Supporting information for: Kirkwood-Buff Derived Alcohol Parameters for Aqueous Carbohydrates and their Application to Preferential Interaction Coefficient Calculations of Proteins

Theresa Cloutier,[†] Chaitanya Sudrik,[†] Hasige A. Sathish,[‡] and Bernhardt L. Trout[*,†]

†*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA*

‡*Formulation Sciences, MedImmune LLC, Gaithersburg, MD, USA*

E-mail: trout@mit.edu

Phone: 617-258-5021

Simulations Details

Table S1: Simulation details for carbohydrate-water simulations used to calculate Kirkwood-Buff integral values.

Carbohydrate	Side length (nm)	Concentration (mol/kg)	n_3	n_1	Time (ns)
glycerol	6.4	0.25	39	8563	235
		0.50	75	8375	135
		0.75	111	8195	165
		1.00	145	8022	165
		1.50	208	7697	135
		2.00	267	7398	125
sorbitol	7.4	0.25	58	12952	135
		0.50	112	12412	135
		1.00	206	11455	145
		1.50	287	10636	105
		2.00	358	9926	85
glucose	7.4	0.25	58	12959	135
		0.50	112	12423	115
		1.00	207	11475	145
		1.50	288	10661	85
		2.00	359	9955	105
sucrose	8.0	0.25	71	15762	145
		0.50	132	14610	135
		1.00	182	10089	135
		1.50	306	11306	135
		2.00	366	10157	105
trehalose	8.0	0.25	56	12475	135
		0.50	132	14610	135
		1.00	182	10089	135
		1.50	306	11306	135
		2.00	366	10157	85

Table S2: Simulation details for preferential interaction coefficient calculation.

Carbohydrate	Protein	Concentration (mol/kg)	n_3	n_1	Time (ns)
glycerol	lysozyme	0.25	59	13005	85
glycerol	lysozyme	0.50	115	12719	75
glycerol	lysozyme	1.00	220	12182	85
glycerol	lysozyme	1.50	316	11690	65
glycerol	lysozyme	2.00	405	11235	105
glycerol	BSA	0.25	154	34253	55
glycerol	BSA	0.50	302	33499	85
glycerol	BSA	1.00	578	32086	55
glycerol	BSA	1.50	832	30788	55
sorbitol	RNase A	0.25	57	12758	85
sorbitol	RNase A	0.50	110	12225	65
sorbitol	RNase A	1.00	203	11283	65
sorbitol	RNase A	1.50	283	10476	115
glucose	lysozyme	0.25	57	12731	85
glucose	lysozyme	0.50	110	12205	85
glucose	lysozyme	1.00	203	11273	85
glucose	lysozyme	1.50	283	10474	85
glucose	BSA	0.25	151	33531	55
glucose	BSA	0.50	290	32146	55
glucose	BSA	1.00	535	29692	55
glucose	BSA	1.50	746	27587	55
glucose	α-Cgn A	0.25	55	12128	85
glucose	α-Cgn A	0.50	105	11627	85
glucose	α-Cgn A	1.00	194	10740	85
glucose	α-Cgn A	1.50	270	9978	85
sucrose	α-Cgn A	0.25	53	11675	135
sucrose	α-Cgn A	0.50	98	10822	135
sucrose	α-Cgn A	1.00	170	9442	85
sucrose	α-Cgn A	1.50	226	8375	85
sucrose	RNase A	0.25	55	12287	85
sucrose	RNase A	0.50	103	11390	85
sucrose	RNase A	1.00	179	9937	85
sucrose	RNase A	1.50	238	8814	105
trehalose	lysozyme	0.25	55	12256	85
trehalose	lysozyme	0.50	102	11360	85
trehalose	lysozyme	1.00	179	9912	85
trehalose	lysozyme	1.50	238	8791	135
trehalose	lysozyme	2.00	285	7898	95
trehalose	RNase A	0.25	55	12287	155
trehalose	RNase A	0.50	103	11390	85
trehalose	RNase A	1.00	179	9937	135
trehalose	RNase A	1.50	238	8814	85

Kirkwood-Buff Integral Calculation and Sensitivity

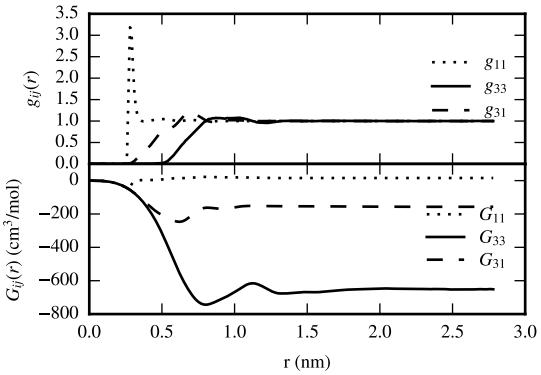

Figure S1: Radial distribution functions $g_{ij}(r)$ and KB integral values $G_{ij}(r)$ for 1 m trehalose using the KBPs. The KB integral value corresponds to the average value of $G_{ij}(r > R)$, where R for this simulation is 2 nm (see methods section).

Table S3: Impact of q and ε on KB Integral Values for 0.5 m Glycerol[a]

q (|e|)		general ε (kJ/mol)		water ε (kJ/mol)		KB integral (cm^3/mol)		
O	H	O	H	O	H	G_{11}	G_{33}	G_{31}
-0.50	0.18	0.45	0.12	0.90	0.30	-15.7 ± 0.1	-287 ± 27	-77 ± 1
-0.50	0.18	0.45	0.12	0.85	0.25	-15.6 ± 0.1	-167 ± 36	-80 ± 2
-0.45	0.13	0.45	0.12	0.90	0.30	-15.6 ± 0.1	-247 ± 50	-80 ± 2
-0.50	0.18	0.55	0.15	0.90	0.30	-15.5 ± 0.1	-156 ± 34	-82 ± 1

[a] The partial charges for all non-alcohol atoms in glycerol are as given in figure 1.

Carbohydrate Clustering

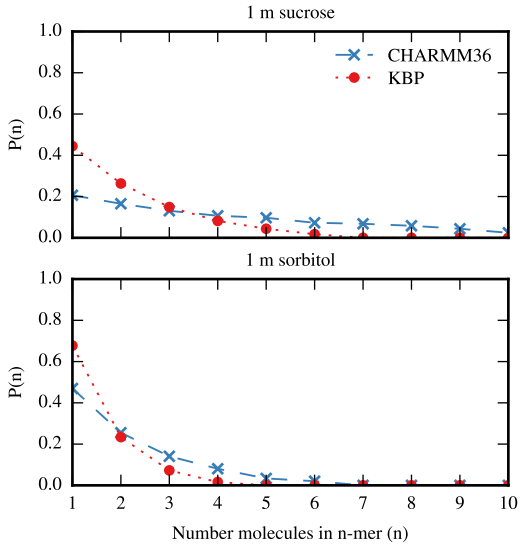

Figure S2: The probability of finding a carbohydrate molecule in a cluster of size n for sucrose and sorbitol. There is a higher probability of a carbohydrate being a monomer in simulations using the KBPs than those using the CHARMM36 parameters.

Carbohydrate Diffusion

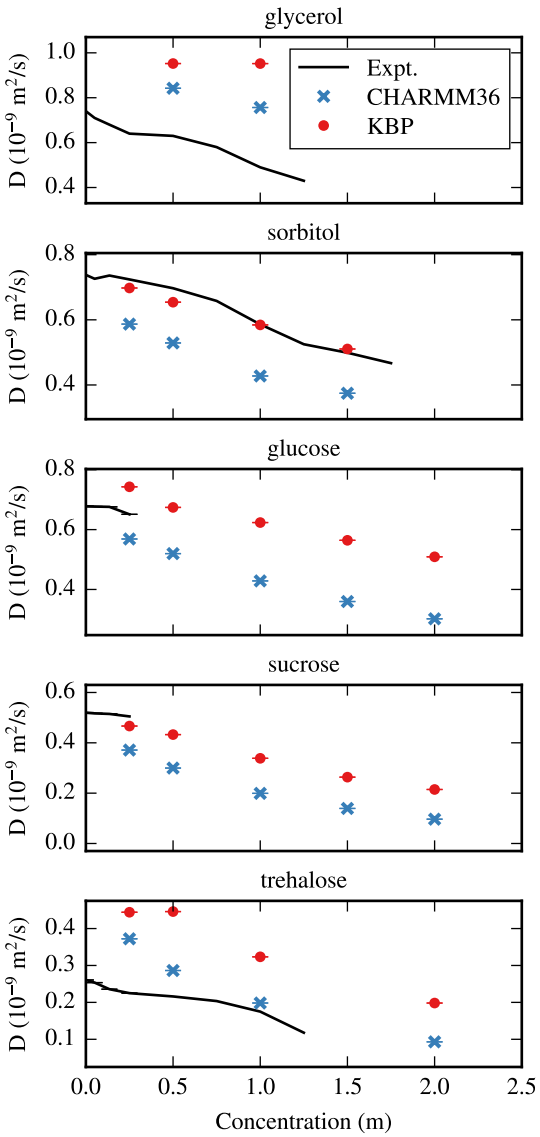

Figure S3: Diffusion coefficients were not included in the development of the KBPs. Generally, the diffusion coefficients with the KBPs are higher than those with the CHARMM36 parameters. For sorbitol, glucose, and sucrose, the results with the KBPs are a similar or better match for experimental data compared to results with the CHARMM36 parameters. Experimental data for glycerol from ref S1, for sorbitol from ref S2, for glucose from ref S3, for sucrose from ref S4, and for trehalose from ref S5.

Diffusion coefficients were calculated according to the methods of Hatcher et al.[S6] Briefly, D_{PBC} was calculated from the slope of the mean square displacement of solute molecules

versus time. Due to the underestimation of water viscosity by the TIP3P model, the diffusion coefficient was calculated according to:

$$D = \left(D_{PBC} + \frac{k_b T \zeta}{6\pi\eta L} \right) \times 0.375 \tag{S1}$$

where k_B is the Boltzmann constant, T the temperature, ζ a constant of 2.837297, η the viscosity, and L the side length of the simulation. The viscosity η is calculated from $\eta = \eta_{TIP3P}\,(1 + 2.5\phi)$ due to the TIP3P model, where $\eta_{TIP3P} = 0.35$ cP. ϕ is the volume fraction of the solute.

Pure Carbohydrate Data

The KBPs were developed for use in simulating carbohydrates in aqueous solution at concentrations below 2 m. For a more complete picture of how the force field performes for pure carbohydrates, some data of pure carbohydrates were calculated, with the results below.

Table S4: Cremer and Pople ring pucker parameter, Q, for glucose.

		Q (Å)
Glucose	Experiment	0.567[a]
	CHARMM36	0.570
	KBP	0.572

[a] Experimental data from ref S7.

Table S5: Crystal cell volumes. Crystals have four molecules per unit cell.

		Volume (Å^3)
Glucose	Experiment	774.2[a]
	CHARMM36	963
	KBP	919
Sucrose	Experiment	735.3[b]
	CHARMM36	1834
	KBP	1711
Trehalose	Experiment	1658[c]
	CHARMM36	1816
	KBP	1724

[a] Experimental data for glucose from ref S7.
[b] Experimental data for sucrose from ref S8.
[c] Experimental data for trehalose from ref S9.

Table S6: Pure component densities.

		ρ (g/cm^3)
Glycerol	Experiment	1.26
	CHARMM36	1.21
	KBP	1.08
Sorbitol	Experiment	1.49
	CHARMM36	1.30
	KBP	1.24
Glucose	Experiment	1.52
	CHARMM36	1.24
	KBP	1.30
Sucrose	Experiment	1.59
	CHARMM36	1.24
	KBP	1.33
Trehalose	Experiment	1.58
	CHARMM36	1.25
	KBP	1.32

References

(S1) D'Errico, G.; Ortona, O.; Capuano, F.; Vitagliano, V. Diffusion coefficients for the binary system glycerol + water at 25 C. A velocity correlation study. *J. Chem. Eng. Data* **2004**, *49*, 1665–1670.

(S2) Sartorio, R.; Wurzburger, S.; Guarino, G.; Borriello, G. Diffusion of polyols in aqueous solution. *J. Soln. Chem.* **1986**, *15*, 1041–1049.

(S3) Ribeiro, A. C.; Ortona, O.; Simoes, S. M.; Santos, C. I.; Prazeres, P. M.; Valente, A. J.; Lobo, V. M.; Burrows, H. D. Binary mutual diffusion coefficients of aqueous solutions of sucrose, lactose, glucose, and fructose in the temperature range from (298.15 to 328.15) K. *J. Chem. Eng. Data* **2006**, *51*, 1836–1840.

(S4) Gosting, L. J.; Morris, M. S. Diffusion studies on dilute solutions by gouy interference. *J. Am. Chem. Soc.* **1949**, *6*, 1998–2006.

(S5) Magazu, S.; Maisano, G.; Migliardo, P.; Middendorf, H. D.; Villari, V. Hydration and transport properties of aqueous solutions of a-a-trehalose. *J. Chem. Phys.* **1998**, *109*, 1170–1174.

(S6) Hatcher, E. R.; Guvench, O.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates, and Inositol. *J. Chem. Theory Comput.* **2009**, *5*, 1315–1327.

(S7) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. Additive Empirical Force Field for Hexopyranose Monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543–2564.

(S8) Beevers, C. A.; McDonald, T. R. R.; Robertson, J. H.; Stern, F. The crystal structure of sucrose. *Acta Crystallogr.* **1952**, *5*, 689–690.

(S9) Taga, T.; Senma, M.; Osaki, K. The crystal and molecular structure of trehalose dihydrate. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **1972**, *B28*, 3258–3263.