Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The combination of molecular dynamics with crystallography for elucidating protein–ligand interactions: a case study involving peanut lectin complexes with T-antigen and lactose

J. V. Pratap,[a] G. M. Bradbrook,[b] G. B. Reddy,[a] A. Surolia,[a] J. Raftery,[b] J. R. Helliwell[b] and M. Vijayan[a]*

[a]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, and [b]Department of Chemistry, University of Manchester, Manchester M13 9PL, England

Correspondence e-mail: mv@mbu.iisc.ernet.in

Received 24 January 2001
Accepted 16 July 2001

Peanut lectin binds T-antigen [Galβ(1–3)GalNAc] with an order of magnitude higher affinity than it binds the disaccharide lactose. The crystal structures of the two complexes indicate that the higher affinity for T-antigen is generated by two water bridges involving the acetamido group. Fresh calorimetric measurements on the two complexes have been carried out in the temperature range 280–313 K. Four sets of nanosecond molecular-dynamics (MD) simulations, two at 293 K and the other two at 313 K, were performed on each of the two complexes. At each temperature, two somewhat different protocols were used to hydrate the complex in the two runs. Two MD runs under slightly different conditions for each complex served to assess the reliability of the approach for exploring protein–ligand interactions. Enthalpies based on static calculations and on MD simulations favour complexation involving T-antigen. The simulations also brought to light ensembles of direct and water-mediated protein–sugar interactions in both the cases. These ensembles provide a qualitative explanation for the temperature dependence of the thermodynamic parameters of peanut lectin–T-antigen interaction and for the results of one of the two mutational studies on the lectin. They also support the earlier conclusion that the increased affinity of peanut lectin for T-antigen compared with that for lactose is primarily caused by additional water bridges involving the acetamido group. The calculations provide a rationale for the observed sugar-binding affinity of one of the two available mutants. Detailed examination of the calculations point to the need for exercising caution in interpreting results of MD simulations: while long simulations are not possible owing to computational reasons, it is desirable to carry out several short simulations with somewhat different initial conditions.

1. Introduction

The relation between structure and thermodynamics is a key question in structural biology. This question has received particular attention in recent years in relation to lectins (Toone, 1994; Chevernak & Toone, 1995; Lemieux, 1996; Loris *et al.*, 1998; Bradbrook *et al.*, 1998, 2000), which are carbohydrate-binding proteins involved in a variety of biological processes (Sharon & Lis, 1989; Rini, 1995; Vijayan & Chandra, 1999). Those from the seeds of leguminous plants are the most extensively studied lectins. A wealth of structural and thermodynamic information is available on them (Schwarz *et al.*, 1993; Mandal *et al.*, 1994; Chevernak & Toone, 1995; Rini, 1995, 1999; Surolia *et al.*, 1996; Drickamer, 1997, 1999; Gupta *et al.*, 1997; Weis *et al.*, 1998; Vijayan & Chandra, 1999). Broad explanations have been offered for the carbo-

hydrate specificities of legume lectins, as revealed by thermodynamic measurements, on the basis of the crystal structures of the relevant lectin–carbohydrate complexes (Elgavish & Shaanan, 1998; Drickamer, 1999; Bouckaert *et al.*, 1999; Rini, 1999; Vijayan & Chandra, 1999; Manoj *et al.*, 2000; Ravishankar *et al.*, 2000), but they have been qualitative and not altogether satisfactory. The first attempt to use molecular dynamics, in addition to the obvious molecular mechanics, to supplement the information provided by X-ray crystallography to address this problem was made by some of us (Bradbrook *et al.*, 1998). Undoubtedly, the most desirable result is the prediction of the free energies of binding of a given lectin and its carbohydrate ligands. The attempt of Bradbrook *et al.* (1998) addressed, as a first step, the problem of explaining the difference in the strengths of the binding of two similar ligands, glucoside and mannoside, to concanavalin A. Although no significant difference was found between the binding enthalpies of mannoside and glucoside, the work suggested that dynamic models are needed to provide a more complete picture. A similar study on the binding of the monosaccharide galactose and the disaccharide *N*-acetyl lactosamine to *Erythrina corallodendron* was subsequently carried out (Bradbrook *et al.*, 2000) in which different starting conditions were generated with a protocol of cycles of heating followed by cooling. Here, we report an attempt to explore the binding of peanut lectin to lactose and the T-antigenic disaccharide Galβ(1–3)GalNAc by means of sets of MD simulations carried out under slightly different conditions in each case. These two sugars are of similar structure, but have very different affinities for the lectin.

Peanut lectin (PNA) is a tetrameric non-glycosylated lectin, M_r = 110 000, with four identical subunits which are arranged in an unusual quaternary structure (Banerjee *et al.*, 1994). Each subunit binds a galactose molecule in its primary binding site. It is highly specific to the tumour-associated T-antigenic determinant [Galβ(1–3)GalNAc, Thomsen–Friedenreich antigen]. The crystal structures of several peanut-lectin complexes, including those with lactose and T-antigen, have been determined (Banerjee *et al.*, 1994, 1996; Ravishankar *et al.*, 1997, 1998, 1999, 2000) to resolutions ranging from 2.8 to 2.25 Å. Therefore, reasonable starting points, including those involving bound waters, are available for MD simulations. Although peanut lectin binds T-antigen with an order of magnitude higher strength than it does lactose, the crystal structures showed that the direct protein–carbohydrate interactions in the complexes essentially are the same. The only significant additional interactions found in the crystals of the PNA–T-antigen complex are two water bridges between the O atom of the acetamido group of the sugar and the lectin (Fig. 1) (Ravishankar *et al.*, 1997, 1999). The complexes provided a telling example of the generation of ligand specificity through water-mediated interactions. However, the result was so unexpected and novel that it appeared necessary to explore it further.

Lectin–carbohydrate interactions are generally enthalpically driven, although some entropy compensation also exists (Toone, 1994). Thus, in the present study, as in the earlier two studies, the analyses have been simplified by considering enthalpies rather than free energies. As described in some detail in earlier publications (Bradbrook *et al.*, 1998, 2000), one in effect calculates the difference ($\Delta\Delta H_{\rm bind}$) between the enthalpies of binding ($\Delta H_{\rm bind}$) of the lectin with the two sugars. Specifically, the enthalpies of the complexes of peanut lectin with T-antigen (CT) and lactose (CL) and of the

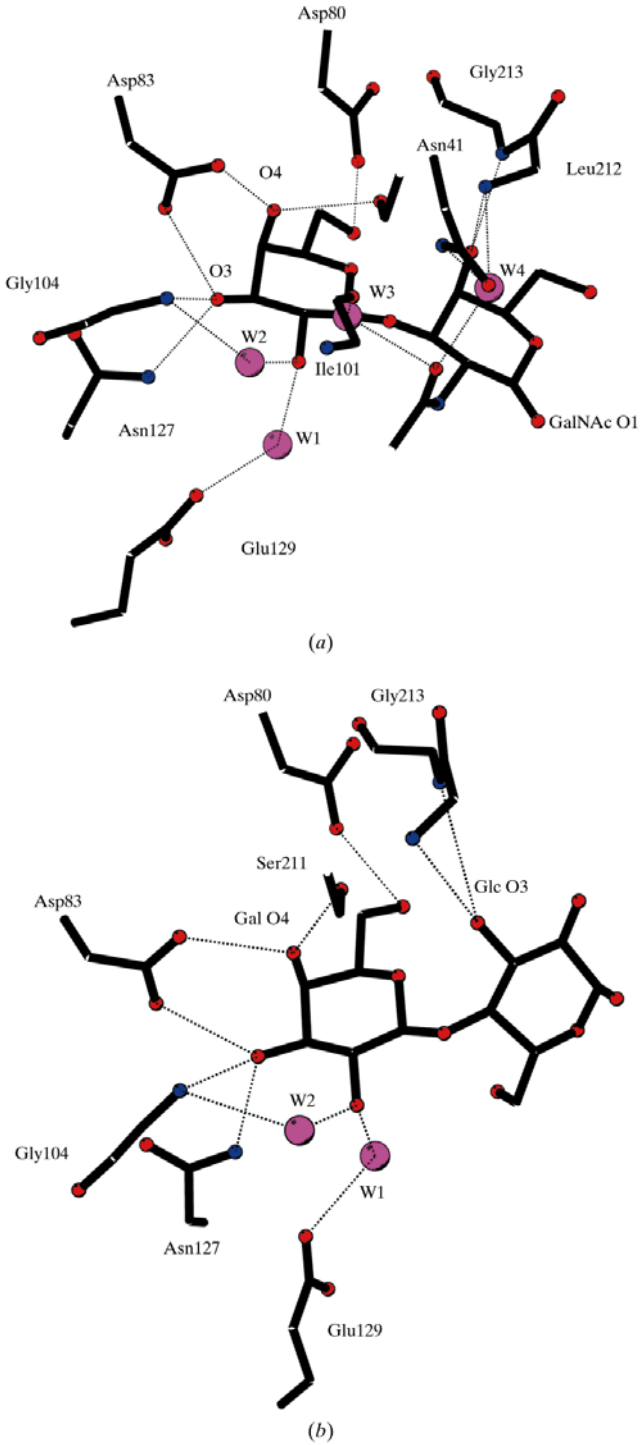

Figure 1
Peanut lectin–sugar interactions observed in the crystal structures of (*a*) the T-antigen complex and (*b*) the lactose complex.

corresponding free hydrated sugars (FT and FL) are estimated using molecular-mechanics and molecular-dynamic calculations, with the assumption that ΔH is approximately equal to the heat released at constant volume. Thus,

$$\Delta\Delta H_{\rm bind} = (H_{\rm CT} - H_{\rm FT}) - (H_{\rm CL} - H_{\rm FL}).$$

The unbound lectin is not explicitly considered, as its contribution will cancel out when the difference is taken. Furthermore, the molecular-dynamics calculations also provide, perhaps more importantly, an ensemble of states. Mini-ensembles of X-ray crystal structures are often available *via* multiple copies in the asymmetric unit and different crystal forms of the same protein (Elgavish & Shaanan, 1998; Ravishankar *et al.*, 1999; Bradbrook *et al.*, 1998), but they are limited and do not correspond to a full dynamical simulation. The additional interactions accessible in MD calculations result in new insights into the modes of binding and provide a better appreciation of the thermodynamic data reported here and elsewhere (Pereira *et al.*, 1976; Neurohr *et al.*, 1982).

In addition to questions concerning the reliability of the force fields used, a major disadvantage of MD calculations is that the present computational power does not permit simulations long enough to ensure that all possible states have been adequately explored. In order to partially alleviate this disadvantage, four separate calculations with somewhat different conditions have been performed on each complex. Specifically, calculations have been performed at two different temperatures. At each temperature, two models which differ somewhat in the way the complexes are hydrated have been used as the starting point (see below). Such multiple calculations under slightly different conditions should also serve as a check on the reliability of the method, which is still in the early stages of development.

2. Materials and methods

Subunit *D* in the PNA crystal structures (PDB codes 2pel and 2tep) was chosen as the starting model for T-antigen as well as lactose simulations, since this subunit is involved in the least number of intertetramer interactions. The corresponding sugar coordinates were chosen for the free sugar simulations. *GAUSSIAN*94 (Frisch *et al.*, 1995) with a 6-31G* basis set was used to calculate the charges for the sugar atoms. The *AMBER*4.1 (Cornell *et al.*, 1995) forcefield was used along with the carbohydrate parameters of Woods *et al.* (1995).

Both the systems were energy minimized prior to the simulations. 20 000 minimization steps were performed. The first 200 steps used the steepest-descent approach which was then followed by the conjugate-gradient algorithm. A dielectric constant of 1.0 was used during minimization as well as dynamics. A cutoff of 12 Å was used for calculating non-bonded interactions. The SANDER module of *AMBER*4.1 was used to carry out MD and static energy minimization.

The following regions of the protein were allowed to move: residues 33–44, 73–86, 97–114, 120–138, 141–147, 207–218 and 237–240. These represent the loops near the sugar-binding region and the metal ions. In the ligands, the two torsion angles in the glycosidic linkage were allowed to vary. In the case of T-antigen, the torsion angles defining the conformation of the acetamido group were also allowed to vary. The inter-metal distance and the metal–water ligand distances were restrained and bonds involving H atoms were constrained using *SHAKE* (van Gunsteren & Berendsen, 1990). The simulations of the complexes were performed at two different temperatures, 293 and 313 K, reflecting the range of available thermodynamic data. At each temperature, two simulations were performed using slightly different hydration protocols. In the first protocol, a random 19 Å cap of TIP3P waters (Jorgensen, 1981) was added, while in the second water molecules located in the crystal structures within 7 Å of any non-carbon sugar heavy atom were added before completing solvation using the TIP3P waters. These two protocols are referred to as CTW1 and CTW2 in the T-antigen complex simulations. The corresponding protocols in the lactose complex simulations are CLW1 and CLW2. The free-sugar simulations were carried out at 313 K including solvation with a 15 Å cap of TIP3P waters. One atom of the sugar was fixed in order to prevent the sugar from drifting towards the solvent boundary. The systems were heated to the required temperature over 80 ps, followed by equilibration for 200 ps, using steps of 0.5 fs. After equilibration, dynamics were run for 1 ns, with data collected each picosecond. The analyses of the MD trajectories were achieved using in-house routines and the CARNAL module of *AMBER*. Energies were estimated at each step in the simulations and in the static calculation using the ANAL module of *AMBER*4.1. For the estimation of $\Delta\Delta H_{\rm bind}$, the MD simulations for the free sugars at 313 K have been used. Although the dynamics were run over a large part of the structure, the enthalpy calculations were limited to an 8 Å sphere centered on the binding site in order to reduce large statistical errors caused by insufficient sampling.

Two O or N atoms were considered as hydrogen bonded if the donor–acceptor distance is less than 3.6 Å and the donor—hydrogen—acceptor angle is greater than 120°. A protein O or N atom and a sugar O or N atom were considered to form a water-mediated interaction if a water O atom was less than 3.6 Å from each of them and if the angle subtended by these atoms at the water O atom is greater than 90°.

Table 1
Calorimetric data on binding of T-antigen and lactose to PNA.

Temp (K)	K_a (10^{-3} M^{-1})	ΔG (kJ mol^{-1})	ΔH (kJ mol^{-1})	$T\Delta S$ (kJ mol^{-1})
T-antigen				
281.3	53.6	−25.5	−49.4	−23.9
289.9	37.9	−25.5	−55.3	−29.7
297.9	20.6	−24.7	−59.0	−34.3
312.9	4.1	−21.8	−69.5	−47.7
Lactose				
280.8	4.2	−19.7	−39.3	−19.7
297.0	2.0	−18.8	−40.6	−21.8
312.5	0.5	−16.3	−41.0	−24.7

Table 2
Direct protein–sugar interactions that occur for at least 5% of the duration in any of the MD runs.

Interactions observed in the crystal structures are highlighted. The numbers correspond to the percentage of the time during which each interaction occurs.

Interaction	293CTW1	293CTW2	313CTW1	313CTW2	All Tant	293CLW1	293CLW2	313CLW1	313CLW2	All Lact
GalO2–Glu129 OE*x*	9.1	—	52.6	39.8	26.2	—	68.8	—	32.5	27.5
GalO2–Asn127 ND2	—	—	8.9	16.9	6.5	—	—	—	—	—
GalO2–Gly104 N	—	—	7.2	—	1.8	37.4	—	—	17.4	13.9
GalO3–Asp83 OD*x*	100	100	98.7	99.9	99.7	100	—	99.9	99.8	75.0
GalO3–Asn127 ND2	49.6	29.2	58.4	—	34.6	—	79.4	—	—	20.5
GalO3–Gly104 N	21.4	10.3	—	—	8.6	47.9	22.9	31.1	66.1	42.0
GalO3–Glu129 OE*x*	—	—	—	—	—	—	75.8	—	—	19.0
GalO4–Asp83 OD*x*	100	100	97.8	99.9	99.5	100	99.5	92.8	100	98.1
GalO4–Ser211 OG	99.8	87.3	13.4	91.4	73.0	95.9	23.8	94.0	95.7	77.4
GalO4–Asn127 ND2	—	—	—	—	—	—	18.2	—	—	4.6
GalO4–Gly104 N	—	—	—	—	—	—	6.9	—	—	1.7
GalO5–Ser211 OG	31.8	49.7	25.3	53.7	40.2	45.3	25.6	48.0	35.2	38.5
GalO5–Gly214 N	14.8	—	—	—	3.7	—	—	—	—	—
GalO6–Asp80 OD*x*	91.6	71.8	***	65.1	57.8	64.3	—	72.0	72.1	52.2
GalO6–Ser211 OG	—	—	34.2	—	8.8	—	6.6	—	—	1.6
GalO6–Gly214 N	—	—	13.2	—	3.7	—	19.3	5.2	—	6.2
GalNAcO1–Tyr130 OH	—	—	—	15.9	4.0	—	—	—	—	—
GalNAcN2–Glu129 OE*x*	—	—	8.3	35.6	11.0	—	—	—	—	—
GalNAcO4–Gly213 N	34.7	22.4	6.7	13.7	19.4	—	—	—	—	—
GalNAcO4–Leu212 N	—	13.4	—	—	5.6	—	—	—	—	—
GalNAcO4–Ser 211 OG	—	—	8.0	—	2.0	—	—	—	—	—
GalNAcO4–Gly103 N	—	—	29.4	—	7.4	—	—	—	—	—
GalNAcO4–Ile101 O	—	—	10.0	—	2.5	—	—	—	—	—
GalNAcO4–Asn41 ND2	—	—	13.5	—	3.4	—	—	—	—	—
GalNAcO6–Gly213 N	—	—	10.3	—	2.8	—	—	—	—	—
GalNAcO6–Leu212 N	—	—	35.8	—	9.1	—	—	—	—	—
GalNAcO6–Asn41 ND2	—	—	15.2	—	3.8	—	—	—	—	—
GalNAcO7–Leu212 N	—	—	—	5.7	1.9	—	—	—	—	—
GlcO2–Gly213 N	—	—	—	—	—	—	50.6	—	—	12.7
GlcO3–Gly213 N	—	—	—	—	—	66.6	70.8	60.3	76.5	68.6
GlcO3–Leu212 N	—	—	—	—	—	—	10.1	—	—	4.0
GlcO6–Tyr125 OH	—	—	—	—	—	—	6.6	—	—	1.4
GlcO6–Glu129 OE*x*	—	—	—	—	—	—	—	—	7.8	1.9

Thermodynamic parameters were determined by isothermal titration calorimetry (ITC) using a Microcal Omega titration calorimeter (Wiseman *et al.*, 1989). ITC experiments were performed in the *c*-value range of $2 < c < 20$, where $c = K_b$[PNA], K_b being the association constant and [PNA] the initial concentration of PNA monomers. This corresponds to a binding regime that is best suited for the accurate measurements of K_b, binding stoichiometry (n) and binding enthalpy change (ΔH_b^0) simultaneously in a single ITC experiment. The PNA, dialyzed exhaustively against PBS [20 m*M* phosphate-buffered (pH 7.4) saline (0.15 *M* NaCl)], was centrifuged to remove any particulate material and its concentration was determined by $A_{280}^{1\%} = 7.7$. Typically, the concentration of PNA monomers ranged from 1–4 m*M*, while that of the sugar varied between 12 and 50 m*M*. Sugar solutions were prepared in the last dialysate of PNA to ensure that the ionic composition of both the protein and the ligand solutions were the same. The PNA solution (1.34 ml) in the ITC cell was titrated with 5.2–6.0 µl additions of the saccharide solution through a rotating syringe at a constant stirring speed of 390 rev min^{-1}. The time duration between the injections was at least 3 min to allow the peak to return to the baseline. The titration of the saccharide solution into the buffer solution alone gave negligible heats of dilution. Nonetheless, for each experiment, the heat of the dilution of the ligand was subtracted from the runs performed with PNA. The ITC data thus obtained were fitted to obtain K_b, n and ΔH_b^0 using the *Origin* program (Wiseman *et al.*, 1989; Yang, 1990). Values of ΔS_b were obtained from

$$\Delta G_b^0 = \Delta H_b^0 - T\Delta S_b,$$

where $\Delta H_b^0 = -RT\ln K_b$, T being the absolute temperature.

3. Results and discussion

3.1. Static calculations

Energy-minimized models of the T-antigen complex, CTW1 and CTW2, and the corresponding lactose complexes, CLW1 and CLW2, were used to estimate the interaction energies. The calculations yielded $\Delta\Delta H_{\rm bind}$ values of −23.4 and −69.9 kJ mol^{-1} in favour of T-antigen, in qualitative agreement with the results of calorimetric measurements (Table 1). The T-antigen molecule is larger than the lactose molecule and, other things being equal, the complex of the former is expected to contain more interaction terms than the latter. Therefore, the result indicating lower interaction energies in the T-antigen complex is not surprising. However, the static calculations represent an average effect and do not provide information on the different possible modes of interaction and

their importance resulting from the mobility of the binding pocket and the ligand.

3.2. MD calculations: general features

Analysis of the molecular-dynamics calculations involved the examination of a total of eight runs: four for each complex, with two starting models at 313 K and two at 293 K. (The numbers 313 and 293 are prefixed to CTW1, CTW2, CLW1 and CLW2 to indicate the temperature at which the MD simulation was performed.) Each run involved 1000 steps. Altogether, 53 direct protein–ligand hydrogen bonds occur at some point during the four runs involving T-antigen. The corresponding number in the runs involving lactose is 32. Those that occur for more than 5% of the time in any given run in the two complexes are listed in Table 2. Also listed is the percentage occurrence averaged over the four runs in each case. In spite of the large number of protein–ligand hydrogen bonds accessed, the average number of such hydrogen bonds at a given time are 6.3 and 6.6, respectively, in the T-antigen and the lactose complexes. The number of hydrogen bonds observed in the crystal structures are 11 and 9, respectively. In both the complexes, lectin–carbohydrate interactions also involve water bridges. Altogether, 310 such water bridges are accessed in the T-antigen complex in the four runs, with an average value of 13.3 at any given time. The corresponding numbers in the lactose complex are 170 and 9.3. The number of water bridges observed in the crystal structures of the T-antigen and the lactose complexes are four and two, respectively.

3.3. Interactions accessed in the simulations

The number of interactions accessed during simulations is understandably larger than those observed in the crystal structures. Furthermore, the two O atoms in the carboxylate groups flip positions and cannot be meaningfully distinguished, a phenomenon also observed in the earlier studies (Bradbrook *et al.*, 1998, 2000).

Four direct interactions are observed in all legume lectin–galactose complexes. They are GalO3···Asp83 OD1, GalO3···Gly104 N, GalO3···Asn127 ND2 and GalO4···Asp83 OD2. It is interesting to examine their occurrence in the simulations. The GalO4···Asp83 OD*x* interaction occurs in all the runs well above 90% of the time. This is also true of GalO3···Asp83 OD*x* except in 293CLW2. The interaction involving Asn127 ND2 occurs for a reasonable fraction of the time except in the two simulations at 313 K. The GalO3···Gly104 N interaction is observed for substantial periods of time in the simulations involving lactose, but not in those involving T-antigen. In addition to the four invariant interactions, the following peanut lectin–galactose interactions are observed in the crystal structures: GalO4···Ser211 OG, GalO5···Ser211 OG and GalO6···Asp80 OD2. The first of these occurs with high frequency in six of the eight runs, but with low frequency in 313CTW1 and 293CLW2. The second occurs with moderate frequency in all the runs. The frequency of occurrence of GalO6···Asp80 OD*x* is moderately high in all runs except 313CTW1 and 293CLW2. GalNAcO4 of T-antigen and GlcO3 of lactose occupy nearly the same position in the crystal structures. Their interaction with Ser211 OG, observed in the crystal structures, does not occur for 5% of the duration or more in any of the eight runs. GlcO3···Gly213 N occurs for more than 50% of the time in each run involving lactose. However, the interaction involving GalNAcO4 occurs with low frequency in all the runs involving T-antigen. GalNAcO4/GlcO3···Leu212 N occurs only very rarely in all the runs.

Thus, the MD simulations of the complexes suggest that the crystallographically observed interactions are formed and broken dynamically with varying residence times. A large number of possible direct lectin–carbohydrate interactions not observed in the crystal structures are also accessed in the simulations. In general, they occur only infrequently, although

Table 3
Residency times of water molecules involved in water bridges that occur for more than 10% of the time during the four simulations.

Interactions observed in the crystal structures are highlighted.

	T-antigen		Lactose	
Interaction	Max (ps)	Average	Max (ps)	Average
GalO2···O Ile101	—	—	21	2.1
GalO2···O Gly102	14	1.6	13	17
GalO2···N Gly103	12	1.9	—	—
GalO2···N Gly104	63	4.7	34	2.9
GalO2···OH Tyr125	7	1.8	—	—
GalO2···O Ser126	28	3.0	—	—
GalO2···ND2 Asn127	33	3.2	18	2.1
GalO2···OE*x* Glu129	29	2.3	30	2.3
GalO2···OH Tyr130	87	4.5	25	2.6
GalO3···N Gly104	15	1.8	29	5.0
GalO3···OE*x* Glu129	57	3.6	52	2.3
GalO5···OD*x* Asp80	55	6.8	—	—
GalO5···OH Tyr125	—	—	6	1.3
GalO5···OG Ser211	—	—	34	4.7
GalO5···N Leu212	25	3.1	—	—
GalO5···N Gly213	152	6.9	57	4.6
GalO5···N Gly214	23	2.6	23	2.5
GalO6···OD*x* Asp80	26	2.1	32	2.9
GalO6···OH Tyr 125	9	1.4	—	—
GalO6···OG Ser211	35	2.4	—	—
GalO6···N Leu212	26	2.6	—	—
GalO6···N Gly213	64	2.6	—	—
GalO6···O Gly213	—	—	7	1.6
GalO6···N Gly214	23	2.7	—	—
GlycO···N Gly103	27	2.2	—	—
GlycO···OG Ser211	19	2.2	51	4.7
GlycO···N Leu212	61	2.2	24	1.8
GlycO···N Gly213	35	4.9	33	4.2
GalNAcN2···OH Tyr125	7	1.4	—	—
GalNAcN2···OE*x* Glu129	16	1.8	—	—
GalNAcO4···ND2 Asn41	18	2.3	—	—
GalNAcO4···OG Ser211	145	6.1	—	—
GalNAcO4···N Leu212	23	2.0	—	—
GalNAcO4···N Gly213	69	4.1	—	—
GalNAcO7···ND2 Asn41	11	1.7	—	—
GalNAcO7···O Ile101	10	1.7	—	—
GalNAcO7···N Gly103	15	2.1	—	—
GalNAc O7···OH Tyr130	20	2.1	—	—
GalNAcO7···N Leu212	29	2.3	—	—
GlcO3···OG Ser211	—	—	47	5.1
GlcO6···OH Tyr125	—	—	16	1.7
GlcO6···ND2 Asn127	—	—	12	2.0
GlcO6···OE*x* Glu129	—	—	13	1.7

Table 4
Five representative modes of interactions in simulations involving the T-antigen complex.

Mode No.	% of time observed	Direct	Water-mediated interactions
I	0.63	GalO3···Asp83 OD*x*, GalO3···Gly104 N, GalO3···Asn127 ND2, GalO4 ···Asp83 OD*x*, GalO4···Ser211 OG, GalO6···Asp80 OD*x*, GncO4···Gly213 N	GalO2···Glu129 OE*x*, GalO2···Gly104 N, GalO5···Asp80 OD*x*
II	1.00	GalO2···Asn127 ND2, GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*, GalO4···Ser211 OG, GalO5···Ser211 OG, GalO6···Asp80 OD*x*	GalO2···Glu129 OE*x*, GalO2···Ser126 O, GalO5···Gly213 N, GalNAcO3···Gly213 N, GalNAcO3···Leu212 N, GalNAcO4···Ser211 OG, GalNAcO4···Leu212 N, GalNAcO4···Gly213 N, GalNAcO7···Leu212 N
III	8.59	GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*, GalO4··· Ser211 OG, GalO6···Asp80 OD*x*	GalO2···Glu129 OE*x*, GalO2···Gly104 N, GalO2···Asn127 ND2, GalO3···Glu129 OE*x*, GalO4···Gly213 N, GalO5···Gly213 N, GalNAcO3···Leu212 N, GalNAcO3···Gly213 N, GalNAcO4···Ser211 OG, GalNAcO4···Leu212 N, GalNAcO4···Gly213 N
IV	1.20	GalO2···Glu129 OE*x*, GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*, GalO4···Ser211 OG	GalNAcO7···Lys112 NZ
V	0.70	GalO2···Glu129 OE*x*, GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*	GalO2···Ser126 O, GalO2···Ser128 N, GalO4···Ser211 N, GalO4···Ser211 OG, GalO4···Gly214 O, GalO5···Ser211 N, GalO5···Leu212 N, GalO5···Gly213 N, GalO5···Gly214 N, GalO5···Gly214 O, GalO6···Ser211 OG, GalO6··· Leu212 N, GalO6···Gly213 N, GalO6···Asp80 OD*x*, GalNAcN2···Glu129 OE*x*, GalNAcO4···Gly213 N, GalNAcO4···Gly214 N, GalNAcO6···Gly214 N, GalNAcO6···Gly213 N

there are a few exceptions. The most extensive interactions not observed in the crystal structures accessed by MD are those involving the side chain of Glu129 with GalO2 and GalO3. In the crystal structures, a water bridge involving this residue with GalO2 is observed. MD simulations suggest that Glu129 OE*x* can also directly interact not only with GalO2, but also with GalO3. The simulations also indicate the possibility of its direct interactions with GalNAcN2 in the T-antigen complex. Asn41 and Ile101 are the other residues which exhibit only water-mediated interactions in the crystal structure and only in the PNA–T-antigen complex. The MD calculations, however, indicate that these two residues can directly interact with GalNAcO4 and GalNAcO6.

In the crystal structure of both the complexes, GalO2 forms water bridges with Glu129 OE1 and Gly104 N. The first of these occurs with high frequency in the simulations. The second also occurs, but much less frequently. The second ring forms water bridges only in the T-antigen complex in the crystal structure. The bridges involve GalNAcO7 on the one hand and Ile101 O, Leu212 N and Asn41 ND2 on the other. All these occur in the simulations, but with low frequency and consistency. However, a large number of water bridges not seen in the crystal structures are accessed in the simulations. In fact, all the O atoms in the sugar molecules appear to be involved in water bridges at some time or other. The acetamido N atom in T-antigen also forms water bridges. Water bridges involving the acetamido group with at least a dozen residues are accessed during the simulations. As expected, water-mediated interactions form and break continuously. The maximum and the average residency times of water molecules involved in water bridges that occur reasonably frequently are listed in Table 3. The average residency time varies between 1.3 and 6.9 ps. The maximum duration for which an interaction is mediated by the same water molecule is 152 ps, being that involving GalO5 and Gly213 N in the T-antigen simulations.

In the crystal structures, there are a total of 15 direct protein–sugar hydrogen bonds and water bridges involving 11 amino-acid residues. The number of direct hydrogen bonds accessed with at least 5% occupancy and water bridges with more than 10% occupancy in any of the runs total 140. A majority of these interactions still involve the 11 residues found to interact with sugars in the crystal structures. The remaining interactions are with 13 further residues, most of which are located near to the 11 key residues (referred to above). Clearly, the atomic movements in the carbohydrate-binding site and in the sugar molecules are not large enough to cause any change in the location of the combining site on the lectin.

As mentioned earlier, the two torsion angles in the glycosidic linkage were allowed to vary during MD simulations in both the complexes. In the T-antigen complex, the additional torsion angles which define the conformation of the acetamido group were also allowed to vary. The average values of the torsion angles in the glycosidic linkages are within 10° of those observed in the crystal structures in 12 of the 16 cases. In the runs involving the lactose complexes, the average value of the second angle deviates by more than 20° in 293CLW2. In the simulations involving the T-antigen complexes, the average values of both the angles deviate substantially in 313CTW2 from those in the crystal structure, while that of the second angle deviates by about 12° in 293CTW2. Understandably, larger deviations are exhibited by the torsion angles that define the conformation of the acetamido group. The scatter is also large in these angles. The scatter in the torsion angles in the glycosidic linkage is in general lower, particularly in the runs involving the lactose complex. The scatter is also temperature independent in this complex; in fact, the highest scatter is observed in 293CLW2. On the contrary, the scatter is

higher in the runs at 313 K in the T-antigen complex. It is particularly pronounced in 313CTW2.

The rigid-body movements of the average positions of the two rings and the substituents in the carbohydrate molecules in the runs from their positions in the crystal structures were also examined. Again, the movements are larger and temperature dependent in T-antigen. Except in 293CLW2, those of the lactose rings are comparatively small. This is nearly so in the simulations of the T-antigen complex at 293 K. However, the movements are in general substantially larger in the 313 K simulations.

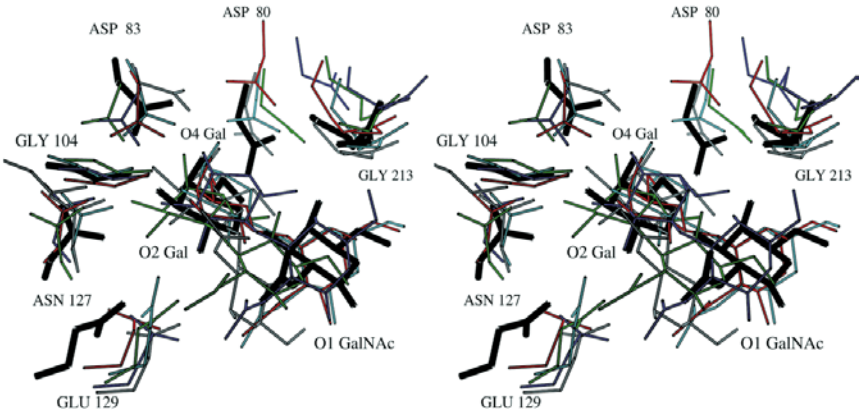

Figure 2
Stereoview of the locations of the representative modes I (red), II (green), III (cyan), IV (magenta) and V (grey) in the T-antigen complex. The location in the crystal structure is shown in black. See text and Table 4 for details.

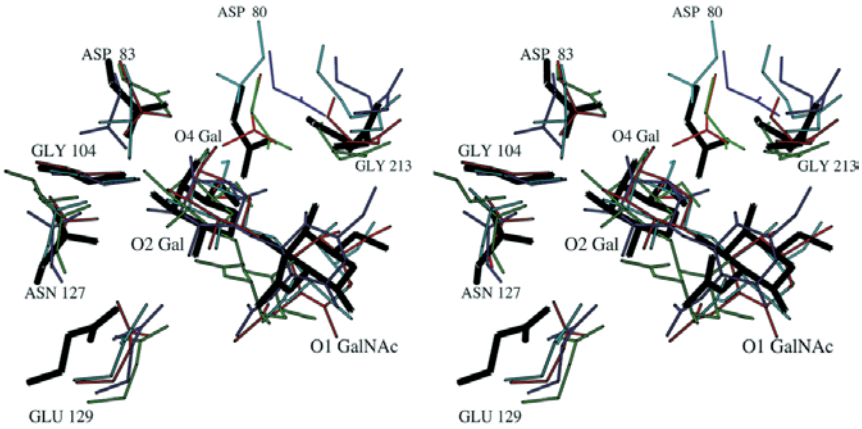

Figure 3
Stereoview of the average positions of the sugar molecule and the surrounding residues in 293CTW1 (red), 293CTW2 (cyan), 313CTW1 (magenta) and 313CTW2 (green). The location in the crystal structure is shown in black.

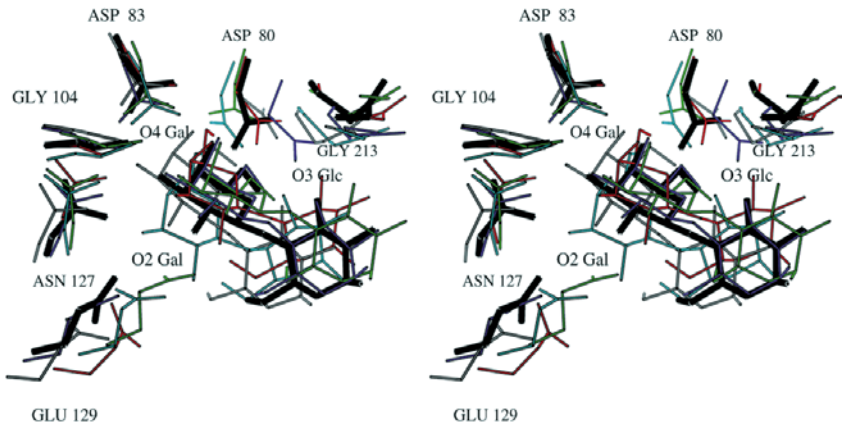

Figure 4
Stereoview of the location of the representative modes I (red), II (green), III (cyan), IV (magenta) and V (grey) in the lactose complex. The location in the crystal structure is shown in black. See text and Table 5 for details.

3.4. Modes of lectin–carbohydrate interactions

On the basis of the simultaneous presence of direct lectin–carbohydrate interactions, 794 unique modes are found in the 4000 snapshots in the four MD runs of the peanut lectin–T-antigen complex. The corresponding number in the lactose complex is 540. None of the unique modes are observed for more than 10% of the time, although there are close relationships among some of the modes. Roughly 25% of the modes are observed only for less than five times. 33 modes occur 0.5% of the time each or more in the T-antigen complex and they account for 52% of the total time of simulations (4 ns). The number of such modes is also 33 in the lactose complex, but they account for 68% of the total time. The water bridges that occur at least 50% of the time when a given mode of interaction exists were identified with the mode. Defined in this manner, each mode consists of a number of direct lectin–carbohydrate interactions along with a certain number of water bridges. The number of direct interactions in the 33 modes vary from three to seven in the T-antigen complex, while the corresponding numbers are three and eight in the lactose complex. The water bridges in the two complexes vary between one and 19, and between three and 12, respectively.

Five typical modes of interactions in the T-antigen complex are listed in Table 4. The positions of the sugar molecule and a few relevant residues in the binding site corresponding to these modes are illustrated in Fig. 2, while Fig. 3 shows the corresponding picture in the average structure of the four T-antigen simulations. Table 5 and Fig. 4 are appropriate for the five typical modes in the lactose complex and Fig. 5 corresponds to the average structures in the four lactose simulations. None of the modes correspond exactly to the two observed in the crystal structures,

Table 5
Five representative modes of interactions in simulations involving the lactose complex.

Mode No.	% of time observed	Direct interactions	Water-mediated interactions
I	7.45	GalO3···Asp83 OD*x*, GalO3···Gly104 N, GalO4···Asp83 OD*x*, GalO4···Ser211 OG, GalO6···Asp80 OD*x*, GlcO3···Gly213 N	GalO2···Glu129 OE*x*, GalO2···Gly104 N, GalO2···Asn127 ND2, GalO3···Glu129 OE*x*
II	4.22	GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*, GalO4···Ser211 OG, GalO5···Ser211 OG, GalO6···Asp80 OD*x*, GlcO3···Gly213 N	GalO2···Glu129 OE*x*, GalO2···Gly104 N, GalO5···Asn127 ND2, GlcO6···Tyr125 OH
III	0.57	GalO2···Gly104 N, GalO3···Asp83 OD*x*, GalO4···Asp83 OD*x*, GalO4···Ser211 OG, GalO5···Ser211 OG, GalO6···Asp80 OD*x*, GlcO3···Gly213 N	GalO2···Glu129 OE*x*, GlcO6···Glu129 OE*x*, GlcO6···Asn127 ND2
IV	0.60	GalO2···Glu129 OE*x*, GalO2···Tyr130 OH, GalO3···Asp83 OD*x*, GalO3···Gly104 N, GalO4···Asp83 OD*x*, GalO4···Ser211 OG	GalO2···Ile101 O, GalO2···Gly102 O, GalO5···Gly213 N, GalO6···Asp80 OD*x*, GlcO3···Ser211 OG, GlcO3···Gly213 N, GlcO4···Ser211 OG, GlcO4···Gly213 N
V	3.90	GalO2···Glu129 OE*x*, GalO3···Asn127 ND2, GalO3···Glu129 OE*x*, GalO4···Asp83 OD*x*,	GalO3···Gly104 N, GalO5···Ser211 OG, GalO5···Leu212 N, GalO5···Gly213 N, GlcO3···Ser211 OG, GlcO4···Ser211 OG, GlcO4···Gly213 N

although some are close to them. In general, the direct interactions and water bridges tend to be complementary in the sense that when the number of direct interactions is small the number of water bridges tend to be large and *vice versa*.

In the representative modes of interactions of T-antigen, mode I is fairly close to that observed in the crystal structure. The two water bridges involving the galactose moiety observed in the crystal structure occur in the mode, but not those involving GalNAcO7. Interactions involving Gly104 N are altogether absent and Asn127 ND2 interacts with O2 instead of O3 in mode II. Direct interactions involving GalNAc are also absent. However, a large number of water bridges, including one each involving GalO2 and GalNAcO7, exist. Among the four invariant direct interactions observed in galactose-binding legume lectins, only the two involving Asp83 OD*x* are found in mode III. Asn127 ND2, however, is involved in a water bridge, while Gly104 N retains the water bridge observed in the crystal structure. Residues 211 and 213 form several water bridges with GalNAcO3 and O4 instead of the direct interactions in the crystal structure. Remarkably, Glu129 OE1 and OE2 form as many as four water bridges with GalO2 and O3. In mode IV, the side chain of Glu129 directly interacts with GalO2 and GalNAcN2. Again, only the two invariant interactions involving Asp83 are observed in this mode. Also, there is surprisingly only one water bridge in the mode. Among the representative modes, mode V contains the least number of direct interactions. This is, however, substantially compensated by a large number of water bridges. Asp80 OD*x*, Ser211 OG, Leu212 N and Gly213 N, which interact directly with the sugar in the crystals, are now involved in water bridges in one way or the other. There are several other water bridges as well.

In the case of lactose, all the direct interactions in modes I and II are those observed in the crystal structure. The water-mediated interactions in the two modes include those found in the crystal structure. Both the modes do not contain a direct interaction involving Asn127 ND2 which is, however, involved in a water bridge in both the cases. Mode III is also similar to that in the crystal structure. Gly104 N interacts in this mode directly with GalO2 instead of interacting through a water bridge as in the crystal structure. The other water bridge observed in the crystal structure occurs in this mode. Two new direct hydrogen bonds involving O2, one with Glu129 OE*x* and the other with Tyr130 OH, are observed in mode IV. Furthermore, the water bridges in this mode do not contain the two observed in the crystal structure. Mode V represents a drastic departure from that observed in the crystal structure. This is the only representative mode in which a direct Asp83 OD*x*···GalO3 interaction does not exist. Modes related to this one are accessed particularly in the MD run 293CLW2. The modes of interactions in the lactose complex exhibit less variety than those in the T-antigen complex. They also, in general, contain a lower number of water bridges.

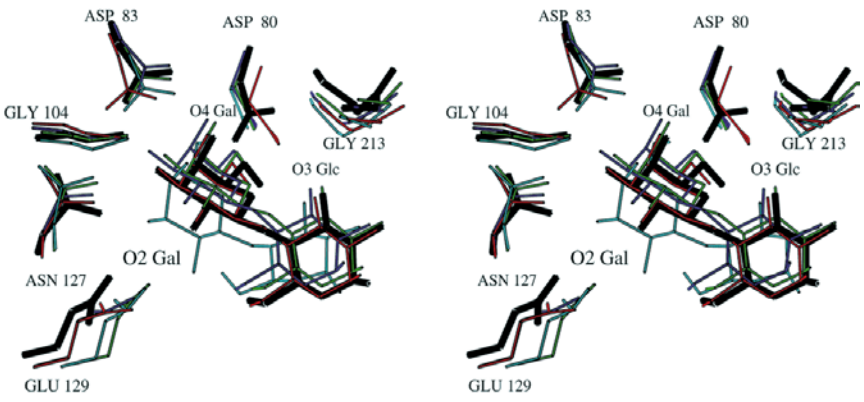

Figure 5
Stereoview of the average positions of the sugar molecule and the surrounding residues in 293CLW1 (magenta), 293CLW2 (cyan), 313CLW1 (red) and 313CLW2 (green). The location in the crystal structure is shown in black.

3.5. Energetics

The enthalpies averaged over each simulation are listed in Table 6. The values of $\Delta\Delta H$ favour the T-antigen complex, in conformity with the experimental values (Table 1), but the standard errors are so

Table 6
Enthalpies (in kJ mol^{-1}) calculated from MD simulations.

System	Sugar–water	Water–internal	Sugar–internal	Sugar–protein	Protein–water	Total
293CW1						
CT	−364 ± 21	—	−180 ± 17	−431 ± 33	−2125 ± 155	−3100 ± 159
CL	−356 ± 17	—	−21 ± 13	−389 ± 21	−2029 ± 138	−2791 ± 29
FT	−640 ± 21	−15627 ± 25	−230 ± 17	—	—	−16498 ± 21
FL	−644 ± 21	−15803 ± 21	−13 ± 13	—	—	−16456 ± 21
$\Delta\Delta H$	−13 ± 38	−176 ± 33	59 ± 29	−42 ± 38	−100 ± 205	—
$\Delta\Delta H_{\rm bind}$ (TANT − LACT)	−272 ± 201					
293CW2						
CT	−305 ± 42	—	−180 ± 13	−448 ± 59	−2172 ± 121	−3105 ± 130
CL	−418 ± 42	—	46 ± 25	−351 ± 17	−1835 ± 96	−2556 ± 100
FT	−640 ± 21	−15627 ± 25	−230 ± 17	—	—	−16498 ± 21
FL	−644 ± 21	−15803 ± 21	−13 ± 13	—	—	−16456 ± 21
$\Delta\Delta H$	109 ± 63	−176 ± 33	−8 ± 38	−92 ± 59	−339 ± 155	—
$\Delta\Delta H_{\rm bind}$ (TANT − LACT)	−506 ± 163					
313CW1						
CT	−364 ± 42	—	−167 ± 16.8	−377 ± 63	−2075 ± 100	−2987 ± 100
CL	−339 ± 13	—	−13 ± 8	−364 ± 25	−185 ± 71	−2548 ± 75
FT	−640 ± 21	−15627 ± 25	−230 ± 17	—	—	−16498 ± 21
FL	−644 ± 21	−15803 ± 21	−13 ± 13	—	—	−16456 ± 21
$\Delta\Delta H$	−29 ± 54	−176 ± 33	42 ±29	−13 ± 67	−222 ± 126	—
$\Delta\Delta H_{\rm bind}$ (TANT − LACT)	−397 ± 126					
313CW2						
CT	−448 ± 46	—	−280 ± 21	−527 ± 42	−1845 ± 80	−3096 ± 96
CL	−347 ± 33	—	38 ± 25	−393 ± 46	−1996 ± 142	−2699 ± 151
FT	−640 ± 21	−15627 ± 25	−230 ± 17	—	—	−16498 ± 21
FL	−644 ± 21	−15803 ± 21	−13 ± 13	—	—	−16456 ± 21
$\Delta\Delta H$	−105 ± 63	−176 ± 33	−96 ± 42	−134 ± 63	151 ± 163	—
$\Delta\Delta H_{\rm bind}$ (TANT − LACT)	−356 ± 180					

large as to render the actual values somewhat meaningless. The same is true of the energetics involved in protein–sugar interactions. However, the fact that energies calculated from MD simulations led to a difference in enthalpy in favour of the correct sugar is in itself satisfying. This progress beyond earlier work is substantially because of a new approach of limiting the sphere size for the area involved in the energetics analysis. However, an examination of the accessed protein–sugar interactions is more useful.

The number of direct interactions in the T-antigen complex at 313 K is substantially higher than at 293 K. This difference is not observed in the lactose complex. The number of direct protein–sugar interactions in 293CTW1, 293CTW2, 313CTW1 and 313CTW2 are 19, 18, 42 and 29, respectively. The corresponding values in the lactose complexes are 19, 22, 22 and 21, respectively. The difference is more pronounced in the case of the water-mediated interactions. The number of water mediated interactions in 293CTW1, 293CTW2, 313CTW1 and 313CTW2 are 65, 85, 138 and 119, respectively. The corresponding numbers in the case of the four lactose simulations are 69, 59, 73 and 82, respectively. These observations are in consonance with the variation of the enthalpy terms as a function of temperature in calorimetric measurements (Table 1).

A comparison of the relevant crystal structures led to the conclusion that the increased affinity of peanut lectin for T-antigen compared with that for lactose was primarily a consequence of the facility in forming useful water bridges involving the acetamido group. The results of the MD simulations corroborate this conclusion. The acetamido group is only very rarely seen as interacting directly with the protein. However, both the O and the N atoms in the group are extensively involved in water bridges. In the crystal structure, only two such water bridges are seen. Understandably, many more are seen in the simulations. However, interestingly, GalNAcO7 is involved in one water bridge or the other 67% of the time. The percentage in the case of GalNAcN2 is 33%. In fact, O7 or N2 or both are involved in one or more bridges 80% of the time. Also, among the 13 water-mediated interactions that are seen for more than 10% of the total time (4 ns), nine involve either of these two atoms. Thus, as suggested earlier (Ravishankar *et al.*, 1997, 1999), PNA–T-antigen interactions illustrate how carbohydrate specificity can be generated by water bridges.

As in the case of the binding of α-galactose and *N*-acetyl lactosamine to the *E. corallodendron* lectin (Bradbrook *et al.*, 2000), the work reported here also appears to present an example for the breaking of 'enthalpy–entropy compensation'. Such compensation, which occurs through the correlation of ΔH and ΔS leading to almost mutual cancellation, thus resulting in only small changes in ΔG, is a well discussed phenomenon (Searle *et al.*, 1995; Cooper, 1999). Thermodynamic measurements and energy calculations clearly show that the ΔH associated with the interaction with PNA is

higher for T-antigen than for lactose. However, the difference in ΔH is not altogether compensated by that in $T\Delta S$. The MD simulations presented here, suggest that the entropy cost of the tighter binding of T-antigen is reduced by its ability to access a wide range of binding motifs. Lactose does not appear to have the same degree of access to alternative binding modes. Thus, the 'breaking' of enthalpy–entropy compensation is a way of understanding why one sugar binds more tightly than another.

3.6. Results of mutational studies

Results of studies involving mutations at two positions are available. In the first case, Leu212 has been mutated to asparagine and alanine (Sharma *et al.*, 1996), while in the second study Glu129 has been mutated to aspartic acid and alanine (Sharma *et al.*, 1998).

Binding affinity of both T-antigen and lactose decreases substantially when Glu129 is changed to Asp. The affinity reduces much more substantially when the residue at this position is Ala. In the crystal structures of both the complexes, the side chain of Glu129 is involved in a water bridge with GalO2. This water bridge could have been retained even if the residue was Asp. Replacement by Ala would of course abolish the water bridge, but the observed reduction in affinity appeared to be more than that which could be expected by the disruption of a single water bridge. The MD simulations, however, provide a rationale for the observed variation in the affinity of the wild-type protein and the mutants for the sugars. In both sets of simulations, Glu129 is involved in more water bridges, however transiently, than observed in the crystal structure. More importantly, the side chain of the residue directly interacts with the sugar for a substantial part of the time. A change in the length of the side chain, as in the case of E129D, would lead to a reduction in the interactions; a change to Ala would altogether abolish them.

The effect of the mutation of Leu212 is more complex. Affinity for lactose is unaffected in L212A, but increases in L212N. On the other hand, affinity for T-antigen is unaffected in L212N, but decreases very substantially in L212A. The crystal structures of the complexes involving the wild-type protein do not provide a satisfactory explanation for these results. Nor do the MD simulations. In fact, a nanosecond dynamic simulation was performed at 313 K with the first hydration protocol with leucine at residue 212 replaced by an alanine. However, this did not result in any significant change in either direct protein–sugar interactions or water-mediated interactions (data not shown). Perhaps the crystal structures of the mutants or longer MD simulations may lead to an explanation.

4. Concluding remarks

Although the standard deviations are high, all the runs lead to $\Delta\Delta H$ values favourable to T-antigen in conformity with the results of the thermodynamic measurements. More importantly, what the simulations do and what static calculations by their very nature cannot do is to access possible additional modes of interactions in the form of an ensemble. They lead to plausible explanations for one of the two mutational studies. The simulations also appear to provide a more realistic picture of water-mediated interactions than do the crystal structures by themselves, in that they continually form and break in the former. The water structure around proteins, including that involving water molecules which directly interact with the protein, is known to be extremely dynamic (Levitt & Park, 1993), an aspect which is brought out with insufficient clarity by crystal structures alone. The MD simulations clearly demonstrate that residency times of individual water molecules are very short even when they are involved in bridging the lectin and the sugar. Significantly, the average number of direct protein–ligand interactions at any given point of time in the simulation is not much different from that observed in the crystal structures. However, as one would intuitively expect when dealing with highly mobile water molecules unconstrained by packed protein molecules, the average number of water bridges is much larger in the MD simulations. At the same time, the results of the simulations corroborate the conclusion that the higher affinity of peanut lectin for T-antigen than that for lactose is primarily because of water bridges involving the acetamido group. The total entropic compensation of the enthalpic advantage in PNA–T-antigen interactions in relation to those involving lactose is prevented by the accessibility of T-antigen to a wider range of binding motifs. The understanding of this 'breaking of enthalpy-entropy compensation' may be useful during the process of structure-based rational drug design, whereby one would design ligands which have a number of possible binding motifs with their target receptor. Perhaps this could be achieved by the addition of functional groups to a lead compound on the basis of potential binding interactions suggested by MD simulations.

A very satisfying, though somewhat unexpected, result of the MD simulations has been their behaviour as a function of temperature. The basic average features of the simulations remain nearly the same at the two temperatures used in the case of the lactose complex. In the case of the T-antigen complex, they show a temperature dependence along the lines indicated by the variation of the thermodynamic parameters as a function of temperature. MD simulations also provide a satisfactory explanation for the sugar-binding properties of one of the available mutants.

Despite the positive results that emanated from the simulations, the problems related to the reliability of the force fields and the duration of MD calculations remain. This is clearly brought out by the results of the simulation involving 293CLW2, which have often been distinctly, perhaps disturbingly, different from those emanating from the other three simulations of the lactose complex. Thus, small differences in the initial conditions could lead to large differences in the outcome, at least in a simulation of short duration. Thus, if a long simulation is not possible owing to computational

problems, it is important to carry out several short simulations of the system with somewhat different initial conditions. Even if it is, it might be preferable to use the available computation time for several simulations from different initial conditions.

To sum up, in addition to providing fresh insights into peanut lectin–sugar interactions and also a rationale for some of the experimental observations, the present work confirms the utility of the MD simulations (Bradbrook *et al.*, 1998, 2000) in exploring protein–ligand interactions, while at the same time indicating the need for exercising caution when interpreting the results of the simulation. It also demonstrates the usefulness of carrying out several short simulations with different starting points.

We thank the Indian National Science Academy and The Royal Society for funding a study visit by GMB to IISc, Bangalore. JRH is very grateful to BBSRC and The Wellcome Trust who provided salary support to GMB on return to Manchester. JRH is also very grateful to the BBSRC who provided funding for computer workstations and molecular graphics as well as partial salary support to JR. MV and AS thank the Department of Science and Technology, India for support. We thank Madhu and Malar for their help during the course of this work. Most of the computations were carried out at the Supercomputer Education and Research Centre at the Indian Institute of Science.

References

Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). *J. Mol. Biol.* **259**, 281–296.
Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanraj, V., Mahanta, S. K., Suguna, K., Surolia, A. & Vijayan, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.
Bouckaert, J., Hamelryck, T., Wyns, L. & Loris, R. (1999). *Curr. Opin. Struct. Biol.* **9**, 572–577.
Bradbrook, G. M., Forshaw, J. R. & Perez, S. (2000). *Eur. J. Biochem.* **267**, 4545–4555.
Bradbrook, G. M., Gleichmann, T., Harrop, S. J., Habash, J., Raftery, J., Kalb (Gilboa), J., Yariv, J., Hillier, I. H. & Helliwell, J. R. (1998). *J. Chem. Soc. Faraday Trans.* **94**, 1603–1611.
Chevernak, M. C. & Toone, E. J. (1995). *Biochemistry*, **34**, 5685–5695.
Cooper, A. (1999). *Curr. Opin. Chem. Biol.* **3**, 557–563.
Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M. Jr, Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). *J. Am. Chem. Soc.* **117**, 5179–5197.
Drickamer, K. (1997). *Structure*, **5**, 465–468.
Drickamer, K. (1999). *Curr. Opin. Struct. Biol.* **9**, 585–590.
Elgavish, S. & Shaanan, B. (1998). *J. Mol. Biol.* **277**, 917–932.
Frisch, M. J. *et al.* (1995). *GAUSSIAN*94. Carnegie, PA, USA: Gaussian Inc.
Gunsteren, W. F. van & Berendsen, H. J. C. (1990). *Angew. Chem. Int. Ed. Engl.* **29**, 992.
Gupta, D., Dam, T. K., Ocarson, S. & Brewer, C. F. (1997). *J. Biol. Chem.* **272**, 6388–6392.
Jorgensen, W. L. (1981). *J. Am. Chem. Soc.* **103**, 335–340.
Lemieux, R. U. (1996). *Acc. Chem. Res.* **29**, 373–380.
Levitt, M. & Park, B. H. (1993). *Structure*, **1**, 223–226.
Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). *Biochim. Biophys. Acta*, **1383**, 9–36.
Mandal, D. K., Kishore, N. & Brewer, C. F. (1994). *Biochemistry*, **33**, 1149–1156.
Manoj, N., Srinivas, V. R., Surolia, A., Vijayan, M. & Suguna, K. (2000). *J. Mol. Biol.* **302**, 1129–1137.
Neurohr, K. J., Bundle, D. R., Young, M. & Mantsch, H. H. (1982). *Eur. J. Biochem.* **123**, 305–310.
Pereira, M. E. A., Kabat, E. A., Lotan, R. & Sharon, N. (1976). *Carbohydr. Res.* **51**, 107–118.
Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. & Vijayan, M. (1997). *Curr. Sci.* **72**, 855–861.
Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. & Vijayan, M. (1999). *Curr. Sci.* **76**, 1393.
Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (2000). *Acta Cryst.* D**55**, 1375–1382.
Ravishankar, R., Surolia, A., Vijayan, M., Lim, S. & Kishi, Y. (1998). *J. Am. Chem. Soc.* **120**, 11297–11303.
Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–557.
Rini, J. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 578–584.
Schwarz, F. P., Puri, K. D. & Surolia, A. (1993). *J. Biol. Chem.* **266**, 24344–24350.
Searle, M. S., Williams, D. H. & Packman, L. C. (1995). *Nature Struct. Biol.* **2**, 999–1006.
Sharma, V., Srinivas, V. R., Adhikari, P., Vijayan, M. & Surolia, A. (1998). *Glycobiology*, **8**, 1007–1012.
Sharma, V., Vijayan, M. & Surolia, A. (1996). *J. Biol. Chem.* **271**, 21209–21213.
Sharon, N. & Lis, H. (1989). *Science*, **246**, 227–234.
Surolia, A., Sharon, N. & Schwarz, F. P. (1996). *J. Biol. Chem.* **271**, 17697–17703.
Toone, E. J. (1994). *Curr. Opin. Struct. Biol.* **4**, 719–728.
Vijayan, M. & Chandra, N. R. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Weis, W. I., Taylor, M. E. & Drickamer, K. (1998). *Immunol. Rev.* **163**, 19–34.
Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). *Anal. Biochem.* **179**, 131–137.
Woods, R. J., Dwek, R. A., Edge, C. J. & Fraser-Reid, B. (1995). *J. Phys. Chem.* **99**, 3832–3846.
Yang, C. P. (1990). *Omega Data in Origin.* Northampton, MA, USA: Microcal Inc.