccharides galactose- β (1-3)-*N*-acetylglucosamine and galactose- β (1-4)-*N*-acetylglucosamine were investigated by constructing contact maps for these two disaccharides while fixing the L-fucose residue in all its allowed orientations sequentially.

The potential energy of the molecules was computed considering non-bonded, electrostatic, and torsional contributions. The form of the functions and the constants used are those reported by Momany et al.¹¹ The fractional charges on various atoms were taken from Yathindra and Rao.¹²

The various milk oligosaccharides can assume a large number of conformations, since many rotational angles are involved in specifying their conformation. A systematic analysis by varying these angles at discrete intervals was not practicable with the available computer facilities. To simplify the calculations, the two basic structures, lacto-*N*-tetraose and lacto-*N*-neotetraose, were studied by the Fletcher-Powell¹³ Davidson¹⁴ minimization procedure. Subsequently, the L-fucose residue was added

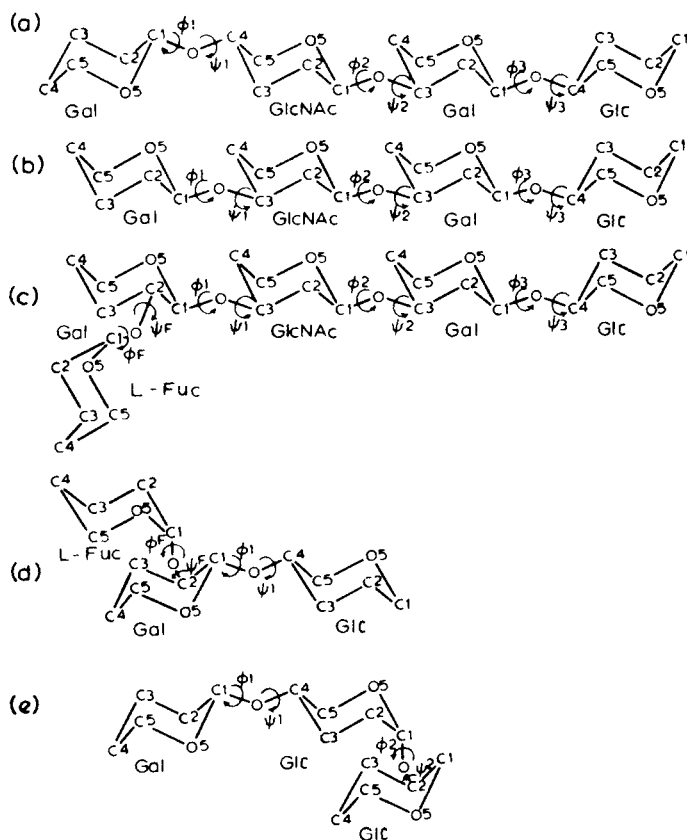


Fig. 1. Numbering of atoms and dihedral angles in the various milk oligosaccharides: (a) lacto-*N*-neotetraose, (b) lacto-*N*-tetraose, (c) lacto-*N*-fucopentaose I, (d) 2'-fucosidolactose, and (e) lactosyl- α (1-2) glucose. The side groups are not shown. Abbreviations: Fuc, fucose; Gal, galactose; Glc, glucose; GlcNAc, glucosamine.

to tetraose and lactose in their minimum-energy conformations ($0-3 \text{ kcal mol}^{-1}$) and the nonreducing ends of the molecules were again minimized. In this way, the minimum-energy conformations of the fucose-containing oligosaccharides were obtained. This approach is justified, since from model-building it can be seen that additions at the nonreducing end of these molecules interfere only with those (ϕ , ψ) angles on either side of the substituent and not with those further removed.

RESULTS AND DISCUSSION

In the disaccharide maps (Figs. 2 and 3), the allowed region in the (ϕ , ψ) plane is restricted to 4% or less of the total area. This is in general agreement with earlier results for oligosaccharides.¹⁵ Figures 2 and 3 also show that the addition of an L-fucose residue to the terminal galactose further restricts the allowed conformations of the disaccharide fragments, galac-

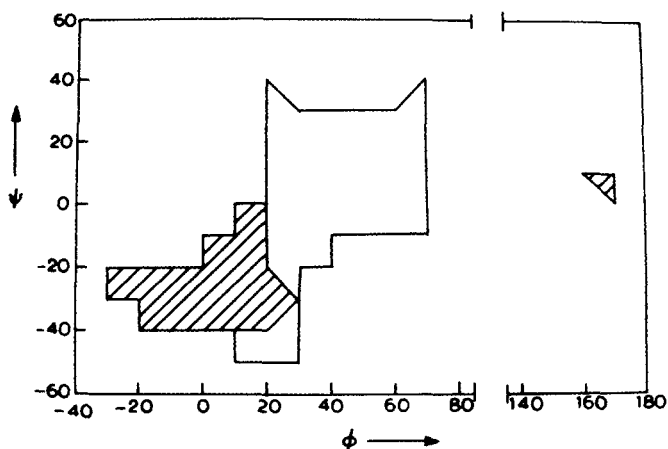


Fig. 2. Steric map for the disaccharide fragment galactose- $\beta(1-4)$ -*N*-acetylglucosamine. Shaded area is disallowed when an L-fucose residue is attached to the C-2 atom of the galactose residue.

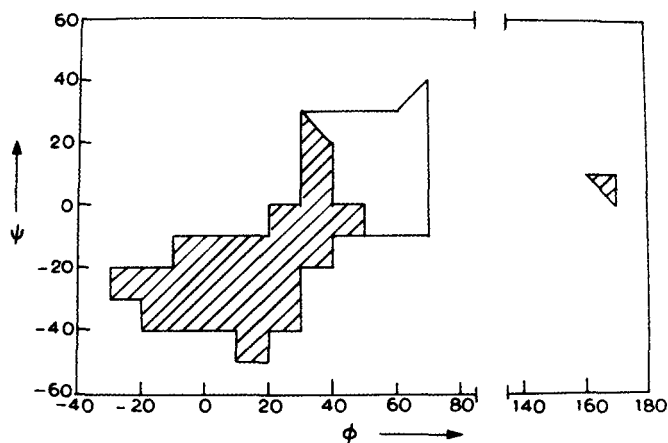


Fig. 3. Steric map for the disaccharide fragment galactose- $\beta(1-3)$ -*N*-acetylglucosamine. Shaded area is disallowed when an L-fucose residue is attached to the C-2 atom of the galactose residue.

tose- $\beta(1-4)$ -*N*-acetylglucosamine and galactose- $\beta(1-3)$ -*N*-acetylglucosamine, suggesting that these trisaccharide moieties are highly restricted.

Tables I and II show that a number of different conformations are possible for the tetrasaccharides, lacto-*N*-neotetraose and lacto-*N*-tetraose. It is interesting to note that in all the conformations with energies less than 3 kcal mol^{-1} , (ϕ_2, ψ_2) for both tetrasaccharides always favor values around $(50^\circ, 10^\circ)$. Two different values are favored in each case for (ϕ_1, ψ_1) , namely, $(50^\circ, 10^\circ)$ and $(-20^\circ, -30^\circ)$. (ϕ_3, ψ_3) can assume three different values, around $(50^\circ, 10^\circ)$, $(-20^\circ, -30^\circ)$, and $(10^\circ, -40^\circ)$. These results indicate that lacto-*N*-neotetraose and lacto-*N*-tetraose can assume very few conformations.

TABLE I
 Starting and Minimized Conformations^a of Lacto-*N*-Neotetraose

Set No.	Angle (deg)						Relative Energy (kcal mol ⁻¹)
	ϕ_1	ψ_1	ϕ_2	ψ_2	ϕ_3	ψ_3	
1	55	1	49	18	53	6	0.00
	60	10	50	10	60	10	
2	55	0	59	-12	53	7	0.23
	60	10	0	-40	60	10	
3	-15	-31	49	19	54	6	0.65
	0	-20	50	10	60	10	
4	55	0	48	18	-14	-32	0.70
	60	10	50	10	0	-20	
5	-4	-35	61	-12	55	3	0.89
	0	-20	0	-40	0	-20	
6	55	0	59	-12	-8	-35	1.01
	60	10	0	-40	0	-20	
7	-15	-31	48	19	-12	-33	1.41
	0	-20	50	10	0	-20	
8	-15	-31	55	3	-4	-36	1.64
	0	-20	0	-40	0	-20	
9	-9	-20	80	-24	57	-1	3.21
	0	-20	180	0	60	10	
10	54	0	44	21	161	8	3.70
	60	10	50	10	180	0	
11	49	9	49	17	162	8	3.91
	60	10	0	-40	180	0	
12	23	-42	94	-31	53	6	3.92
	60	10	180	0	60	10	
13	160	8	48	19	53	6	4.27
	180	0	50	10	60	10	
14	53	4	95	-31	18	-46	4.32
	60	10	180	0	0	-20	
15	-15	-31	45	21	161	8	4.37
	0	-20	50	10	180	0	
16	160	8	57	3	55	3	4.43
	180	0+0	-40	60	10		
17	0	-37	95	-30	19	-45	4.86
	0	-20	180	0	0	-20	

^a In each set the second row is the starting conformation and the first row is the minimized conformation.

It is also interesting to note from Tables I and II that the values of (ϕ_i , ψ_i) in the region (160°, 10°) increase the conformational energy significantly. The increase in energy is much higher when (ϕ_2 , ψ_2) assumes values in this region, in contrast to (ϕ_1 , ψ_1) and (ϕ_3 , ψ_3).

Tables III and IV show that only one set of values around (50°, 10°) is favored for (ϕ_1 , ψ_1) in the trisaccharide fragment L-Fuc- α (1-2)-gal- β (1-3)glcNAc of lacto-*N*-fucopentaose I and in 2'-fucosidolactose (Figs. 2 and 3). Thus, the additional L-fucose residue only restricts the favored orientation of the galactose residue and does not push it into a totally new conformation. For both these fucose-containing trisaccharides, the preferred value for (ϕ_F , ψ_F) is around (40°, 25°).

TABLE II
 Starting and Minimized Conformations^a of Lacto-*N*-Tetraose

Set No.	Angle (deg)						Relative Energy (kcal mol ⁻¹)
	ϕ_1	ψ_1	ϕ_2	ψ_2	ϕ_3	ψ_3	
1	54	3	49	19	52	6	0.00
	60	10	50	10	60	10	
2	-21	-27	48	19	54	5	0.08
	0	-20	50	10	60	10	
3	54	-1	50	15	52	9	0.10
	60	10	0	-40	60	10	
4	-23	-24	51	17	54	8	0.17
	0	-20	0	-40	60	10	
5	55	3	47	19	-15	-31	0.71
	60	10	50	10	0	-20	
6	-23	-26	49	18	19	-43	0.77
	0	-20	50	10	0	-20	
7	54	4	56	11	-3	-36	0.93
	60	10	0	-40	0	-20	
8	-23	-24	57	6	4	-40	1.12
	0	-20	0	-40	0	-20	
9	162	6	48	19	53	6	2.86
	180	0	50	10	60	10	
10	161	6	59	-10	53	5	3.10
	180	0	0	-40	60	10	
11	54	4	96	-31	54	5	3.51
	60	10	180	0	60	10	
12	162	6	48	20	19	-46	3.62
	180	0	50	10	0	-20	
13	54	4	44	21	161	8	3.64
	60	10	50	10	180	0	
14	-22	-27	95	-31	52	8	3.66
	0	-20	180	0	60	10	
15	-21	-27	45	21	161	8	3.81
	0	-20	50	10	180	0	
16	49	11	52	18	161	8	3.83
	60	10	0	-40	180	0	
17	160	7	58	-10	18	-45	3.83
	180	0	0	-40	0	-20	
18	50	-12	157	10	56	10	4.20
	180	0	180	0	60	10	
19	55	6	96	-31	17	-45	4.26
	60	10	180	0	0	-20	
20	-23	-27	96	-31	19	-44	4.34
	0	-20	180	0	0	-20	

^a In each set the second row is the starting conformation and the first row is the minimized conformation.

Table V shows that for lactosyl- α (1-2)-glucose the values of (ϕ_1, ψ_1) fall in the two regions near ($50^\circ, 10^\circ$) and ($-20^\circ, -30^\circ$). This is not different from the result obtained for (ϕ_1, ψ_1) in lacto-*N*-neotetraose. For conformations falling within 3 kcal mol⁻¹ of minimum (ϕ_2, ψ_2) can assume three different values—around ($-40^\circ, -20^\circ$), ($-20^\circ, 45^\circ$), and ($20^\circ, 30^\circ$).

TABLE III
Starting and Minimized Conformations^a of the L-Fuc- α (1-2)-gal- β (1-3) glcNAc Fragment of Lacto-*N*-Fucopentaose I

Set No.	Angle (deg)				Relative energy (kcal mol ⁻¹)
	ϕ_F	ψ_F	ϕ_1	ψ_1	
1	39	25	65	2	0.00
	30	20	60	20	
2	-22	-21	68	12	4.60
	-20	-20	60	0	
3	30	-10	139	18	5.28
	40	25	160	0	

^a In each set the second row is the starting conformation and the first row is the minimized conformation.

TABLE IV
Starting and Minimized Conformations^a of the 2'-Fucosidolactose

Set No.	Angle (deg)				Relative Energy (kcal mol ⁻¹)
	ϕ_F	ψ_F	ϕ_1	ψ_1	
1	37	24	63	-6	0.00
	30	20	60	20	
2	-17	-25	56	-1	2.06
	-20	-10	60	20	

^a In each set the second row is the starting conformation and the first row is the minimized conformation.

TABLE V
Starting and Minimized Conformations^a of the Lactosyl- α (1-2)-Glucose

Set No.	Angle (deg)				Relative Energy (kcal mol ⁻¹)
	ϕ_1	ψ_1	ϕ_2	ψ_2	
1	55	3	-40	-19	0.00
	60	10	-40	-30	
2	-16	-30	-40	-19	0.83
	-20	-20	-40	-30	
3	55	2	-21	44	1.50
	60	10	-30	30	
4	55	2	17	30	1.60
	60	10	20	20	
5	16	-42	-21	44	2.24
	-20	-20	-30	30	
6	-16	-30	17	30	2.40
	-20	-20	20	20	
7	160	8	-41	-19	4.15
	160	10	-40	-30	

^a In each set the second row is the starting conformation and the first row is the minimized conformation.

The most favored conformations for lacto-*N*-neotetraose and lacto-*N*-tetraose are shown in Fig. 4. Lacto-*N*-tetraose takes up a "curved" conformation, with the CH₂OH groups of the two galactose residues and

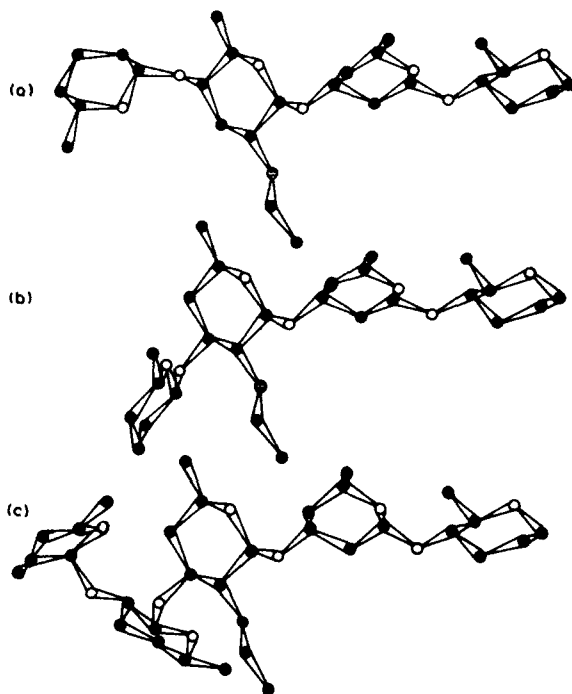


Fig. 4. Projections of (a) lacto-*N*-neotetraose in minimum-energy conformation, (b) lacto-*N*-tetraose in minimum energy, and (c) lacto-*N*-fucopentaose I in conformation 5 kcal mol⁻¹ above the minimum (Table III, set 3). Symbols: ●, carbon; ○, oxygen; ⊙, nitrogen.

of the *N*-acetylglucosamine residue lying on the same side of the molecule. The change from a $\beta(1-3)$ - to a $\beta(1-4)$ -linkage at the nonreducing end results in the approximately "straight" conformation favored for lacto-*N*-neotetraose in which the CH₂OH group of the nonreducing galactose residue falls on the opposite side to the CH₂OH groups of the *N*-acetylglucosamine and the other galactose residue.

It is interesting to note that the minimum-energy conformation of lacto-*N*-tetraose agrees with the model proposed by Kabat,^{2,4} while the one for lacto-*N*-neotetraose differs. In proposing his models, Kabat imposed the restriction that in both molecules the three CH₂OH groups should fall on the same side. Such an arrangement leads to the minimum-energy conformation for lacto-*N*-tetraose, whereas for lacto-*N*-neotetraose it leads to a conformation having 4 kcal mol⁻¹ above the minimum (Table I, set 13). Such a conformation would thus be highly unlikely to occur in solution.

It is interesting to note that the shapes of these two tetrasaccharides differ mainly in the disposition of the terminal galactose residue, which is important in binding. If the galactose residue alone is important for binding, then both tetrasaccharides should have the same activity. Experiments clearly reveal⁴ that lacto-*N*-neotetraose is a much better inhibitor

of the cross-reaction of blood group P1 fractions with Type XIV pneumococcal antipolysaccharide as compared to lacto-*N*-tetraose. This suggests that besides the nature of the terminal residue, the overall shape of the tetrasaccharide is also important. Figure 5 is a schematic representation of the two tetrasaccharides in the antibody binding site. It is not meant to suggest that this is the maximum size for such a binding site. In fact, there is evidence to suggest that an antibody binding site can accommodate up to six hexoses.¹⁶ It is interesting to note that for the inactive lacto-*N*-tetraose, site A is unoccupied.

As mentioned above, the addition of the L-fucose residue to the C2 atom of the terminal nonreducing galactose residue of lacto-*N*-tetraose does not alter the favored conformation of the last galactose residue in lacto-*N*-fucopentaose I. Hence, the observed increase in activity cannot be attributed to better positioning of the galactose residue in the active site. However, from Fig. 4 it can be seen that when ϕ_1 is around 180° (Table III, set 3), the L-fucose residue can be partially superimposed on the terminal galactose residue of the highly active lacto-*N*-neotetraose and thus partially occupies site A (Fig. 5). Such a value for ϕ_1 ($\sim 180^\circ$) leads to an increase in energy of about 5 kcal mol^{-1} over the minimum. This, and the fact that the L-fucose residue does not exactly superimpose on the galactose residue (since the orientation of the hydroxyl groups are different) would explain its low activity compared to lacto-*N*-neotetraose. It thus seems that the inactivity of lacto-*N*-tetraose is not due to steric interference of the $\beta(1-3)$ -linked terminal galactose residue with the active site, but rather to the absence of the sugar residue at site A.

Projections of the minimum-energy conformations of the disaccharides galactose $\beta(1-4)$ -*N*-acetylglucosamine, galactose- $\beta(1-3)$ -*N*-acetylglucos-

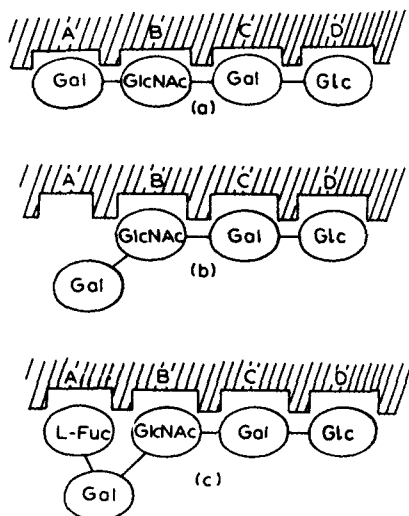


Fig. 5. Schematic representation of (a) lacto-*N*-neotetraose, (b) lacto-*N*-tetraose, and (c) lacto-*N*-fucopentaose I in the antibody binding site. Abbreviations as in Fig. 1.

amine, and lactose are shown in Fig. 6. It can be seen that all are similar in overall shape in spite of the differences in the linkage. The change in the linkage affects particularly the orientations of the side groups of the second residue. In galactose- $\beta(1-4)$ -*N*-acetylglucosamine, the acetamido group of *N*-acetylglucosamine is placed on the side opposite to the CH₂OH group of the galactose residue, whereas in galactose $\beta(1-3)$ -*N*-acetylglucosamine, the acetamido group falls on the same side as the CH₂OH group of the galactose residue. In other words, the change in the linkage causes the orientation of the CH₂OH and acetamido groups of the *N*-acetylglucosamine residue to interchange. Lactose, on the other hand, differs from galactose- $\beta(1-4)$ -*N*-acetylglucosamine in the nature of the substituent at the C-2 atom of the glucose residue. Since the overall shape of these disaccharides is similar, these minor differences in the side groups or their exchange of positions should not interfere with their fitting in the binding site. This explains the comparable inhibiting activities of the disaccharides.⁴ Since the shape and the nature of the residue at the nonreducing end (galactose) in these disaccharides is similar to the terminal disaccharide

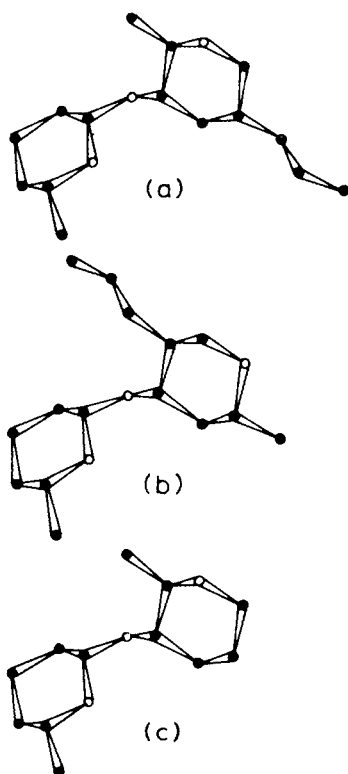


Fig. 6. Projections of the minimum-energy conformations of the disaccharides (a) galactose- $\beta(1-4)$ -*N*-acetylglucosamine, (b) galactose- $\beta(1-3)$ -*N*-acetylglucosamine, and (c) lactose. Symbols as in Fig. 4.

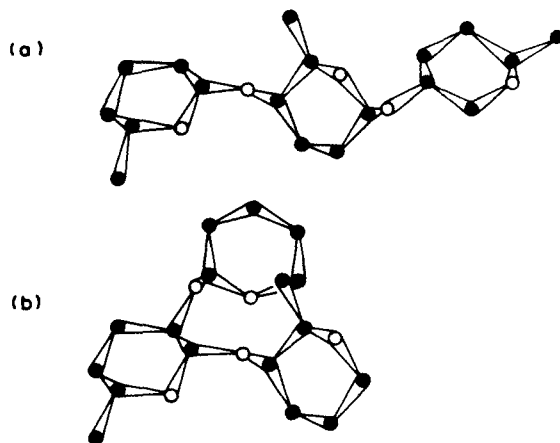


Fig. 7. Projections of the minimum-energy conformations of (a) lactosyl- $\alpha(1-2)$ -glucose and (b) 2'-fucosidolactose. Symbols as in Fig. 4.

moiety of the highly active inhibitor, lacto-*N*-neotetraose, it is likely that these disaccharides will bind in regions A and B of the binding site (Fig. 5). However, the higher activity of galactose- $\beta(1-4)$ -*N*-acetylglucosamine over lactose and galactose- $\beta(1-3)$ -*N*-acetylglucosamine suggests that the acetamido group in the galactose- $\beta(1-4)$ -*N*-acetylglucosamine also plays a significant role in binding.

Figure 7 shows the minimum-energy conformations of lactosyl- $\alpha(1-2)$ -glucose and 2'-fucosidolactose. It is interesting to note that substitution of the glucose residue at the reducing end of lactose does not substantially alter the chain direction, although the glucose residue does not superimpose on the galactose residue, which would normally occupy site C, but is slightly displaced. Such an orientation of the glucose residue apparently does not sterically interfere with the binding of lactosyl- $\alpha(1-2)$ -glucose to the antibody binding site. On the other hand, substitution of the L-fucose residue at the nonreducing end of lactose alters the contour drastically [Fig. 7(b)], and hence 2'-fucosidolactose cannot be recognized by the antibody binding site. This explains the observed activity of lactosyl- $\alpha(1-2)$ -glucose and the inactivity of 2'-fucosidolactose.⁴

The present work indicates that the earlier belief that antibodies are specific for a particular linkage cannot be applied in a general way, especially in the case of the disaccharides, for which the linkage does not affect the overall shape.

References

1. Kabat, E. A. (1976) *Structural Concepts in Immunology and Immunochemistry*, 2nd ed., Holt, Rinehart and Winston, New York, Chap. 6.
2. Kabat, E. A. (1976) in *Human Blood Groups*, Fifth International Convocation on Immunology, Buffalo, N.Y., Karger, Basel, 1977, pp. 236-245.
3. Kabat, E. A. (1978) *J. Supramol. Struct.* **8**, 79-88.
4. Kabat, E. A. (1962) *Arch. Biochem. Biophys.* (Suppl. 2), 181-186.

5. Rovis, L., Kabat, E. A., Pereira, M. E. A. & Feizi, T. (1973) *Biochemistry* **12**, 5355-5360.
6. Ramachandran, G. N. & Sasisekharan, V. (1968) *Adv. Protein Chem.* **23**, 283-437.
7. Arnott, S. & Scott, W. E. (1972) *J. Chem. Soc., Perkin Trans.* **2**, 324-335.
8. Corey, R. B. & Pauling, L. (1953) *Proc. R. Soc. London, Ser. B* **141**, 10-20.
9. Sundarajan, P. R. & Rao, V. S. R. (1970) *Biopolymers* **9**, 1239-1247.
10. Yathindra, N. & Rao, V. S. R. (1970) *Biopolymers* **9**, 783-787.
11. Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. (1975) *J. Phys. Chem.* **79**, 2361-2381.
12. Yathindra, N. & Rao, V. S. R. (1972) *Carbohydr. Res.* **25**, 256-260.
13. Fletcher, R. & Powell, M. J. D. (1963) *Comput. J.* **6**, 163-168.
14. Davidson, W. C. (1959) AEC Research and Development Report, ANL 5990.
15. Sathyanarayana, B. K. (1971) Ph.D. thesis, University of Madras, Madras, India.
16. Kabat, E. A. (1960) *J. Immunol.* **84**, 82-85.

Received December 19, 1979

Accepted March 21, 1980