Treating Solute Polarization in Non-polarizable MD Simulations: Carbohydrates

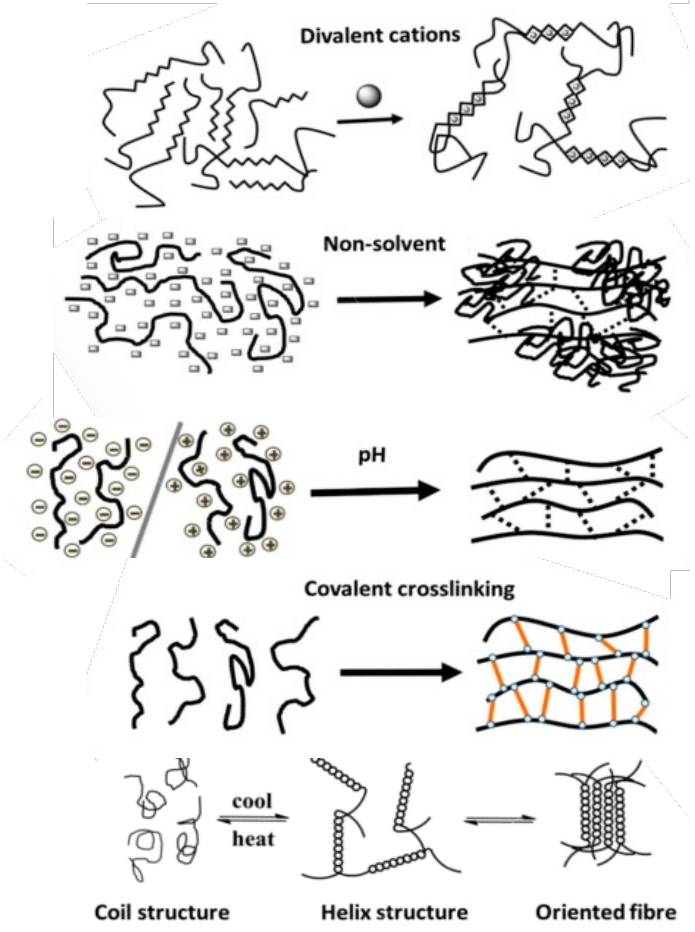
Xiaocong Wang[1,2], Samuel Holmes[1], Luke Newman[3,4], Jessica Nash[5], Taylor Barnes[5], Daniel Crawford[6], Valerie Welborn[3,4],

Robert J. Woods[1,7]

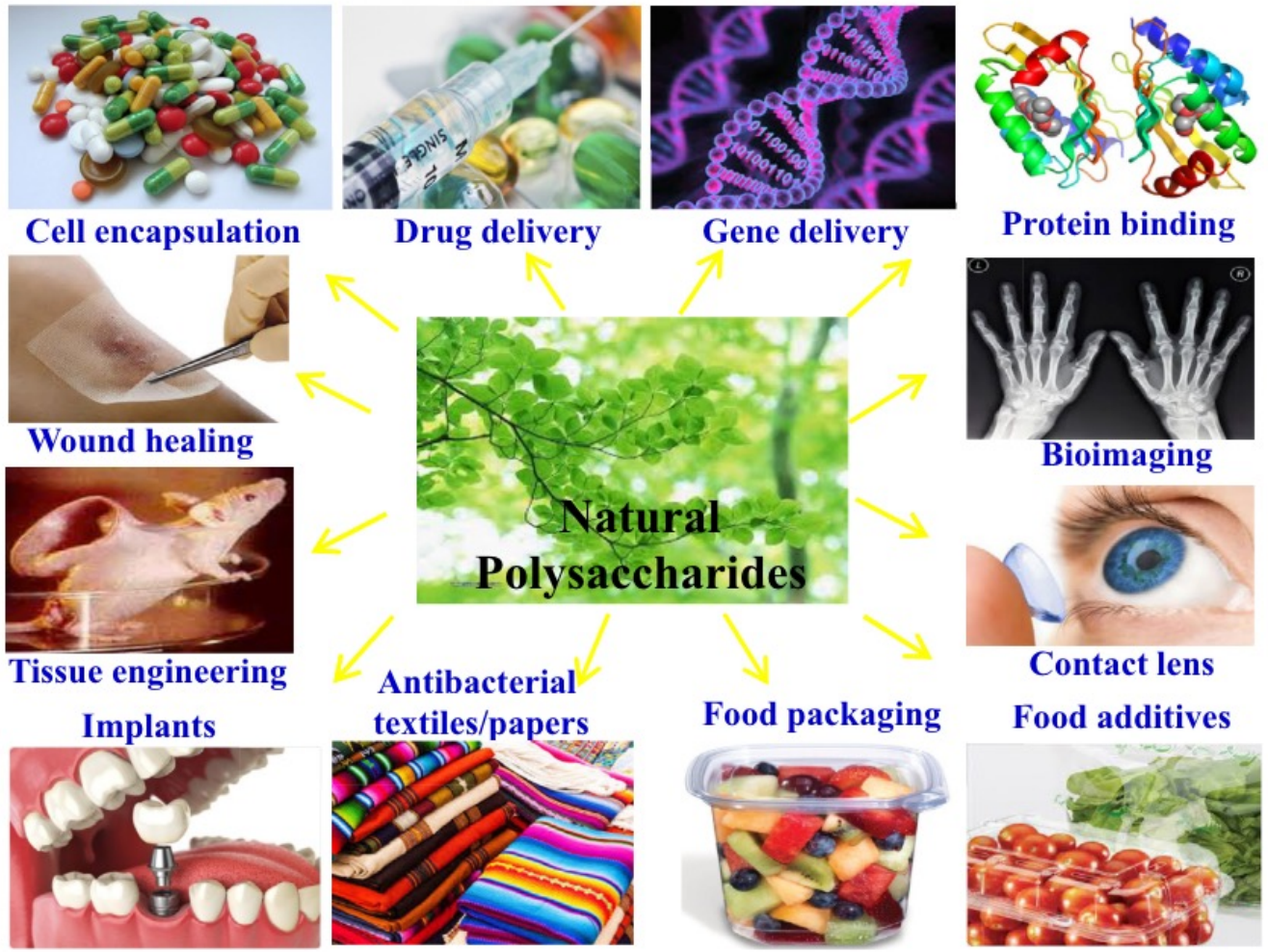
[1]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602
[2]Hubei Key Laboratory of Agricultural Bioinformatics, Huazhong Agricultural University, Wuhan, Hubei, 430070, China
[3]Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States
[4]Macromolecules Innovation Institute, Virginia Tech, Blacksburg, Virginia 24061, United States
[5]Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States
[6]Molecular Sciences Software Institute, Blacksburg, Virginia 24060, United States;
[7]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

GlycoMIP
NSF Materials Innovation Platform

Polysaccharide Gels: Biodegradable, Sustainable, Biotolerable, Designable?

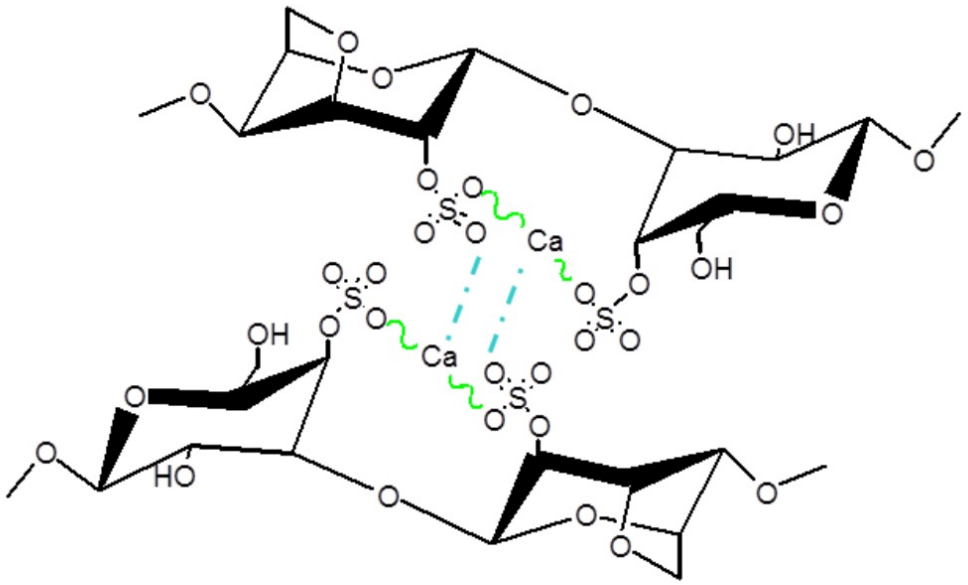


Ganesan et al.: Review on the Production of Polysaccharide Aerogel Particles



Shariatinia : Pharmaceutical applications of natural polysaccharides

Can We Predict Polysaccharide Gel Formation?



Tako: "The Principle of Polysaccharide Gels"

Mechanism: "The gelation occurs because of the formation of intra- and inter-molecular associations, where … the sugar residues of the polysaccharides contribute to hydrogen bonding or van der Waals forces of attraction."

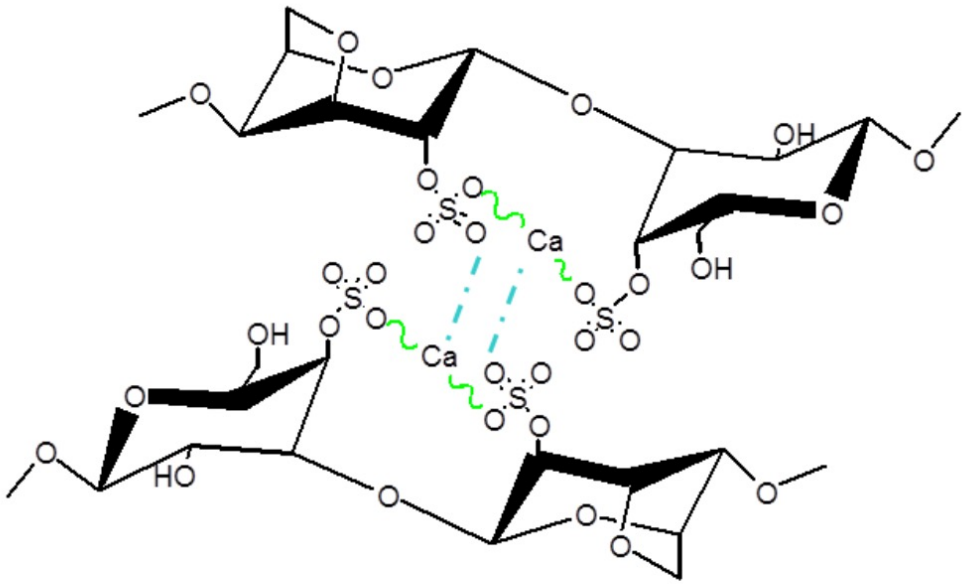
"…might be caused by salting-out effect which might take place because of the rigidity of the molecular chains of agarose through the formation of intra- and inter-molecular association…"

- extremely generic definition
- what about entropy?

Requirements for Modeling Carbohydrate Gel Formation:

- Accurate carbohydrate force field
- Fast MD simulation method (large solvated intermolecular complexes)
- Accurate energies for water-ion-solute interactions

Can We Predict Polysaccharide Gel Formation?



Tako: "The Principle of Polysaccharide Gels"

Mechanism: "The gelation occurs because of the formation of intra- and inter-molecular associations, where … the sugar residues of the polysaccharides contribute to hydrogen bonding or van der Waals forces of attraction."

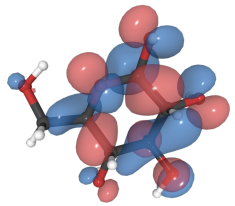
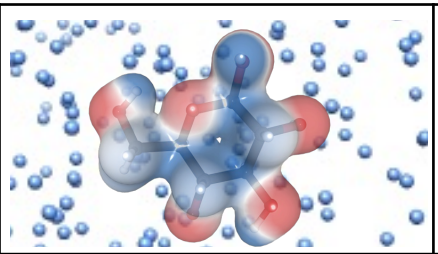
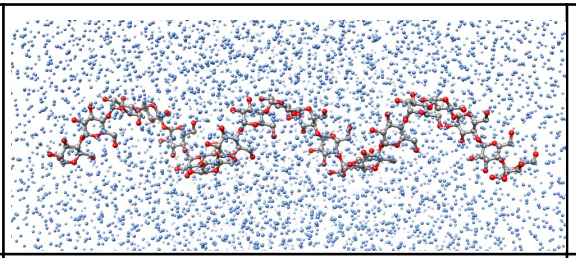
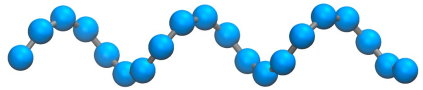
"…might be caused by salting-out effect which might take place because of the rigidity of the molecular chains of agarose through the formation of intra- and inter-molecular association…"

- extremely generic definition
- what about entropy?

Requirements for Modeling Carbohydrate Gel Formation:

- Accurate carbohydrate force field ?
- Fast MD simulation method (large solvated intermolecular complexes) ✓
- Accurate energies for water-ion-solute interactions ?

Considerations and Challenges in Modeling

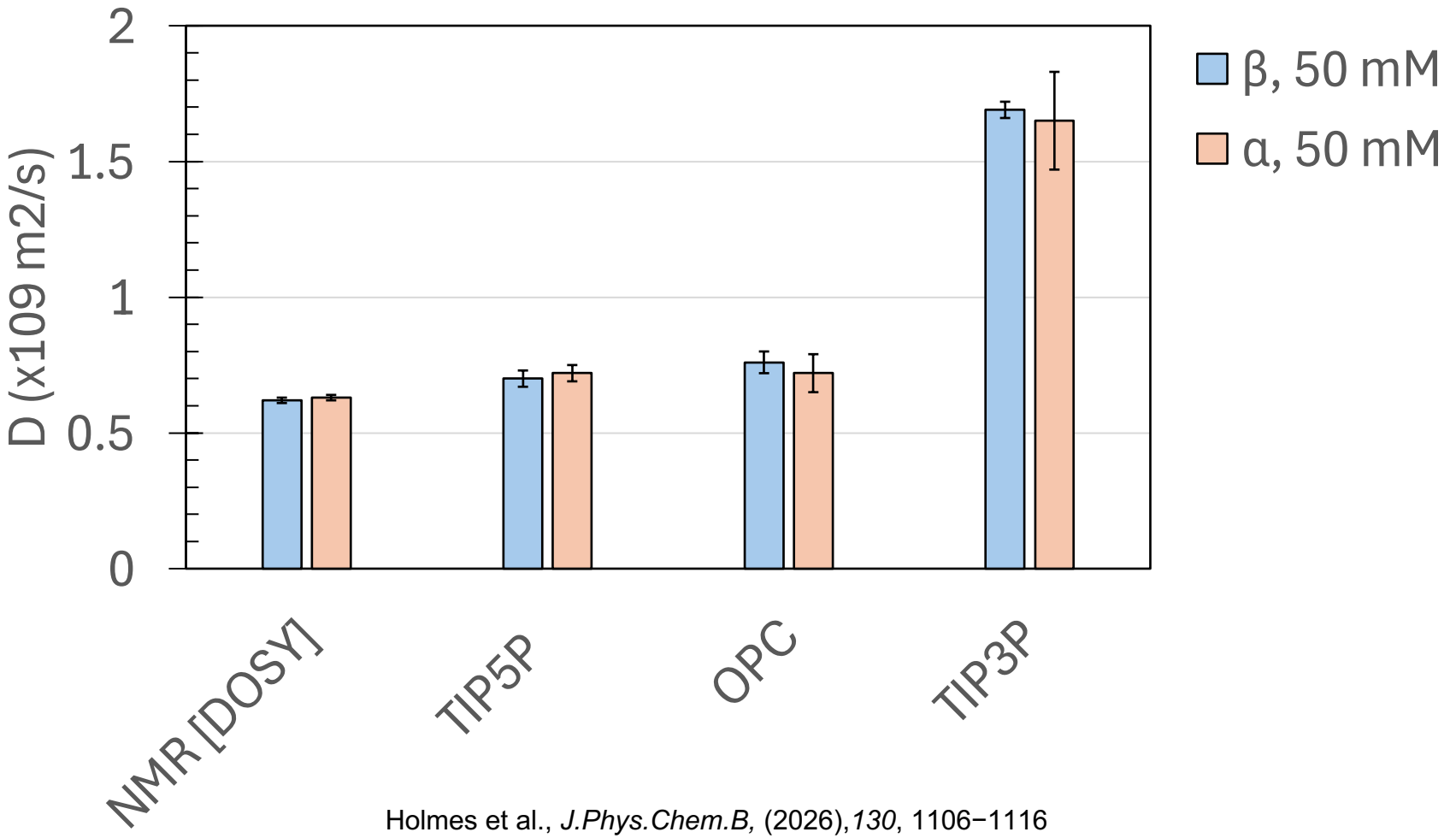
				
	QM/MD	Polarizable MD	Classical MD	Coarse Grained MD
Time Scale	< 1 ns	< 100 ns	< 10 μs	< 100 ms
System size (atoms)	< 1 K	< 10 K	< 1M	< 1M

Challenges – accuracy, system size, water, computer time (target < 1 month per simulation)

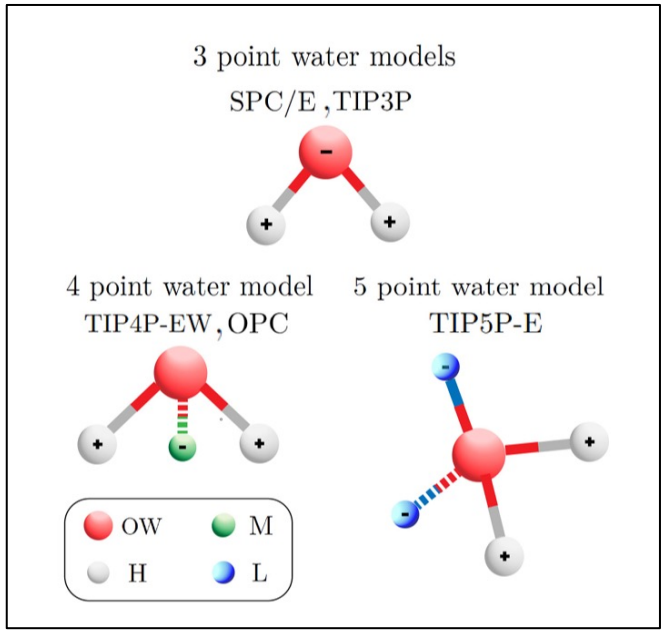
Synergy between modeling methods:

- Quantum mechanical (QM) data → guide force field development for polarizable and classical MD simulations
- Data from polarizable and classical MD simulations → guide coarse grained MD development
  → provide ensembles for QM calculations

Effect of water model on the calculated diffusion coefficient for Glucopyranose

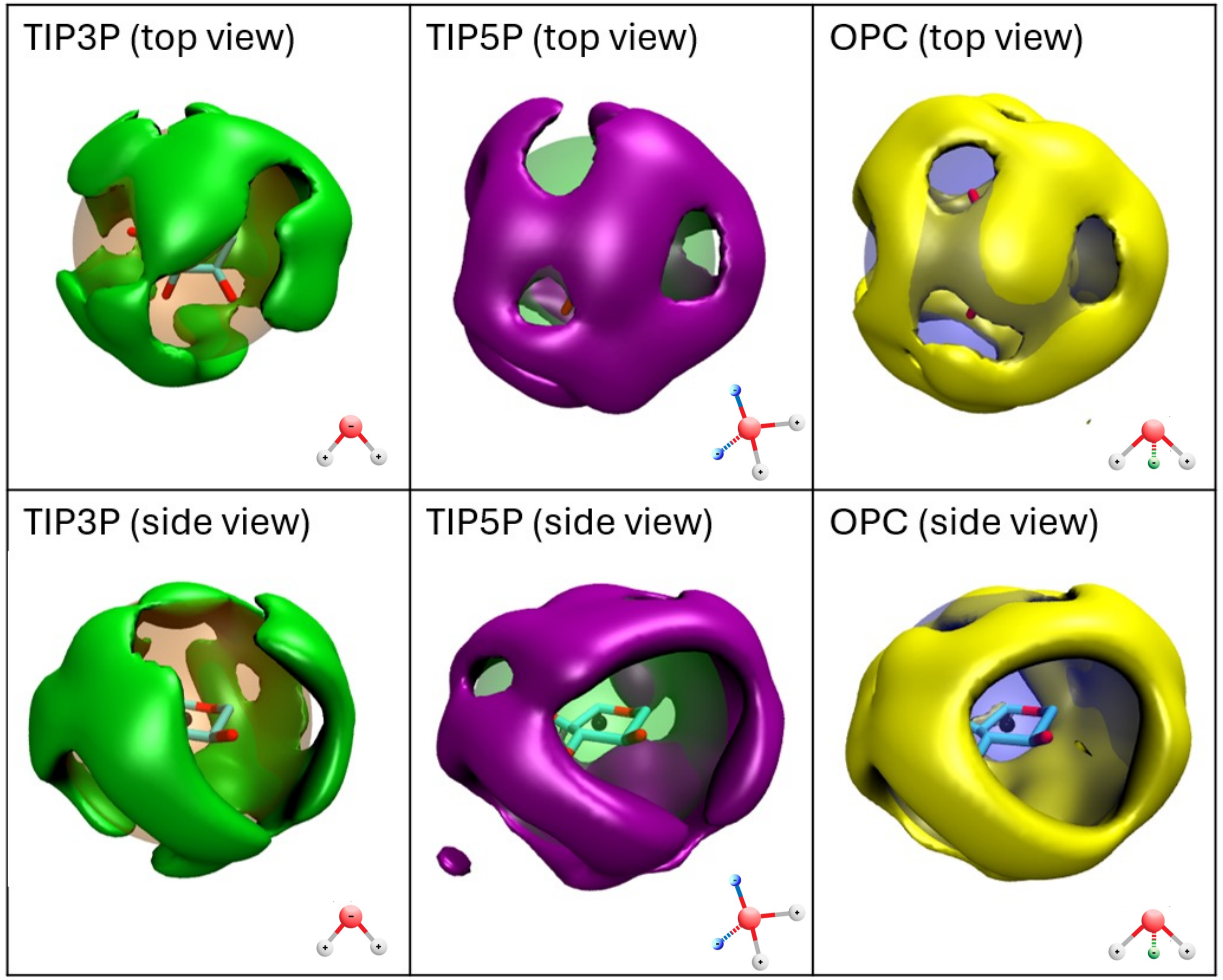


Holmes et al., *J.Phys.Chem.B,* (2026),*130*, 1106−1116



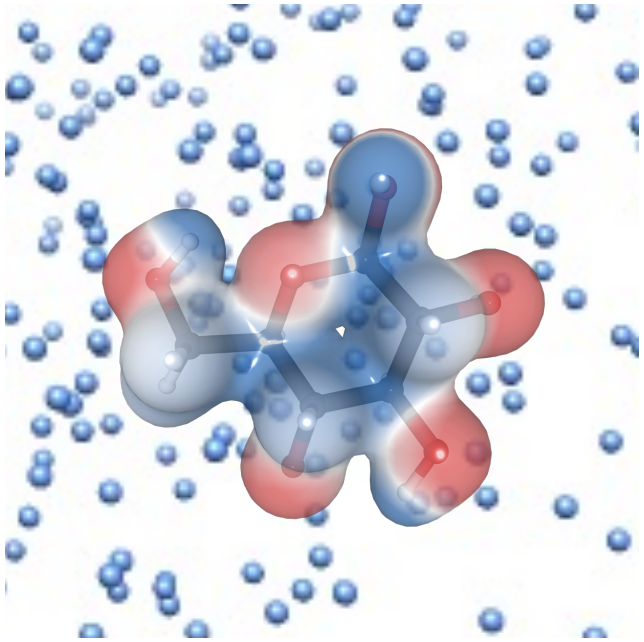
Shevy et al. *J. Chem. Phys.* (2026), *165*, 084101

Water occupancy maps for the first solvation shell

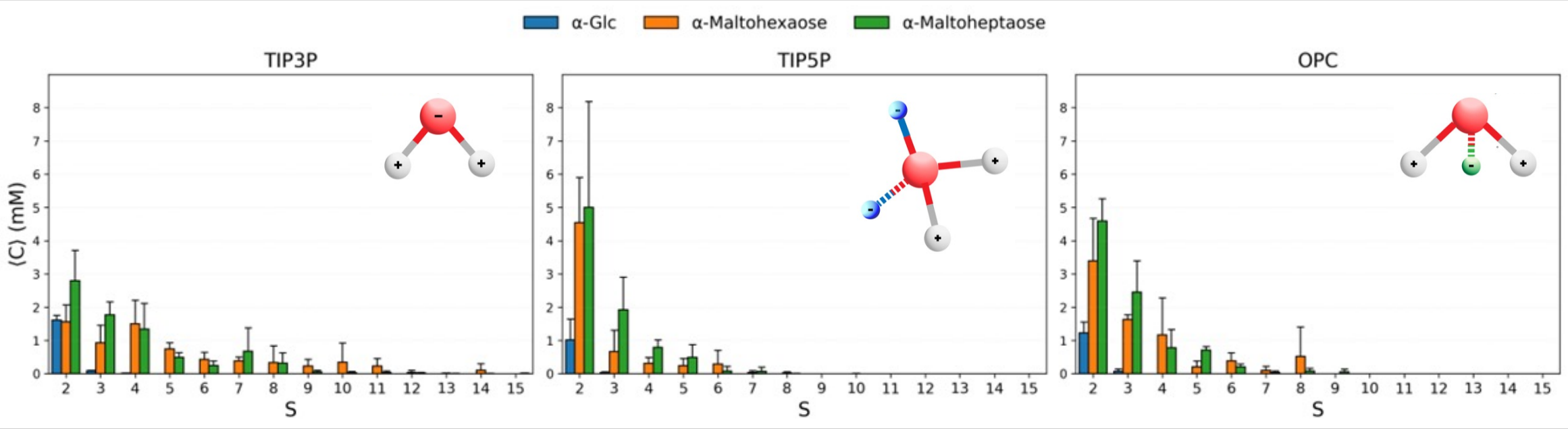


OPC and TIP5P models gave rise to comparable water structure at the carbohydrate surface in the vicinity of the first solvation shell

TIP3P interactions appeared less well-ordered



Exploring Aggregation in Carbohydrate Solutions



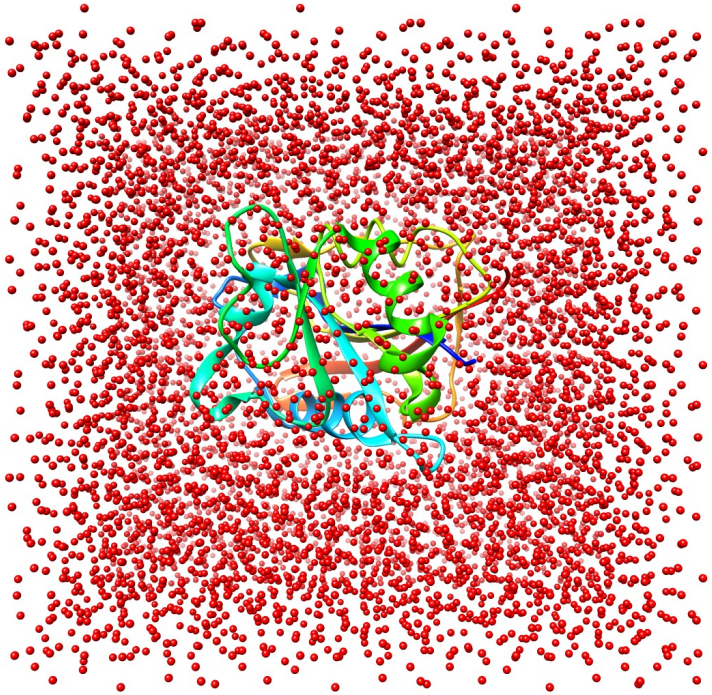
- **TIP3P**, aggregates as large as 15-mers were detected
- **TIP5P and OPC** the maximum aggregate sizes were approximately 8- and 9-mers.
- **All water models** a dimer was the predominant aggregate



$(Glc)_2$ $(Glc)_3$ $(Glc)_4$ $(Glc)_5$

Why Not Just Use a Polarizable Force Field?

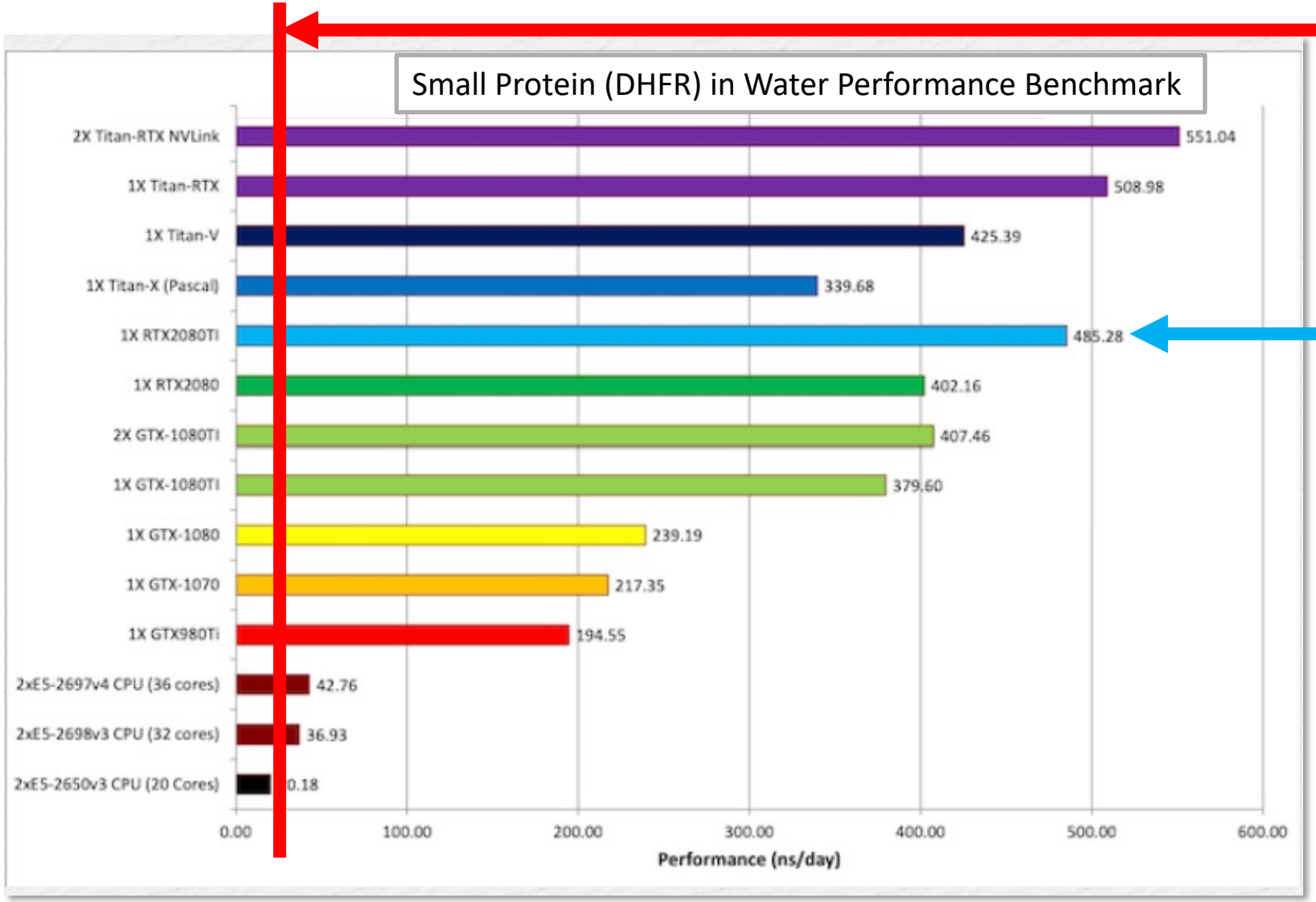
- *Small* protein test case
- human dihydrofolate reductase (DHFR)
- 159 residues
- 7279 waters
- 23,588 atoms total



Time scales of approximately 1μs are the current standard for classical (non-polarizable) MD simulations

The time to achieve 1μs varies tremendously with system size

Polarizable Force Fields are Still Far Too Slow



Tinker-HP*: 22.52 ns/day

*Adjoua, O., et al. (2021). "Tinker-HP: Accelerating Molecular Dynamics Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields Using GPUs and Multi-GPU Systems." J Chem Theory Comput **17**(4): 2034-2053.

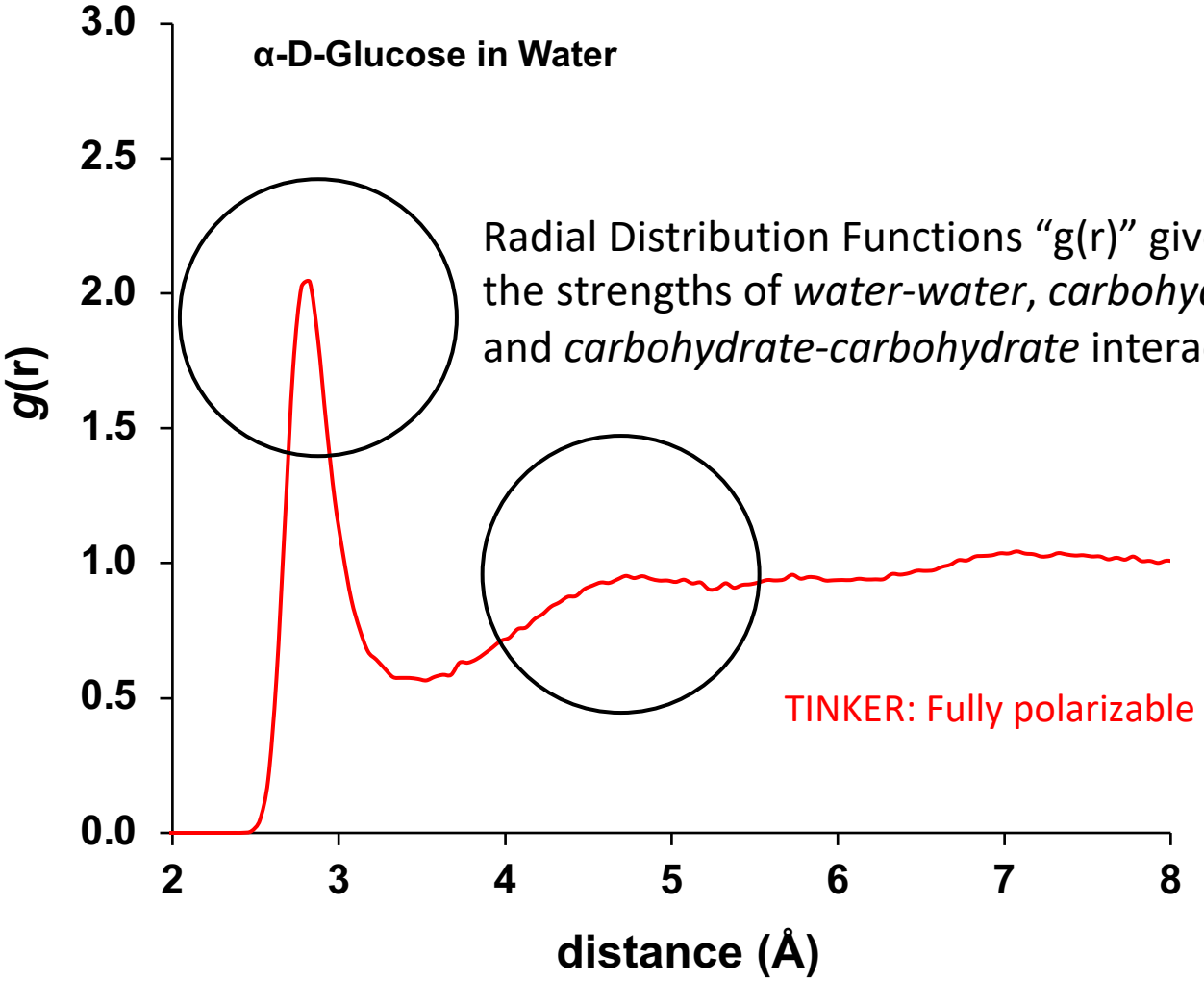
AMBER 23: 485.28 ns/day

Wall Clock time to achieve **1 µs**

- Amber: 2.1 days
- Tinker-HP: 44.4 days

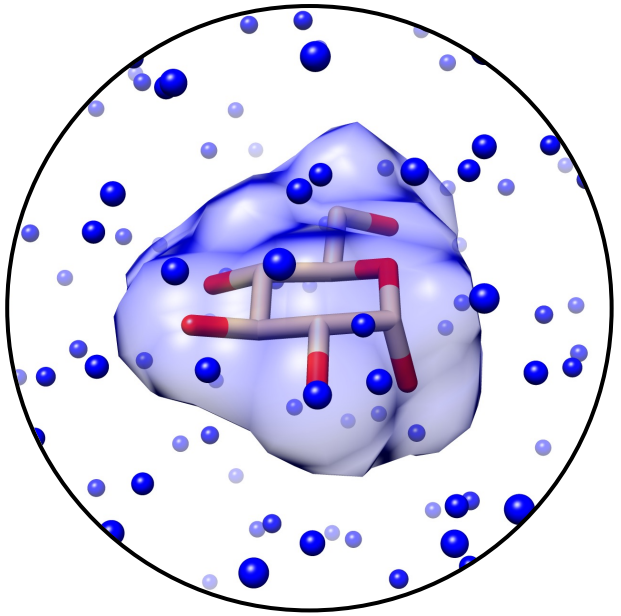
https://ambermd.org/GPUPerformance.php
AMBER23, DHFR, NPT, dT = 2fs, single GPU

Assessing the Properties of Carbohydrate-Water Interactions



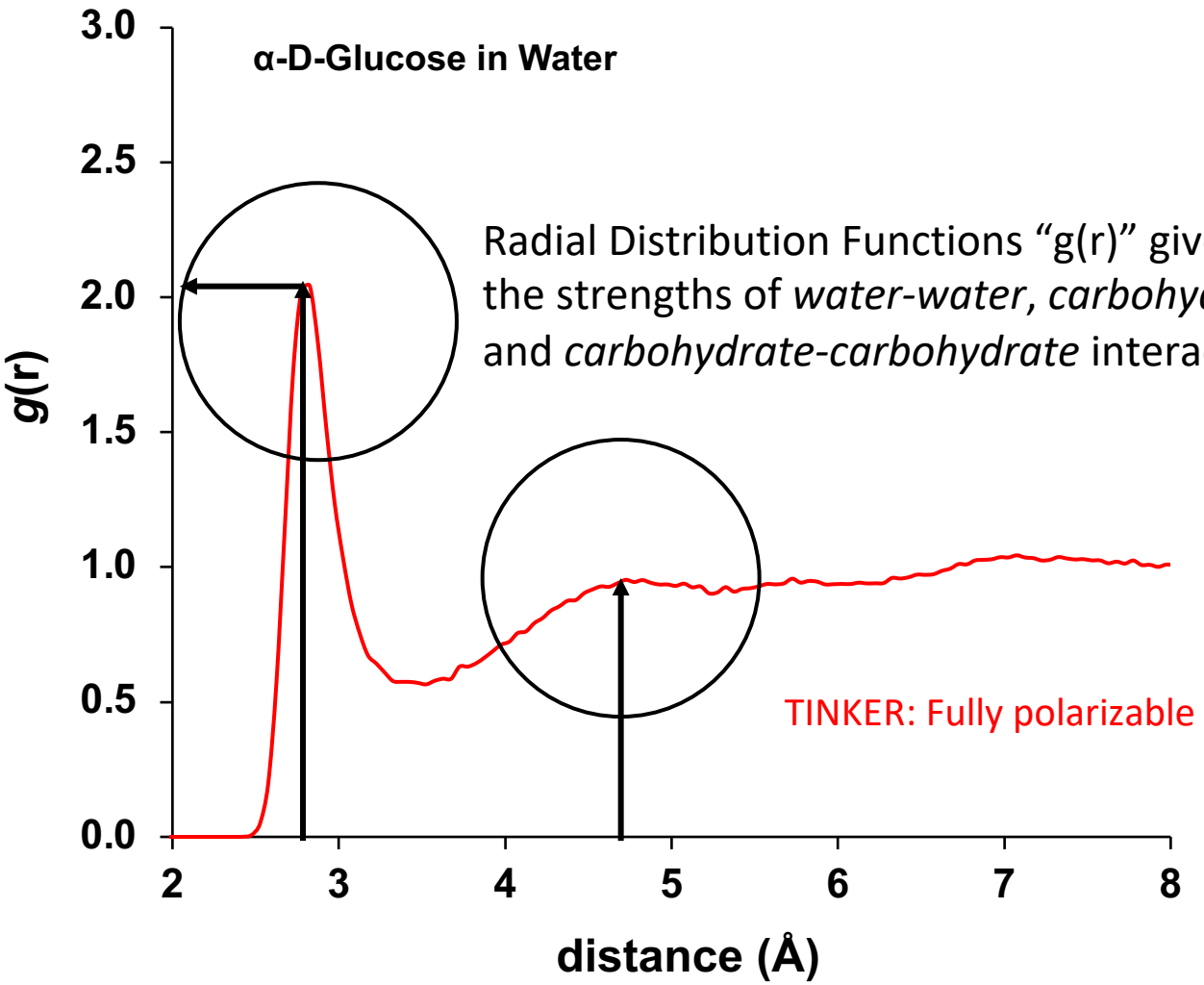
Radial Distribution Functions "g(r)" give insight into the strengths of *water-water*, *carbohydrate-water*, and *carbohydrate-carbohydrate* interactions

TINKER: Fully polarizable



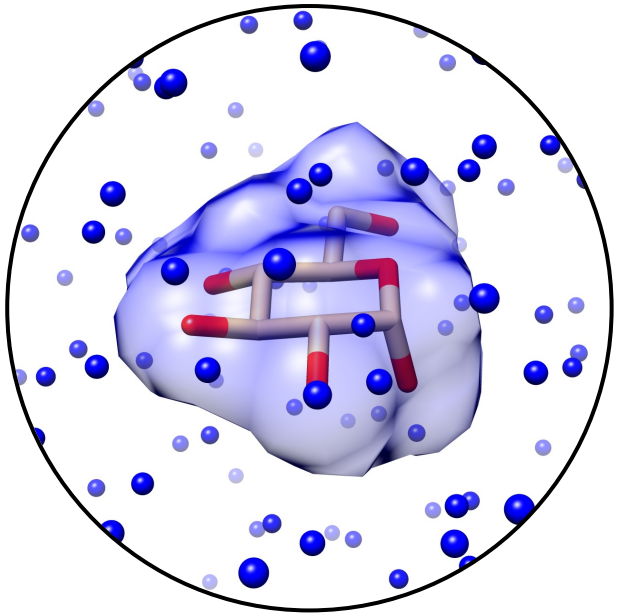
Multiscale modeling with experimental data will allow us to develop more accurate models of these large complex systems

Assessing the Properties of Carbohydrate-Water Interactions



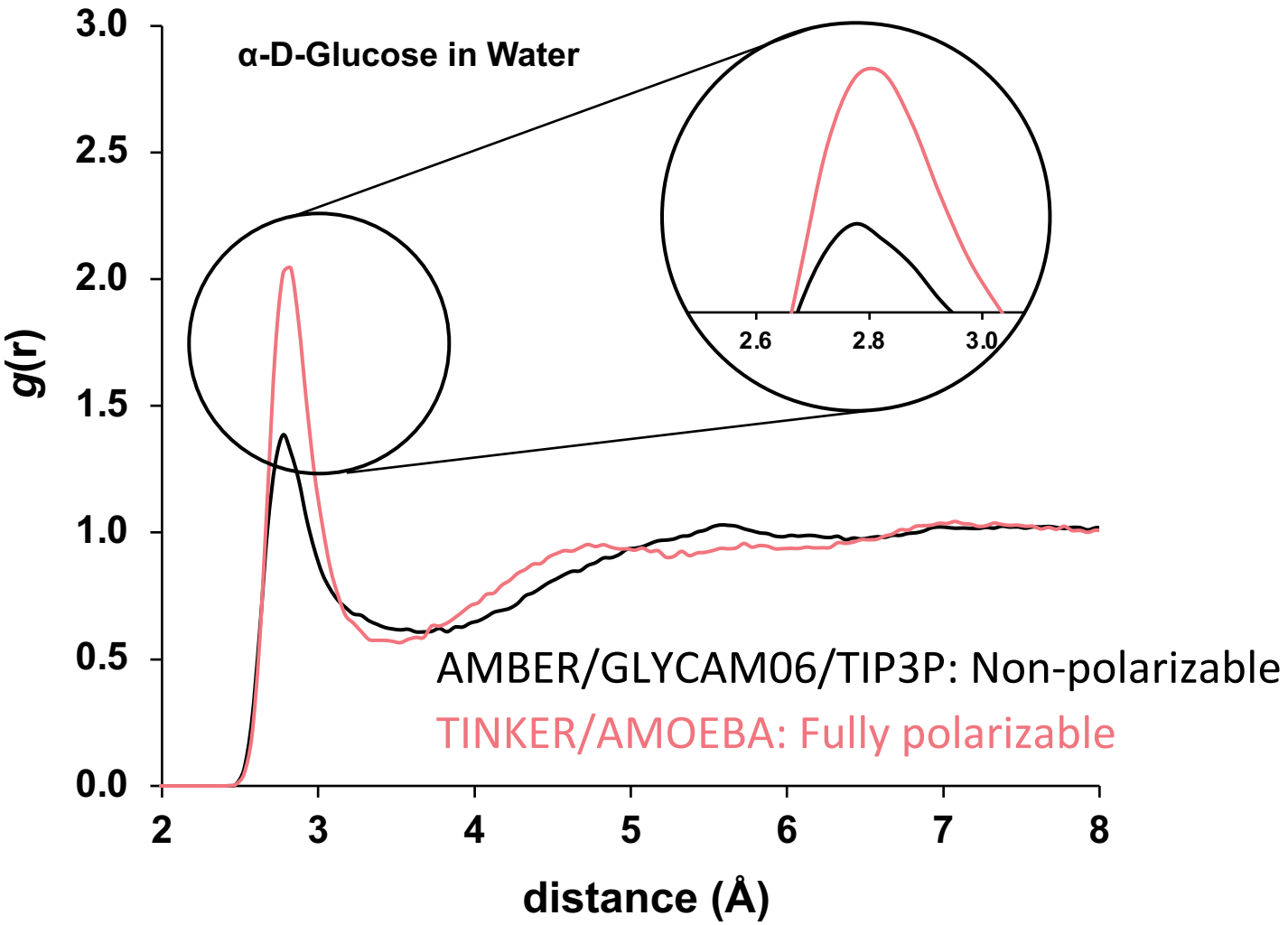
Radial Distribution Functions "g(r)" give insight into the strengths of *water-water*, *carbohydrate-water*, and *carbohydrate-carbohydrate* interactions

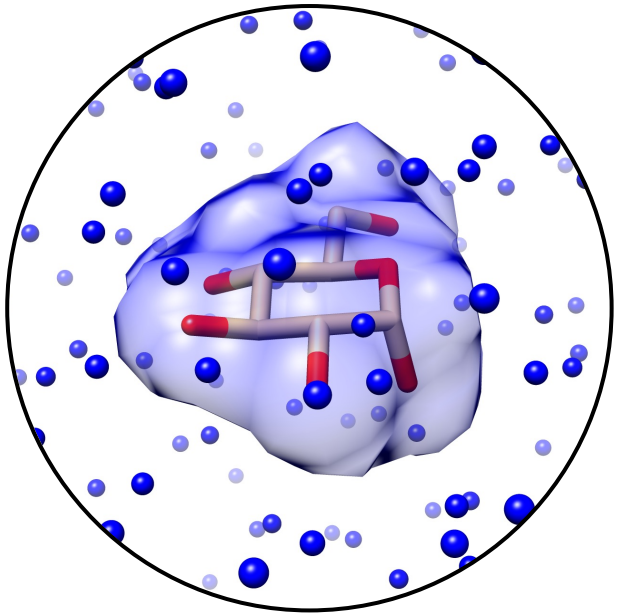
TINKER: Fully polarizable



Multiscale modeling with experimental data will allow us to develop more accurate models of these large complex systems

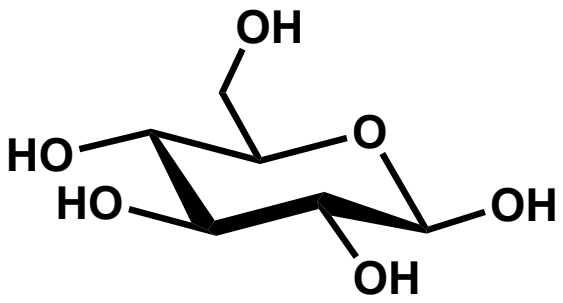
AMBER/GLYCAM (non-polarized) *vs* AMOEBA (polarizable)

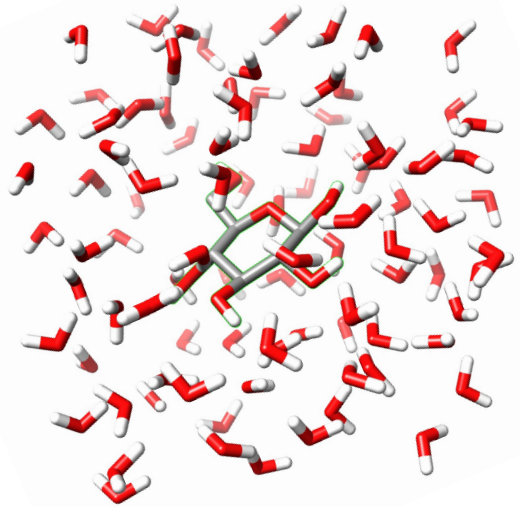




- Is there a problem with the AMBER/GLYCAM force field?
- Under-polarized?

Pre-polarized Partial Atomic Charges: Derivation



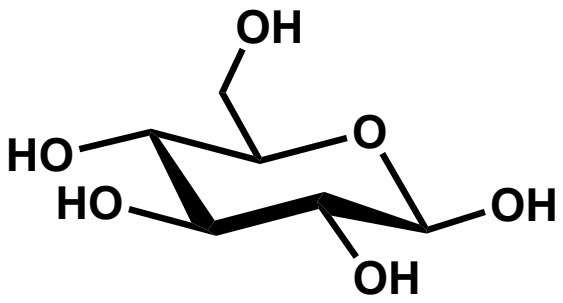


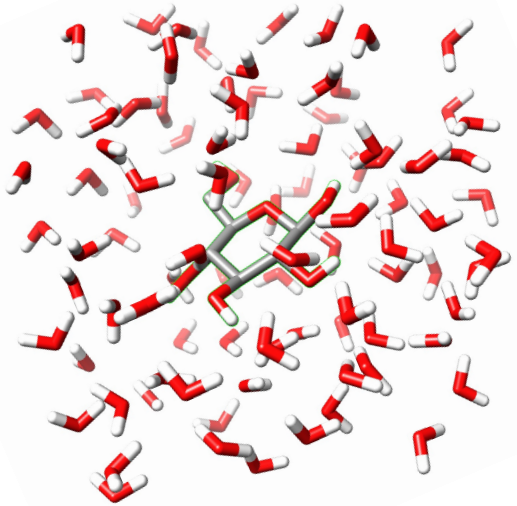
1) Compute the QM dipole moments for the monosaccharides *in vacuum*

2) Compute the QM dipole moments for the monosaccharides surrounded by explicit water molecules

3) Perform RESP charge fitting to the solvated QM electrostatic potentials

Pre-polarized Partial Atomic Charges: Derivation



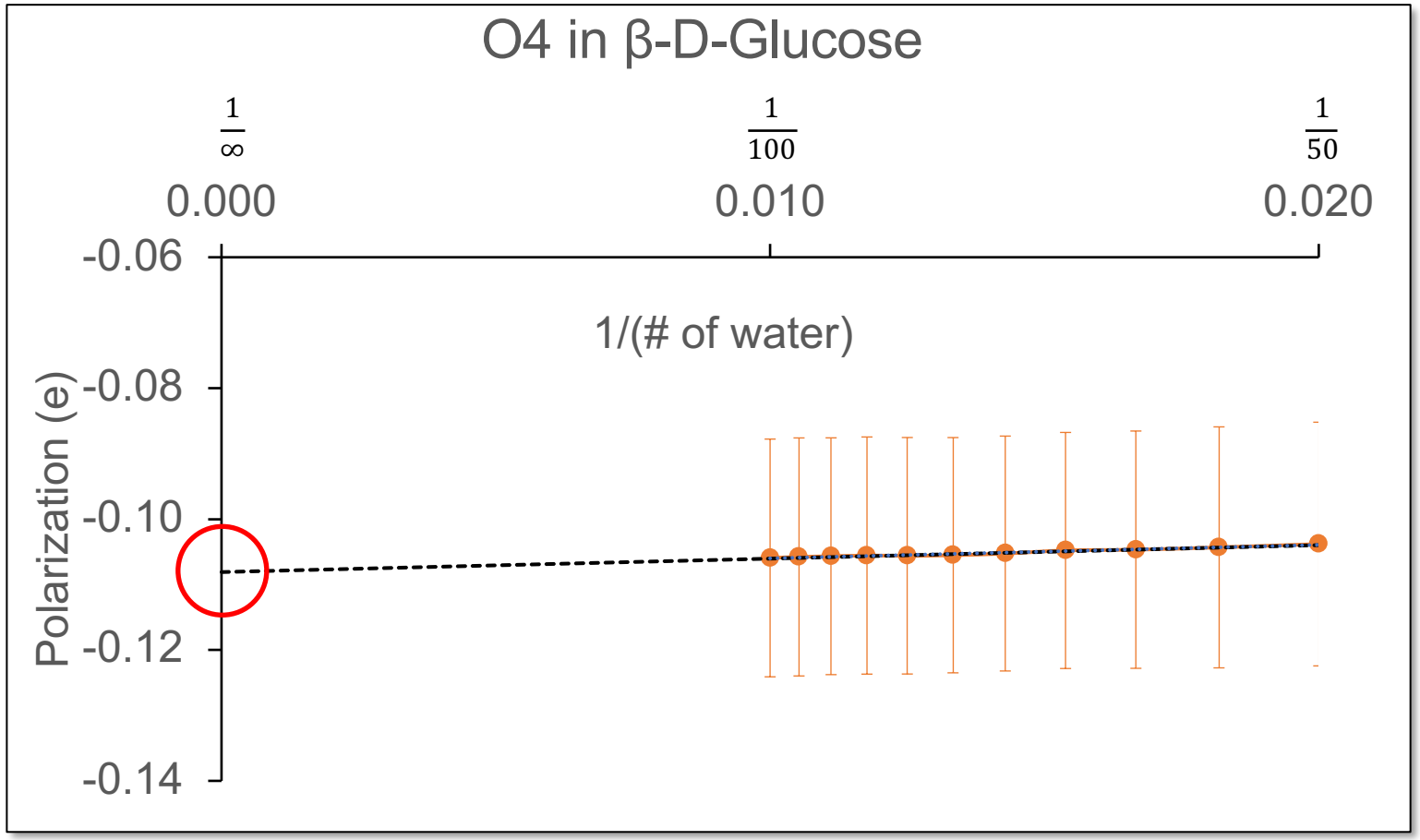


1) Compute the QM dipole moments for the monosaccharides *in vacuum*
   - Basis set dependence? → HF/6-31G*, B3LYP/cc-pVTZ
2) Compute the QM dipole moments for the monosaccharides surrounded by explicit water molecules
   - How to generate the water configurations? → Classical MD simulation
   - How many waters? → nearest 50, 60, 70, 80, 90, and 100 waters, then extrapolate to infinite dilution
   - How many configurations? → 200 MD frames per water concentration
3) Perform RESP charge fitting to the solvated QM electrostatic potentials

   RESP atomic partial charges (q_j) were derived by minimizing the error (χ^2_{esp}) between the classical Coulomb ESPs ($\hat{V}_i$), and the quantum mechanical ESPs (V_i) computed on a grid of points around the solute in the gas phase:

$$\chi^2_{\text{resp}} = \sum_i \left(V_i - \hat{V}_i\right)^2 + a \sum_j \left(q_j\right)^2$$

Estimating Charge Polarization at Infinite Dilution



Interpolate the dipole and ESP charges to infinite dilution

Each data point is an average from 200 MD frames evenly extracted from a 20 ns MD simulation.

Water Significantly Enhances Monosaccharide Dipoles

Table 1. SCF dipole moments[a,b] for monosaccharides with and without explicit water electrostatics effects.

Water Model	Gas Phase	Infinite Dilution	Enhancement	Gas Phase	Infinite Dilution	Enhancement
	α-Gal			**β-Gal**		
TIP3P	2.95 ± 0.02	4.26 ± 0.06	44%	3.48 ± 0.11	4.94 ± 0.16	42%
OPC	3.02 ± 0.11	4.48 ± 0.17	48%	3.37 ± 0.04	4.83 ± 0.06	43%
	α-Glc			**β-Glc**		
TIP3P	2.57 ± 0.03	3.78 ± 0.03	47%	2.79 ± 0.03	4.02 ± 0.04	44%
OPC	2.60 ± 0.00	3.95 ± 0.05	52%	2.86 ± 0.09	4.27 ± 0.21	49%
	α-Man			**β-Man**		
TIP3P	2.55 ± 0.08	3.73 ± 0.10	46%	2.87 ± 0.05	4.20 ± 0.01	46%
OPC	2.63 ± 0.06	3.99 ± 0.06	52%	2.88 ± 0.04	4.27 ± 0.05	48%
	α-Xyl			**β-Xyl**		
TIP3P	2.29 ± 0.03	3.35 ± 0.09	46%	2.36 ± 0.09	3.44 ± 0.15	46%
OPC	2.29 ± 0.13	3.45 ± 0.20	50%	2.32 ± 0.03	3.52 ± 0.06	52%

[a]Computed at the B3LYP/cc-pVTZ level of theory. [b]Standard deviations from three independent MD simulations.

Water:
- dipole increases from 1.85 D to approximately 2.9 D going from gas phase to liquid.
- 36% enhancement.

Methanol:
- dipole increases from 1.70 D to 2.6 – 2.7 D going from gas phase to liquid.
- 36% enhancement.

Dipole enhancement:
- mean-field polarization from the bulk dielectric of the solvent
- local polarization through direct hydrogen bond formation.

Water Significantly Enhances Monosaccharide Dipoles

Table 1. SCF dipole moments[a,b] for monosaccharides with and without explicit water electrostatic effects.

Water Model	Gas Phase	Infinite Dilution	Enhancement	Gas Phase	Infinite Dilution	Enhancement
	α-Gal			**β-Gal**		
TIP3P	2.95 ± 0.02	4.26 ± 0.06	44%	3.48 ± 0.11	4.94 ± 0.16	42%
OPC	3.02 ± 0.11	4.48 ± 0.17	48%	3.37 ± 0.04	4.83 ± 0.06	43%
	α-Glc			**β-Glc**		
TIP3P	2.57 ± 0.03	3.78 ± 0.03	47%	2.79 ± 0.03	4.02 ± 0.04	44%
OPC	2.60 ± 0.00	3.95 ± 0.05	52%	2.86 ± 0.09	4.27 ± 0.21	49%
	α-Man			**β-Man**		
TIP3P	2.55 ± 0.08	3.73 ± 0.10	46%	2.87 ± 0.05	4.20 ± 0.01	46%
OPC	2.63 ± 0.06	3.99 ± 0.06	52%	2.88 ± 0.04	4.27 ± 0.05	48%
	α-Xyl			**β-Xyl**		
TIP3P	2.29 ± 0.03	3.35 ± 0.09	46%	2.36 ± 0.09	3.44 ± 0.15	46%
OPC	2.29 ± 0.13	3.45 ± 0.20	50%	2.32 ± 0.03	3.52 ± 0.06	52%

[a]Computed at the B3LYP/cc-pVTZ level of theory. [b]Standard deviations from three independent MD simulations.

Water:
- dipole increases from 1.85 D to approximately 2.9 D going from gas phase to liquid.
- 36% enhancement.

Methanol:
- dipole increases from 1.70 D to 2.6 – 2.7 D going from gas phase to liquid.
- 36% enhancement.

Dipole enhancement:
- mean-field polarization from the bulk dielectric of the solvent
- local polarization through direct hydrogen bond formation.

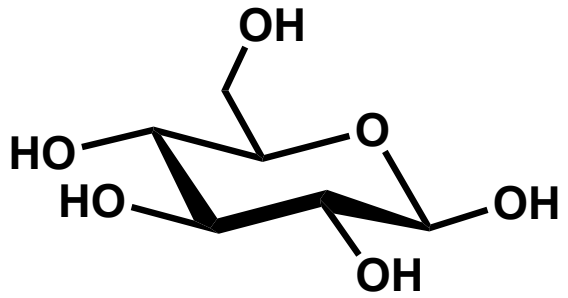
Experimental Data: Gal: 4.6 - 5.3 D
Glc: 3.8 - 4.9 D
Man: 4.8 D

Water Significantly Enhances RESP Atomic Charges

Table 2. Atomic partial charges[a] for α-Gal computed with and without explicit water electrostatics for the TIP3P and OPC water models.

Atom-type RESP Charges						
Atom	**TIP3P**			**OPC**		
	Gas Phase	**Infinite Dilution**	**Enhancement**	**Gas Phase**	**Infinite Dilution**	**Enhancement**
HO	0.392 ± 0.003	0.480 ± 0.007	22%	0.391 ± 0.003	0.488 ± 0.004	25%
OH	-0.631 ± 0.008	-0.761 ± 0.009	21%	-0.627 ± 0.007	-0.774 ± 0.009	23%
C	0.285 ± 0.027	0.336 ± 0.017	18%	0.283 ± 0.019	0.341 ± 0.019	21%
O5	-0.518 ± 0.005	-0.612 ± 0.007	18%	-0.515 ± 0.005	-0.619 ± 0.006	20%

[a]From RESP fitting to SCF ESP values computed at the B3LYP/cc-pVTZ level of theory.



Charge Polarization Weakly Affects Monosaccharide Diffusion

Table 4. Diffusion coefficients[a] computed from MD simulations performed in the TIP3P, OPC, and TIP5P water models with and without monosaccharide charge polarization.

Monosaccharide Charge Set	MD Water Model					
	OPC	TIP3P	TIP5P	OPC	TIP3P	TIP5P
	α-Gal			β-Gal		
GLYCAM06	0.72 ± 0.02	1.76 ± 0.02	0.73 ± 0.01	0.72 ± 0.04	1.79 ± 0.02	0.68 ± 0.03
Gas-phase (TIP3P Ensemble)[b]	0.78 ± 0.07	1.84 ± 0.03	0.76 ± 0.02	0.80 ± 0.04	1.88 ± 0.03	0.79 ± 0.04
Gas-phase (OPC Ensemble)[c]	0.77 ± 0.04	1.76 ± 0.05	0.79 ± 0.04	0.79 ± 0.06	1.86 ± 0.05	0.80 ± 0.02
Polarized (TIP3P Ensemble)[d]	0.68 ± 0.03	1.57 ± 0.02	0.69 ± 0.03	0.64 ± 0.03	1.57 ± 0.11	0.67 ± 0.04
Polarized (OPC Ensemble)[e]	0.68 ± 0.07	1.50 ± 0.07	0.64 ± 0.04	0.67 ± 0.03	1.50 ± 0.11	0.68 ± 0.03

[a] 10^{-9} m^2/s, at 50 mM. [b] Solute coordinates extracted from a solvated MD simulation in TIP3P water. [c] Solute coordinates extracted from a solvated MD simulation in OPC water. [d] Solute coordinates and solvent partial charge positions extracted from a solvated MD simulation in TIP3P water. [e] Solute coordinates and solvent partial charge positions extracted from a solvated MD simulation in OPC water.

Gas-phase and pre-polarized charge models show very similar diffusion rates for monosaccharides in water

Theoretical Water Model Strongly Affects Monosaccharide Diffusion

Diffusion coefficients[a] computed from MD simulations performed in the TIP3P, OPC, and TIP5P water models with and without monosaccharide charge polarization.

Monosaccharide Charge Set	MD Water Model					
	OPC	TIP3P	TIP5P	OPC	TIP3P	TIP5P
	α-Gal			β-Gal		
GLYCAM06	0.72 ± 0.02	1.76 ± 0.02	0.73 ± 0.01	0.72 ± 0.04	1.79 ± 0.02	0.68 ± 0.03
Gas-phase (TIP3P Ensemble)[b]	0.78 ± 0.07	1.84 ± 0.03	0.76 ± 0.02	0.80 ± 0.04	1.88 ± 0.03	0.79 ± 0.04
Gas-phase (OPC Ensemble)[c]	0.77 ± 0.04	1.76 ± 0.05	0.79 ± 0.04	0.79 ± 0.06	1.86 ± 0.05	0.80 ± 0.02
Polarized (TIP3P Ensemble)[d]	0.68 ± 0.03	1.57 ± 0.02	0.69 ± 0.03	0.64 ± 0.03	1.57 ± 0.11	0.67 ± 0.04
Polarized (OPC Ensemble)[e]	0.68 ± 0.07	1.50 ± 0.07	0.64 ± 0.04	0.67 ± 0.03	1.50 ± 0.11	0.68 ± 0.03

[a] 10^{-9} m^2/s, at 50 mM. [b] Solute coordinates extracted from a solvated MD simulation in TIP3P water. [c] Solute coordinates extracted from a solvated MD simulation in OPC water. [d] Solute coordinates and solvent partial charge positions extracted from a solvated MD simulation in TIP3P water. [e] Solute coordinates and solvent partial charge positions extracted from a solvated MD simulation in OPC water.

Theoretical water models heavily impact monosaccharide diffusion rates

Theoretical Water Model and Monosaccharide Charge Model Strongly Affect Solute Aggregation

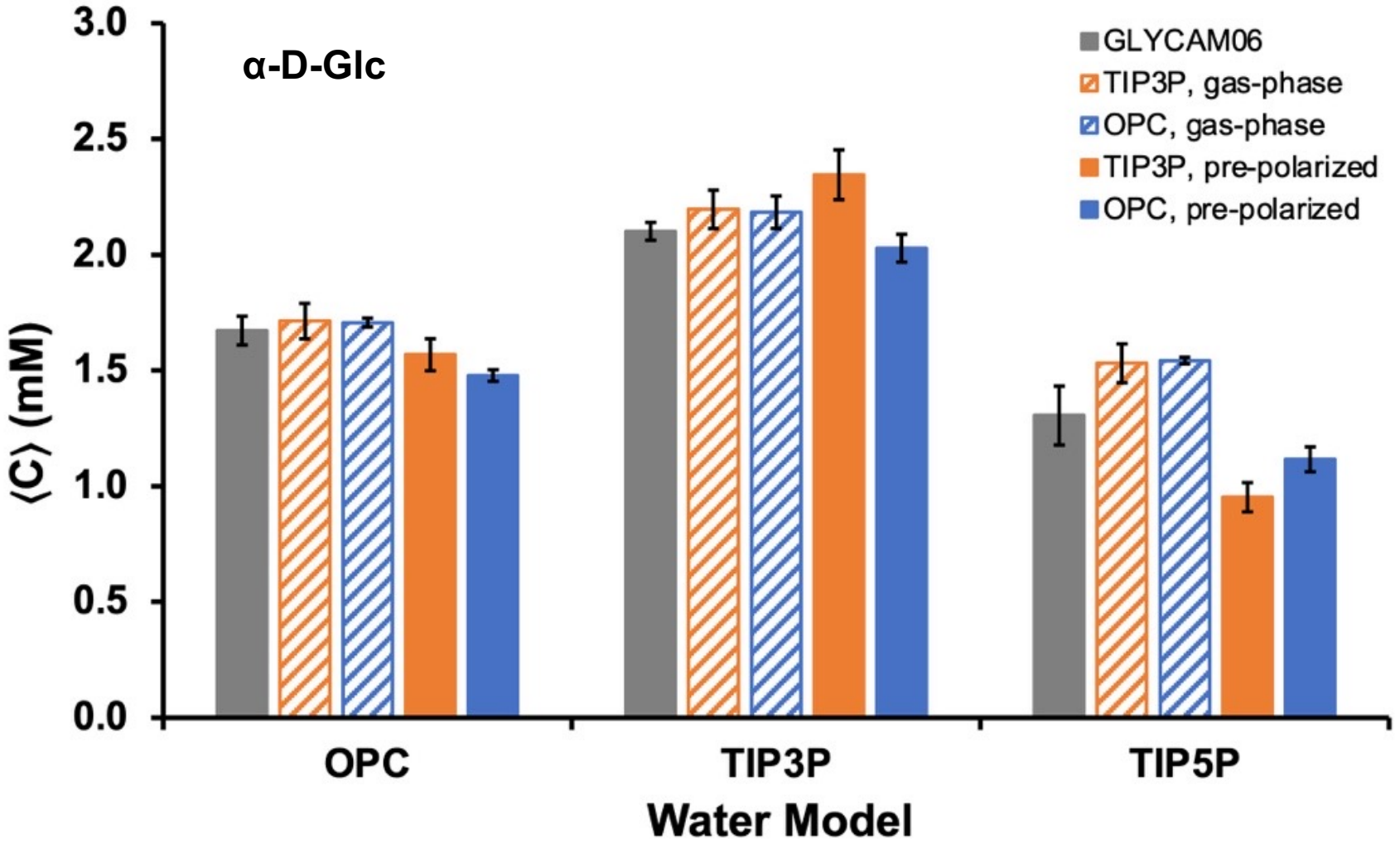
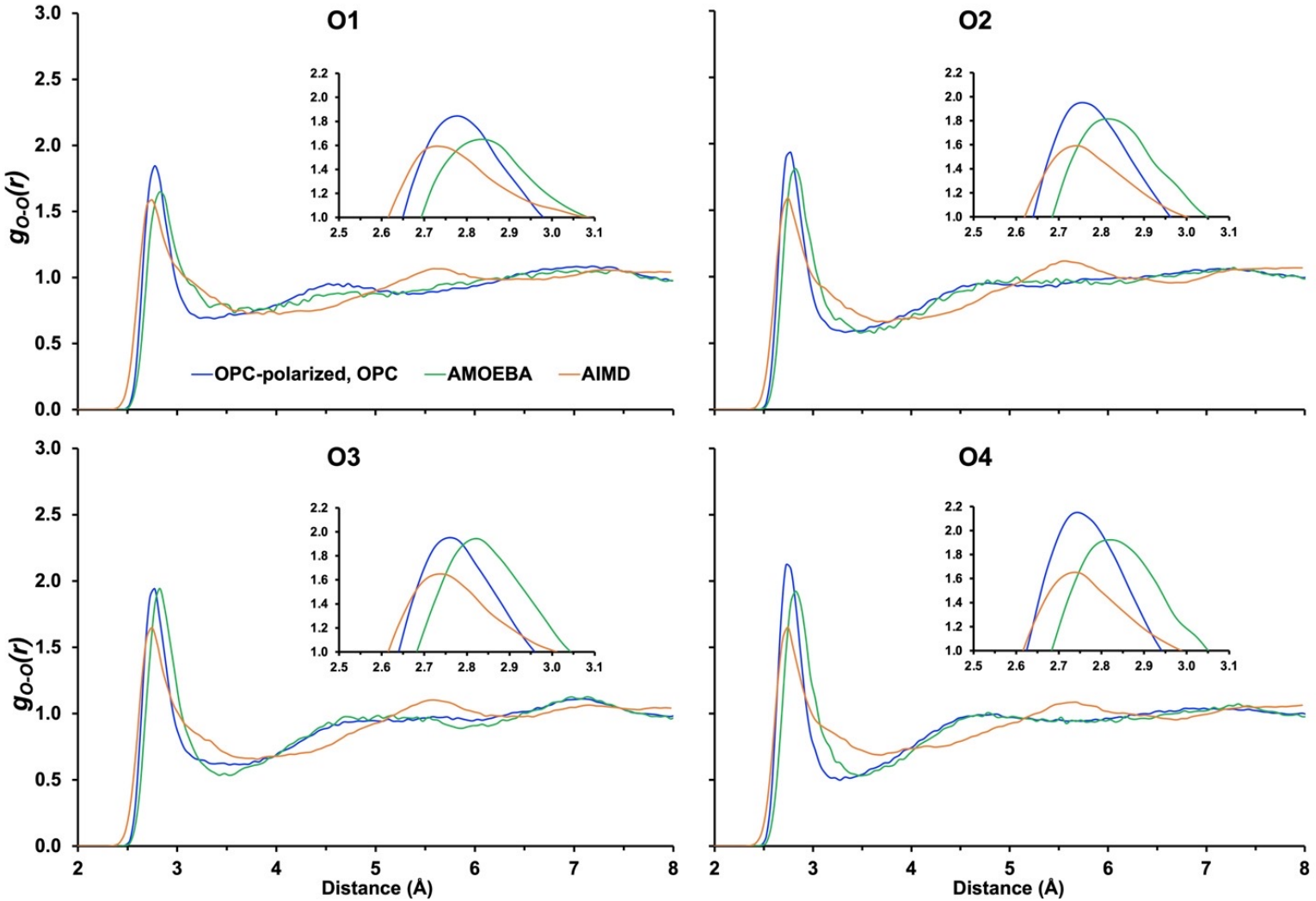
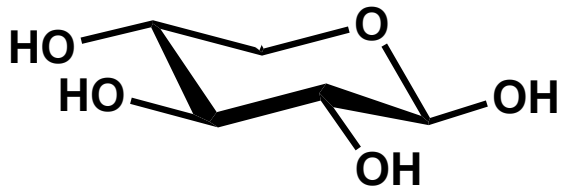


Figure 1. Concentration (mM) of dimers of α-Glc in 50 mM solutions, for three water models and five charge models. Averages and standard deviations were derived from three independent simulations under each condition.

Radial Distribution Functions





MD Configurations:
Pre-Polarized (AMBER):

- 439 waters
- 3 x 20,000 ps
- OPC water model

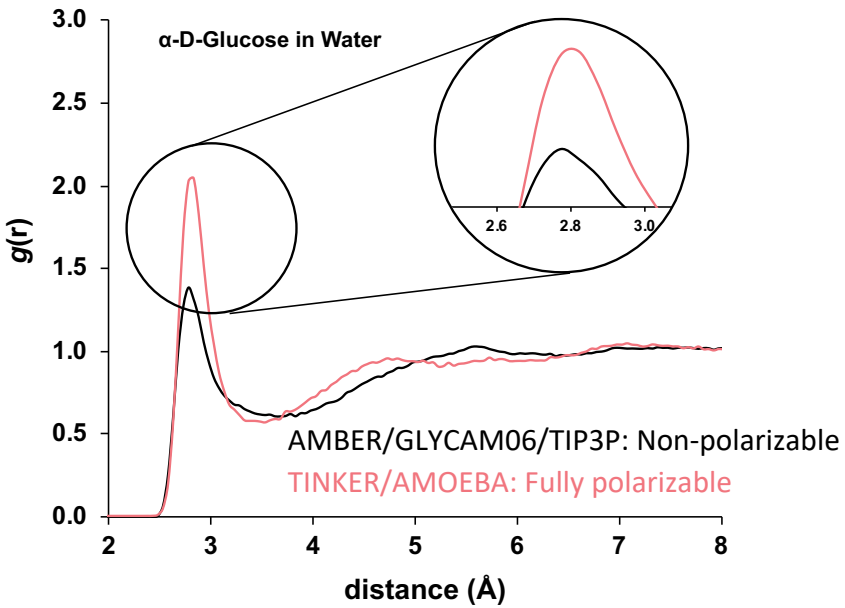
Fully Polarized (AMOEBA):

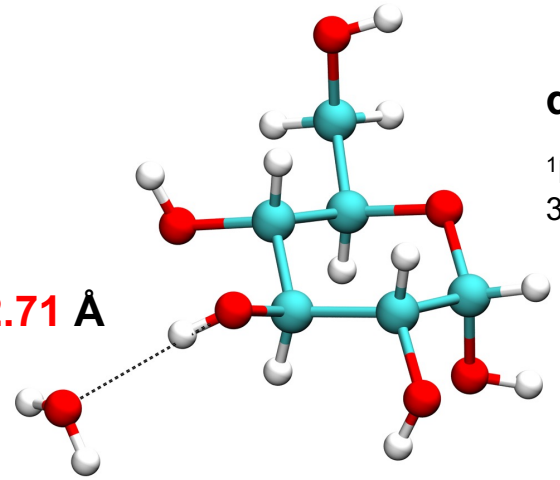
- 889 waters
- 1 x 30,000 ps
- AMOEBA water model

QM (AIMD):

- 238 waters
- 10 x 100 ps
- CP2K for forces and LAMMPS for MD
- SCC-DFTB basis set

Water-Sugar Distances in Monosaccharide Crystals





R_{OO} = 2.71 Å

α-D-Glucose monohydrate[1]

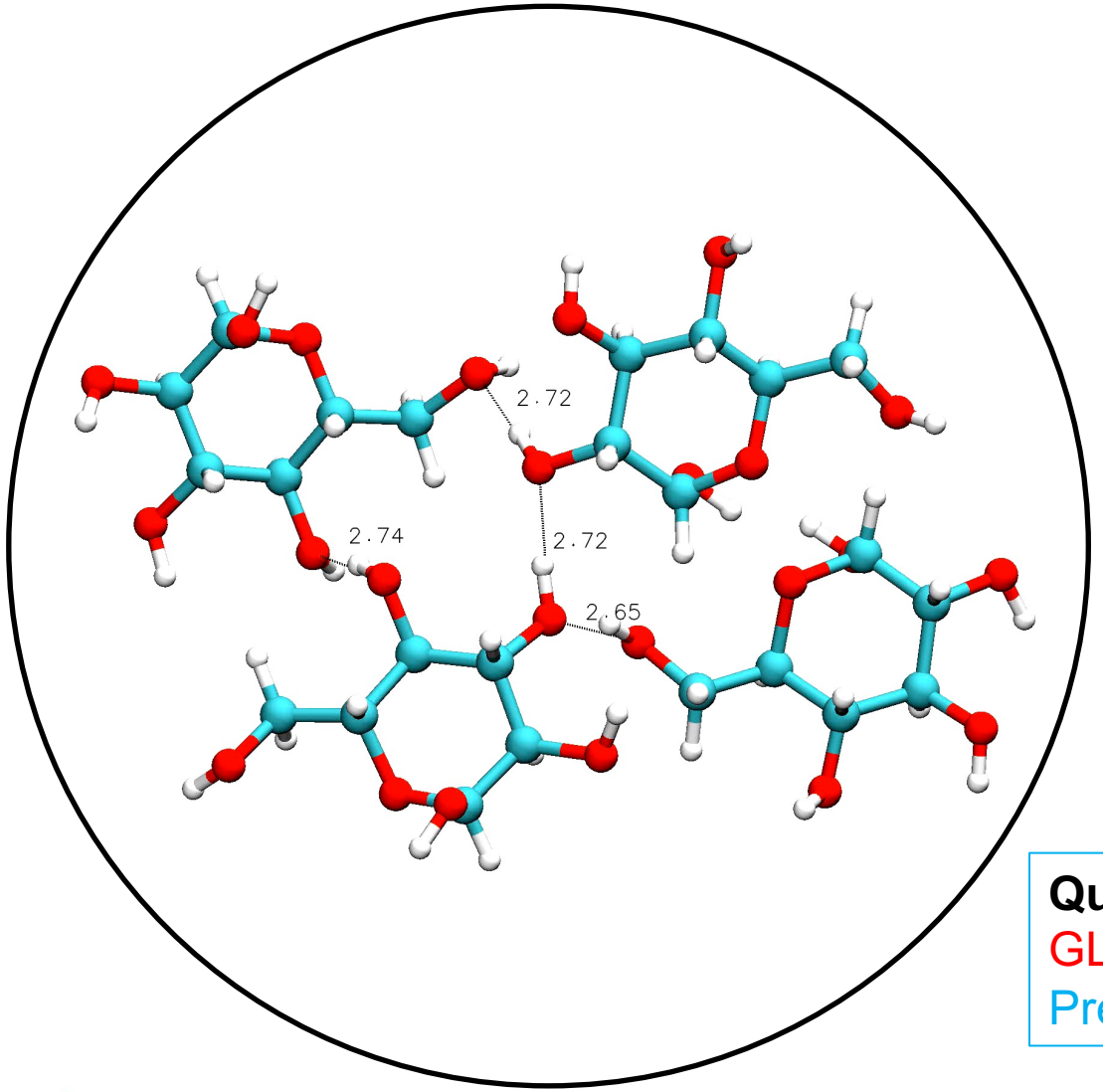
[1]Hough, E. et al., *Acta Cryst. B,* (1973), *29*, 365-367.

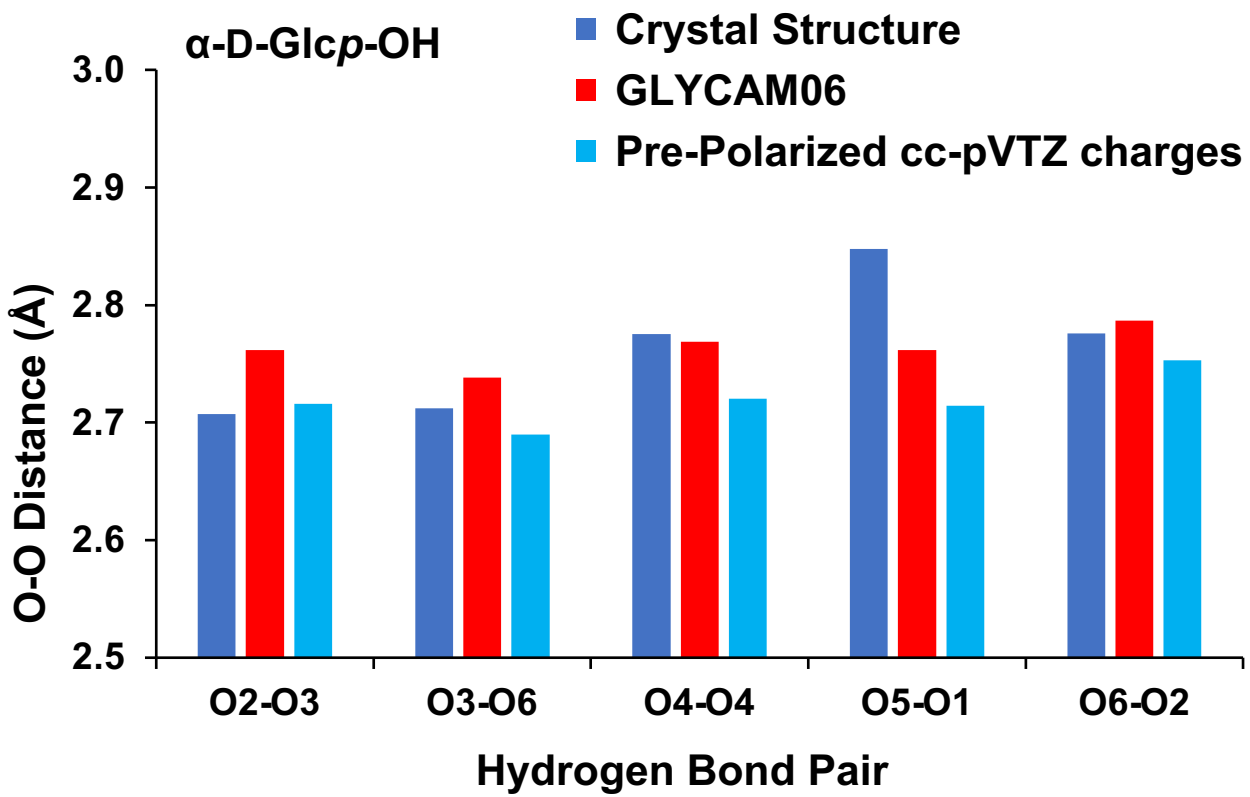
First hydration shell properties for hydroxyl oxygen atoms in β-Xyl.

Atom		O1			O2			O3			O4	
Method	g_1^{max}	r_1^{max},a	N_w	g_1^{max}	r_1^{max}	N_w	g_1^{max}	r_1^{max}	N_w	g_1^{max}	r_1^{max}	N_w
Pre-polarized[b]	1.84	2.78	3.5	1.93	2.78	3.3	1.94	2.78	3.4	2.12	2.73	3.3
AMOEBA	1.65	2.83	3.6	1.81	2.83	3.4	1.94	2.83	3.4	1.92	2.83	3.4
AIMD	1.59	2.75	3.8	1.59	2.75	3.6	1.65	2.75	3.6	1.65	2.75	3.6

[a]In Å. [b]In OPC water.

Condensed Phase Crystal Simulations





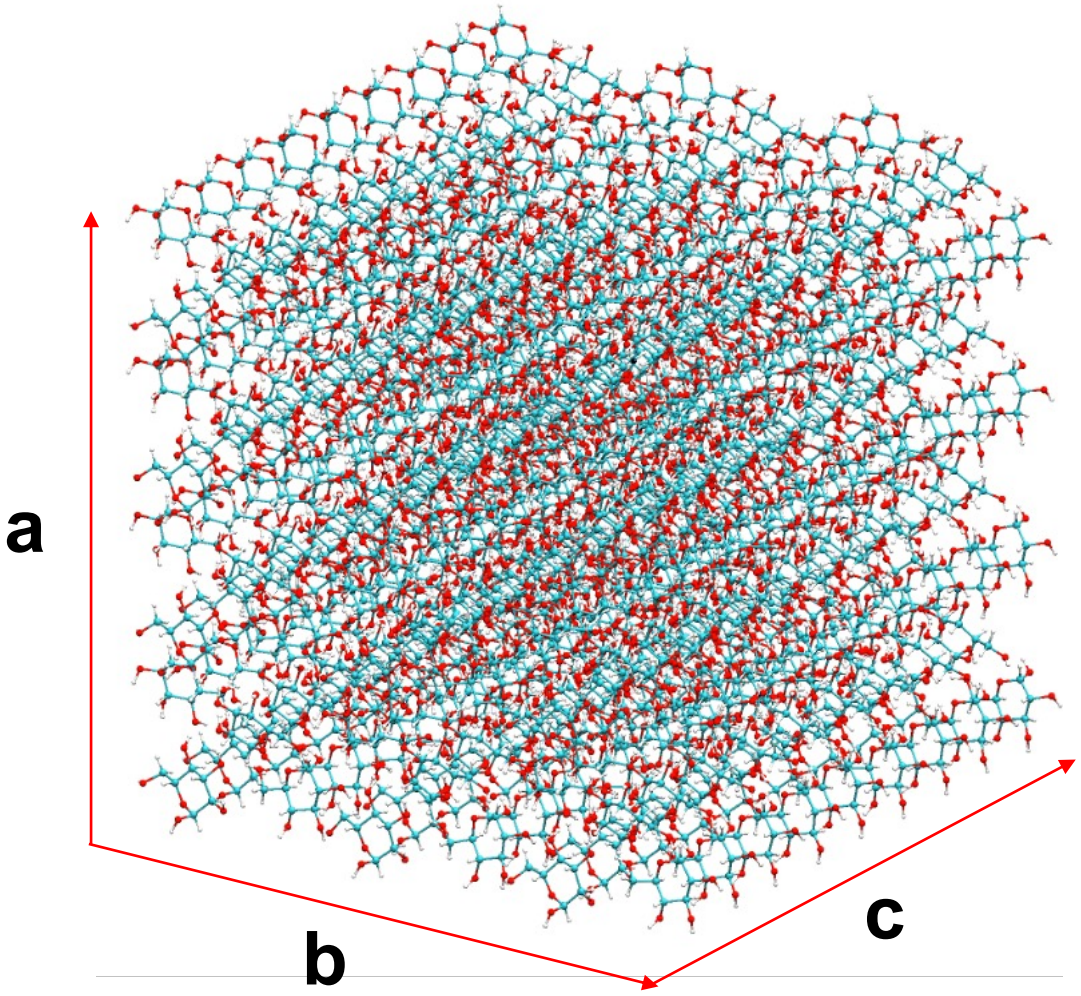
Qualitative conclusions:

GLYCAM06 – slightly over-estimates O-O distances

Pre-polarized cc-pVTZ slightly underestimates O-O distances

GlycoMIP
NSF Materials Innovation Platform

Condensed phase MD simulations



Crystal dimensions (a, b, c) were stable with both non-polarized and pre-polarized charge sets.

α-D-Glucose

Dimensions (Å)	Expt.	GLYCAM06	HF-6-31G*, Pre-Polarized
a	44.6	45.3	44.8
b	41.5	42.1	41.6
c	44.8	44.3	44.1
Volume (Å^3)	**82,719**	**84,636**	**82,248**

Qualitative conclusions:

GLYCAM06 – slightly over-estimates unit cell dimensions

Pre-polarized HF-6-31G* looks good

Pre-polarized cc-pVTZ in progress

Conclusions

Solvent-induced polarization:

- enhanced monosaccharide dipoles by 42-52%,
- resulted in values that were generally consistent with experiment.

Classical MD simulations with pre-polarized atomic charges:

- reproduced the QM SCF dipole moments to within 2.2%.
- led to quantitative agreement with experimental diffusion coefficients.
- not recommended for use TIP3P water, due primarily to its unrealistically low viscosity.
- RDFs agreed well with the results from fully polarizable MD simulations with the AMOEBA force field.
- appear to correct weaknesses previously observed for carbohydrate-carbohydrate interactions.

Next:

- Examine performance in simulations of carbohydrate-protein complexes
- Release new GLYCAM force field (GLYCAM26)

Acknowledgments

University of Georgia:

Xiaocong Wang – QM calculations and MD simulations, GLYCAM-Hybrid force field

Samuel Holmes – Diffusion simulations and NMR Measurements

Virginia Tech:

Valerie Welborn, Luke Newman – Fully-polarizable (AMOEBA) MD simulations

Daniel Crawford, Jessica Nash, Taylor Barnes – Quantum MD (AIMD) simulations

This work was supported by GlycoMIP, a National Science Foundation Materials Innovation Platform funded through Cooperative Agreement DMR-1933525.