SUPPORTING INFORMATION: Development of a PointNet for Detecting Morphologies of Self-Assembled Block Oligomers in Atomistic Simulations

Zhengyuan Shen,[†,‡,¶] Yangzesheng Sun,[‡,¶] Timothy P. Lodge,[†,‡] and J. Ilja Siepmann[*,†,‡,¶]

†*Department of Chemical Engineering and Materials Science, University of Minnesota, 421 Washington Avenue SE, Minneapolis, Minnesota 55455-0132, United States*

‡ *Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431, United States*

¶ *Chemical Theory Center, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431, United States*

E-mail: siepmann@umn.edu

PointNet Model and Data Availability

The implementation of the PointNet models and sample data files are freely available from:

`https://github.com/donshen/pointnet.phasedetection`

Table S1. Details of the simulated systems used to generate the point clouds shown in Figure 3

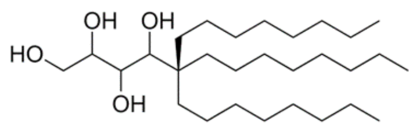
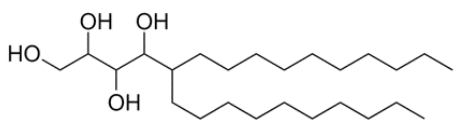
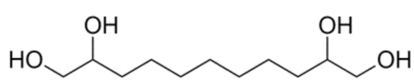
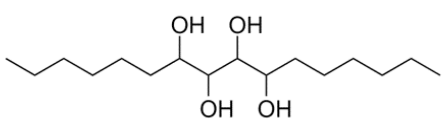
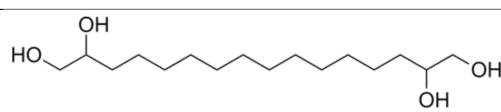
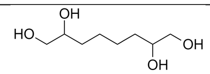
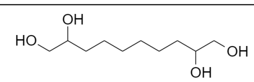
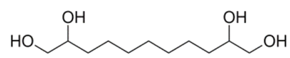
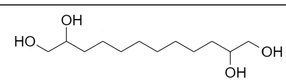
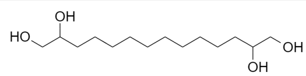
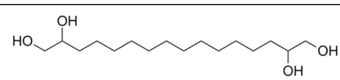
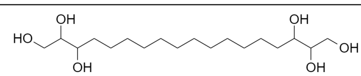
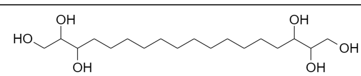
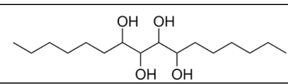
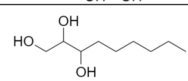
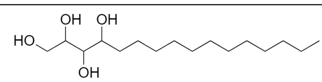
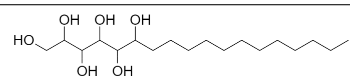
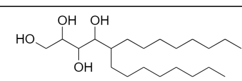
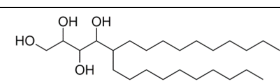
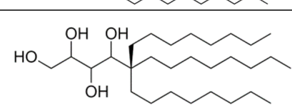
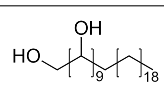
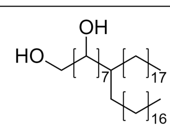
Compound	Structure	System Size	T_{SIM}/K	Morphology
$A_4B(B_8)_3$		594	460	BCC
$A_4B(B_{10})_2$		500	490	HPC
$A_2B_7A_2$		1000	430	LAM
$B_6A_4B_6$		1504	440	HPL
$A_2B_{12}A_2$		1000	520	DIS

Table S2. Details of all the simulated systems to generate point clouds for all morphologies other than NET structures.

Compound	Structure	System Size	T_{SIM}/K	Morphology
$A_2B_4A_2$		1000	400	DIS
$A_2B_6A_2$		1000 1000	400 460	LAM DIS
$A_2B_7A_2$		1000 1000	400 460	LAM DIS
$A_2B_8A_2$		1000 & 8000 1000	400 460	LAM DIS
$A_2B_{10}A_2$		1000 1000	400 460	LAM DIS
$A_2B_{12}A_2$		1000 1000	430 490	LAM DIS
$A_3B_{12}A_3$		1000 1000	460 550	LAM DIS
$A_3B_{12}A_3$		1000 1000	460 550	LAM DIS
$B_6A_4A_6$		1504 1504	440 490	HPL DIS
A_3B_6		1000 1000	400 460	LAM DIS
A_4B_{12}		1000 1000	490 550	LAM DIS
A_6B_{12}		1000 1000	580 670	LAM DIS
$A_4B(B_8)_2$		500 & 4000 500	490 550	HPC DIS
$A_4B(B_{10})_2$		500 500	490 550	HPC DIS
$A_4B(B_8)_3$		594 594	460 550	BCC DIS
$A_{10}B_{20}$		600	640	LAM
$A_8B(B_{18})_2$		600	610	HPC

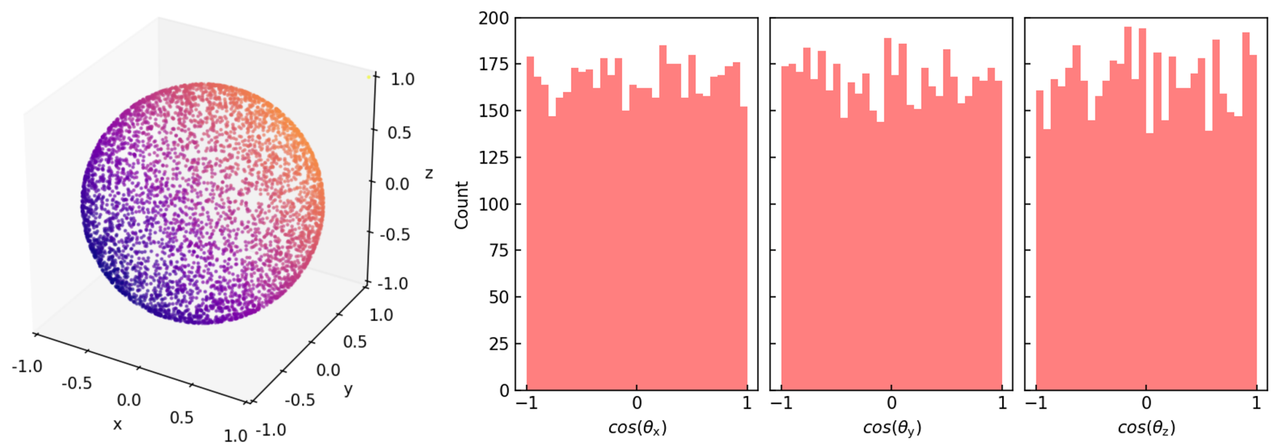


Figure S1. (Left) Unit vectors obtained by multiplying 5000 random rotation matrices $\boldsymbol{M}_{\mathrm{rot}}$ to the vector [1,0,0]. (Right) Distributions of the cosines of angles between the rotated vectors and the three Cartesian axes.

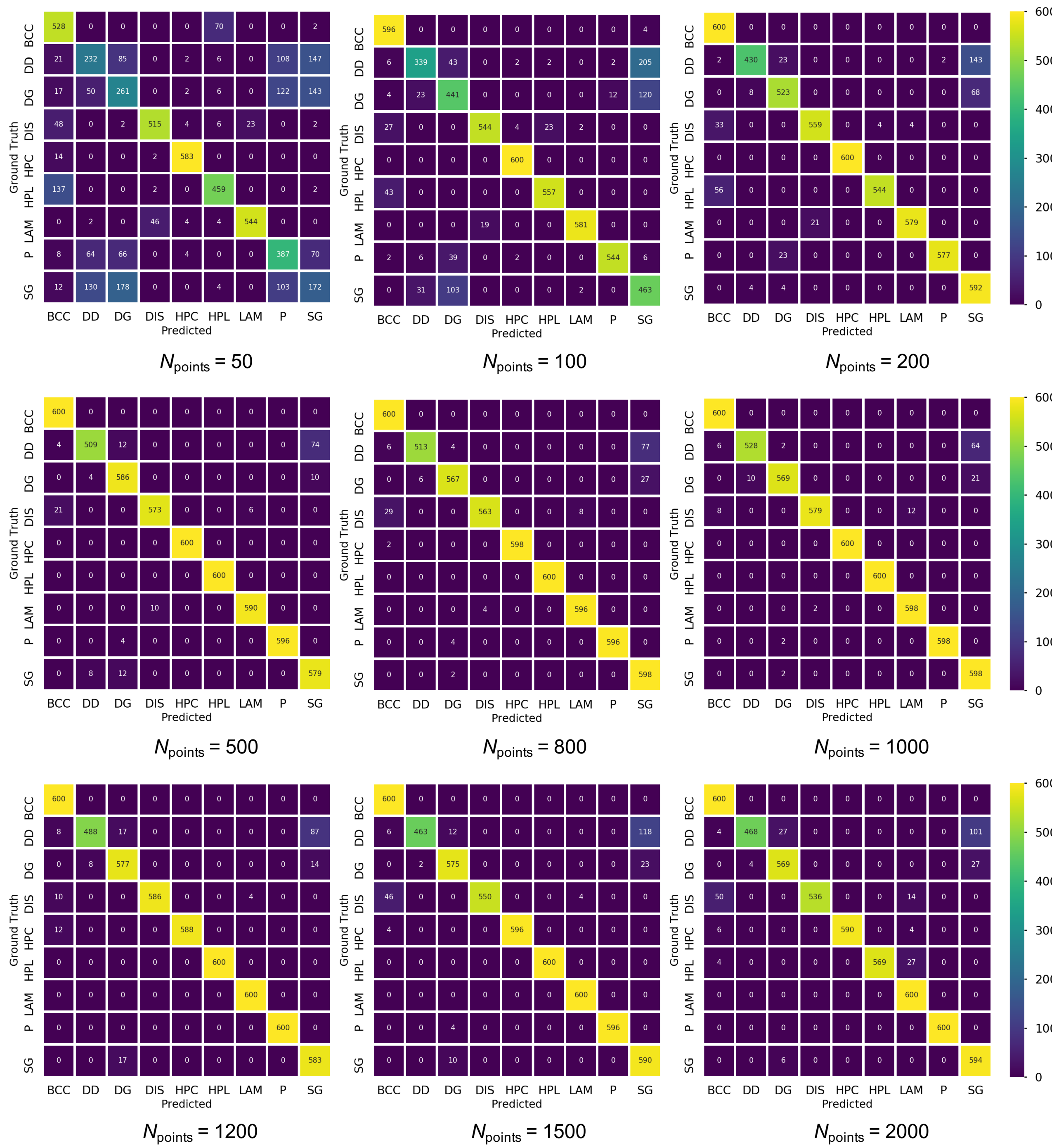


Figure S2. Confusion matrices for the test data achieved with model A for different sizes of input point clouds.

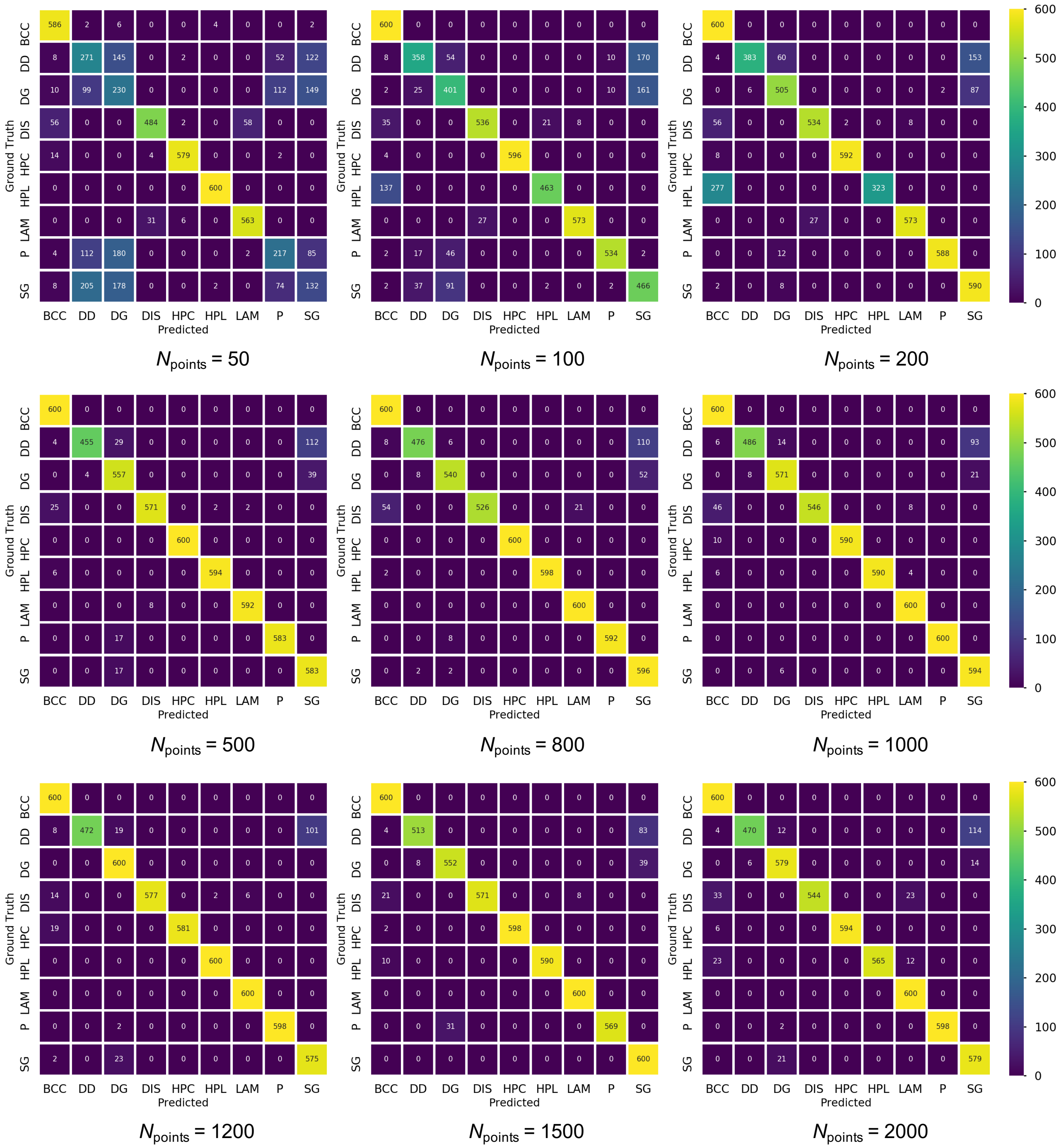


Figure S3. Confusion matrices for the test data achieved with model B for different sizes of input point clouds.

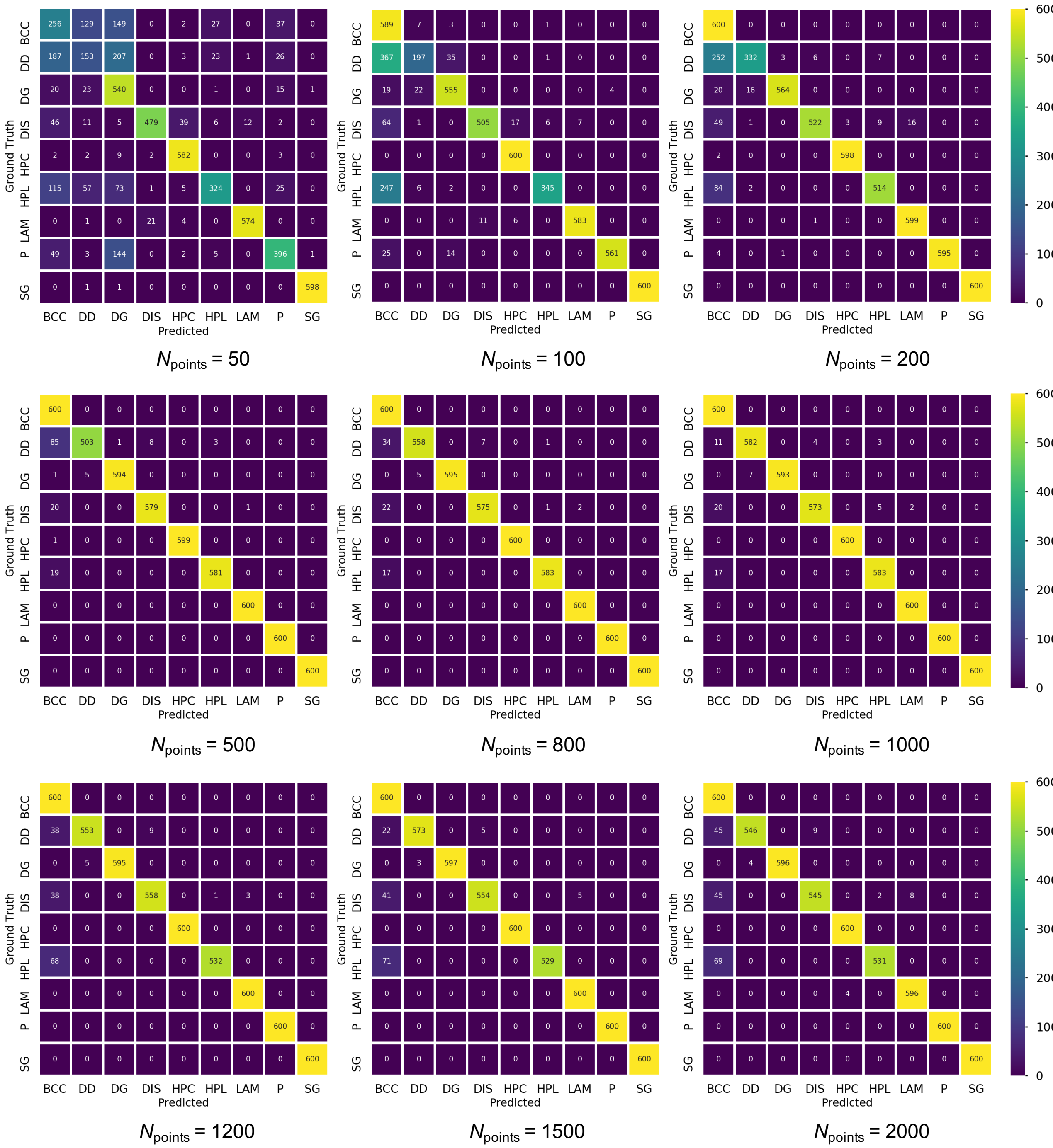


Figure S4. Confusion matrices for the test data achieved with model C for different sizes of input point clouds.

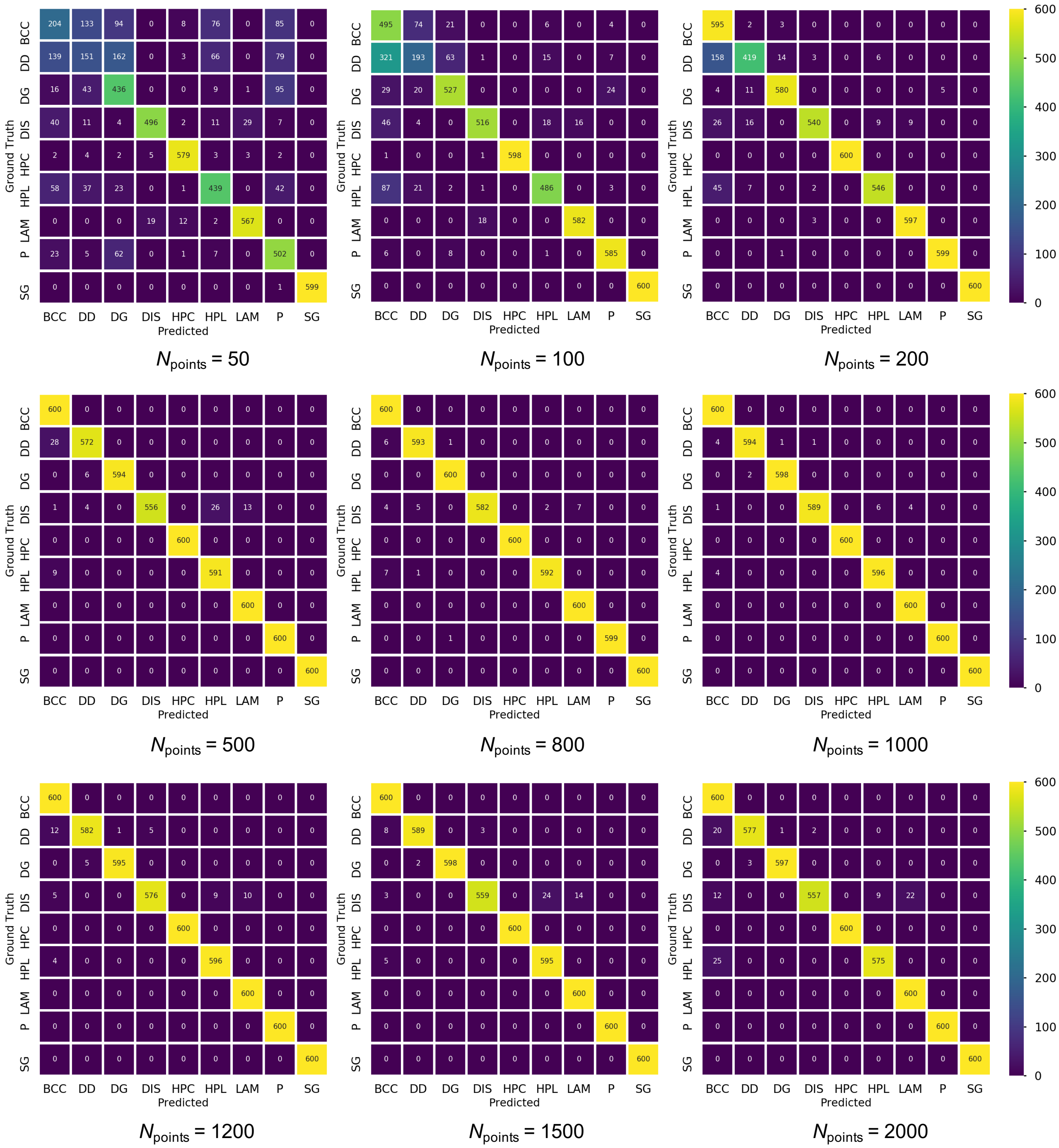


Figure S5. Confusion matrices for the test data achieved with model D for different sizes of input point clouds.

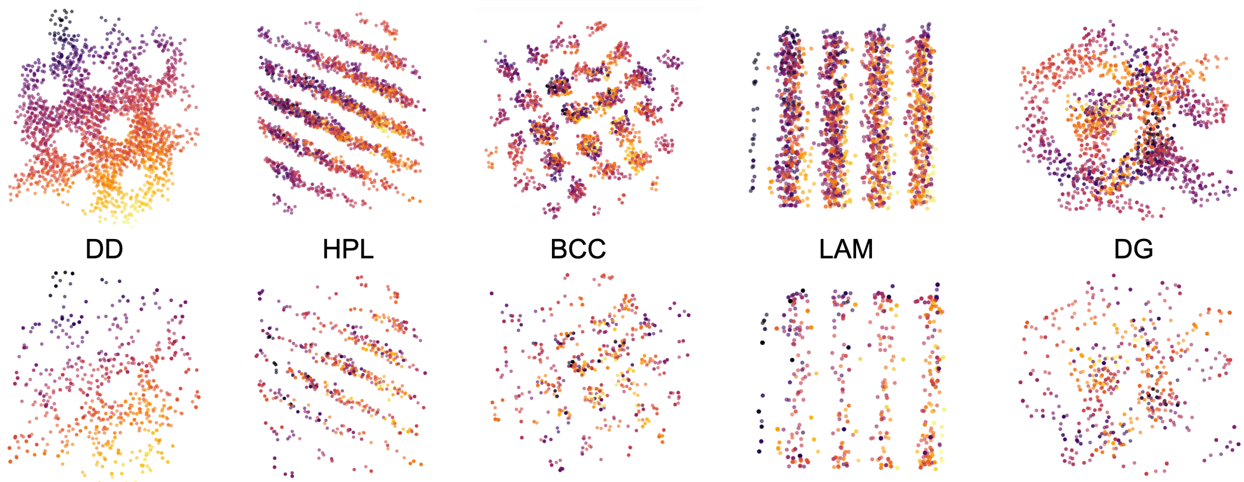


Figure S6. Examples of input point clouds with $N_{\text{points}} = 2000$ (top) and the critical points with highest contributions to the max-pooled features (bottom). To guide the eye, the point clouds are colored from yellow to purple according to a point's average values of its x, y, and z coordinates.

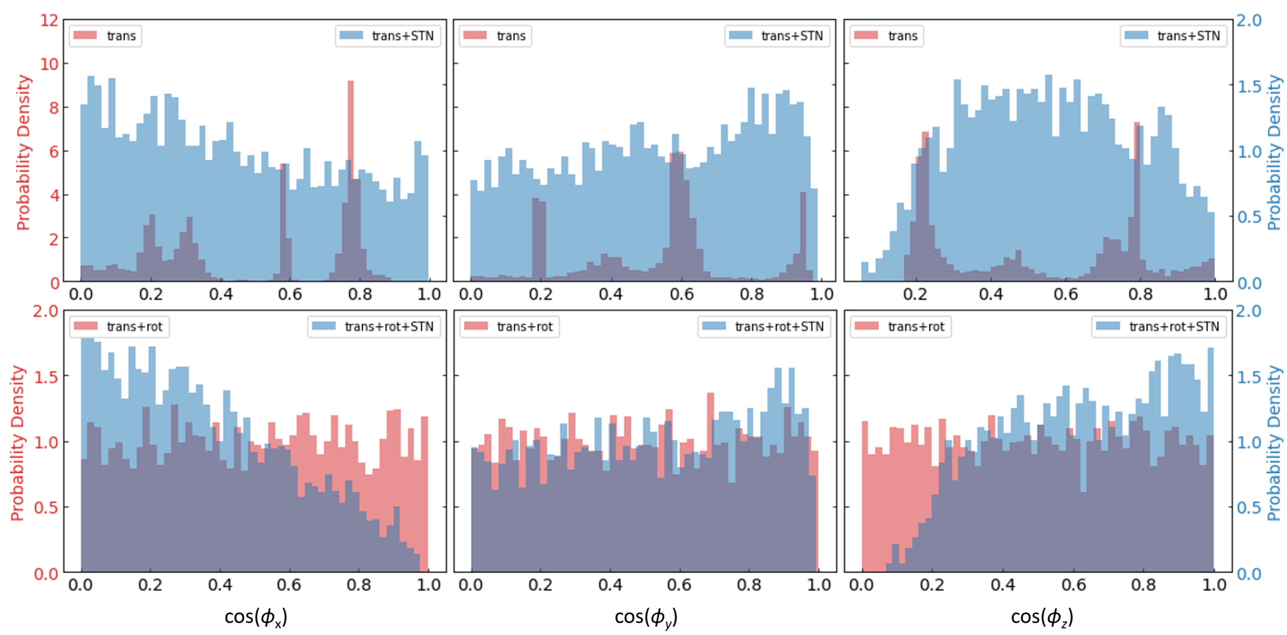


Figure S7. Orientational distributions of the vector perpendicular to the lamellar plane for the systems with LAM morphology: (top row) models A and B; (bottom row) models C and D.

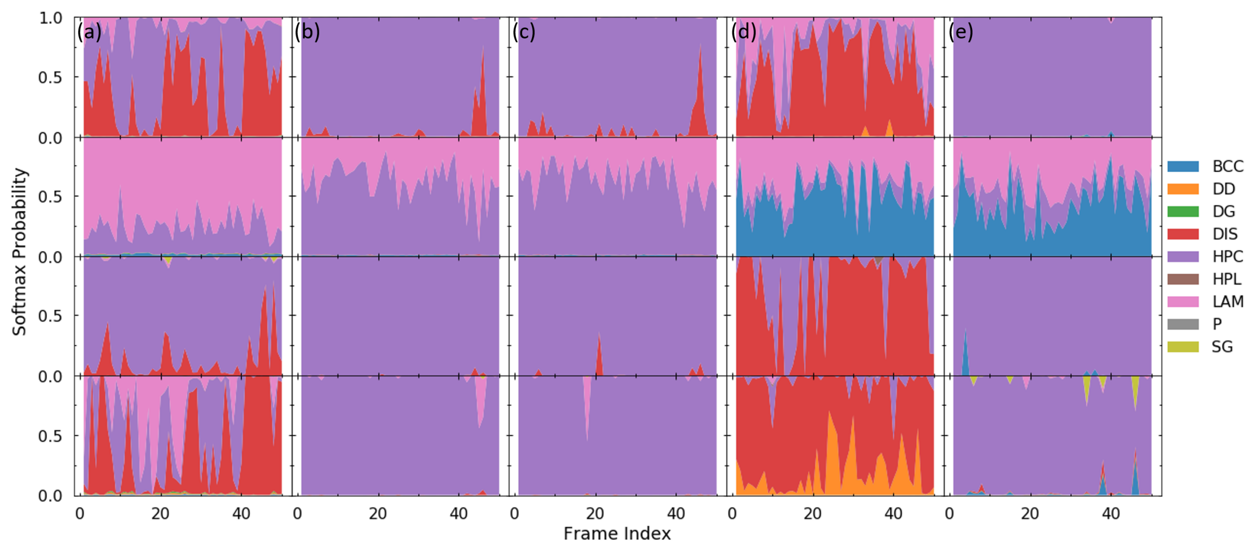


Figure S8. Stack plots of the predicted softmax classification probabilities obtained with models A, B, C, and D (top to bottom) for 50 frames taken at 10-ns intervals during a 500-ns MD trajectory. Data in columns (a)-(c) are for a 600-molecule $A_8B(B_{18})_2$ system with 1000 oxygen atoms for the point cloud selected at random from (a) all oxygen atoms, (b) only from oxygen atoms in positions 1, 3, 5, and 7 of the polar group, and (c) only from oxygen atoms in positions 2, 5, and 8 of the polar group. Data in columns (d) and (e) are for a 8000-molecule system of the $A_4B(B_8)_2$ block oligomer with 1000 oxygen atoms for the point cloud selected at random from (d) all oxygen atoms, and (e) only from oxygen atoms located in a subvolume with linear dimensions of $L_x/2$, $L_y/2$, and $L_z/2$, where L_x, L_y, and L_z are the box lengths for the entire simulation box.