Electronic Supplementary Material (ESI) for Journal of Materials Chemistry C.

Design of Polyimides with Targeted Glass Transition Temperature Using Graph Neural Network

Haoke Qiu, [ab] Xuepeng Qiu,[bc] Xuemin Dai,[**c] Zhao-yan Sun [*ab]

[a] State Key Laboratory of Polymer Physics and Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, China. E-mail: zysun@ciac.ac.cn (Z.Y.S.)

[b] School of Applied Chemistry and Engineering, University of Science and Technology of China, Hefei 230026, China

[c] CAS Key Laboratory of High-Performance Syntheic Rubber and its Composite Materials, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, China. E-mail: xmdai@ciac.ac.cn (X.M.D.)

1 GNN Model

The model framework is all based on Keras, TensorFlow, and DeepChem and the model schematic is shown as Figure 4. The model first obtains atomic-level features, including atomic information and bonding information between atoms. Then the atomic-level features are aggregated to form the features of the whole molecule, and then the description of the molecule is realized, which is shown in **Figure S1**.

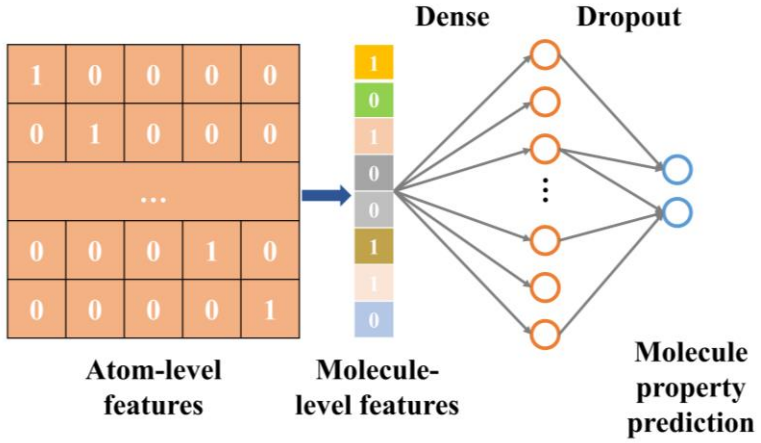

Fig S1. Atomic-level features, including atomic information and bonding information between atoms are collected, and then aggregated to form the features of the whole molecule.

The parameters of the model were then fine-tuned, mainly for the number of convolution layers, the size of convolution layers and the number of fully connected layers. We conducted five repeated experiments for each model to reduce experimental contingency, and recorded the time cost under each parameter and the inference performance on the train set and the test set. The specific results are shown in **Table S1**. We find that when the number of convolution layers is 1, the information of atoms is not completely passed, resulting in poor inference ability of the model. However, when the number of convolution layers is 3, the message passing between atoms is too sufficient, so that the difference between each atom is small, which is not conducive to the inference ability of the model. Thus, we fine-tuned our model after

setting the number of convolution layers to 2 and finally got the optimal model with R^2, RMSE, MAE equaling to 0.86, 24.83, 18.26, respectively. For brevity, one can access the training process for global nodes and classification models on https://github.com/HKQiu/PPP-1_PredictionTg4Polyimides/tree/main/src.

Table S1. Fine-tuned results of GNN model for regression.

Model parameters	Time cost /min	Train metrics (R^2, RMSE, MAE)	Test metrics (R^2, RMSE, MAE)
Graph conv. layers: 1			0.74
convolution size: 32	-	-	34.32
dense layer size: 64			21.55
Graph conv. layers: 2		0.96 (+/- 0.00)	0.82 (+/- 0.02)
convolution size: 64*64	49.114	13.25 (+/- 0.61)	28.68 (+/- 1.62)
dense layer size: 128		8.12 (+/- 0.56)	19.35 (+/- 1.04)
Graph conv. layers: 2		0.94 (+/- 0.01)	0.81 (+/- 0.01)
convolution size: 32*32	38.442	14.84 (+/- 0.71)	29.19 (+/- 1.01)
dense layer size: 128		10.17 (+/- 0.69)	19.83 (+/- 0.48)
Graph conv. layers: 2		0.92 (+/- 0.00)	0.79 (+/- 0.04)
convolution size: 32*32	48.424	17.82 (+/- 0.32)	31.13 (+/- 3.17)
dense layer size: 64		9.44 (+/- 0.39)	20.58 (+/- 1.81)
Graph conv. layers: 3		0.87 (+/- 0.04)	0.72 (+/- 0.06)
convolution size: 64*64*64	65.130	22.76 (+/- 3.73)	35.71 (+/- 3.91)
dense layer size: 128		13.07 (+/- 2.31)	22.84 (+/- 2.79)
Best Model	-	-	**0.86** **24.83** **18.26**

2 Materials

Pyromellitic dianhydride (PMDA), 4,4'-oxydiphthalic anhydride (ODPA), 3,3',4,4'-biphenyl tetracarboxylic diandhydride (BPDA), 1,4-bis(3,4-dicarboxyphenoxy)benzene dianhydride (HQDPA), 4,4'-(hexafluoroisopropylidene)diphthalic anhydride (6FDA), and 2,2-bis [4-(3,4dicarboxyphenoxy) phenyl] propane dianhydride (BPADA) were purchased from Shanghai Chemical Reagent Plant and dried at 180 ℃ prior to use. 4,4'-oxydianiline (ODA), *m*-phenylenediamine (MPD), 4,4'-bis(3-aminophenoxy)biphenyl (BAPB), 2,2'-bis(trifluoromethyl)benzidine (TFMB) and 9,9-Bis(4-aminophenyl)fluorene (BAF) were purchased from Chang-zhou Sunlight Medical Raw Material Co. Ltd. and used directly. N,N-dimethylacetamide (DMAc) and other reagents were obtained from Tianjin Fu Chen Chemicals Reagent Factory without further purification prior to use.

3 Synthesis of polymers

3.1 Synthesis of Polyamic acid

A series of PAA solutions were successfully synthesized by the copolymerization of different diamine with equimolar amount of dianhydride in DMAc with 12 wt% solid concentration. PI-A is used as an example to illustrate the synthesis process of poly-amic acid. First, ODA (20.02 g, 0.1 mol) and DMAc were placed in a 500-ml three-necked flask and mechanically stirred at room temperature under nitrogen at-mosphere. After complete dissolution of ODA, PMDA (21.81 g, 0.1 mol) were grad-ually added into the solution. The reaction was carried out at 25 °C for 3 h. The corresponding inherent viscosities of PAA solutions were listed in Table 1.

3.2 Synthesis of Polyimide (A-H)

Polyimide powder was obtained by chemical imidization, and the experiment was carried out in nitrogen atmosphere. First, the previously obtained PAA solution was

followed by adding isoquinoline (12.92 g, 0.1 mol, 1:1 mol ratio to dianhydride) , and then acetic anhydride (25.52 g, 0.25 mol,) was added in batches for cyclodehydration. After 2 hours, the polyimide precipitated as a yellow powder under the action of rapid stirring. The precipitates were collected through filtration, subsequently washed repeatedly through hot ethanol, and dried under vacuum at 150 ℃ for 12 h.

Polyimide powder was also obtained by chemical imidization. First, the previously obtained PAA solution was followed by adding isoquinoline (12.92 g, 0.1 mol, 1:1 mol ratio to dianhydride), and then acetic anhydride (25.52 g, 0.25 mol,) was added in batches for cyclodehydration. After the completion of the reaction, the polyimide solution was rapidly poured into absolute ethanol. The precipitates were obtained through filtration, subsequently washed repeatedly through hot ethanol, and dried under vacuum at 150 ℃ for 12 h.

4 Characterization

The inherent viscosities (η_{inh}) of PAA solutions were measured with an Ubbelohde viscometer at a concentration of 0.5 g·dL^{-1} in DMAc at 30 °C. Differential scanning calorimetry (DSC) analysis was performed using a calorimeter (Q20 DSC, TA instruments), and the sample mass was approximately 5–10 mg.

The analysis was carried out at a heating rate of 10 ℃/min and a cooling rate of 20 ℃/min under N_2 atmosphere, over a temperature range from room temperature to 400 ℃. The middle of thermal transition dip in DSC curves a re regarded as T_g (as shown in **Table S2**) obtained from the second heating sc an and the original DSC data can be accessed at https://github.com/HKQiu/PPP-1_PredictionTg4Polyimides/tree/main/DSC%20data.

Table S2. Inherent viscosity of the PAA solutions of corresponding PI.

Sample	η_{inh} (dL/g)	T_g (℃)
PI-A	2.04	360
PI-B	1.78	262

PI-C	0.92	265
PI-D	0.71	329
PI-E	0.86	247
PI-F	1.08	218
PI-G	--	269
PI-H	1.35	239

Note that PI-G was synthesized using thermal imidization method.