Electronic Supplementary Material for PCCP

Carbohydrate-Protein recognition probed by density functional theory and *ab initio* calculations including dispersive interactions.

Rajesh K. Raju, Anitha Ramraj, Mark A. Vincent, Ian H. Hillier and Neil A. Burton

School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, UK.

Table S1. Relative energies (kcal mol^{-1}) and O-H vibrational frequencies (cm^{-1}) for fucose monomers (DFT-D/TZV2D).

Monomer	Relative Energy[a]	O-H Stretch			
		OH1	OH2	OH3	OH4
A	0.16	3673	3678	3656	3631
B	0.00	3628	3661	3662	3632
C	1.63	3643	3656	3634	3631

[a] with zero point and thermodynamic corrections, at 298 K.

Table S2. C-H frequencies (cm^{-1}) for fucose monomers (DFT-D/TZV2D).

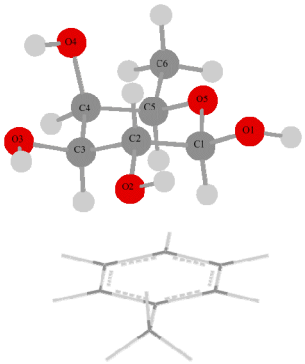
Monomer	C-H Stretch							
	CH1	CH2	CH3	CH4	CH5	CH6(a)	CH6(b)	CH6(c)
A	2853	2973	2924	2904	2879	2980	3040	3062
B	2920	2946	2926	2909	2890	2980	3042	3062
C	2874	3002	2890	2912	2877	2980	3042	3062

Table S3. Structures of α-methyl glucose-toluene complexes at the MP2/6-31G** level. U, L and S label Upper, Lower and Side-on structures, Lower being on the same side as the OMe group of the sugar. M, C and CC label the arrangement of the C(2), C(3) and C(4) hydroxyl groups of the sugar, M for mixed (meaning no pattern), C for all clockwise and CC for all counter clockwise. T, GP amd GM label the orientation of the exogenous side chain oxygen atom. For T structures the O-C(5)-C(6)-O dihedral angle is approximately trans, whilst for GP and GM structures the dihedral angle is gauche (P for counter clockwise and M for clockwise). The final value is the O-C(6)….C(T)-CH_3 dihedral angle (unsigned) where C(T) is the tertiary carbon of toluene and CH_3 is the methyl group of toluene, and gives the relative orientation of the sugar and toluene Finally the relative energies of each complex to the lowest are given in kcal mol^{-1}.

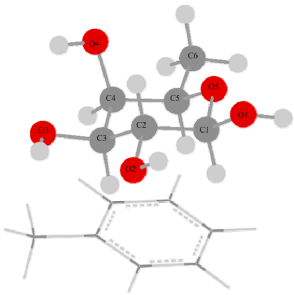
U-M-T-116	4.91	L-M-T-59	4.38
U-M-T-71	4.54	L-M-T-55	3.08
S-CC-GP-127	1.88	L-M-GP-40	5.41
U-M-GP-42	3.50	L-M-T-7	1.48
S-M-GP-44	4.74	L-M-GM-33	2.93
U-M-T-37	3.15	L-CC-GM-3	0.31
U-M-GP-121	3.06	L-CC-GM-41	1.36
U-M-GP-166	2.50	L-CC-T-6	0.00
U-M-GP-42	1.82	L-CC-GP-104	1.82
U-CC-T-148	2.44	L-CC-GP-105	2.18
U-CC-T-31	1.98	L-CC-GP-133	1.81
U-CC-GM-145	2.49	L-CC-GP-171	2.12
U-CC-GP-156	1.30	L-CC-GP-27	1.29
U-CC-GP-48	1.29	L-CC-GP-2	0.39
U-CC-GP-93	2.02	L-CC-GP-42	3.13
S-CC-GP-83	4.78	L-M-T-97	4.98
U-CC-GP-45	3.63	L-M-T-149	2.07
U-M-GP-160	3.55	L-C-T-117	5.52
U-C-T-151	6.60	S-CC-GP-58	4.39
U-CC-GM-26	2.12		

Table S4. The relative energies (kcal mol^{-1}) of the three main sugar conformations of α-methyl glucose at the MP2/6-31G** and, in parentheses, at the DFT-D/TZV2D level.

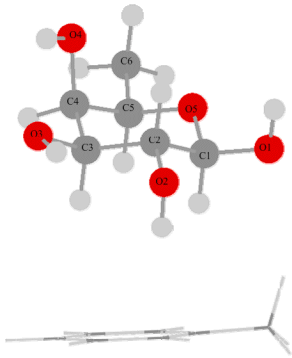
Sugar conformation		
	Energy	Free energy
A (T)	0.09 (0.24)	0.29 (0.45)
B (G-)	0.01 (0.00)	0.32 (0.00)
C (G+)	0.00 (0.13)	0.00 (0.08)



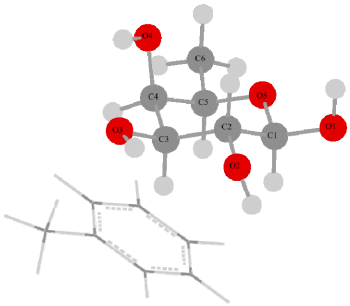
1



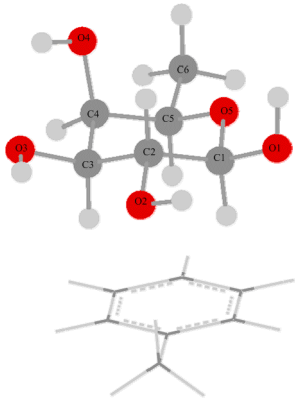
2



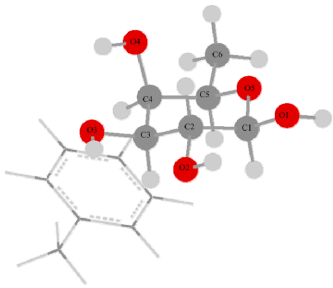
3



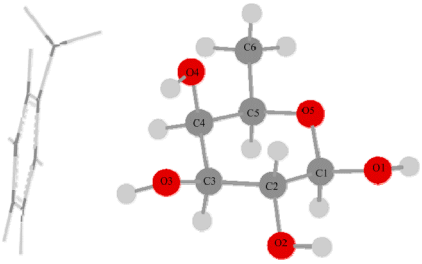
4



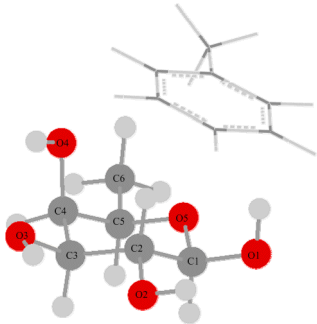
5



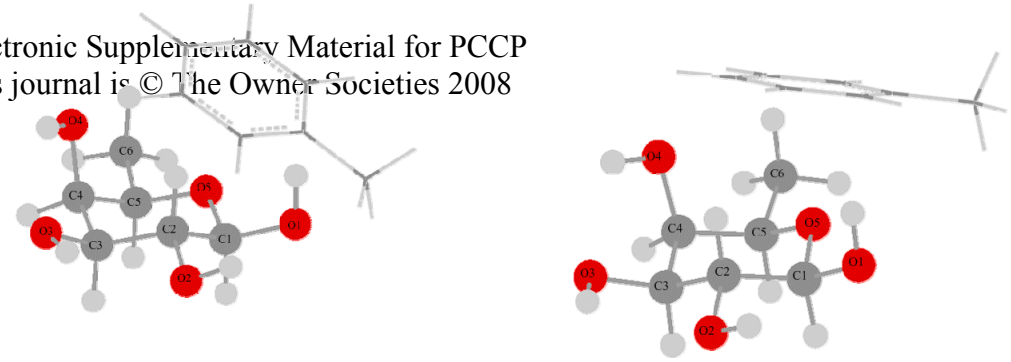
6



7

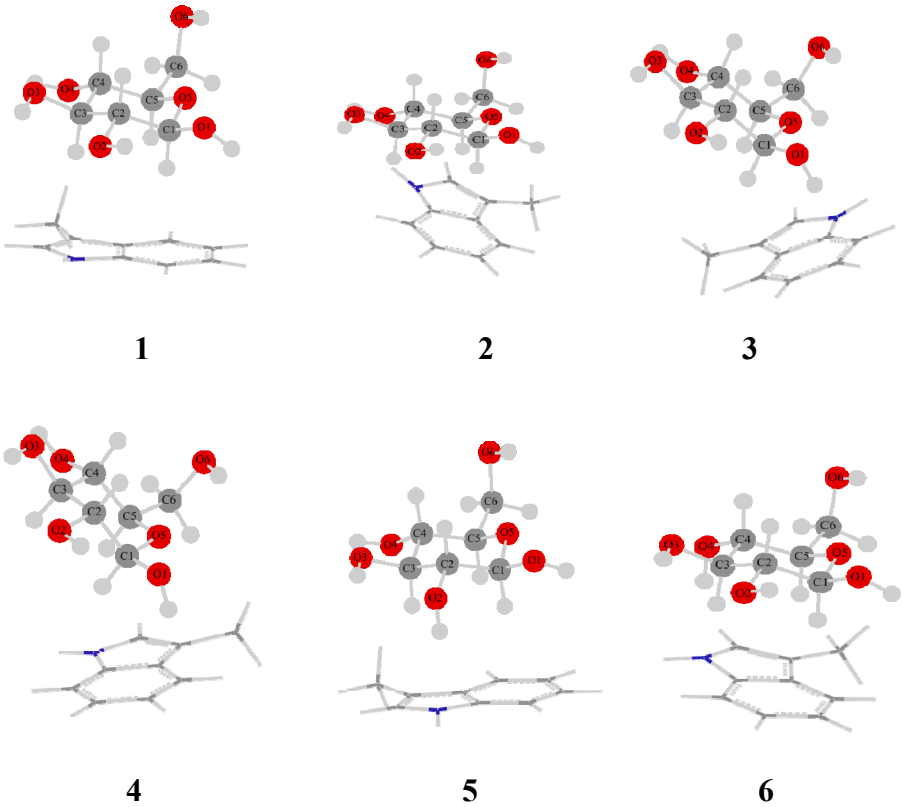


8



9 **10**

Fig. S1 Optimized structures of fucose-toluene complexes at DFT-D/TZV2D level.



1 **2** **3**

4 **5** **6**

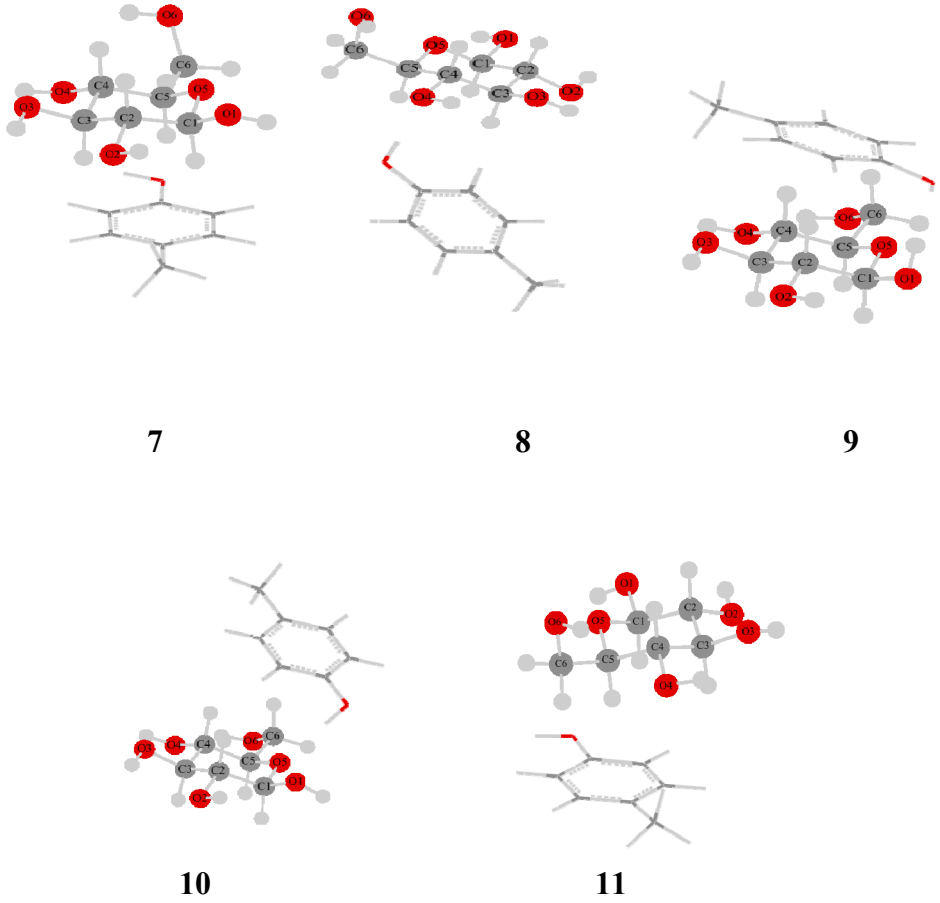


Fig. S2. Optimized structures of glucose–3-methylindole and –*p*-hydroxytoluene complexes

The procedure for extrapolating the MP2 energy to the complete basis set (CBS) limit .

The extrapolation to the basis set limit requires calculations at both MP2 and Hartree Fock (HF) levels using two basis sets of different sizes, here taken to be aug-cc-pVDZ and aug-cc-pVTZ .

The corresponding energies are substituted into the following equations of Helgaker and co-workers [1].

$$E^{HF}(X) = E^{HF}(CBS) + Ae^{-\alpha X}$$

$$E^{corr}(X) = E^{corr}(CBS) + BX^{-3}$$

Here E^{HF} and E^{corr} are the Hartree Fock and correlation (MP2) energies, X=2 for a double-zeta basis set (aug-cc-pVDZ) and X=3 for a triple-zeta basis set (aug-cc-pVTZ), and α equals 1.43. These equations are then solved to give $E^{corr}(CBS)$ and $E^{HF}(CBS)$, which are then combined to give the required CBS total energy.

The individual energies for the two complexes (au) are:

Complex	HF	MP2
Benzene and β-Galactose		
aug-cc-pVDZ	-914.182064	-917.081990
aug-cc-pVTZ	-914.386055	-917.867677
Benzene		
aug-cc-pVDZ	-230.729362	-231.545350
aug-cc-pVTZ	-230.779954	-231.746515
β-Galactose		
aug-cc-pVDZ	-683.455682	-685.526994
aug-cc-pVTZ	-683.609096	-686.110627

Fucose and Toluene		
aug-cc-pVDZ	-878.355166	-881.216459
aug-cc-pVTZ	-878.551763	-881.978908
Fucose		
aug-cc-pVDZ	-608.586920	-610.463857
aug-cc-pVTZ	-608.723880	-610.987078
Toluene		
aug-cc-pVDZ	-269.769561	-270.738300
aug-cc-pVTZ	-269.829187	-270.976177

1 A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, Chem. Phys. Letters, 1998, **286**, 243.