Comparison of Three Ionic Liquid Tolerant Cellulases by Molecular Dynamics

Vance Jaeger, Patrick Burney, and Jim Pfaendtner*

Department of Chemical Engineering, University of Washington, Seattle, Washington 98105, United States

*Correspondence can be sent to jpfaendt@uw.edu

This work included analysis of 54 molecular dynamics simulations. The simulated systems are composed of: the enzymes *Trichoderma viride* (VIR), *Thermogata maritime* (TMA), and *Pyrococcus horikoshii* (PHO); each in water with 0, 15 or 50 wt% [EMIM][OAc]; and run at both 310 and 353 K. Each of these 18 systems are simulated in triplicate to increase statistics and help to verify observations. The list of simulations is provided in Table S1. Root-mean-square displacement (RMSD) was used to track departure from the experimental reference structure for each of the 54 simulations, but the number of systems made the figures cluttered. For this reason we have relegated plots of the RMSD from each trajectory to this supporting document and used a table of summarizing statistics in the main article. Bezier curve fitting is also used to help distinguish each replicate trajectory in Figues S1 and S2 below.

Table S1: A list of the conditions of each of the 54 simulations performed. VIR, PDB: 3QR3. TMA, PDB: 3AMD. PHO, PDB: 3AXX.

Enzyme	Temp. (K)	IL (wt%)	Time (ns)	Replicas	pH
VIR	310	0	500	3	4.5
VIR	310	15	500	3	4.5
VIR	310	50	500	3	4.5
VIR	353	0	500	3	4.5
VIR	353	15	500	3	4.5
VIR	353	50	500	3	4.5
TMA	310	0	500	3	4.8
TMA	310	15	500	3	4.8
TMA	310	50	500	3	4.8
TMA	353	0	500	3	4.8
TMA	353	15	500	3	4.8
TMA	353	50	500	3	4.8
PHO	310	0	500	3	6.4
PHO	310	15	500	3	6.4
PHO	310	50	500	3	6.4
PHO	353	0	500	3	6.4
PHO	353	15	500	3	6.4
PHO	353	50	500	3	6.4

Table S2: List of charged GLU and ASP residues. If not listed, GLU or ASP is protonated. All LYS and ARG are charged. All HIS are uncharged. Residue indices are from the respective PDB files.

VIR	GLU	232, 259
	ASP	20, 43, 54, 79, 87, 152, 195, 208, 223, 225, 239, 285, 299
TMA	GLU	7, 27, 37, 78, 99, 105, 106, 127, 144, 151, 152, 172, 180, 188, 202, 209, 212, 215, 234, 239, 240, 253, 263, 275
	ASP	4, 29, 36, 40, 70, 77, 103, 120, 124, 161, 226, 227, 261, 296, 305
PHO	GLU	42, 57, 58, 87, 108, 137, 163, 169, 173, 174, 182, 234, 333, 342, 368, 399
	ASP	46, 88, 120, 126, 131, 145, 170, 175, 197, 214, 228, 262, 282, 302, 317, 321, 352, 364, 382, 385, 392, 393, 400, 410

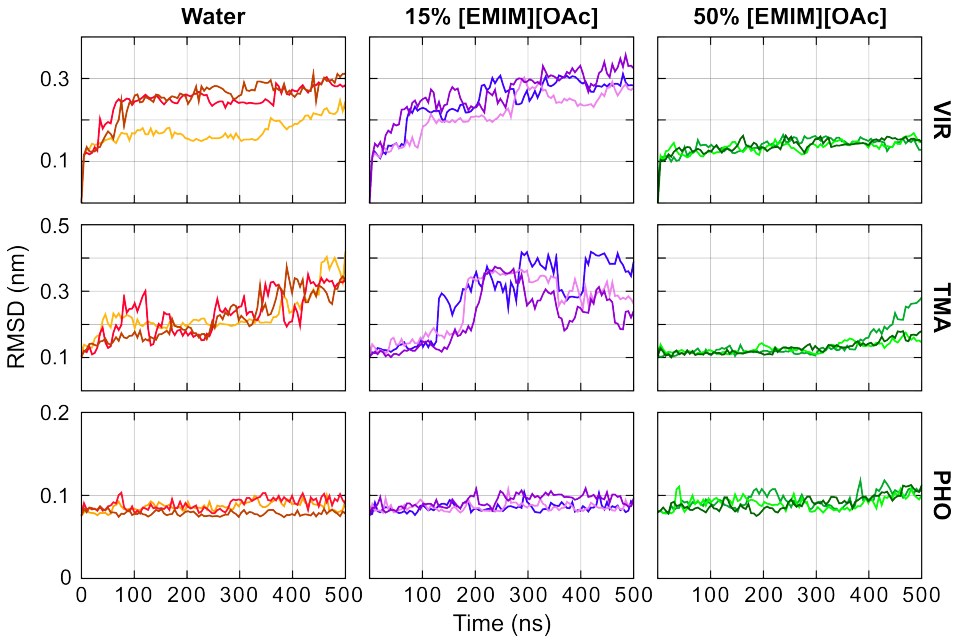

Figure S1: Bezier curve fitting of the Cα RMSD for each of the simulated systems at 310 K. The RMSD was calculated from (after alignment to) the corresponding experimental structure for each enzyme. Note the change in y-axis scale between each row.

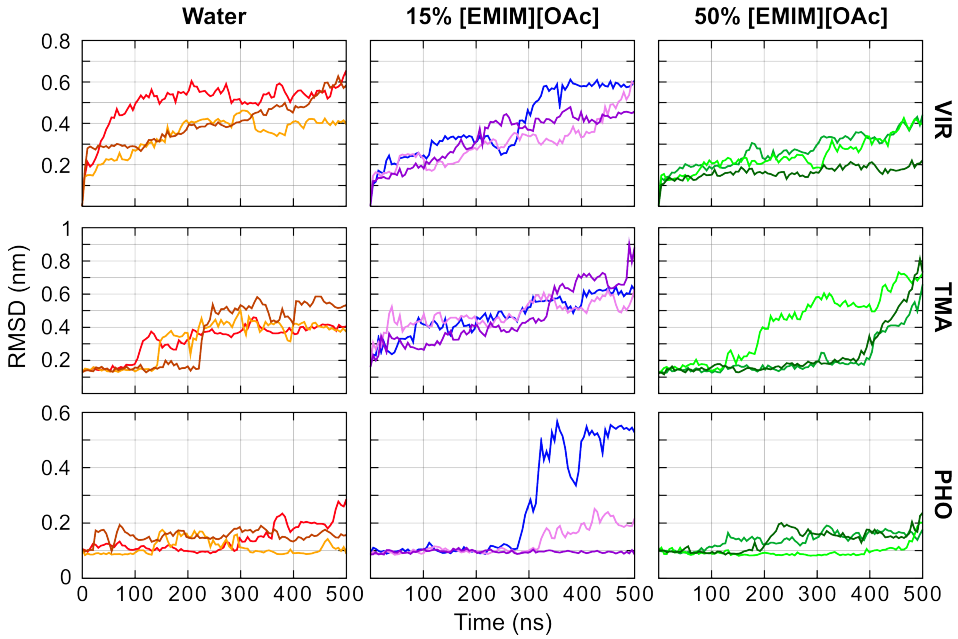

Figure S2: Bezier curve fitting of the Cα RMSD for each of the simulated systems at 353 K. The RMSD was calculated from (after alignment to) the corresponding experimental structure for each enzyme. Note the change in y-axis scale between each row.

RMSD was calculated using the g_rms tool in Gromacs 4.6 (1). These data are for the full sequence, whereas those presented in the main document for TMA do not include the first 15 N-terminal carbons. The reference structure for VIR was obtained by modifying GLU53 to ASP in the *Trichoderma reesei* (PDB: 3QR3) (2) x-ray structure in order to match the known sequence for T. viride endoglucanase (3). The reference structures for TMA and PHO were modeled using x-ray structures (PDB: 3AMD, 3AXX) (4, 5).

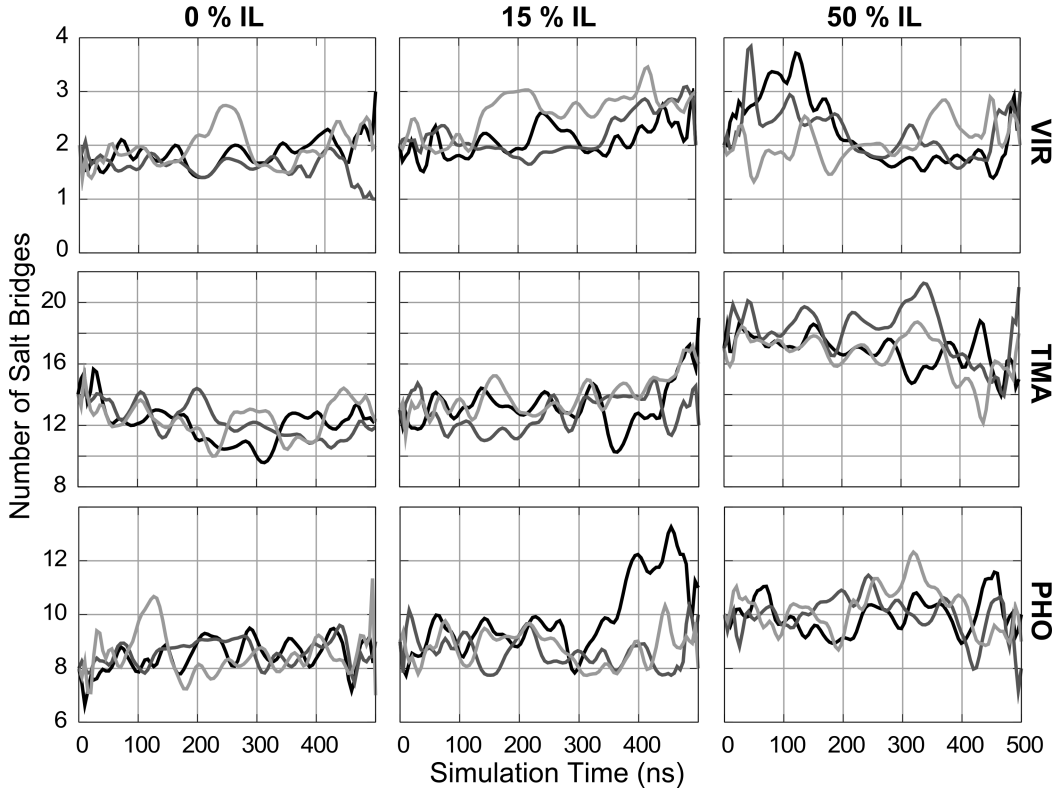

Figure S3: Bezier curve fitting of number of salt bridges for each of the simulated systems at its optimum temperature. Note the change in y-axis scale.

Salt bridges are calculated by measuring the distance between all salt-bridge-forming pairs of oxygens and nitrogens over the trajectory. This simplifies the calculation compared to calculating each oxygen-hydrogen pair. Plots begin after minimization. The intensity of each salt bridge, *S*, between atom *i* and *j* is calculated by using the sigmoidal function below. When compared to a hard cutoff, this method gives a smoother curve that helps with visualization. The intensities are summed over the whole protein. Bezier smoothing was used to smooth the plots. This removes some detail, but it separates the plots so that they can be distinguished across the three replicas. The smoothing removes the more extreme values, especially if the extremes are short-lived.

$$S = \frac{1 - \left(\frac{r_{ij}}{3.7}\right)^{20}}{1 - \left(\frac{r_{ij}}{3.7}\right)^{40}}$$

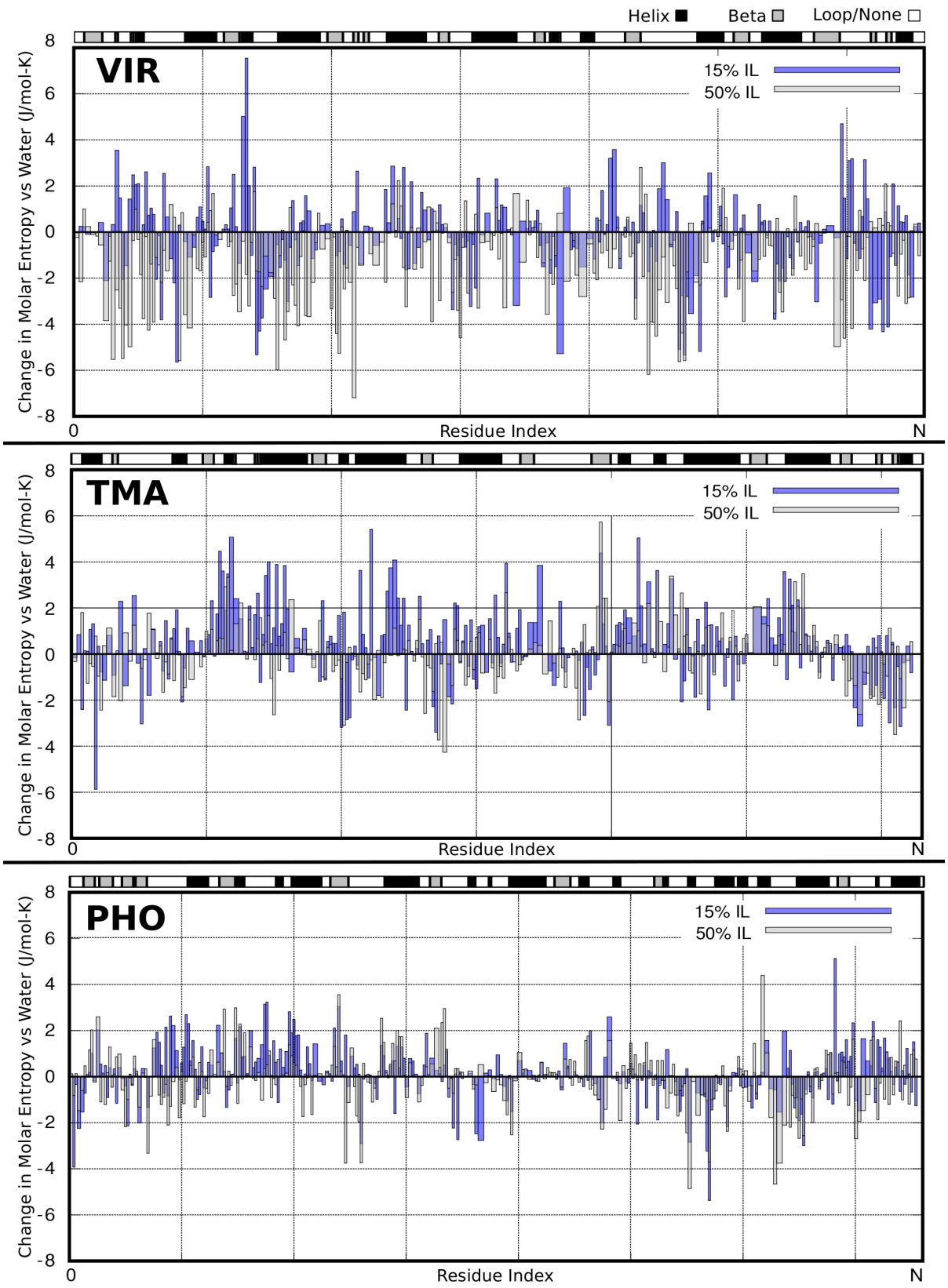

Figure S4: Relative statistical entropy of sidechains χ_1 angles. Wide bars compensate for alanines that do not have a χ_1 angle. Secondary structure is displayed at the top.

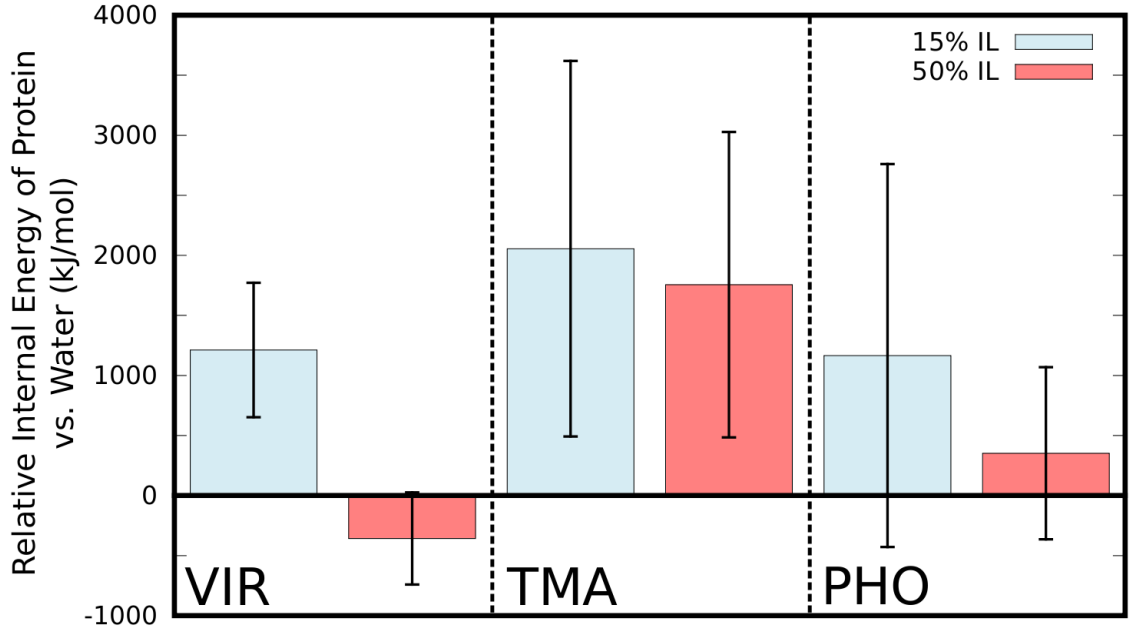

Figure S5: Relative internal energy of the protein averaged over the last half of the trajectories. Error bars are standard one deviation of the data.

Supporting References

1. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput. 4:435-447.
2. Lee, T. M., M. F. Farrow, F. H. Arnold, and S. L. Mayo. 2011. A structural study of Hypocrea jecorina Cel5A. Protein Sci. 20:1935-1940.
3. Li, X. H., P. Zhang, M. X. Wang, F. Zhou, F. A. Malik, H. J. Yang, R. Bhaskar, J. B. Hu, C. G. Sun, and Y. G. Miao. 2011. Expression of Trichoderma viride endoglucanase III in the larvae of silkworm, Bombyx mori L. and characteristic analysis of the recombinant protein. Mol. Biol. Rep. 38:3897-3902.
4. Wu, T. H., C. H. Huang, T. P. Ko, H. L. Lai, Y. H. Ma, C. C. Chen, Y. S. Cheng, J. R. Liu, and R. T. Guo. 2011. Diverse substrate recognition mechanism revealed by Thermotoga maritima Cel5A structures in complex with cellotetraose, cellobiose and mannotriose. Biochim. Biophys. Acta 1814:1832-1840.
5. Kim, H. W., and K. Ishikawa. 2011. Functional analysis of hyperthermophilic endocellulase from Pyrococcus horikoshii by crystallographic snapshots. Biochem. J. 437:223-230.