Contribution of C-H…π Interactions to the Affinity and Specificity of Carbohydrate Binding Sites

Petety V. Balaji*

Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

Abstract: X-ray crystallographic studies have shown that C-H groups of saccharides interact with aromatic amino acid residues in binding sites in proteins. Such C-H…pi interactions have been shown to be dominated by London's dispersion interactions. The strength of interaction depends on a number of factors: (i) the nature of aromatic residue and saccharide, (ii) the form (acyclic, pyranose or furanose) and conformation of the saccharide, (iii) the functional groups present in modified saccharides (e.g., N-acetyl group), and (iv) the mutual position-orientation of the interacting moieties; this, in turn, determines the number of interacting C-H groups and the geometry of interaction. The strength of interaction also depends on the surrounding medium and on the microenvironment in a protein's binding sites. A variety of experimental techniques such as isothermal titration calorimetry, turbidity measurements, NMR spectroscopy, infrared ion depletion spectroscopy and fluorescence spectroscopy have been used to investigate C-H…pi interactions. Quantum chemical calculations of saccharide - aromatic systems have shown that their interaction is stabilizing. The interaction energy ranges between 1 and 12 kcal/mol and is thus comparable in strength to a conventional hydrogen bond. However, experiments that mimic double mutant cycles are needed to be designed, and performed, to determine the contribution of C-H…pi interactions to the affinity of glycans to proteins and the glycan specificity of proteins.

Keywords: Pyranose, furanose, dispersion interaction, apolar patch, stacking.

Non-Covalent Interactions

All non-covalent interactions are essentially electrostatic in nature. They are categorized into different types such as dipolar, hydrogen bonding and van der Waals for convenience of study. Several experimental and computational methods have been developed, and used extensively, to probe these non-covalent interactions. Knowledge gained from such studies is useful to understand enthalpy changes and, in turn, free energy changes associated with biological processes.

Carbohydrate - Protein Interactions

All the chemical entities, viz., molecules and trace elements, that constitute a biological system are essential for that system. Abundance, absence, shortage or modification of even one of the chemical entities may lead to abnormalities in the system. However, some molecules are deemed more important than others merely because of the range of biological processes they participate in. Carbohydrates and proteins are two such biological molecules. Non-covalent interactions between these molecules include both the conventional and non-conventional interactions. The former are mediated primarily by the hydroxyl groups of saccharides and the latter, by the C-H groups.

Aromatic Residues in Saccharide-Binding Sites

X-ray crystallographic studies have shown that one or more of the aromatic residues tryptophan, tyrosine, phenylalanine and histidine are present in the saccharide-binding site(s) of several proteins [1]; protein engineering studies have shown that these residues play a role in determining the affinity of the glycan to protein and the saccharide-specificity of the protein (Table **1**). In proteins such as maltoporin, the aromatic residues form a layer on which the maltotriose slides through [2].

Participation of C-H Groups and Aromatic Residues in Hydrogen Bonding Interactions

Aromatic amino acids and C-H of methylene and methyl groups are often treated as apolar and thought to participate only in van der Waals interactions. However, it is known since the early 1950s that aromatic compounds can form hydrogen bonds with such molecules as chloroform and methanol-*d* [3,4]. Participation of aromatic residues and of C-H groups in hydrogen bonding interactions is now reported in many processes including folding and recognition of biomolecules [5-9].

C-H…π Interactions

Experimental and theoretical studies have shown that the nature of the C-H...π interaction is different when the C-H group is from an alkane / alkene as compared to when it is from an alkyne [10-15]. The former is of van der Waals type with a small contribution from electrostatic interaction whereas the latter is a hydrogen bond type with a large contribution from electrostatic interaction. The latter are also referred to as 'activated' C-H...π interaction and these are characterized by an upfield (blue) shift of the stretching vibrational frequency of the hydrogen atom [16]. Ab initio calculations of fucose-benzene complex showed that electrostatic contribution to the C-H...π interaction is relatively small because of which the orientational dependence is rather weak [17].

Interactions of Saccharides with Proteins

A monosaccharide can interact with an aromatic residue through (i) O-H...π, O-H...O, O...H-N or C-H...O hydrogen bonds and/or (ii) C-H...π interactions (the atom/group before the "..." symbol is from the monosaccharide and after the symbol are from the aromatic residue). The latter are generally referred to as stacking interactions [18]. X-ray crystallographic studies have shown that only in some of the carbohydrate-binding proteins the aromatic residue interacts with the saccharide through C-H...π stacking. In others, the orientation of the aromatic ring with respect to the pyranose ring is such that stacking interaction is not possible; they interact through conventional hydrogen bonding interactions e.g., Trp62 in the complex of *Escherichia coli* maltodextrin / maltose-binding protein with maltose (PDB id 1ANF) [19].

Stacking Interactions

The interaction between an aromatic residue and a saccharide is often described as stacking interaction. 'Stacking' as such does not denote any non-covalent interaction; it merely indicates that the two interacting moieties are 'piled' one on top of another. A recent compilation of data from experimental and computational studies on the energetics of stacking interactions between various types of molecules shows that the nature and extent of contribution of the

*Address correspondence to this author at the Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India; Tel: +91-22-2576 7778; Fax: +91-22-2572 3480; E-mail: balaji@iitb.ac.in

1570-193X/11 $58.00+.00

Table 1. Some of the Experimental Studies Which Demonstrated the Contribution of Aromatic Residues to Affinity and Specificity of Carbohydrate Binding Sites

- *Saccharomycopsis fibuligera* α-amylase [67]
  The Trp, Leu or Asn mutants at Y83 show higher transglycosylation activity and reduced hydrolysis activity
- *Dolichos biflorus* lectin DBL [68]
  L127F mutation increases affinity for galactose
- *Maackia amurensis* isoectins MAL and MAH [69]
  Y45, Y131 and Y136+Y221 interact with NeuAc, Gal and Glc, respectively, of sialyllactose. Y221 in MAL is replaced by Ala in MAH and this residue is responsible for the difference in the fine specificity of the isolectins
- Family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A [70]
  An aromatic residue at position 72 and Trp at 54 are required for the effective binding of crystalline cellulose to the carbohydrate-binding module
- Human lysozyme [71]
  Comparison of wild-type and Tyr63Leu mutant showed that the stacking interaction with Tyr63 is essential for positioning the substrate in productive binding mode
- Pokeweed (*Phytolacca americana*) lectin PL-D2 [72]
  Methyl of the N-acetyl group stacks against His72 (A chain)
- *Escherichia coli* β-galactosidase [73]
  Site-specific substitution of Trp999 by Phe, Tyr, Gly or Leu led to greatly decreased affinities for substrates / inhibitors that bind in the 'shallow' mode. These mutants had less affinity for glucose also (as a substrate for the transgalactoside reaction)
- Lewis α3- and α3/4-fucosyltransferases [74]
  Trp in the acceptor binding motif: α3/4-activity; Arg instead of Trp, only α3-activity
- *Escherichia coli* maltooligosaccharide-specific LamB-channel (maltoporin) [2]
  Mutation of one or more of Y6, Y41, W74, F229, W358 and W420 led to a substantial decrease of the stability constant of maltopentaose binding.
- Jacalin [75]
  Tyr78 stacks against mannose
- Human salivary α-amylase [76]
  Aromatic residues in the seconary saccharide-binding site play an important role in bacterial binding and starch hydrolytic activity
- *Coprinopsis cinerea* CGL3 [77]
  R81W mutant binds to lactose, instead of chitooligosaccharides

different types of non-covalent interactions to stacking varies depending on the interacting molecular species [20]. A systematic study using 'dangling' end bases showed that base stacking interactions in nucleic acids is a combination of dispersion, dipolar and hydrophobic interactions [21]. Such a systematic dissection of the stacking interaction between an aromatic amino acid and a monosaccharide has not been carried out so far. Dispersion interaction will certainly be an important contributor [22], since the stacked saccharide and aromatic residue will be in close proximity of each other. However, the contribution from dipolar interaction may depend on the nature of the aromatic amino acid: benzene (cf. phenylalanine) does not have a dipole whereas indole (cf. tryptophan) has [23]. Recent high-level ab initio calculations [$E_{CCSD(T)(limit)}$] of fucose-benzene complex indicated that the dispersion interaction has a major contribution to the interaction energy; the contribution from electrostatic interactions is relatively small [17].

Hydrophobic effect also will contribute to stacking interaction between an aromatic residue and a saccharide. The extent of contribution will be context dependent: if the aromatic residue is an independent molecule in an aqueous solution, then the contribution of the hydrophobic effect can be expected to be substantial (cf. dangling end experiments on self-complementary oligonucleotides) [21]. If the aromatic residue is part of the binding site of a protein, then the contribution from the hydrophobic effect will depend on the extent of solvent accessibility of the aromatic residue. As for the saccharide, when this 'polar and soluble' saccharide stacks against an aromatic residue, the solvent accessibility of the C-H groups will be the determining factor.

Electrostatic Potential of Aromatic Residues

The electrostatic potentials of indole, phenol and benzene [24] are illustrative of those of tryptophan, tyrosine and phenylalanine. In these molecules, the electrostatic potential is symmetric with respect to the plane of the aromatic ring and thus the 'stacking' saccharide can approach the aromatic ring from either of the two sides of the ring. The electrostatic potential surfaces of benzene, phenol and the 6-membered ring of indole are quite similar to each other. The presence of nitrogen in the 5-membered ring of indole makes the potential near this atom relatively positive but overall, the available surface area for saccharide to interact through C-H...π interactions is more in indole compared to that in either benzene or phenol.

Data from experimental studies on the interaction of wild type and mutant AcAMP2 with chitooligosaccharides illustrate the influence of (i) the surface area available for stacking and (ii) the electrostatic potential on the strength of C-H...π interactions [25]. AcAMP2 is a 30-residue peptide which resembles hevein domains and these domains are small (30-45 residues), plant proteins that bind to chitin. It has been found that the residues Phe18, Tyr20 and Tyr27 of AcAMP2 interact with the bound saccharide through van der Waals interactions and C-H...π stacking. The binding of chitotriose to wild-type AcAMP2 and the mutants Phe18Trp, Phe18 Naphthylalanine and Phe18Pff+Tyr20Pff (Pff=4-fluorophenylalanine) was investigated by NMR spectroscopy and thermodynamic analyses. It was found that larger the aromatic group, the higher the association constant and the binding enthalpy. The binding affinity was found to decrease in the protein which has fluorinated aromatic residue. This is not surprising considering the effect of the electron-withdrawing fluorine substituents on the electrostatic potential of the aromatic ring [24].

Apolar Patch in the Pyranose Ring form of Saccharides

Most of the studies on C-H...π interactions in protein-carbohydrate complexes pertain to the pyranose ring form of the saccharide. This ring form can be thought to have an a-face and a b-face. Viewing the pyranose ring (i) along an imaginary axis that is parallel to the axially oriented groups and (ii) in such a manner that the ring atom numbering is clockwise, the top part (or the near side) of the pyranose ring is considered to be the a-face and the bottom part (or the far side) to be the b-face (Fig. **1**) [18]. The nature of

atoms / functional groups that constitute the a- and b-faces depends on (i) the configuration of the ring carbon atoms and (ii) the conformation of the pyranose ring.

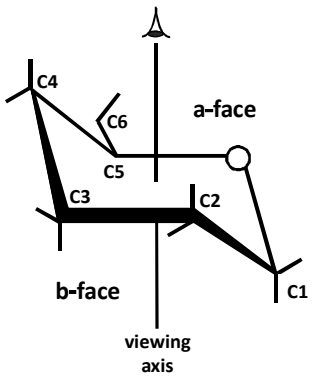

Fig. (1). Schematic of an aldohexose showing the a-face and b-face of the pyranose ring form.

The b-face of β-D-galactopyranose in the $^4C_1$ conformation has an apolar patch formed by the C-H groups of the pyranose ring (Fig. **2a**). This patch may be further extended by the C-H groups of the exocyclic hydroxymethyl group; the number of C-H groups participating in the apolar patch depends on the conformation of the -CH_2OH group. In comparison, the apolar patch in the b-face of α-D-glucopyranose is less pronounced, even though the conformations of the pyranose ring and -CH_2OH group are same (Fig. **2b**). This is obviously due to the changes in the configurations of C1 and C4 atoms.

The conformation of the exocyclic group becomes irrelevant vis-à-vis the apolar patch in 6-deoxysugars such as fucose (6-deoxygalactose). In Nature, the D-enantiomer of galactose is prevalent whereas for fucose, it is the L-enantiomer. The α-anomer of L-fucopyranose prefers the $^1C_4$ conformation and in this, the apolar patch is in the a-face (Fig. **2c**). This apolar patch is similar to that in the b-face of β-D-galactopyranose (Fig. **2a**) except for the absence of the C1-H, due to the α-anomeric configuration of fucose.

As mentioned above, the b-face of β-D-galactopyranose has an apolar patch. The -OH groups at C1, C2 and C3 are equatorially oriented (Fig. **2a**). Hence, the a-face has only C2-H and C4-OH. In contrast, D-glucopyranose has apolar patches on both the a-face (Fig. **2d**) and b-face (Fig. **2b**). The number of C-H groups constituting the apolar patch, however, depends on the configuration of the anomeric carbon atom and the conformation of the -CH_2OH group.

The α-anomer of D-mannopyranose is more common in biological systems than the β-anomer. In this saccharide, the C3-H and C5-H atoms on the b-face seem to make an apolar patch, especially if the -CH_2OH group is in the *gg* conformation (Fig. **2e**). However, the axially oriented C1-OH group also becomes part of the same b-face. In the a-face, C1-H, C4-H and C6-H (only one of the two) are the only C-H groups (Fig. **2f**).

Effect of Changes in Configuration on Stacking Propensity

It can thus be seen that the location and extent of the apolar patch is very much related to the nature of the saccharide (cf. glucose, galactose, mannose), anomeric configuration (cf. α- and β-glucopyranose) and the conformation of the -CH_2OH group. This apolar patch, in turn, determines how the saccharide can interact with the aromatic residue through C-H...π stacking. Thus, an aromatic residue can interact with (i) galactose through the b-face (Fig. **2a**), (ii) glucose through both the a-face or b-face (Figs. **2b** and **2d**), (iii) mannose either through C4-H in the a-face (Fig. **2f**) or through the -CH_2OH group, and (iv) fucose through the a-face (Fig. **2c**). The two experimental studies mentioned below illustrate this relationship between the configuration / conformation of the pyranose ring and its stacking preferences.

The molecular complexes of toluene formed at low temperature with the α- and β-anomers of glucose, galactose and fucose were investigated using infrared ion depletion spectroscopy [22]. The spectrum for the complex of α-MeGlc (methylglucopyranose) was found to be completely different from that of β-MeGlc. α-MeGlc interacts with toluene through O-H...π interactions. On the other hand, β-MeGlc interacts through C-H...π interactions but it was not possible to determine if the a-face or the b-face is involved. However, the interactions were found to be through C-H...π in the case of α-MeFuc (a-face), β-MeFuc (a-face) and α-MeGal (b-face) [22].

The interactions of methoxy and acetyl derivatives of galactose and mannose with benzene were investigated using microcalorimetry and NMR spectroscopy [26]. It was found that methyl 2,3,4,6-tetra-O-methyl-α-D-galactopyranoside interacted with benzene through C-H...π interactions whereas the corresponding α-anomer or the derivative which has acetyl groups instead of methyl groups did not interact at all.

C-H…π Stacking and Specificity

A saccharide interacts with amino acids other than the aromatic residue in the binding site of a protein. These interactions, which are primarily through the -OH groups, determine the placement of the saccharide with respect to the aromatic residue within the binding site. Thus, the bound saccharide has to optimize not only C-H...π stacking but also the interactions of the -OH groups. The 'mode' in which a saccharide interacts with an aromatic residue can be specified by defining the position-orientation (PO) of the saccharide relative to the aromatic residue [27]. The POs of the bound saccharide in the binding site of Glc/Man-specific (Glc-POs) and Gal-specific (Gal-POs) were used to calculate the interaction energies of saccharide - aromatic residue analogue complexes using quantum chemical methods [28]. It was found that in Glc-POs, the interaction energies of galactose and glucose with the aromatic residue analogue are comparable to each other. In contrast, in Gal-POs, the interaction energies of glucose are unfavourable. This implies that the architecture of galactose-binding site is such that the C-H...π stacking is unfavourable for binding glucose. Thus, the differences in the mode of interaction of the aromatic residue due to the location and composition (i.e., the number of C-H groups) of the apolar patch in the saccharide can be related to saccharide specificity.

The aromatic residue seems to be especially important for affinity and specificity in galactose-binding proteins (see, for example, [29-34]. An analysis of the 3-D structures of 18 galactose-specific proteins belonging to seven non-homologous protein families (i.e., proteins which are not evolutionarily related) showed that the presence of an aromatic residue in the binding site is one of the key features shared by these proteins [35].

Less Preferred Conformations of the Pyranose Ring

Monosaccharides such as glucose, galactose and mannose are generally taken to be in the pyranose ring form and in the preferred $^4C_1$ conformation. While this is true, as evidenced by the numerous x-ray crystallographic studies of mono- and di-saccharides, and protein-saccharide complexes [36], other conformations are also possible for the pyranose ring.

Besides the two chair forms, the pyranose ring can exist in any of the various boat, skew-boat and half-chair forms [37,38]. The pyranose ring is in the half-chair conformation in the active sites of lysozyme [39,40] and *Bacillus circulans* D229N-E257Q double mutant CGTase [41]. Transitions between the chair and boat forms have been implicated to take place when amylose is subjected to tensile stress [42]. The uronic acid derivative of idose (i.e., having a

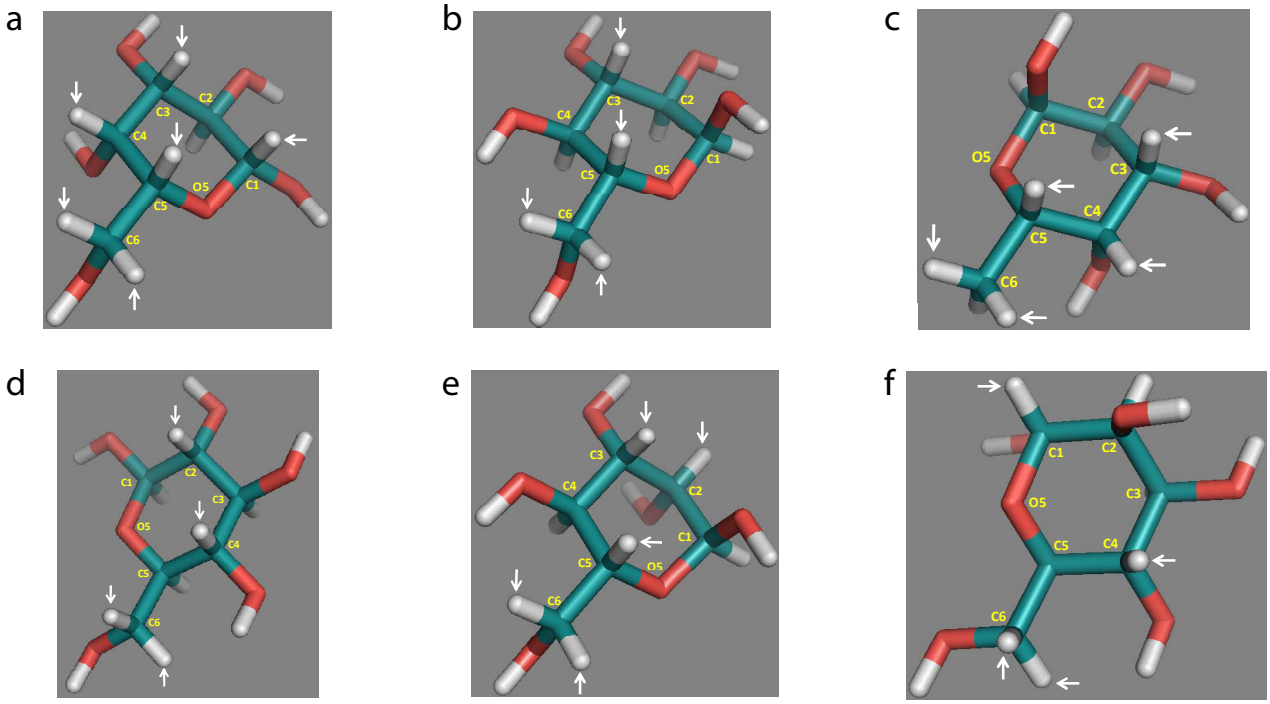

Fig. (2). Stick representations depicting the b-face of β-D-galactopyranose **(a)**, b-face of α-D-glucopyranose **(b)**, a-face of α-L-fucopyranose **(c)**, a-face of β-D-glucopyranose **(d)**, b-face of α-D-mannopyranose **(e)** and a-face of α-D-mannopyranose **(f)**. The pyranose rings are in the preferred conformations: $^1C_4$ for fucose and $^4C_1$ for all others. The conformation of the exocyclic -CH_2OH group is *gg*, *gg*, *gt*, *gg* and *gt* in **(a)**, **(b)**, **(d)**, **(e)** and **(f)**, respectively; in this notation, the letters *g* (gauche) and *t* (trans) specify the conformation of the exocyclic oxygen atom with respect to the ring oxygen and C4 atoms, respectively, of the pyranose ring. The conformations of the -OH groups have been chosen randomly. In this figure and Figure **3**, (i) the molecules were rendered using Open-Source PyMol 0.99rc6 using coordinates that are generated in-house based on standard geometry; (ii) only some of the atoms have been labelled; and (iii) the carbon, oxygen and hydrogen atoms are coloured teal blue, red and white, respectively.

-COOH group instead of -CH_2OH at C5) is a constituent of glycosaminoglycans. Because of its peculiar configuration, iduronic acid exists as a mixture of different chair and skew-boat conformations [43]. This, in turn, influences the conformation of the glycosaminoglycans. The iduronic acid moiety in heparan sulfate fragment bound to antithrombin is in the ^2S_O conformation [44]. An interesting example of the 'co-existence' of both the $^4C_1$ and $^1C_4$ conformations of glucopyranose is found in one of the crystal structures of permethyl β-cyclodextrin [45]. The apolar patch consisting of C-H groups in the non-chair conformations of the pyranose ring, will obviously, depend on the configuration and conformation of the saccharide.

Spatial Disposition of C-H Groups in the Furanose Ring Form

Other than the pyranose ring form, monosaccharides can exist in acyclic, furanose and heptanose forms also. Of these, the five-membered furanose ring form is more common. Galactofuranose has been reported to be present in prokaryotes and lower eukaryotes (see, for example, [46]). A furanose ring can exist in either envelope or twist conformation [47] and can interconvert between some of these conformations [48]. In hexofuranoses, the exocyclic group at C4 is -CHOH-CH_2OH, unlike pentofuranoses (just the -CH_2OH group). Depending upon the conformation of this moiety, its C-H groups will either be towards the a- face or b-face (Fig. **3**).

A recent x-ray crystallographic study of the furanose-specific sugar transport protein YtfQ from *Escherichia coli* showed that the furanose ring form can also stack against the aromatic residues. In this complex, a single molecule of galactofuranose was found to be sandwiched between Trp16 and Phe170 (PDB id 2VK2) [49]. The conformation of the bound galactofuranose can not be ascertained from this structure; the furanose ring atoms along with the C2-O and C3-O atoms are essentially coplanar. However, the configuration of C2 and C3 atoms in galactose is such that in the furanose form, the C2-H (a-face) and C3-H (b-face) atoms are trans to each other, and hence, are placed on opposite sides of the furanose ring (Fig. **3**). C1-H will be on the same side of the furanose ring as C2-H, if the configuration of the anomeric carbon is α.

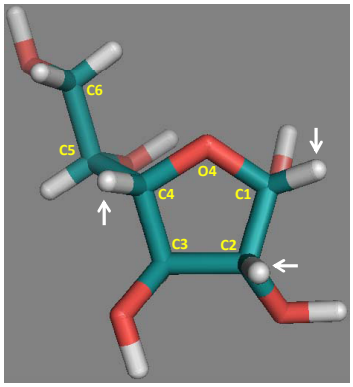

Fig. (3). Stick representation depicting the a-face of α-D-galactofuranose. The conformations of the furanose ring (^2E) and of the -OH and exocyclic -CHOH-CH_2OH groups have been chosen randomly.

With the aim of probing C-H...π interactions between monosaccharides and aromatic compounds, the effect of monosaccharides as co-solvents in increasing the solubility of adenine was investigated by a combination of experimental and computations methods [50]. Based on molecular mechanics calculations, it was inferred that the saccharides interact with adenine through C-H...π interactions, thereby preventing their aggregation through vertical stacking. The ability of the saccharides to increase the solubility of adenine followed this order: Galactose = lactose > sucrose > glucose = maltose > ribose > fructose. The weak co-solvent effect of

fructose was attributed to the absence of apolar patch in β-D-fructofuranose [50].

Taken together, the data from the x-ray crystallographic study of the YtfQ-Galf complex (stacking possible for galactofuranose) [49] and the study on the effect of monosaccharides on the solubility of adenine (stacking not possible for fructofuranose) [50] show that the configuration and conformation are key determinants of C-H...π stacking involving the furanose form also.

C-H…π Interactions in Modified Saccharides

The enormous structural diversity observed for glycans in Nature arises due to the possibility of (i) alternative anomeric linkages and (ii) using different -OH groups for forming glycosidic linkages. In addition, diversity also arises from the various types of functional group modifications, one of which is the replacement of an -OH group by an N-acetyl group. The $-CH_3$ in the acetyl moiety of 2-acetylamino-2-deoxy-saccharides also can interact with an aromatic residue through C-H...π interactions, even though it can not be part of the apolar patch on either the a- or b-face. An example of this can be seen in the complex of *Urtica dioica* agglutinin with tri-N-acetylchitotriose (Tyr30 and Tyr76; PDB id 1EHH) [51].

Strength of C-H…π Stacking in Model Systems

Quantum chemical calculations of saccharide - aromatic systems have been carried out to quantify the strength of C-H...π interactions. It is to be noted that the PO of the saccharide relative to the aromatic residue determines the number of interacting C-H groups and the geometry of interaction. The interaction energies for different saccharide - aromatic residue analogue complexes have been calculated using quantum chemical methods [27,28,52-57]. In these studies, the energies were computed for POs that are observed in various protein-carbohydrate complexes. The level of theory, basis sets and the POs used in these studies are not the same. The interaction energy was found to be stabilizing and to vary between approximately 1 and 12 kcal/mol. These values are comparable to those obtained by conventional hydrogen bonds. However, C-H...π stacking interactions (dominant dispersion component) are weakly directional unlike the conventional hydrogen bonds (dominant electrostatic component) [17].

The interaction energies of tryptophan with different saccharides (tetraacetyl β-D-Glc, tetraacetyl β-D-Gal, triacetyl N-acetyl-β-D-glucosamine, N-acetyl-β-D-glucosamine and tetramethyl-β-D-glucose) and of tetraacetyl β-D-Glc with different aromatic residues (Trp, naphthalene, phenylalanine and cyclohexane as a control) in the context of a β-hairpin have been determined by NMR spectroscopy: the magnitude varied between -0.5 to -0.8 kcal/mol [58]. These values are lower than those obtained from computational methods (see above). This difference can be attributed to the presence of solvent in the experimental study, since the energies calculated by the computational methods are for the gas phase. It is known that even conventional hydrogen bond energies are reduced in presence of solvent water (see, for example, [59]).

Strength of C-H…π Stacking in a Protein Environment

The binding of xylans and xylooligosaccharides $(Xyl)_n$, n=2 to 6 to two enzymatically inactive xylanases has been investigated by isothermal titration calorimetry and the thermodynamic data have been interpreted in light of the 3-D structures available from x-ray crystallographic studies [60]. Xylose is an aldopentose and is in the pyranose form in xylans. Xylopyranose is like glucopyranose except that it has a hydrogen atom in place of the exocyclic $-CH_2OH$ group of glucose. The stacking interaction between xylopyranose and the aromatic residues Tyr and Trp were found to contribute about -100 to -150 cal/mol.K each to ΔC_p. However, the strength of the C-H...π stacking is not known from this study. In this context, it is instructive to consider the outcome of the experimental studies performed to determine the strength of other types of non-covalent interactions in protein environment.

Numerous studies have investigated the strength of conventional hydrogen bonds and cation...π interactions both in gas phase and in an aqueous medium. The general consensus appears to be that the hydrogen bond becomes weaker in presence of water. However, there are fewer studies on the determination of the contribution of hydrogen bonds [61,62] and cation...π interactions [63,64] in a protein environment. Overall, these studies indicate that the strength of the same type of hydrogen bond may vary depending upon the geometry and context in which the hydrogen bond is formed. A double-mutant cycle analyses of bacteriorhodopsin showed that most hydrogen bonding interactions in membrane proteins are modestly stabilizing; it is suggested that this may be due to sub-optimal hydrogen bonding geometries, adverse entropic costs and increase in the local dielectric constant [65]. The strength of the Arg138-Glu52 and Arg138-Glu128 salt bridges in the gate region of the OmpA ion-channel were found to be -5.6 and -0.6 kcal/mol, respectively [66]. Collectively, the outcome of these studies suggest that the contribution of C-H...π stacking, especially vis-à-vis conventional hydrogen bonding, to the binding affinity of a saccharide to protein may vary. Experiments which mimic the double mutant cycle analysis (used for measuring the interaction energy between two amino acid side chains) have to be designed and executed to gain further insights into the contribution of C-H...π interactions to the affinity of glycans to proteins and the glycan specificity of proteins.

ACKNOWLEDGEMENT

I have benefited a lot from the many discussion meetings I have had with Manju, Sasidhar, Sujatha and Sunoj and I am very grateful to them. I am thankful to the Council for Scientific and Industrial Research, India for the financial support. I am also thankful to Indian Institute of Technology Bombay for providing me the opportunity to teach and for providing all the research facilities. Special thanks to Manju for reading the draft version of this manuscript; her suggestions were very useful.

REFERENCES

[1] Malik, A.; Ahmad, S. Sequence and structural features of carbohydrate binding in proteins and assessment of predictability using a neural network. *BMC Struct. Biol.*, **2007**, *7*, 1.

[2] Denker, K.; Orlik, F.; Schiffler, B.; Benz, R. Site-directed mutagenesis of the greasy slide aromatic residues within the LamB (maltoporin) channel of *Escherichia coli*: effect on ion and maltopentaose transport. *J. Mol. Biol.*, **2005,** *352*, 534-550.

[3] Jones, L.H.; Badger, R.M. The frequency of the third harmonic O-H band of methanol in various solvents, and the hydrogen bonding of methanol with aromatic hydrocarbons. *J. Am. Chem. Soc.*, **1951,** *73*, 3132-3134.

[4] Tamres, M. Aromatic compounds as donor molecules in hydrogen bonding. *J. Am. Chem. Soc.*, **1952**, *74*, 3375-3378.

[5] Babu, M.M.; Singh, S.K.; Balaram, P. A C–H…O hydrogen bond stabilized polypeptide chain reversal motif at the C terminus of helices in proteins. *J. Mol. Biol.*, **2002**, *322*, 871-880.

[6] Gu, Y.; Kar, T.; Scheiner, S. Fundamental properties of the CH…O interaction: is it a true hydrogen bond? *J. Am. Chem. Soc.*, **1999**, *121*, 9411-9422.

[7] Mandel-Gutfreund, Y.; Margalit, H.; Jernigan, R.L.; Zhurkin, V.B. A role for CH…O interactions in protein-DNA recognition. *J. Mol. Biol.*, **1998**, *277*, 1129-1140.

[8] McGaughey, G.B.; Gagne, M.; Rappe, A.K. π-Stacking interactions. Alive and well in proteins. *J. Biol. Chem.*, **1998**, *273*, 15458–15463.

[9] Tewari, A.K.; Dubey, R. Emerging trends in molecular recognition: utility of weak aromatic interactions. *Bioorg. Med. Chem.*, **2008**, *16*, 126-143.

[10] Fujii, A.; Morita, S.-i.; Miyazaki, M.; Ebata, T.; Mikami, N. A molecular cluster study on activated CH/π interactions: infrared spectroscopy of aromatic molecule-acetylene clusters. *J. Phys. Chem. A*, **2004**, *108*, 2652-2658.

[11] Morita, S.-i.; Fujii, A.; Mikami, N.; Tsuzuki, S. Origin of the attraction in aliphatic C-H/π interactions: infrared spectroscopic and theoretical characterization of gas-phase clusters of aromatics with methane. *J. Phys. Chem. A*, **2006**, *110*, 10583-10590.

[12] Ramos, C.; Winter, P.R.; Stearns, J.A.; Zwier, T.S. The spectroscopic consequences of C-H...π H-bonding: C_6H_6-$(C_4H_2)_n$ clusters with n = 1 and 2. *J. Phys. Chem. A*, **2003**, *107*, 10280-10287.

[13] Shibasaki, K.; Fujii, A.; Mikami, N.; Tsuzuki, S. Magnitude and nature of interactions in benzene-X (X = ethylene and acetylene) in the gas phase: significantly different CH/π interaction of acetylene as compared with those of ethylene and methane. *J. Phys. Chem. A*, **2007**, *111*, 753-758.

[14] Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. The magnitude of the CH/π interaction between benzene and some model hydrocarbons. *J. Am. Chem. Soc.*, **2000**, *122*, 3746-3753.

[15] Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Fujii, A. Magnitude and directionality of the interaction energy of the aliphatic CH/p interaction: significant differences from the hydrogen bond. *J. Phys. Chem. A*, **2006**, *110*, 10163-10168.

[16] Hermansson, K. Blue-shifting hydrogen bonds. *J. Phys. Chem. A*, **2002**, *106*, 4695-4702.

[17] Tsuzuki, S.; Uchimaru, T.; Mikami, M. Magnitude and nature of carbohydrate-aromatic interactions: ab initio calculations of fucose-benzene complex. *J. Phys. Chem. B*, **2009**, *113*, 5617–5621.

[18] Sundari, C.S.; Balasubramanian, D. Hydrophobic surfaces in saccharide chains. *Prog. Biophys. Mol. Biol.*, **1997**, *67*, 183-216.

[19] Quiocho, F.A.; Spurlino, J.C.; Rodseth, L.E. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure*, **1997**, *5*, 997-1015.

[20] Hobza, P. Stacking interactions. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 2581-2583.

[21] Guckian, K.M.; Schweitzer, B.A.; Ren, R.X.-F.; Sheils, C.J.; Tahmassebi, D.C.; Kool, E.T. Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.*, **2000**, *122*, 2213-2222.

[22] Stanca-Kaposta, E.C.; Gamblin, D.P.; Screen, J.; Liu, B.; Snoek, L.C.; Davis, B.G.; Simons, J.P. Carbohydrate molecular recognition: a spectroscopic investigation of carbohydrate–aromatic interactions. *Phys. Chem. Chem. Phys.*, **2007**, *9*, 4444-4451.

[23] Snoek, L.C.; Kroemer, R.T.; Hockridge, M.R.; Simons, J.P. Conformational landscapes of aromatic amino acids in the gas phase: infrared and ultraviolet ion dip spectroscopy of tryptophan. *Phys. Chem. Chem. Phys.*, **2001**, *3*, 1819-1826.

[24] Mecozzi, S.; West, A.P. Jr.; Dougherty, D.A. Cation-π interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 10566-10571.

[25] Chavez, M.I.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J.L.; Asensio, G.; Muraki, M.; Canada, F.J.; Jimenez-Barbero, J. On the importance of carbohydrate–aromatic interactions for the molecular recognition of oligosaccharides by proteins: NMR studies of the structure and binding affinity of AcAMP2-like peptides with non-natural naphthyl and fluoroaromatic residues. *Chem. Eur. J.*, **2005**, *11*, 7060-7074.

[26] Ramirez-Gualito, K.; Alonso-Rios, R.; Quiroz-Garcia, B.; Rojas-Aguilar, R.; Diaz, D.; Jimenez-Barbero, J.; Cuevas, G. Enthalpic nature of the CH/π interaction involved in the recognition of carbohydrates by aromatic compounds, confirmed by a novel interplay of NMR, calorimetry, and theoretical calculations. *J. Am. Chem. Soc.*, **2009**, *131*, 18129–18138.

[27] Sujatha, M.S.; Sasidhar, Y.U.; Balaji, P.V. Energetics of galactose– and glucose–aromatic amino acid interactions: implications for binding in galactose-specific proteins. *Protein Sci.*, **2004**, *13*, 2502–2514.

[28] Sujatha, M.S.; Sasidhar, Y.U.; Balaji, P.V. Insights into the role of the aromatic residue in galactose-binding sites: MP2/6-311G++** study on galactose- and glucose-aromatic residue analogue complexes. *Biochemistry*, **2005**, *44*, 8554-8562.

[29] Fondeur-Gelinotte, M.; Lattard, V.; Gulberti, S.; Oriol, R.; Mulliert, G.; Coughtrie, M.W.; Magdalou, J.; Netter, P.; Ouzzine, M.; Fournel-Gigleux, S. Molecular basis for acceptor substrate specificity of the human β1,3-glucuronosyltransferases GlcAT-I and GlcAT-P involved in glycosaminoglycan and HNK-1 carbohydrate epitope biosynthesis, respectively. *Glycobiology*, **2007**, *17*, 857-867.

[30] Iobst, S.T.; Drickamer, K. Binding of sugar ligands to a Ca^{2+}-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis. *J. Biol. Chem.*, **1994**, *269*, 15512–15519.

[31] Nakamura, T.; Tonozuka, T.; Ide, A.; Yuzawa, T.; Oguma, K.; Nishikawa, A. Sugar-binding sites of the HA1 subcomponent of *Clostridium botulinum* type C progenitor toxin. *J. Mol. Biol.*, **2008**, *376*, 854-867.

[32] Pratap, J.V.; Jeyaprakash, A.A.; Rani, P.G.; Sekar, K.; Surolia, A.; Vijayan, M. Crystal structures of artocarpin, a *Moraceae* lectin with mannose specificity, and its complex with methyl-α-D-mannose: implications to the generation of carbohydrate specificity. *J. Mol. Biol.*, **2002**, *317*, 237-247.

[33] Rini, J.M. Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.*, **1995**, *24*, 551–577.

[34] Zhang, Y.; Deshpande, A.; Xie, Z.; Natesh, R.; Acharya, K.R.; Brew, K. Roles of active site tryptophans in substrate binding and catalysis by α1,3 galactosyltransferase. *Glycobiology*, **2004**, *14*, 1295-1302.

[35] Sujatha, M.S.; Balaji, P.V. Identification of common structural features of binding sites in galactose-specific proteins. *Proteins*, **2004**, *55*, 44-65.

[36] Rao, V.S.R.; Qasba, P.K.; Balaji, P.V.; Chandrasekaran, R. *Conformation of Carbohydrates*. Harwood Academic Publishers: The Netherlands, Chapters 3 and 4, **1998**.

[37] Strauss, H.L.; Pickett, H.M. Conformational structure, energy, and inversion rates of cyclohexane and some related oxanes. *J. Am. Chem. Soc.*, **1970**, *92*, 7281-7290.

[38] Cremer, D.; Pople, J.A. General definition of ring puckering coordinates. *J. Am. Chem. Soc.*, **1975**, *97*, 1354-1358.

[39] Strynadka, N.C.J.; James, M.N.G. Lysozyme revisited: crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *J. Mol. Biol.*, **1991**, *220*, 401-424.

[40] Vocadlo, D.J.; Davies, G.J.; Laine, R.; Withers, S.G. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature*, **2001**, *412*, 835-838.

[41] Uitdehaag, J.C.; Mosi, R.; Kalk, K.H.; van der Veen, B.A.; Dijkhuizen, L.; Withers, S.G.; Dijkstra, B.W. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. *Nat. Struct. Biol.*, **1999**, *6*, 432-436.

[42] Marszalek, P.E.; Oberhauser, A.F.; Pang, Y.-P.; Fernandez, J.M. Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature*, **1998**, *396*, 661-664.

[43] Rao, V.S.R.; Balaji, P.V.; Qasba, P.K. Controversial iduronate ring conformation. *Glycobiology*, **1995**, *5*, 273-276.

[44] Dementiev, A.; Petitou, M.; Herbert, J.M.; Gettins, P.G. The ternary complex of antithrombin-anhydrothrombin-heparin reveals the basis of inhibitor specificity. *Nat. Struct. Mol. Biol.*, **2004**, *11*, 863-867.

[45] Caira, M.R.; Griffith, V.J.; Nassimbeni, L.R.; van Oudtshoorn, B. Unusual $^1C_4$ conformation of a methylglucose residue in crystalline permethyl-β-cyclodextrin monohydrate. *J. Chem. Soc., Perkin Trans. 2*, **1994**, 2071-2072.

[46] Richards, M.R.; Lowary, T.L. Chemistry and biology of galactofuranose-containing polysaccharides. *ChemBioChem*, **2009**, *10*, 1920-1938.

[47] Altona, C.; Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.*, **1972**, *94*, 8205–8212.

[48] Cremer, D.; Pople, J.A. Molecular orbital theory of the electronic structure of organic compounds. XXIII. Pseudorotation in saturated five-membered ring compounds. *J. Am. Chem. Soc.*, **1975**, *97*, 1358-1367.

[49] Horler, R.S.; Mueller, A.; Williamson, D.C.; Potts, J.R.; Wilson, K.S.; Thomas, G.H. Furanose-specific sugar transport: characterization of a bacterial galactofuranose-binding protein. *J. Biol. Chem.*, **2009**, *284*, 31156-31163.

[50] Maresca, M.; Derghal, A.; Carravagna, C.; Dudin, S.; Fantini, J. Controlled aggregation of adenine by sugars: physicochemical studies, molecular modelling simulations of sugar–aromatic CH–π stacking interactions, and biological significance. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 2792–2800.

[51] Harata, K.; Muraki, M. Crystal structures of *Urtica dioica* agglutinin and its complex with tri-N-acetylchitotriose. *J. Mol. Biol.*, **2000**, *297*, 673-681.

[52] Fernandez-Alonso, M.C.; Canada, F.J.; Jimenez-Barbero, J.; Cuevas, G. Molecular recognition of saccharides by proteins. Insights on the origin of the carbohydrate-aromatic interactions. *J. Am. Chem. Soc.*, **2005**, *127*, 7379-7386.

[53] Raju, R.K.; Ramraj, A.; Hillier, I.H.; Vincent, M.A.; Burton, N.A. Carbohydrate–aromatic π interactions: a test of density functionals and the DFT-D method. *Phys. Chem. Chem. Phys.*, **2009**, *11*, 3411 - 3416.

[54] Sharma, R.; McNamara, J.P.; Raju, R.K.; Vincent,M.A.; Hillier, I.A.; Morgado, C.A. The interaction of carbohydrates and amino acids with aromatic systems studied by density functional and semi-empirical molecular orbital calculations with dispersion corrections. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 2767-2774.

[55] Spiwok, V.; Lipovova, P.; Skalova, T.; Buchtelova, E.; Hasek, J.; Kralova, B. Role of CH/π interactions in substrate binding by *Escherichia coli* β-galactosidase. *Carbohydr. Res.*, **2004**, *339*, 2275-2280.

[56] Spiwok, V.; Lipovova, P.; Skalova, T.; Vondrackova, E.; Dohnalek, J.;Hasek, J.; Kralova, B. Modelling of carbohydrate–aromatic interactions: ab initio energetics and force field performance. *J. Comput.-Aided Mol. Des.*, **2006**, *19*, 887-901.

[57] Sujatha, M.S.; Sasidhar, Y.U.; Balaji, P.V. MP2/6-311++G(d,p) study on galactose–aromatic residue analogue complexes in different position-orientations of the saccharide relative to aromatic residue. *J. Mol. Struct. (Theochem)*, **2007**, *814*, 11-24.

[58] Laughrey, Z.R.; Kiehna, S.E.; Riemen, A.J.; Waters, M.L. Carbohydrate-π interactions: what are they worth? *J. Am. Chem. Soc.*, **2008**, *130*, 14625-14633.

[59] Shen, S.-Y.; Yang, D.-Y.; Selzle, H. L.; Schlag, E.W. Energetics of hydrogen bonds in peptides, *Proc. Natl. Acad. Sci. U.S.A.*, **2003**, *100*, 12683-12687.

[60] Zolotnitsky, G.; Cogan, U.; Adir, N.; Solomon, V.; Shoham, G.; Shoham, Y. Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 11275-11280.

[61] Chapman, E.; Thorson, J.S.; Schultz, P.G. Mutational analysis of backbone hydrogen bonds in staphylococcal nuclease. *J. Am. Chem. Soc.*, **1997**, *119*, 7151-7152.

[62] Thorson, J.S.; Chapman, E.; Schultz, P.G. Analysis of hydrogen bonding strengths in proteins using unnatural amino acids. *J. Am. Chem. Soc.*, **1995**, *117*, 9361-9362.

[63] Prajapati, R.S.; Sirajuddin, M.; Durani, V.; Sreeramulu, S.; Varadarajan, R. Contribution of cation-π Interactions to protein stability. *Biochemistry*, **2006**, *45*, 15000-15010.

[64] Wei, Y.; Horng, J.-C.; Vendel, A.C.; Raleigh, D.P.; Lumb, K.J. Contribution to stability and folding of a buried polar residue at the CARM1 methylation site of the KIX domain of CBP. *Biochemistry*, **2003**, *42*, 7044-7049.

[65] Joh, N.H. Min, A.; Faham, S.; Whitelegge, J.P.; Yang, D.; Woods, V.L. Jr.; and Bowie, J.U. Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature*, **2008**, *453*, 1266-1270.

[66] Hong, H.; Szabo, G.; Tamm, L.K. Electrostatic couplings in OmpA ion-channel gating suggest a mechanism for pore opening. *Nat. Chem. Biol.*, **2006**, *2*, 627-635.

[67] Matsui, I.; Yoneda, S.; Ishikawa, K.; Miyairi, S.; Fukui, S.; Umeyama, H.; Honda, K. Roles of the aromatic residues conserved in the active center of *Saccharomycopsis* α-amylase for transglycosylation and hydrolysis activity. *Biochemistry*, **1994**, *33*, 451-458.

[68] Hamelryck, T.W.; Loris, R.; Bouckaert, J.; Dao-Thi, M.-H.; Strecker, G.; Imberty, A.; Fernandez, E.; Wyns, L.; Etzler, M.E. Carbohydrate binding, quaternary structure and a novel hydrophobic binding site in two legume lectin oligomers from *Dolichos biflorus*. *J. Mol. Biol.*, **1999**, *286*, 1161-1177.

[69] Imberty, A.; Gautier, C.; Lescar, J.; Perez, S.; Wyns, L.; Loris, R. An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J. Biol. Chem.*, **2000**, *275*, 17541-17548.

[70] McLean, B.W.; Bray, M.R.; Boraston, A.B.; Gilkes, N.R.; Haynes, C.A.; Kilburn, D.G. Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. *Protein Eng.*, **2000**, *13*, 801-809.

[71] Muraki, M.; Harata, K.; Sugita, N.; Sato, K.-I. Protein−carbohydrate interactions in human lysozyme probed by combining site-directed mutagenesis and affinity labeling. *Biochemistry*, **2000**, *39*, 292-299.

[72] Hayashida, M.; Fujii, T.; Hamasu, M.; Ishiguro, M.; Hata, Y. Similarity between protein–protein and protein–carbohydrate interactions, revealed by two crystal structures of lectins from the roots of pokeweed. *J. Mol. Biol.*, **2003**, *334*, 551–565.

[73] Huber, R.E.; Hakda, S.; Cheng, C.; Cupples, C.G.; Edwards, R.A. Trp-999 of β-galactosidase (*Escherichia coli*) is a key residue for binding, catalysis, and synthesis of allolactose, the natural *Lac* operon inducer. *Biochemistry*, **2003**, *42*, 1796-1803.

[74] Dupuy, F.; Germot, A.; Julien, R.; Maftah, A. Structure/function study of Lewis α3- and α3/4-fucosyltransferases: the α1,4 fucosylation requires an aromatic residue in the acceptor-binding domain. *Glycobiology*, **2004**, *14*, 347-356.

[75] Jeyaprakasha, A.A.; Jayashree, G.; Mahanta, S.K.; Swaminathan, C.P.; Sekar, K.; Surolia, A.; Vijayan, M. Structural basis for the energetics of jacalin–sugar interactions: promiscuity versus specificity. *J. Mol. Biol.*, **2005**, *347*, 181-188.

[76] Ragunath, C.; Manuel, S.G.; Venkataraman, V.; Sait, H.B.; Kasinathan, C.; Ramasubbu, N. Probing the role of aromatic residues at the secondary saccharide-binding sites of human salivary alpha-amylase in substrate hydrolysis and bacterial binding. *J. Mol. Biol.*, **2008**, *384*, 1232-1248.

[77] Walti, M.A.; Walser, P.J.; Thore, S.; Grunler, A.; Bednar, M.; Kunzler, M.; Aebi, M. Structural basis for chitotetraose coordination by CGL3, a novel galectin-related protein from *Coprinopsis cinerea*. *J. Mol. Biol.*, **2008**, *379*, 146-159.

Received: April 15, 2010 Revised: June 14, 2010 Accepted: June 16, 2010