Supporting information for

Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions

Chiho Kim[1], Anand Chandrasekaran[1], Tran Doan Huan[2], Deya Das[1], Rampi Ramprasad[1,*]

[1] School of Materials Science and Engineering, Georgia Institute of Technology, 771 Ferst Dr. NW, Atlanta, GA 30332, USA

[2] Department of Materials Science and Engineering and Institute of Materials Science, University of Connecticut, 97 North Eagleville Rd., Storrs, Connecticut 06269-3136, USA

*E-mail: rampi.ramprasad@mse.gatech.edu

Figure S1. Learning curves constructed from the RMSE of the machine learning models. For each model, data was obtained from 100 independent runs.

Figure S2. Building blocks implemented in Polymer Genome for constructing polymer repeat units

Figure S1. Learning curves constructed from the RMSE of the machine learning models. For each model, data was obtained from 100 independent runs.

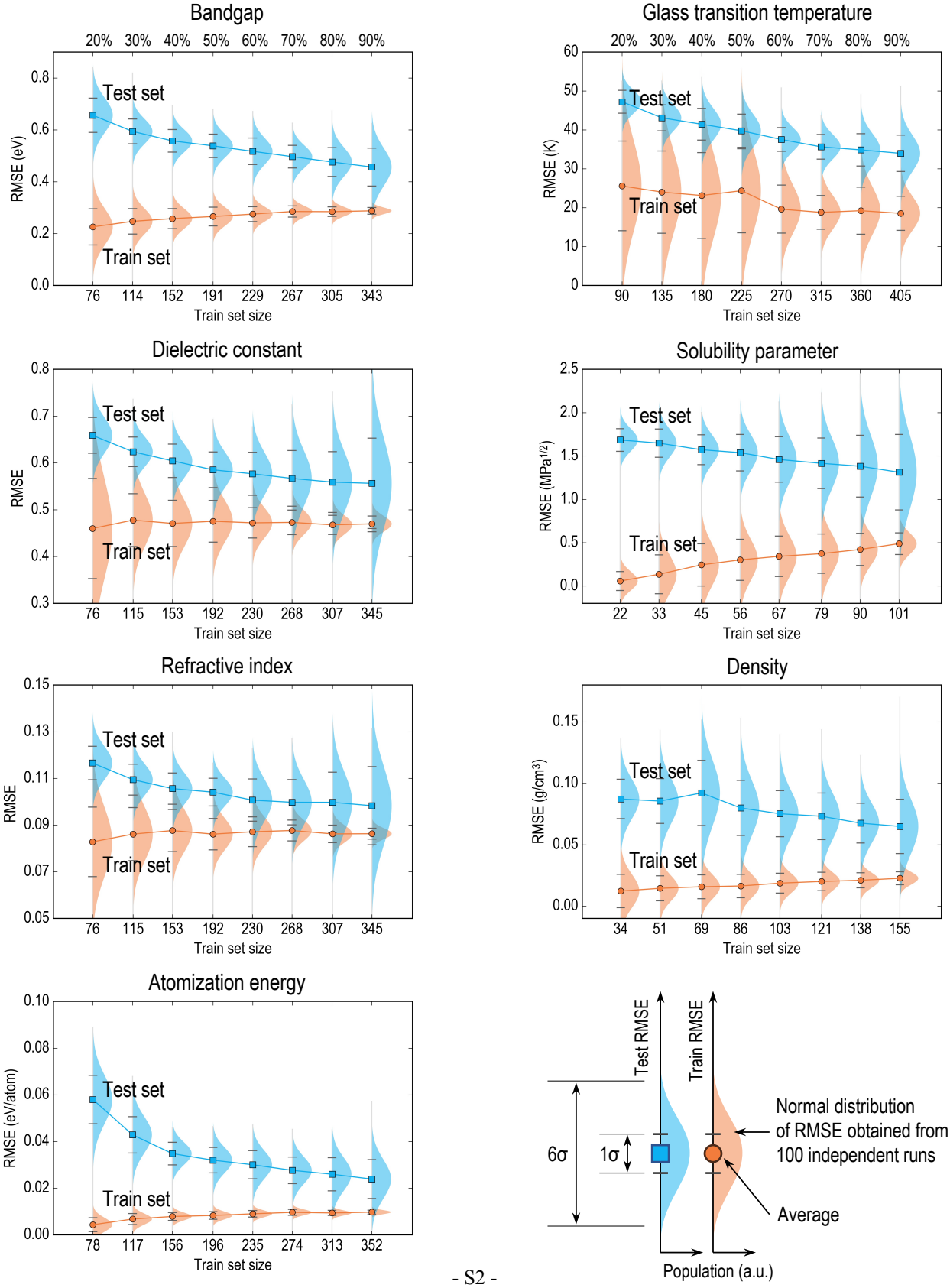
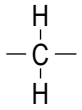
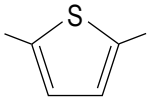
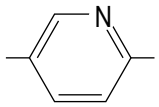
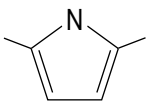
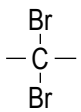
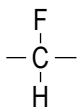
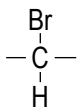
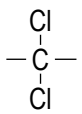
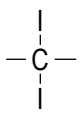
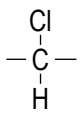
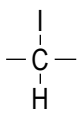


Figure S2. Building blocks implemented in Polymer Genome for constructing polymer repeat units

Chemical formula	Geometry	Chemical formula	Geometry
-CH2-		**-O-**	
***-CH-**		**-C6H4-**	
-CS-		**-C4H2S-**	
-CO-		**-C5H3N-**	
-NH-		**-C4H3N-**	
-CF2-		**-CBr2-**	
***-CF-**		***-CBr-**	
-CHF-		**-CBrH-**	
-CCl2-		**-CI2-**	
***-CCl-**		***-CI-**	
-CClH-		**-CIH-**	

*Must be paired for forming a double bond