PNAS
www.pnas.org

# Supplementary Information for

## A Combined Molecular Dynamics and Neural Network Method for Predicting Protein Antifreeze Activity

**Daniel J. Kozuch, Frank H. Stillinger, and Pablo G. Debenedetti**

**Pablo G. Debenedetti.**
**E-mail: pdebene@princeton.edu**

### This PDF file includes:

Supplementary text
Fig. S1
Caption for Database S1
References for SI reference citations

### Other supplementary materials for this manuscript include the following:

Database S1



www.pnas.org/cgi/doi/10.1073/pnas.1814945115

## Supporting Information Text

### Additional Simulation Details

Temperature was maintained at 265 K to mimic sub-freezing conditions using the v-rescale thermostat (1). All systems were first energy minimized using steepest decent, and then equilibrated for a short period of 100 ps at 1 bar using the Berendsen barostat (2). Pressure was maintained at 1 bar using the Parrinello-Rahman barostat (3) during sampling. The first 30 ns of each trajectory were discarded to allow the proteins to relax and adopt stable solution state conformations. A time step of 2 fs was used and trajectories were saved every 10 ps. Bonds were constrained using the LINCS algorithm (4), and electrostatics were handled by particle mesh Ewald summation (5).

### Computational Efficiency

Computations were performed on Intel Xeon Processor E5-2680 v4 nodes with 14 cores/28 threads and 4 nVidia P100 GPUs. Using GROMACS 2016.4 with GPU acceleration, four protein systems (each containing over 25,000 atoms) were executed per node achieving  100 ns/day. Given our 8 node system and using a sampling time of 100 ns as is described in the main text, roughly 220 proteins can be scanned per week.

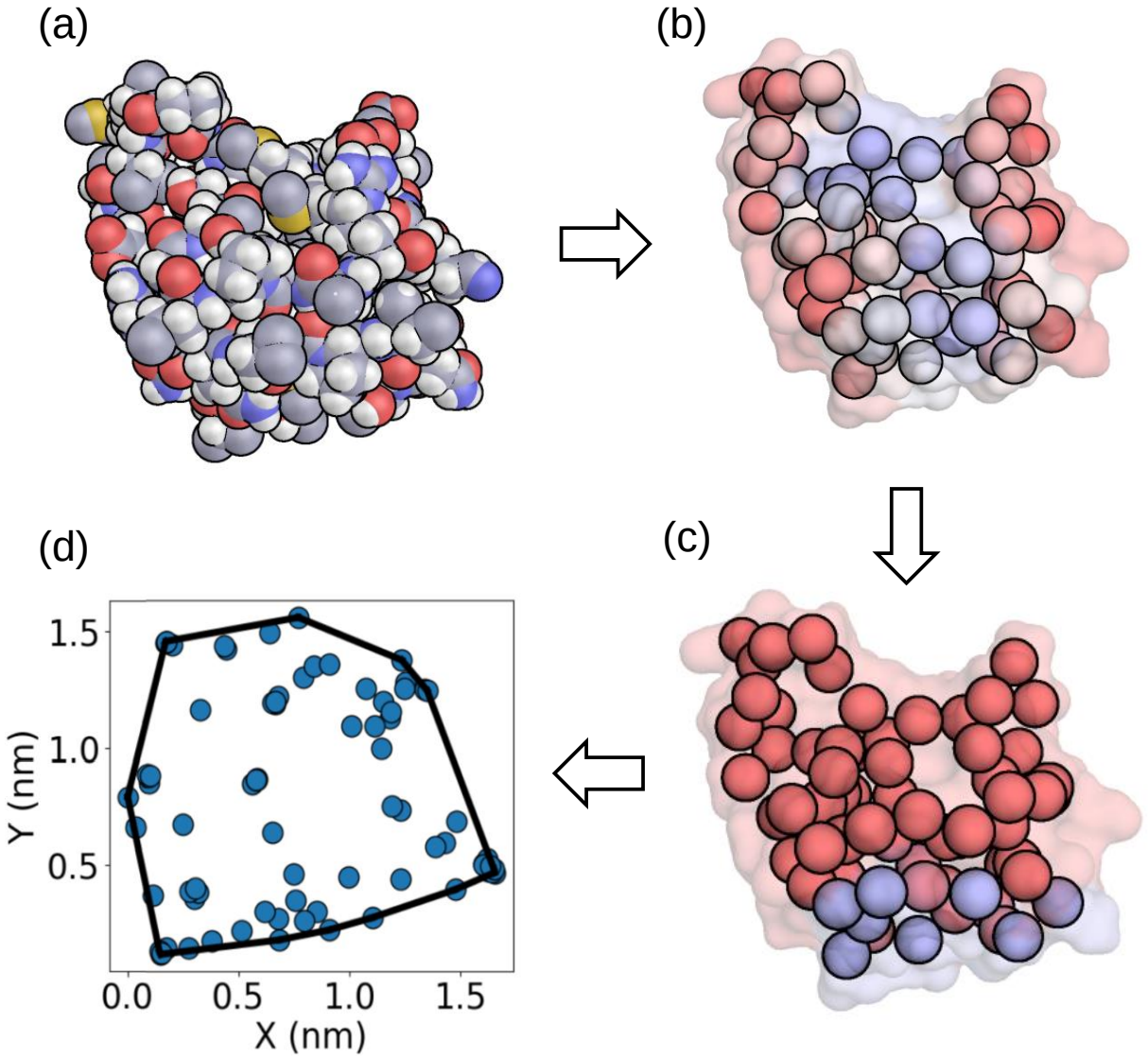


**Fig. S1.** This schematic illustrates the process for scoring a potential ice binding face. (a) A set of average coordinates is obtained for all atoms in the protein. Carbon shown in gray, oxygen in red, nitrogen in blue, hydrogen in white, and sulfur in yellow. (b) The atomic coordinates are reduced to a set of residue coordinates using the geometric center of the surface accessible atoms in each residue, and a hydrogen bond lifetime, $L_i$, is assigned to each residue coordinate. Red corresponds to short $L_i$ and blue to long $L_i$ (c) A set of coordinates $S$ in $S_{all}$ is selected as the potential IBS (shown in blue) and all other residues are designated as the NBS (shown in red). The detailed underlying procedure for this step is given in the main body (Automated detection of the IBS). From these sets, the average hydrogen bond lifetimes $L_{S,B}$ (binding face) and $L_{S,N}$ (non-binding face) are calculated. See the text for details. (d) The atomic coordinates for the surface accessible atoms of the residues in the IBS (blue points) are projected onto plane $P$. A convex hull (black line) is drawn to encompass all the points. The area of the IBS, $A_S$ is evaluated as the area of this convex hull. The set $S$ that maximizes the expression: $A_S^* + (L_{S,B} - L_{S,N})^*$ is then selected as the final ice binding face.

**Additional data table S1 (neural_network_grid.txt)**

A supplementary file containing a grid of points (different $\{A, L_B, L_N\}$ combinations) evaluated with the neural network described in the main text is attached. This dataset can be used to approximate the behavior of the neural network.

## References


1. Bussi G, Donadio D, Parrinello M (2008) Canonical sampling through velocity-rescaling.
2. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 81(8):3684–3690.
3. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* 52(12):7182–7190.
4. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18(12):1463–1472.
5. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics* 103(19):8577–8593.