Supplementary Information (SI) for Journal of Materials Chemistry A.

Supporting Information for

On-Demand Design of Materials with Enhanced Dielectric Properties *via* Machine Learning-Assisted Materials Genome Approach

Shujian Cao[1#], *Zhentao Zhang*[1#], *Siqin Song*[1], *Haoxiang Lan*[1], *Liang Gao*[1]*, *Jiaping Lin*[1]*, *Chongyin Zhang*[2], *and Wanyu Tang*[1]

[1] Shanghai Key Laboratory of Advanced Polymeric Materials, Key Laboratory for Ultrafine Materials of Ministry of Education, Frontiers Science Center for Materiobiology and Dynamic Chemistry, School of Materials Science and Engineering, East China University of Science and Technology, Shanghai, 200237, China

[2] Shanghai Aerospace Equipments Manufacturer Co., Ltd., Shanghai, 200245, China

Contents

S1. Data sources of dielectric properties

We constructed an extensive dataset comprising dielectric performance data of various polymers. The sources of the collected dielectric performance data are primarily from 70 articles, 3 handbook publications, the PolyInfo database, and the AI plus Polymers database as below.

[1] J. S. Ho and S. G. Greenbaum, *ACS Appl. Mater. Interfaces*, 2018, **10**, 29189.

[2] R. Ma, A. F. Baldwin, C. Wang, I. Offenbach, M. Cakmak, R. Ramprasad and G. A. Sotzing, *ACS Appl. Mater. Interfaces*, 2014, **6**, 10445.

[3] J. Pan, K. Li, S. Chuayprakong, T. Hsu and Q. Wang, *ACS Appl. Mater. Interfaces*, 2010, **2**, 1286.

[4] J. Wei, Z. Zhang, J. K. Tseng, I. Treufeld, X. Liu, M. H. Litt and L. Zhu, *ACS Appl. Mater. Interfaces*, 2015, **7**, 5248.

[5] A. F. Baldwin, R. Ma, A. Mannodi-Kanakkithodi, T. D. Huan, C. Wang, M. Tefferi, J. E. Marszalek, M. Cakmak, Y. Cao, R. Ramprasad and G. A. Sotzing, *Adv. Mater.*, 2015, **27**, 346.

[6] C. Wu, A. A. Deshmukh, Z. Li, L. Chen, A. Alamri, Y. Wang, R. Ramprasad, G. A. Sotzing and Y. Cao, *Adv. Mater.*, 2020, **32**, 2000499.

[7] S. Wu, W. Li, M. Lin, Q. Burlingame, Q. Chen, A. Payzant, K. Xiao and Q. M. Zhang, *Adv. Mater.*, 2013, **25**, 1734.

[8] M. Zhang, B. Li, J. J. Wang, H. B. Huang, L. Zhang and L. Q. Chen, *Adv. Mater.*, 2021, **33**, 2008198.

[9] Y. Zhang, Z. Geng, S. Niu, S. Zhang, J. Luan and G. Wang, *Adv. Ind. Eng. Polym. Res.*, 2020, **3**, 175.

[10] C. Zhou, K. K. Guo, H. M. Qi, Y. P. Zhu and F. Wang, *Adv. Mat. Res*, 2015, **1120**, 550.

[11] T. Nagaraja, S. T. Dadami, S. R. Manohara and B. Angadi, *AIP Conf. Proc.*, 2019, **2142**, 070023.

[12] D. Yan, Q. Yang, Y. Zhou, X. Shen and F. Huang, *Chinese. J. Polym. Sci*, 2016, **46**, 43.

[13] Q. Li, F. Yao, Y. Liu, G. Zhang, H. Wang and Q. Wang, *Annu. Rev. Mater. Res.*, 2018, **48**, 219.

[14] Y. Wang, X. Zhou, M. Lin and Q. M. Zhang, *Appl. Phys. Lett.*, 2009, **94**, 202905.

[15] S. Wu, M. Lin, Q. Burlingame and Q. M. Zhang, *Appl. Phys. Lett.*, 2014, **104**, 072903.

[16] S. Chisca, I. Sava, V. E. Musteata and M. Bruma, presented in part at the CAS 2011 Proceedings (2011 International Semiconductor Conference), 17-19 October, 2011.

[17] S. Ishaq, F. Kanwal, S. Atiq, S. Noreen, M. Moussa, U. Azhar and D. Losic, *Ceram. Int.*, 2020, **46**, 24649.

[18] W. Wan, M. Tao, H. Cao, Y. Zhao, J. Luo, J. Yang and T. Qiu, *Ceram. Int.*, 2020, **46**, 13862.

[19] Y. Wang, Z. Li, C. Wu and Y. Cao, *Chem. Eng. J.*, 2020, **401**, 126093.

[20] M. Ma, C. Li, X. Liu, Q. Yuan and F. Huang, *ChemistrySelect*, 2020, **5**, 1146.

[21] S. Chisca, V. E. Musteata, I. Sava and M. Bruma, *Eur. Polym. J.*, 2011, **47**, 1186.

[22] E. N. Demidova, V. M. Matyuk, A. I. Drachev, A. B. Gil'man, A. M. Maksimov and V. E. Platonov, *High Energy Chem.*, 2009, **43**, 56.

[23] H. Chen, H. Xin, J. Lu, J. Tang, Q. Yuan and F. Huang, *High Perform. Polym.*, 2016, **29**, 595.

[24] H. Yan, T. Li, M. Zhang and S. Feng, *High Perform. Polym.*, 2014, **26**, 618.

[25] L. Mandelcorn and R. L. Miller, presented in part at the IEEE 35th International Power Sources Symposium, 22-25 June, 1992.

[26] I. Treufeld, D. H. Wang, B. A. Kurish, L.-S. Tan and L. Zhu, *J. Mater. Chem. A*, 2014, **2**, 20683.

[27] B. Knapp and P. A. Kohl, *J. Appl. Polym. Sci.*, 2014, **131**, 40828.

[28] A. Kumar, D. Ahmad, K. Patra and M. Hossain, *J. Appl. Polym. Sci.*, 2021, **138**, 50991.

[29] W. Li, M. Chen, Y. Yang, D. Yuan, Y. Ren and X. Cai, *J. Appl. Polym. Sci.*, 2019, **136**, 47447.

[30] W. L. Qu and T. M. Ko, *J. Appl. Polym. Sci.*, 2001, **82**, 1642.

[31] H. Tong, A. Ahmad, J. Fu, H. Xu, T. Fan, Y. Hou and J. Xu, *J. Appl. Polym. Sci.*, 2019, **136**, 47883.

[32] Y. Chen, K. Guo, F. Wang, Y. Zhu and H. Qi, *Journal of East China University of Science and Technology*, 2020, **46**, 368.

[33] Y. Thakur, M. Lin, S. Wu and Q. M. Zhang, *J. Electron. Mater.*, 2016, **45**, 4721.

[34] S. Wu, Q. Burlingame, Z. Cheng, M. Lin and Q. M. Zhang, *J. Electron. Mater.*, 2014, **43**, 4548.

[35] I. Murakami, H. Ochiai, K. Tominaga and M. Kajiwara, *J. Inorg. Organomet. Polym.*, 1992, **2**, 255.

[36] Y. Liu, J. Hua and W. Sang, *J. Macromol. Sci. Part B-Phys.*, 2006, **39**, 349.

[37] Z. Li, Gregory M. Treich, M. Tefferi, C. Wu, S. Nasreen, S. K. Scheirey, R. Ramprasad, G. A. Sotzing and Y. Cao, *J. Mater. Chem. A*, 2019, **7**, 15026.

[38] R. Ma, V. Sharma, A. F. Baldwin, M. Tefferi, I. Offenbach, M. Cakmak, R. Weiss, Y. Cao, R. Ramprasad and G. A. Sotzing, *J. Mater. Chem. A*, 2015, **3**, 14845.

[39] F. Wen, L. Zhang, P. Wang, L. Li, J. Chen, C. Chen, W. Wu, G. Wang and S. Zhang, *J. Mater. Chem. A*, 2020, **8**, 15122.

[40] Y. Chen, X. You and S. Deng, *J. Mater. Sci. Chem. Eng.*, 2018, **06**, 1.

[41] N. M. Renukappa, Siddaramaiah and R. D. Sudhakar Samuel, *J. Mater. Sci.: Mater. Electron.*, 2006, **18**, 635.

[42] B. Natesan, N. K. Karan, M. B. Rivera, F. M. Aliev and R. S. Katiyar, *J. Non-Cryst. Solids*, 2006, **352**, 5205.

[43] S. Diaham and M. L. Locatelli, *J. Phys. D: Appl. Phys.*, 2013, **46**, 185302.

[44] Q. Wang, Z. Pu, X. Zheng, Y. Tian, X. Li and J. Zhong, *J. Polym. Res.*, 2020, **27**, 328.

[45] D. H. Wang, B. A. Kurish, I. Treufeld, L. Zhu and L. S. Tan, *J. Polym. Sci. Pol. Chem.*, 2014, **53**, 422.

[46] D. H. Wang, J. K. Riley, S. P. Fillery, M. F. Durstock, R. A. Vaia and L. S. Tan, *J. Polym. Sci. Pol. Chem.*, 2013, **51**, 4998.

[47] H. Tong, J. Fu, A. Ahmad, T. Fan, Y. Hou and J. Xu, *Macromol. Mater. Eng.*, 2019, **304**, 1800709.

[48] D. Xu, W. Xu, T. Seery, H. Zhang, C. Zhou, J. Pang, Y. Zhang and Z. Jiang, *Macromol. Mater. Eng.*, 2020, **305**, 1900820.

[49] A. F. Baldwin, T. D. Huan, R. Ma, A. Mannodi-Kanakkithodi, M. Tefferi, N. Katz, Y. Cao, R. Ramprasad and G. A. Sotzing, *Macromolecules*, 2015, **48**, 2422.

[50] Z. Zhang, M. H. Litt and L. Zhu, *Macromolecules*, 2018, **51**, 1967.

[51] Y.-F. Zhu, Z. Zhang, M. H. Litt and L. Zhu, *Macromolecules*, 2018, **51**, 6257.

[52] S. Kashi, R. K. Gupta, T. Baum, N. Kao and S. N. Bhattacharya, *Mater. Des.*, 2016, **109**, 68.

[53] Z. Zhang, J. Zheng, K. Premasiri, M.-H. Kwok, Q. Li, R. Li, S. Zhang, M. H. Litt, X. P. A. Gao and L. Zhu, *Mater. Horizons*, 2020, **7**, 592.

[54] N. Venkat, T. D. Dang, Z. Bai, V. K. McNier, J. N. DeCerbo, B. Tsao and J. T. Stricker, *Mater. Sci. Eng. B-Adv.*, 2010, **168**, 16.

[55] Y. Thakur, R. Dong, M. Lin, S. Wu, Z. Cheng, Y. Hou, J. Bernholc and Q. M. Zhang, *Nano Energy*, 2015, **16**, 227.

[56] Q. Li, L. Chen, M. R. Gadinski, S. Zhang, G. Zhang, U. Li, E. Iagodkine, A. Haque, L. Q. Chen, N. Jackson and Q. Wang, *Nature*, 2015, **523**, 576.

[57] Q. Li, J. Liu, X. Zhang, S. Tan, J. Lu and Z. Zhang, *Phys. Chem. Chem. Phys.*, 2019, **21**, 15712.

[58] R. G. Lorenzini, W. M. Kline, C. C. Wang, R. Ramprasad and G. A. Sotzing, *Polymer*, 2013, **54**, 3529.

[59] B. Nandan, L. D. Kandpal and G. N. Mathur, *Polymer*, 2003, **44**, 1267.

[60] J. Wei, T. Ju, W. Huang, J. Song, N. Yan, F. Wang, A. Shen, Z. Li and L. Zhu, *Polymer*, 2019, **178**, 121688.

[61] X. Yin, Y. Qiao, M. R. Gadinski, Q. Wang and C. Tang, *Polym. Chem.*, 2016, **7**, 2929.

[62] T. Zhu, Q. Yu, W. Zheng, R. Bei, W. Wang, M. Wu, S. Liu, Z. Chi, Y. Zhang and J. Xu, *Polym. Chem.*, 2021, **12**, 2481.

[63] S. Uzaki, K. Adachi and T. Kotaka, *Polym. J.*, 1988, **20**, 221.

[64] J. Yang, Z. Tang, H. Yin, Y. Liu, L. Wang, H. Tang and Y. Li, *Polymers*, 2019, **11**, 776.

[65] X. Peng, W. Xu, L. Chen, Y. Ding, T. Xiong, S. Chen and H. Hou, *React. Funct. Polym.*, 2016, **106**, 93.

[66] T. Yang, W. Xu, X. Peng and H. Hou, *RSC Adv.*, 2017, **7**, 23309.

[67] P. C. Irwin, D. Q. Tan, Y. Cao, N. Silvi, M. Carter, M. W. Rumler and C. Garet, *SAE Int. J. Aerosp.*, 2008, **1**, 817.

[68] T. Zhang, X. Chen, Y. Thakur, B. Lu, Q. Zhang, J. Runt and Q. M. Zhang, *Sci. Adv.*, 2020, **6**, eaax6622.

[69] T. P. Nguyen, V. H. Tran and S. Lefrant, *Synth. Met.*, 1995, **69**, 443.

[70] H. Xin, H. Chen, J. Lu, Q. Yuan and F. Huang, *Thermosetting Resin*, 2017, **32**, 42-47.

[71] J. Bicerano, *Prediction of Polymer Properties*, 3rd Edition edn., 2002.

[72] J. E. Mark, *Physical Properties of Polymers Handbook*, Springer, 2007.

[73] J. E. Mark, *Polymer Data Handbook*, Oxford University Press, 2009.

[74] Polymer Database (PoLyInfo), https://polymer.nims.go.jp/, (accessed February 2022).

[75] AI plus Polymer, https://polymergenome.ecust.edu.cn, (accessed February 2023).

S2. Model performance using inputs without dimensionality reduction

It is feasible to directly take all the Mordred descriptors as model inputs for ML modeling. Here, we attempted this operation and found that the dimensionality reduction of descriptors is necessary for training ML models with high accuracy. Taking Gaussian process regression (GPR) and support vector regression (SVR) modeling processes as examples, the Mordred descriptors with or without dimensionality reduction serve as model inputs for training the models of dielectric constant ε, dielectric loss tanδ, and T_{d5}, respectively. For the 100 models trained in each case, the distributions of the R^2 values for the test sets and the consumption time for predicting one structure are presented in Figure S1 and S2, respectively. As can be seen, for the three kinds of targeted properties, the ML models based on the inputs after dimensionality reduction have a higher R^2 and a lower consumption time. For example, for the GPR and SVR models of tanδ, if the dimensionality reduction operation is not conducted for the Mordred descriptors, the R^2 of the ML models is difficult to exceed 0.8, and the consumption time for predicting one structure increases significantly.

As we know, feature engineering is an action of extracting features from raw data. The exact and sufficient features can reduce the difficulty of modeling, and therefore enable the ML models to output results with high quality. Therefore, the dimensionality reduction processing of descriptors, such as the PCA and correlation-based feature engineering methods adopted in this work, can achieve more reliable and effective feature engineering, resulting in the rapid and accurate prediction performance for the ML models of ε, tanδ, and T_{d5}.

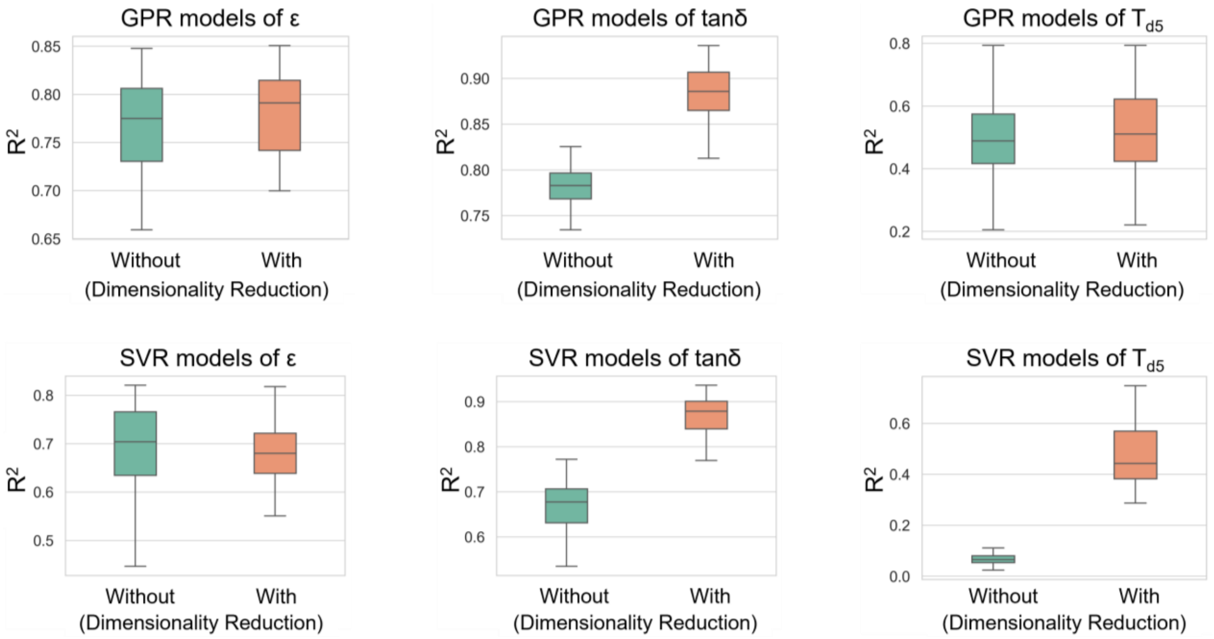


Figure S1. Comparison of the model training results using the inputs with and without dimensionality reduction. The box plots show the distributions of the R^2 values of the test set for the GPR and SVR models of ε, tanδ, and T_{d5}. The green and orange boxes correspond to the results of trained models with and without dimensionality reduction, respectively.

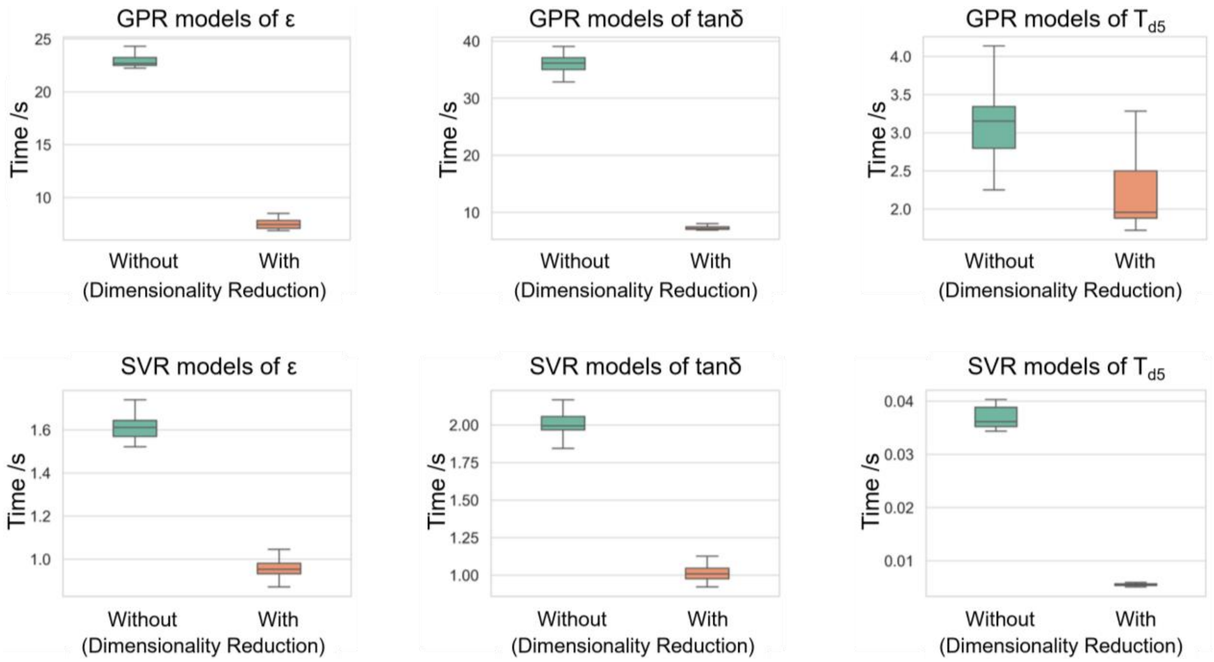


Figure S2. Comparison of the consumption time for predicting one structure when adopting the GPR and SVR models of ε, tanδ, and T_{d5}. The green and orange boxes correspond to the results of trained models with and without dimensionality reduction, respectively.

S3. Model performance using various PCA retaining explainabilities

Here, we compared the prediction performance of the trained ML models using the PCA results from the retaining explainability of 85%, 90%, and 95% as inputs. As shown in Figure S3a, retaining 90% of explainability in the GPR modeling for ε, the trained prediction models exhibit a higher R^2 on the test set. When training the SVR models for T_{d5}, choosing 90% or 95% of explainability can enable the trained models to show an R^2 exceeding 0.8 on the test set (Figure S3b). A low PCA retaining explainability leads to the missing of important features, while a high explainability can cause a complex feature space and increase the possibility of model overfitting. Thus, the retaining explainability of 90% or 95% in the present work is reasonable.

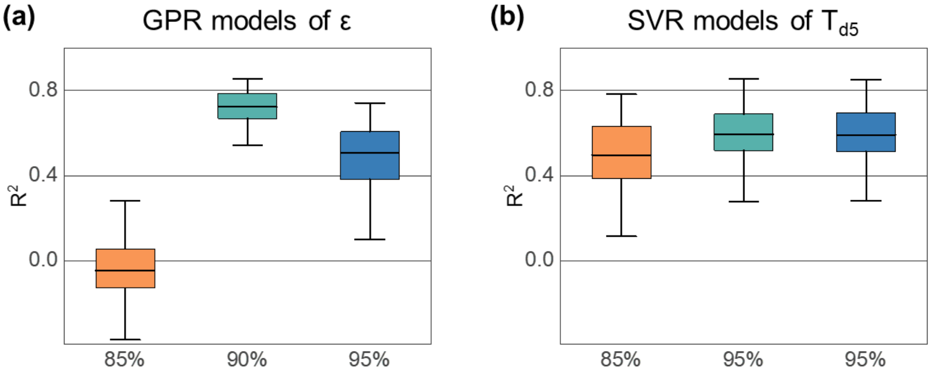


Figure S3. Comparison of the training results using the PCA results from the retaining explainability of 85%, 90%, and 95% as inputs. The box plots show the distributions of the R^2 values of the test set for (a) GPR models of ε and (b) SVR models of T_{d5}.

S4. Model performance using various ML modeling methods

For each targeted property, we adopted at least five ML modeling methods for training, including GPR, SVR, decision tree regression (DTR), *etc.* Under different training parameters, the predictive performance of 100 models trained by the above ML modeling methods was evaluated. Then, the ML modeling methods suitable for each targeted property were selected according to their performance.

For example, after performing GPR, SVR, and DTR modeling on the datasets of ε, tanδ, and T_{d5}, respectively, the predictive performance of these trained models is shown in Figure S4. For the training results of ε and tanδ, the GPR models show the R^2 value exceeding 0.8 on the test set, while the SVR and DTR models perform poorly on the test set (Figure S4a-b). For the training results of T_{d5}, both the GPR and SVR models have a good prediction performance on the test set, that is, R^2 can reach 0.8. However, the R^2 value of the training set for the GPR models is close to 1.0, which is significantly higher than that of the test set, indicating a risk of model overfitting. Therefore, the ML prediction models on the datasets of ε and tanδ were established by using the GPR modeling method, while the ML model of T_{d5} was trained by adopting the SVR modeling method, as indicated by the marks in Figure S4.

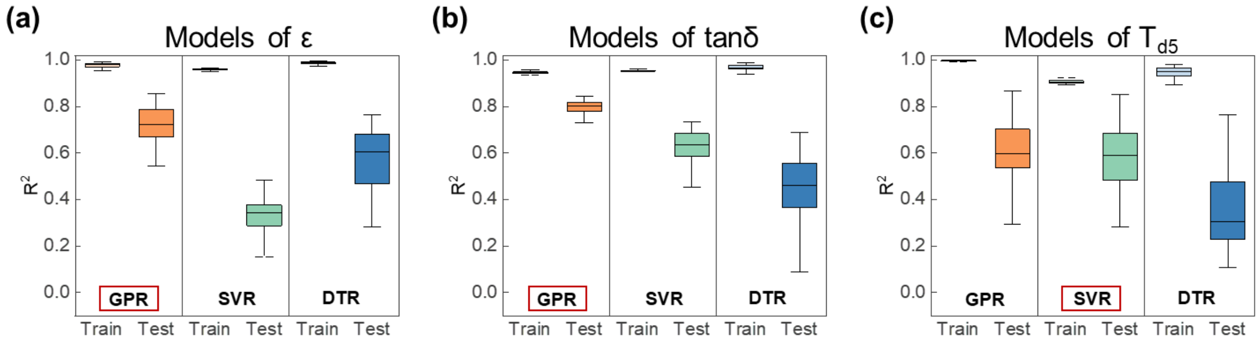


Figure S4. Comparison of the training results using GPR, SVR, and DTR modeling methods on the dataset of (a) ε, (b) tanδ, and (c) T_{d5}, respectively.

S5. Selection and combination of kernels during the GPR training process

During the GPR training process, the Matern kernel is combined with the constant kernel and the white noise kernel, aiming to improve the prediction performance of the GPR models and avoid overfitting. Compared with the radial basis function kernel, the Matern kernel is more flexible. By using different smoothing parameters, the smoothness degree of the regression results can be adjusted.

For example, during the GPR training processes of ε and tanδ, 100 GPR models were trained using kernel parameters with *nu* values of 0.5, 1.5, and 2.5, respectively. The model performance on the training set and the test set is shown in Figure S5. As can be seen from the figure, when the model kernel parameter *nu*=0.5 or 1.5, the performance of the GPR models in the training set and test set shows significant differences, indicating the overfitting of models. We selected the Matern kernel with *nu* = 2.5 (*i.e.*, M52 kernel), which is suitable for modeling the physical processes with certain irregularities. Compared with the Matern kernel with *nu* =1.5 or 0.5, the M52 kernel can capture more local variations, thereby, it not only fits the data features better, but also effectively avoids model overfitting and improves the generalization ability.

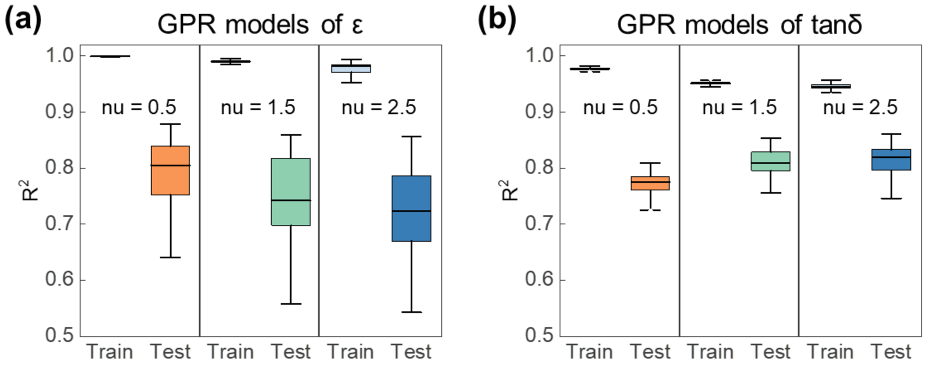


Figure S5. Comparison of the training results using the various kernel parameters of GPR modeling methods on the datasets of (a) ε and (b) tanδ, respectively.

In addition, the product combination of the constant kernel and the Matern kernel enables the GPR model to automatically learn the global scaling of the signal amplitude. The white noise kernel is further combined, achieving the consideration of data noise during the modeling process. Furthermore, the method of maximizing logarithmic marginal likelihood was adopted for the kernel parameter space, achieving the optimization of all parameters of the kernel function.

S6. The 5-fold cross-validation results for the models of targeted properties

Here, the 5-fold cross-validation method has been utilized to evaluate the performance of the trained models of ε, tanδ, and T_{d5}, respectively. The standard deviations from the 5-fold cross-validation for each property are presented in Table S1. It can be seen from Table S1 that the trained models exhibit low standard deviations on the train and validation folds. This indicates that the models have a high prediction accuracy and generalization ability, and model overfitting has been avoided.

Table S1. The standard deviations from the 5-fold cross-validation for each property.

		Model of ε	Model of tanδ	Model of T_{d5}
R^2	Train_std	0.0026	0.0046	0.0091
	Validation_std	0.0398	0.0293	0.0940
MAE	Train_std	0.0058	0.0037	0.8932
	Validation_std	0.0069	0.0123	4.2277

S7. Correlation-based feature engineering for the dielectric loss

The purpose of feature engineering is to capture the structural features of polymers. We utilized SMILES strings to represent the repeating unit structures and employed the RDKit software package to calculate Mordred descriptors for the chemical structures in the datasets. [S1, S2] It is imperative to reduce the dimensionality of the descriptors, obtaining a few descriptors that retain as much molecular information as possible for model training. However, due to the complex relationships between features of tanδ and the test conditions (*i.e.*, temperature and frequency), the utilization of the principal component analysis (PCA) dimensionality reduction for all the descriptors results in the loss of some chemical structure feature information that is related to tanδ.

To solve this problem, we extended the correlation-based feature analysis method to the dimensionality reduction and feature extraction in the present work. This method is commonly used in the feature engineering of molecular structures, such as extracting the descriptors of molecular solubility and establishing the corresponding ML prediction model. [S3, S4] In such a method, the Pearson correlation coefficient (r) is utilized as the metric. If the absolute value of r between two features is high ($r \geq 0.9$), indicating that the two features can replace each other, one of them can be discarded. Using the correlation-based feature engineering method, we selected the top 30 descriptors that have a higher correlation with tanδ from all the descriptors. Subsequently, the descriptors that can be mutually substituted among the top 30 descriptors were further eliminated, remaining only 17 descriptors that are enough to represent molecular structural features. The 17 molecular Mordred descriptors after correlation-based feature engineering for training the tanδ model are listed in Table S2.

Table S2. Mordred descriptors used for training the dielectric loss model.

Name	Description
SMR_VSA6	MOE MR VSA Descriptor 6 ($2.75 < x < 3.05$)
RotRatio	rotatable bonds ratio
NssCH2	number of $ssCH_2$
ATSC3Z	centered moreau-broto autocorrelation of lag 3 weighted by atomic number
AATS3dv	averaged moreau-broto autocorrelation of lag 3 weighted by valence electrons
BIC2	2-ordered bonding information content
SLogP	Wildman-Crippen LogP
fMF	molecular framework ratio
nS	number of S atoms
BalabanJ	Balaban's J index
C3SP2	SP_2 carbon bound to 3 other carbons
AATSC1Z	averaged and centered moreau-broto autocorrelation of lag 1 weighted by atomic number
nHBDon	number of hydrogen bond donor
NaaCH	number of aaCH
MATS1Z	moran coefficient of lag 1 weighted by atomic number
TMPC10	10-ordered total path count
Xc-5d	5-ordered Chi cluster weighted by sigma electrons

Figure S6 presents the analysis results of the intercorrelations among the aforementioned 17 descriptors and their correlation with tanδ. The heatmap illustrates the degree of correlation, ranging from blue (indicating negative correlation) to red (indicating positive correlation). Deeper shades of blue signify a higher negative correlation between the descriptors, while deeper shades of red indicate a higher positive correlation. Red lines denote descriptors positively correlated with tanδ, whereas blue lines denote descriptors negatively correlated with tanδ. There are several features with a high correlation with tanδ. For instance, SLogP, which stands for Synthetic Logarithm of Partition Coefficient, serves as an indicator for the hydrophobicity (lipophilicity) of compounds. There exists a negative correlation between SLogP and tanδ, implying that the introduction of lipophilic groups can lead to a reduction in tanδ. Compounds with lipophilic properties tend to contain fewer polar groups, resulting in a decrease in tanδ. The SMR_VSA series of descriptors is utilized to quantify the contributions of different atom types within a molecule to its molecular refraction. In particular, the SMR_VSA6 exhibits a positive correlation with tanδ, indicating that a high value of SMR_VSA6 signifies the presence of a greater number of atomic or functional groups that contribute significantly to molecular refraction. The existence of these groups leads to an increased tanδ of the molecule. Utilizing the 17 descriptors after correlation-based feature engineering, the prediction model of tanδ was established. Compared to the PCA dimensionality reduction method, the correlation-based feature engineering method can offer good interpretability and achieve an accurate prediction of tanδ.

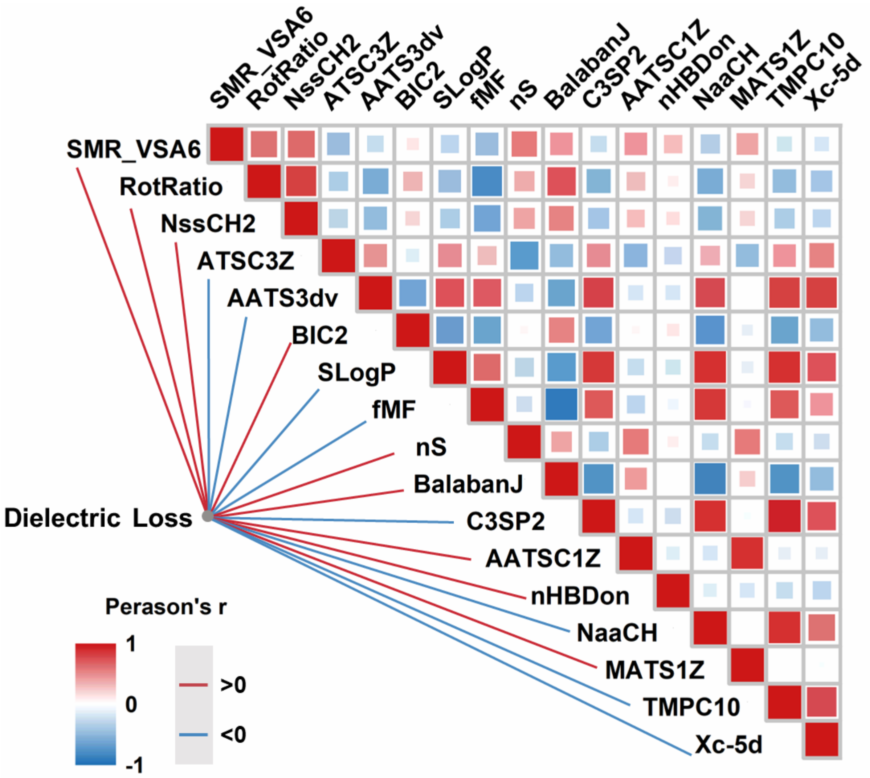
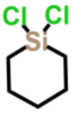


Figure S6. Correlation analysis of descriptors used for training the model of tanδ.

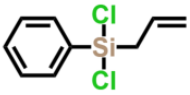
S8. Chemical structures of 130 silane genes

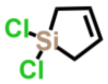
We collected 130 types of silane genes from the open-source PubChem database.[S5] The molecular structures of the 130 types of silane genes are shown below, where the gene number (No.), the chemical abstracts service (CAS) identifier, or the compound ID (CID) in the database are also provided.

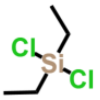

No. 1
CAS: 2406-33-9

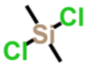

No. 2
CAS: 2406-34-0

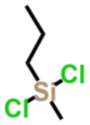

No. 3
CAS: 2351-33-9

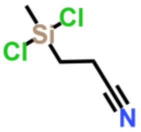

No. 4
CAS: 7719-03-1

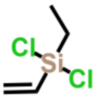

No. 5
CAS: 872-46-8

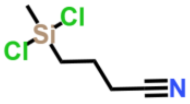

No. 6
CAS: 1719-53-5

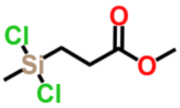

No. 7
CAS: 75-78-5

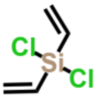

No. 8
CAS: 4518-94-9

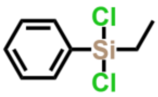

No. 9
CAS: 1071-21-2


No. 10
CAS: 10138-21-3

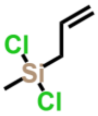

No. 11
CAS: 1190-16-5

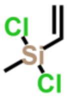

No. 12
CAS: 18163-34-3


No. 13
CAS: 1745-72-8

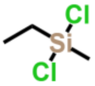

No. 14
CAS: 1125-27-5

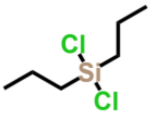

No. 15
CAS: 80-10-4

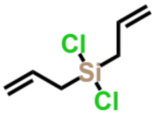

No. 16
CAS: 1873-92-3


No. 17
CAS: 124-70-9


No. 18
CAS: 18395-90-9

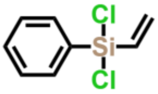

No. 19
CAS: 4525-44-4

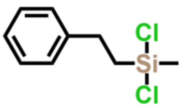

No. 20
CAS: 2295-24-1


No. 21
CAS: 3651-23-8

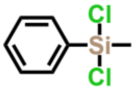

No. 22
CAS: 7751-38-4

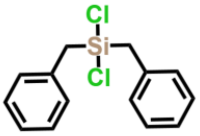

No. 23
CAS: 18236-89-0


No. 24
CAS: 7719-02-0

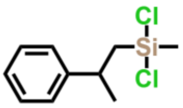

No. 25
CAS: 772-65-6


No. 26
CAS: 139147-73-2

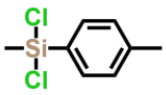

No. 27
CAS: 149-74-6

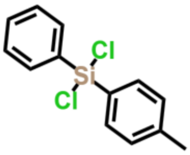

No. 28
CAS: 18414-36-3

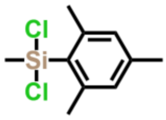

No. 29
CAS: 872-46-8

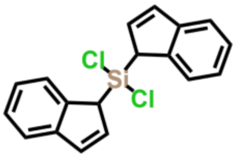

No. 30
CAS: 13617-28-2

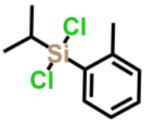

No. 31
CAS: 18030-58-5

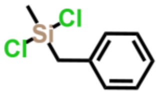

No. 32
CAS: 25898-37-7

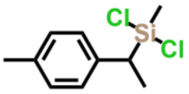

No. 33
CAS: 13788-41-5

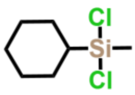

No. 34
CAS :120217-95-0

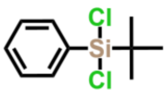

No. 35
CAS: 167776-66-1

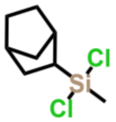

No. 36
CAS: 824934-48-7

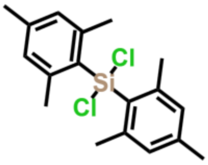

No. 37
CAS: 18236-82-3


No. 38
CAS: 718635-97-3

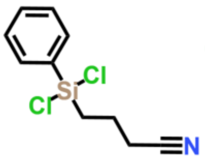

No. 39
CAS: 5578-42-7

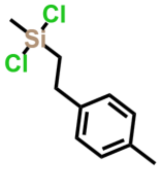

No. 40
CAS: 17887-41-1


No. 41
CAS: 18301-58-1

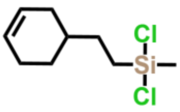

No. 42
CAS: 5599-27-9

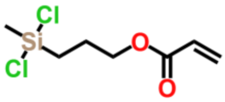

No. 43
CAS: 18414-38-5

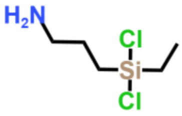

No. 44
CAS: 1078-96-2


No. 45
CAS: 17887-45-5


No. 46
CAS: 18035-74-0

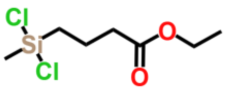

No. 47
CAS: 17864-93-6

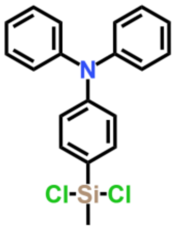

No. 48
CAS: 71550-63-5

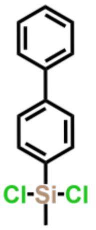

No. 49
CAS: 93917-77-2

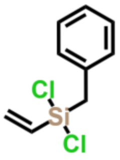

No. 50
CAS: 18042-67-6

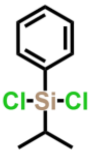

No. 51
CAS: 151236-68-9

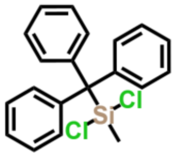

No. 52
CAS: 18276-28-3

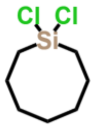

No. 53
CAS: 183798-64-3

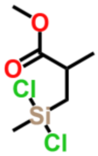

No. 54
CAS: 148718-31-4


No. 55
CAS: 148718-31-4

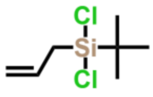

No. 56
CAS: 790234-74-1


No. 57
CAS: 256343-28-9

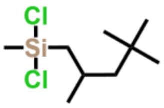

No. 58
Compound CID:
22156374

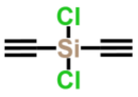

No. 59
CAS: 4074-06-0

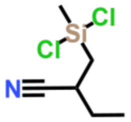

No. 60
CAS: 18028-96-1


No. 61
CAS: 216320-36-4

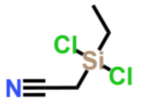

No. 62
CAS: 18147-18-7

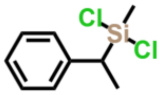

No. 63
CAS: 17869-31-7

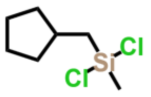

No. 64
CAS: 145289-11-8


No. 65
CAS: 876343-97-4


No. 66
CAS: 18290-79-4

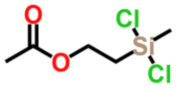

No. 67
CAS: 18243-57-7

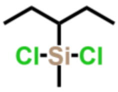

No. 68
CAS: 7282-39-5

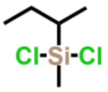

No. 69
CAS: 268217-73-8

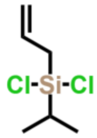

No. 70
CAS: 14579-04-5

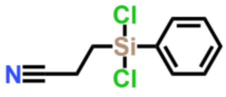

No. 71
CAS: 18395-92-1


No. 72
CAS: 5578-41-6


No. 73
Compound CID:
89655867

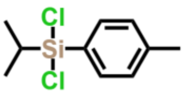

No. 74
CAS: 18379-40-3

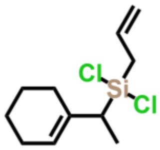

No. 75
CAS: 200114-13-2

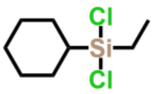

No. 76
CAS: 1077-57-2

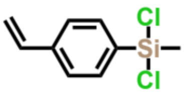

No. 77
CAS: 93037-56-0

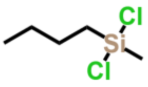

No. 78
CAS: 273934-04-6

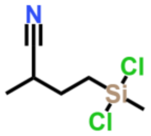

No. 79
CAS: 790234-76-3


No. 80
CAS: 168270-65-3

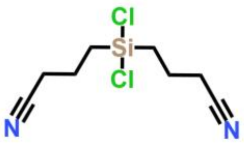

No. 81
CAS: 18290-58-9


No. 82
CAS: 1009-46-7

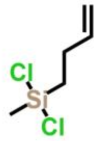

No. 83
CAS: 18147-23-4


No. 84
CAS: 71550-62-4

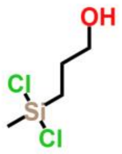
No. 85
CAS: 58066-88-9



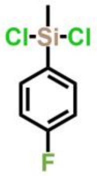
No. 86
CAS: 1071-17-6



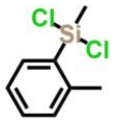
No. 87
CAS: 15983-86-5



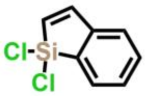
No. 88
CAS: 35153-79-8



No. 89
CAS: 1422-90-8



No. 90
CAS: 18236-56-1

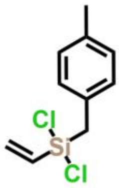


No. 91
CAS: 57831-43-3

No. 92
CAS: 57831-43-3

No. 93
CAS: 57831-43-3

No. 94
CAS: 80-10-4

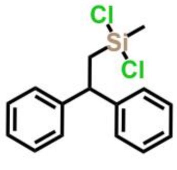


No. 95
CAS: 20151-48-8

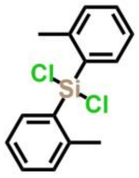
No. 96
CAS: 31101-72-1

No. 97
CAS: 17998-62-8

No. 98
CAS: 18676-65-8

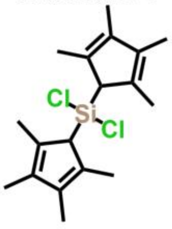


No. 99
CAS: 53888-97-4

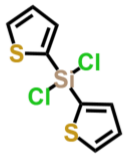


No. 100
CAS: 18414-45-4

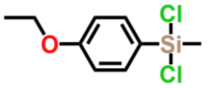
No. 101
CAS: 89486-31-7



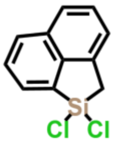
No. 102
CAS: 89597-04-6



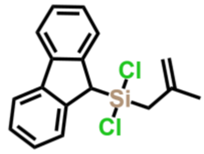
No. 103
CAS: 64102-92-7



No. 104
CAS: 114502-16-8

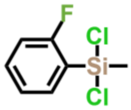


No. 105
CAS: 13093-11-3

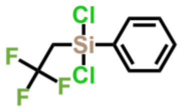


No. 106
Compound CID: 164069027

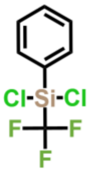
No. 107
CAS: 91920-10-4



No. 108
CAS: 91747-40-9



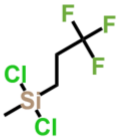
No. 109
CAS: 148337-94-4



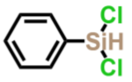
No. 110
CAS: 183790-11-6

No. 111
CAS: 1085700-82-8

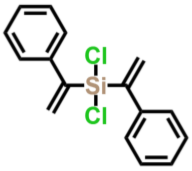
No. 112
CAS: 422-96-8



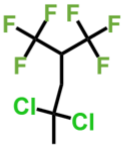
No. 113
CAS: 675-62-7



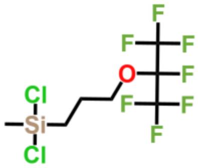
No. 114
CAS: 1631-84-1



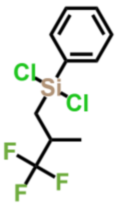
No. 115
CAS: 225940-04-5



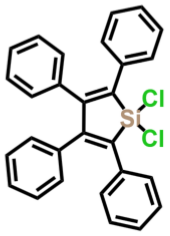
No. 116
CAS: 101931-61-7



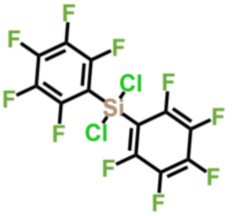
No. 117
CAS: 20006-68-2



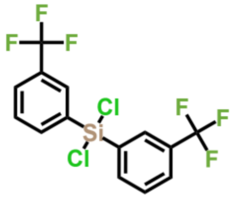
No. 118
Compound CID: 163873224



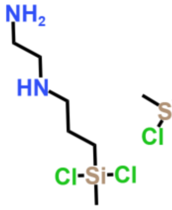
No. 119
CAS: 79343-32-1



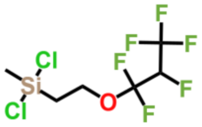
No. 120
CAS: 20160-45-6



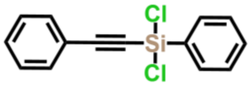
No. 121
CAS: 565-11-7



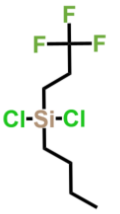
No. 122
CAS: 827627-58-7



No. 123
Compound CID:
164074177

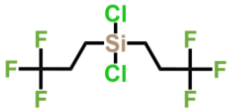


No. 124
CAS: 122056-52-4

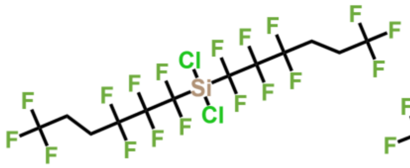


No. 125
CAS: 81347-61-7

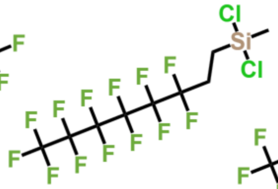
No. 126
CAS: 38436-16-7



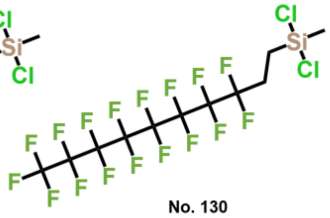
No. 127
CAS: 665-98-5



No. 128
Compound CID: 53431090



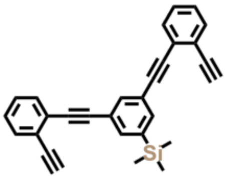
No. 129
CAS: 73609-36-6

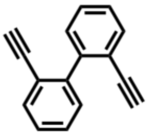


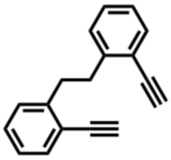
No. 130
CAS: 3102-79-2

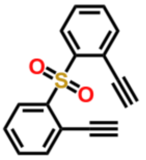
S9. Chemical structures of 287 diynes genes

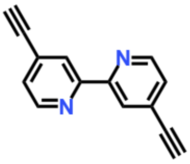
We collected 287 types of diynes genes from the open-source PubChem database.[S5] The molecular structures of the 287 types of diynes genes are shown below, where the gene number (No.), the CAS identifier or CID in the database are also provided.

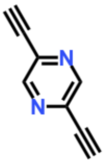

No. 1
CAS: 1132838-29-9

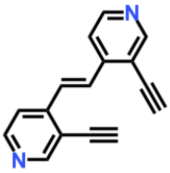

No. 2
CAS: 18442-29-0

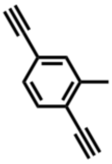

No. 3
CAS: 27499-66-7

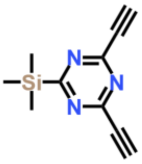

No. 4
CAS: 2417378-12-0

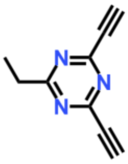

No. 5
CAS: 133810-44-3

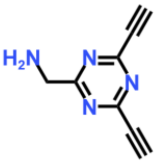

No. 6
CAS: 649774-86-7

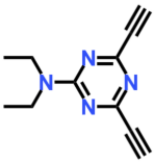

No. 7
CAS: 2417378-34-6

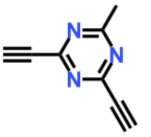

No. 8
CAS: 97009-24-0

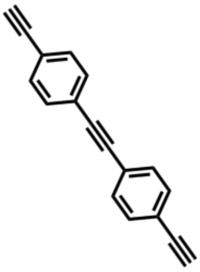

No. 9
CAS: 2417377-25-2

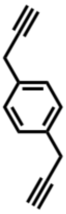

No. 10
CAS: 2417377-33-2

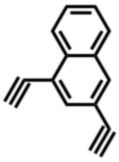

No. 11
CAS: 2417377-30-9

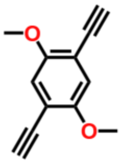

No. 12
CAS: 2417377-54-7

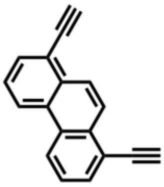

No. 13
CAS: 2417377-33-2

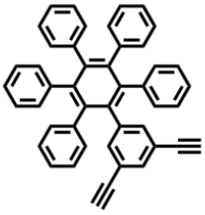

No. 14
CAS: 153295-62-6

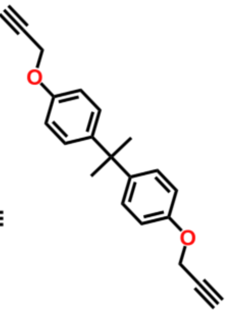

No. 15
CAS: 1350451-48-7

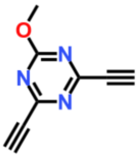

No. 16
CAS: 2417377-03-6

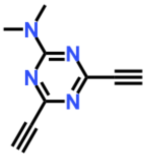

No. 17
CAS: 74029-40-6

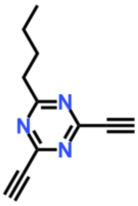

No. 18
CAS: 2417378-47-1

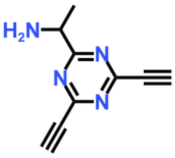

No. 19
CAS: 1203793-93-4

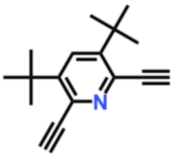

No. 20
CAS: 22235-02-5

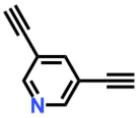

No. 21
CAS: 2417377-14-9

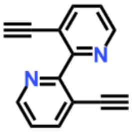

No. 22
CAS: 2417377-37-6

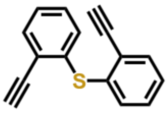

No. 23
CAS: 2417377-52-5

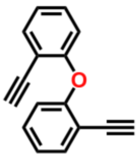

No. 24
CAS: 2417377-93-4

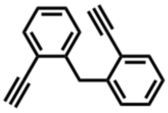

No. 25
CAS: 2417378-15-3

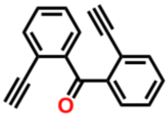

No. 26
CAS: 67227-90-1


No. 27
CAS: 2417378-40-4

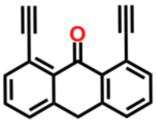

No. 28
CAS: 2417377-23-0


No. 29
CAS: 189620-06-2

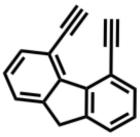

No. 30
CAS: 189620-04-0



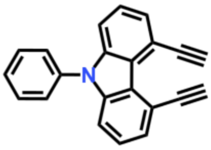
No. 31
CAS: 2417377-17-2



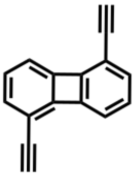
No. 32
CAS: 2417377-73-0



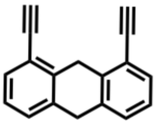
No. 33
CAS: 2417378-09-5



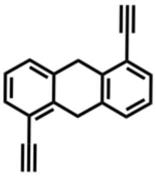
No. 34
CAS: 2417377-58-1



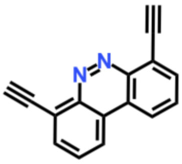
No. 35
CAS: 2417377-77-4



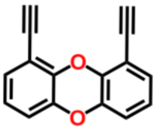
No. 36
CAS: 2417377-68-3



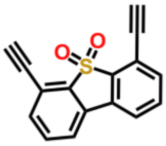
No. 37
CAS: 2417377-68-3



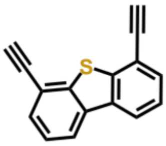
No. 38
CAS: 2417378-41-5



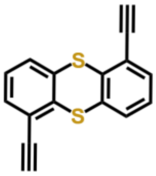
No. 39
CAS: 2417377-45-6



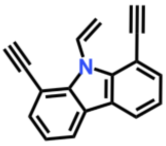
No. 40
CAS: 2417378-06-2



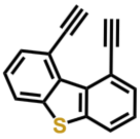
No. 41
CAS: 2409061-50-1



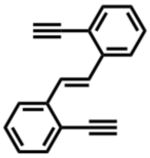
No. 42
CAS: 2417377-96-7



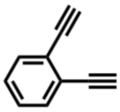
No. 43
CAS: 2417377-83-2



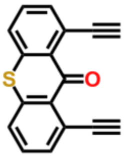
No. 44
CAS: 2417377-60-5



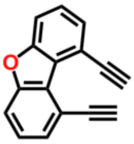
No. 45
CAS: 27559-94-0



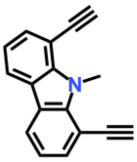
No. 46
CAS: 21792-52-9



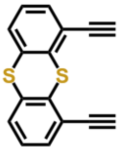
No. 47
CAS: 2417377-51-4



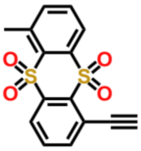
No. 48
CAS: 2417377-64-9



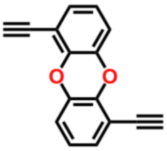
No. 49
CAS: 2417378-18-6



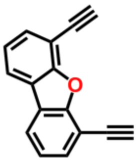
No. 50
CAS: 2417378-30-2



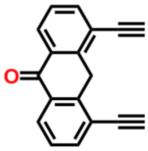
No. 51
CAS: 2417377-94-5



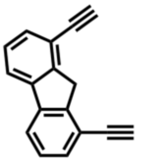
No. 52
CAS: 2417377-31-0



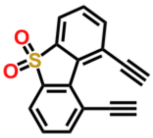
No. 53
CAS: 209254-89-7



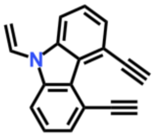
No. 54
Compound CID: 101795911



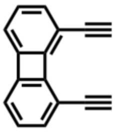
No. 55
CAS: 2417377-79-6



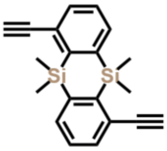
No. 56
CAS: 2417377-76-3



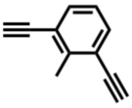
No. 57
CAS: 2417378-20-0



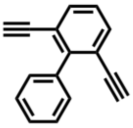
No. 58
CAS: 1397958-88-1



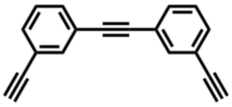
No. 59
CAS: 2417378-00-6



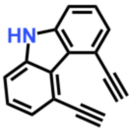
No. 60
CAS: 514790-50-2



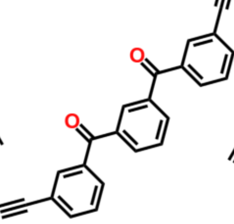
No. 61
CAS: 1293326-48-3



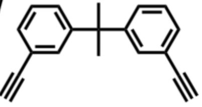
No. 62
CAS: 914673-80-6



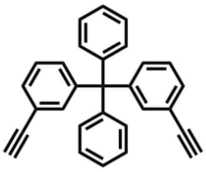
No. 63
CAS: 2417378-02-8



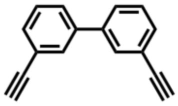
No. 64
CAS: 2417378-29-9



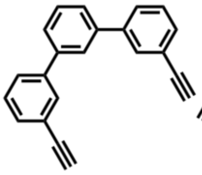
No. 65
CAS: 2417377-19-4



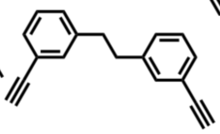
No. 66
CAS: 2417377-61-6



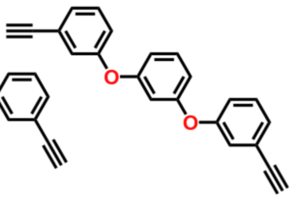
No. 67
CAS: 1266361-97-0



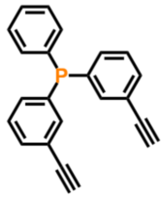
No. 68
CAS: 1404221-85-7



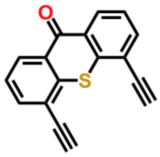
No. 69
CAS: 2417377-26-3



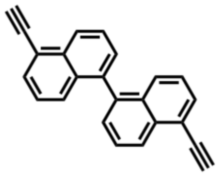
No. 70
CAS: 75142-99-3



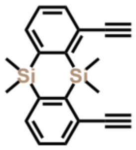
No. 71
CAS: 1240453-65-9



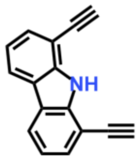
No. 72
CAS: 2417377-62-7



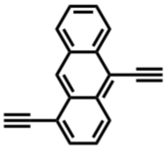
No. 73
CAS: 2417378-24-4



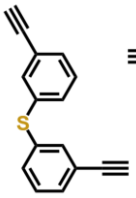
No. 74
CAS: 2417378-33-5



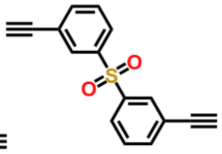
No. 75
CAS: 2093168-56-8



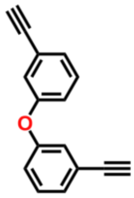
No. 76
CAS: 2417378-49-3



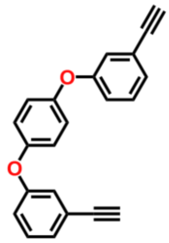
No. 77
CAS: 2417377-16-1



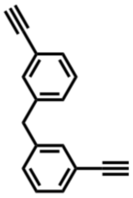
No. 78
CAS: 2266612-77-3



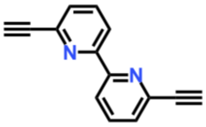
No. 79
CAS: 2097055-11-1



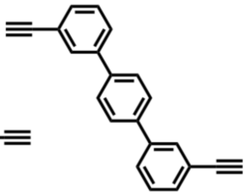
No. 80
CAS: 2417377-07-0



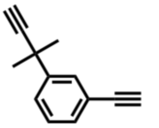
No. 81
CAS: 2097055-12-2



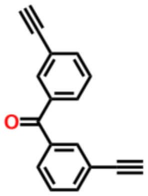
No. 82
CAS: 133810-42-1



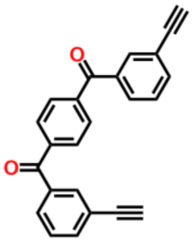
No. 83
CAS: 2417378-46-0



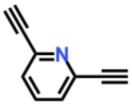
No. 84
CAS: 2229630-09-3



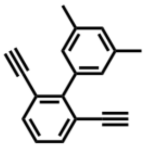
No. 85
CAS: 130249-12-6



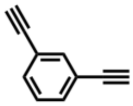
No. 86
CAS: 2417377-20-7



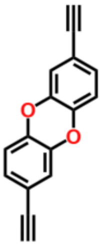
No. 87
CAS: 75867-46-8



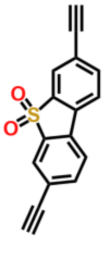
No. 88
CAS: 2149634-67-1



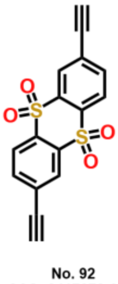
No. 89
CAS: 1785-61-1



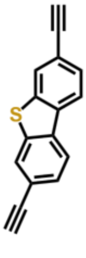
No. 90
CAS: 2417377-53-6



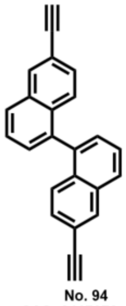
No. 91
CAS: 1426055-67-5



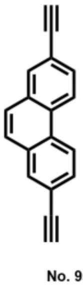
No. 92
CAS: 2417378-03-9



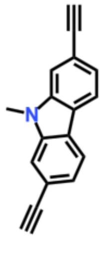
No. 93
CAS: 1629749-70-7



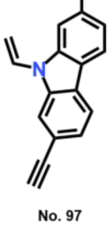
No. 94
CAS: 2417378-11-9



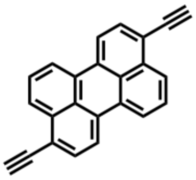
No. 95
CAS: 2417378-11-9



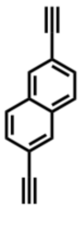
No. 96
CAS: 2417377-90-1



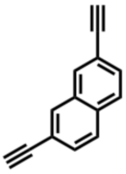
No. 97
CAS: 2417378-27-7



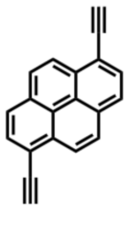
No. 98
CAS: 2126816-74-6



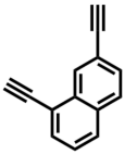
No. 99
CAS: 507276-82-6



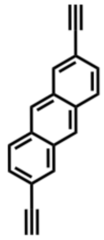
No. 100
CAS: 113705-27-4



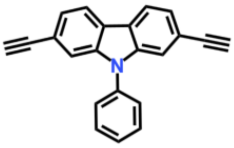
No. 101
CAS: 173678-77-8



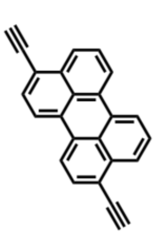
No. 102
CAS: 147672-94-4



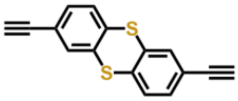
No. 103
CAS: 960223-15-8



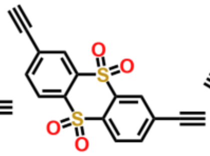
No. 104
CAS: 2417377-65-0



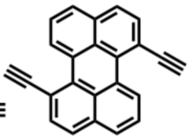
No. 105
CAS: 2126816-75-7



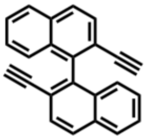
No. 106
CAS: 2417377-80-9



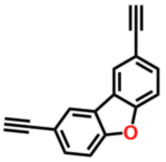
No. 107
CAS: 2417377-85-4



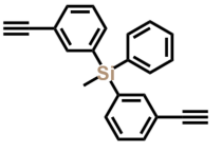
No. 108
CAS: 2417378-23-3



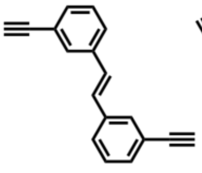
No. 109
CAS:765911-39-5



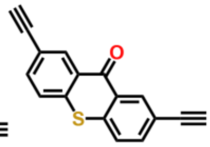
No. 110
CAS: 2348521-77-5



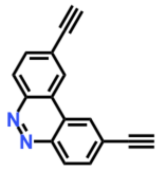
No. 111
CAS:2417377-13-8



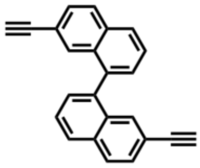
No. 112
CAS: 1610942-28-3



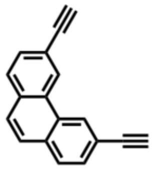
No. 113
CAS: 2417377-48-9



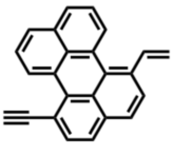
No. 114
CAS: 2417378-39-1



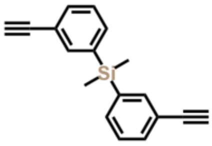
No. 115
CAS:2417378-14-2



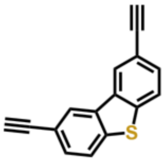
No. 116
CAS: 918778-82-2



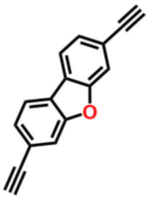
No. 117
CAS: 227079-98-3



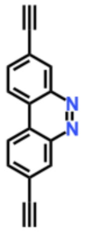
No. 118
CAS:2417377-21-8



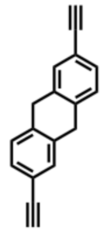
No. 119
CAS: 951791-07-4



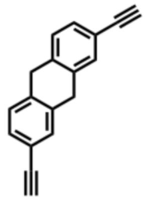
No. 120
CAS: 1629749-69-4



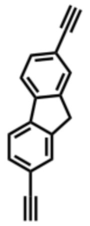
No. 121
CAS: 2417378-35-7



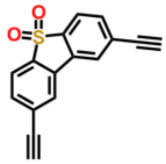
No. 122
CAS: 2417377-59-2



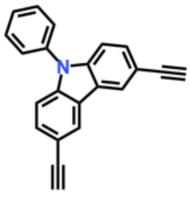
No. 123
CAS:627892-84-6



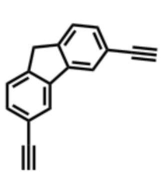
No. 124
CAS: 94463-11-3



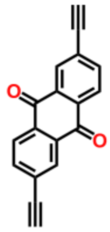
No. 125
CAS: 951791-09-6



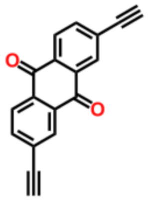
No.126
CAS: 530097-70-2



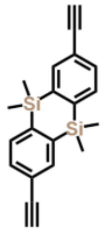
No. 127
CAS:627892-82-4



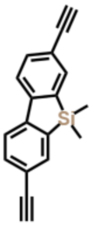
No. 128
CAS:914477-80-8



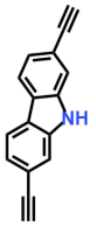
No. 129
CAS:1276668-76-8



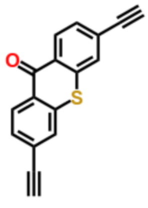
No. 130
CAS:2417377-95-6



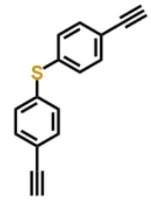
No. 131
CAS: 2417377-95-6



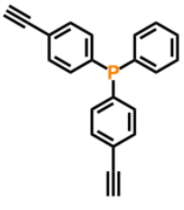
No. 132
CAS: 909342-66-1



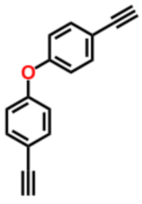
No. 133
CAS: 2417377-70-7



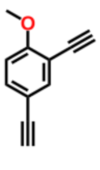
No. 134
CAS:27594-21-4



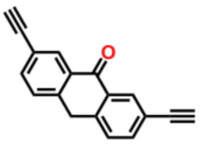
No. 135
CAS: 1240453-62-6



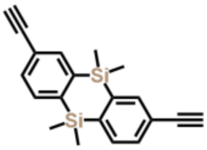
No. 136
CAS: 21368-80-9



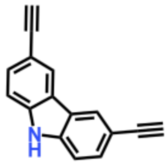
No. 137
CAS: 30427-54-4



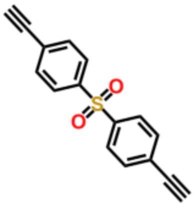
No. 138
CAS: 2417377-56-9



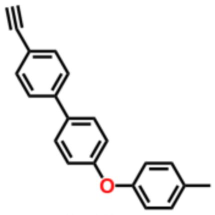
No. 139
CAS2417377-81-0



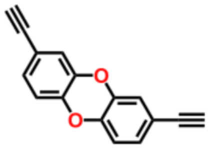
No. 140
CAS: 909342-65-0



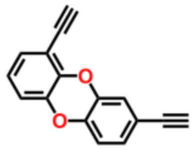
No. 141
CAS: 27594-30-5



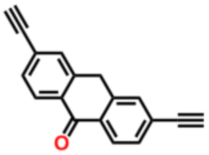
No. 142
CAS: 7535-21-9



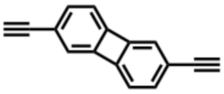
No. 143
CAS: 2417377-55-8



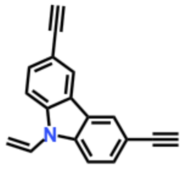
No. 144
CAS: 2417377-43-4



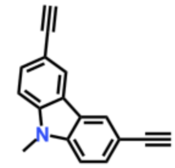
No. 145
CAS: 2417377-57-0



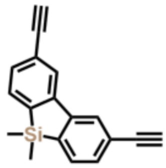
No. 146
CAS:2417378-05-1



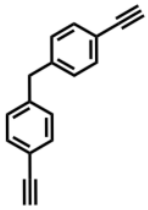
No. 147
CAS: 2417377-78-5



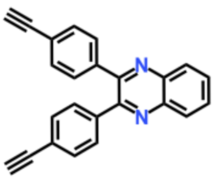
No. 148
CAS: 357437-78-6



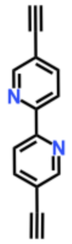
No. 149
CAS: 2417377-22-9



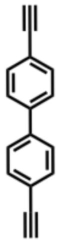
No. 150
CAS: 6140-83-6



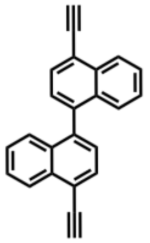
No. 151
CAS: 2417378-38-0



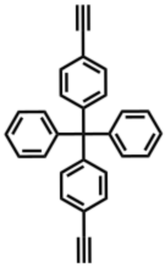
No. 152
CAS: 162318-26-5



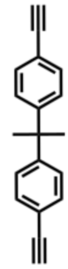
No. 153
CAS: 38215-38-2



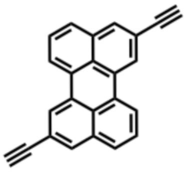
No. 154
CAS: 2417378-19-7



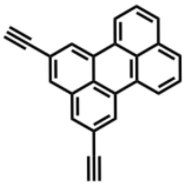
No. 155
CAS: 2417377-63-8



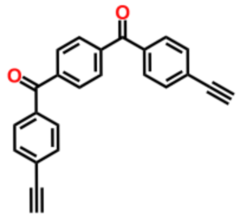
No. 156
CAS: 6140-85-8



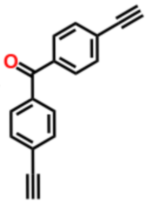
No. 157
CAS: 2417378-13-1



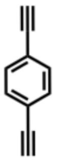
No. 158
CAS: 2417378-25-5



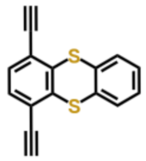
No. 159
CAS: 2417377-08-1



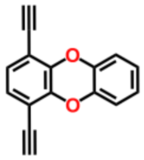
No. 160
CAS: 147613-78-3



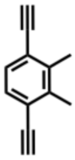
No. 161
CAS: 935-14-8



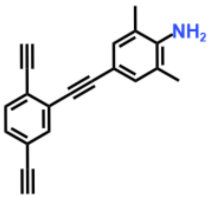
No. 162
CAS: 2417378-22-2



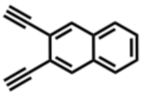
No. 163
CAS: 2417377-71-8



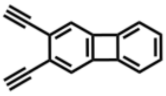
No. 164
CAS: 158599-68-9



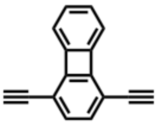
No. 165
CAS: 856000-46-9



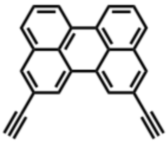
No. 166
CAS: 214406-87-8



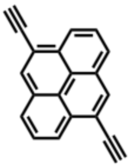
No. 167
CAS: 87710-28-9



No. 168
CAS: 2417378-26-6

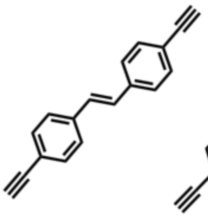


No. 169
CAS: 2417378-10-8

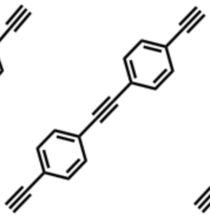


No. 170
CAS: 2417378-43-7

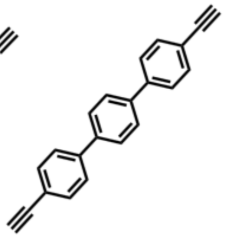
No. 171
CAS: 258531-35-0



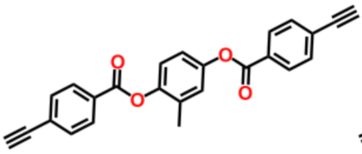
No. 172
CAS: 497140-53-1

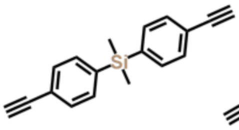


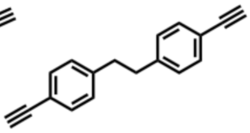
No. 173
CAS: 153295-62-6

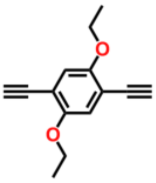


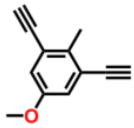
No. 174
CAS: 47230-46-6

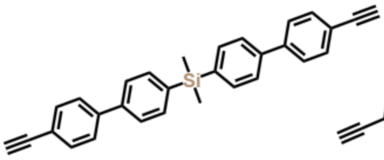

No. 175
CAS: 145289-67-4

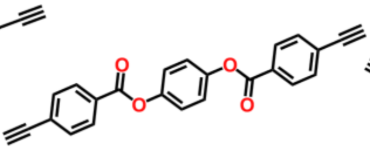

No. 176
CAS: 638164-82-6

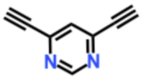

No. 177
CAS: 6128-15-0

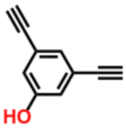

No. 178
CAS: 1126410-21-6

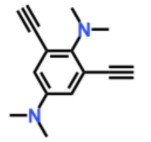

No. 179
CAS: 2417377-74-1

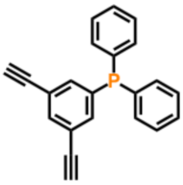

No. 180
CAS: 2417377-92-3

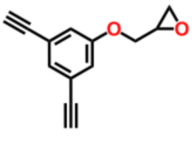

No. 181
CAS: 123233-13-6

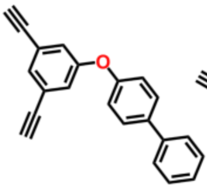

No. 182
CAS: 1018473-95-4

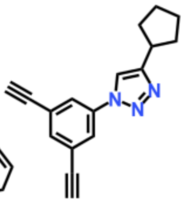

No. 183
CAS: 918826-08-1

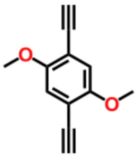

No. 184
CAS: 2417377-10-5

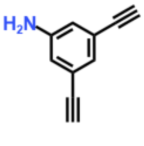

No. 185
CAS: 1047979-43-0

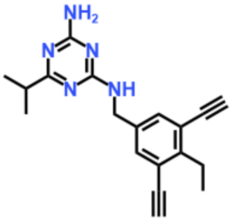

No. 186
CAS: 2229009-38-3

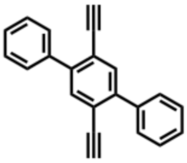

No. 187
CAS: 2417377-89-8

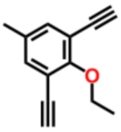

No. 188
CAS: 1273553-78-8

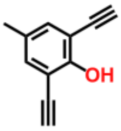

No. 189
CAS: 74029-40-6

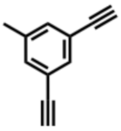

No. 190
CAS: 402956-36-9

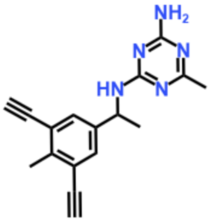

No. 191
CAS: 1142094-64-1

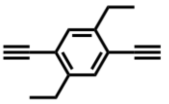

No. 192
CAS: 871248-64-5

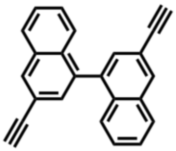

No. 193
CAS: 463296-71-1

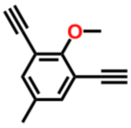

No. 194
CAS: 1245255-95-1

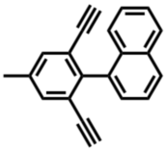

No. 195
CAS: 165047-95-0

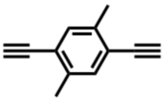

No. 196
CAS: 1141845-48-8

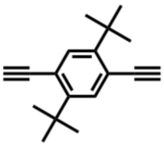

No. 197
CAS: 81833-17-2

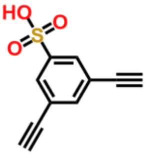

No. 198
CAS: 2417378-16-4

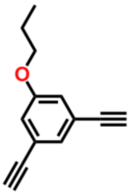

No. 199
CAS: 1612866-70-2

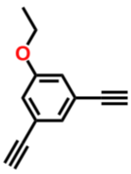

No. 200
CAS: 1636150-03-2

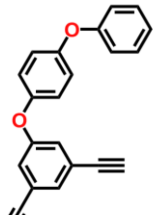

No. 201
CAS: 75867-45-7

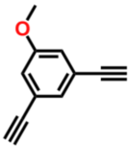

No. 202
CAS: 629165-10-2

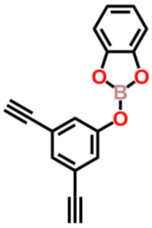

No. 203
CAS: 2058051-39-9

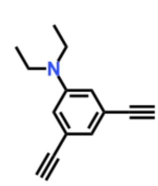

No. 204
CAS: 2251049-43-9


No. 205
CAS: 2417377-18-3

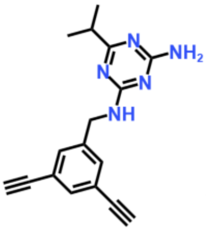

No. 206
CAS: 2417377-98-9


No. 207
CAS: 1542157-70-9

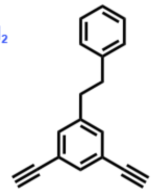

No. 208
CAS: 2417377-39-8



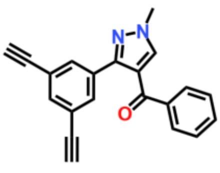
No. 209
CAS: 852944-44-6



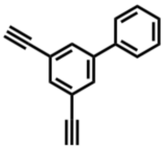
No. 210
CAS: 1142094-58-3



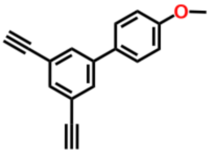
No. 211
CAS: 2417377-50-3



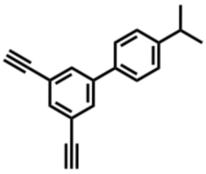
No. 212
CAS: 1853345-89-7



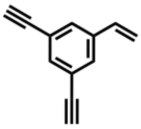
No. 213
CAS: 1623191-70-7



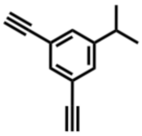
No. 214
CAS: 1778690-09-7



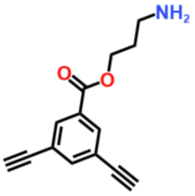
No. 215
CAS: 2417378-31-3



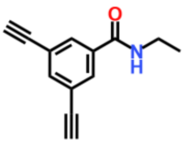
No. 216
CAS: 1220113-59-6



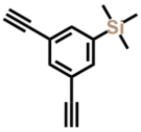
No. 217
CAS: 2417377-97-8



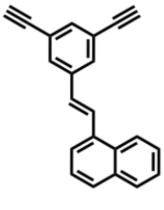
No. 218
CAS: 2229009-43-0



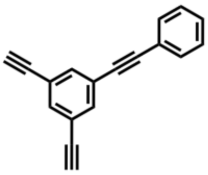
No. 219
CAS: 2058051-41-3



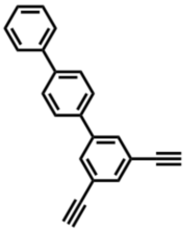
No. 220
CAS: 127182-87-0



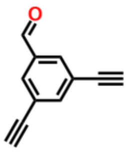
No. 221
CAS: 2417378-37-9



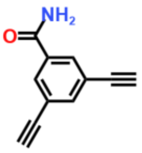
No. 222
CAS: 1669445-40-2



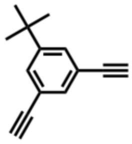
No. 223
CAS: 2417378-44-8



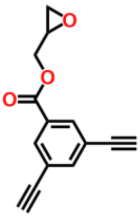
No. 224
CAS: 153390-76-2



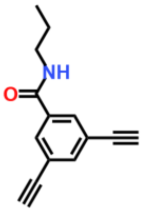
No. 225
CAS: 2058051-58-2



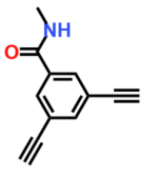
No. 226
CAS: 524728-78-7



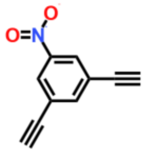
No. 227
CAS: 2229009-32-7



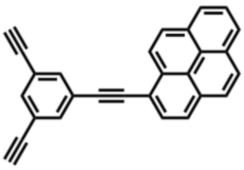
No. 228
CAS: 2417378-08-4



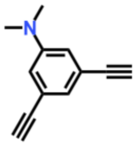
No. 229
CAS: 2058051-40-2



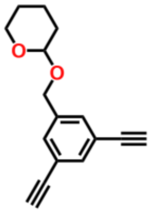
No. 230
CAS: 756499-81-7



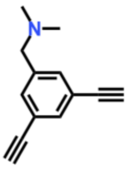
No. 231
CAS: 1431045-48-5



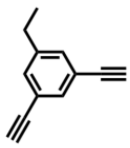
No. 232
CAS: 2417377-09-2



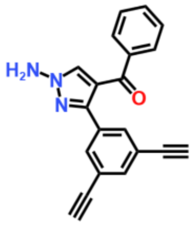
No. 233
CAS: 2088280-03-7



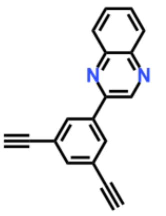
No. 234
CAS: 2249739-99-7



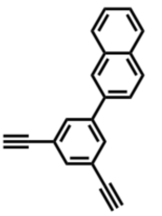
No. 235
CAS: 2417377-29-6



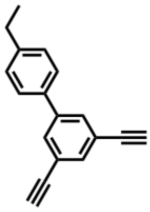
No. 236
CAS: 1853764-46-1



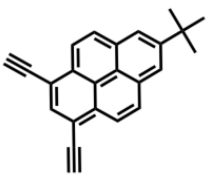
No. 237
CAS: 1486564-39-9



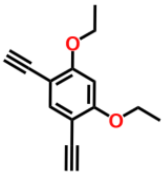
No. 238
CAS: 2417378-42-6



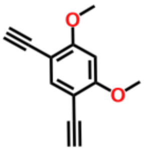
No. 239
CAS: 2417378-31-3



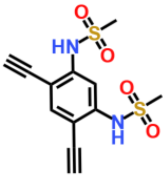
No. 240
CAS: 1613036-64-8



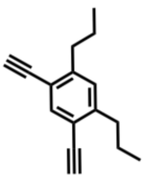
No. 241
CAS: 19300-10-8



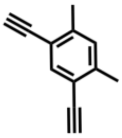
No. 242
CAS: 19300-09-5



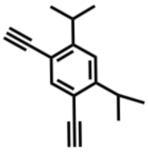
No. 243
CAS: 1159200-18-6



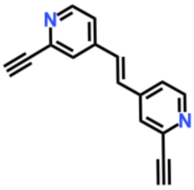
No. 244
CAS: 2417377-69-4



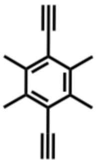
No. 245
CAS: 1379822-09-9



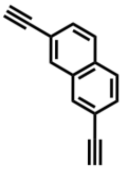
No. 246
CAS: 2417378-17-5



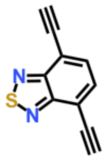
No. 247
CAS: 2417378-36-8



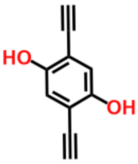
No. 248
CAS: 81833-15-0



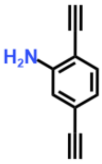
No. 249
CAS: 113705-27-4



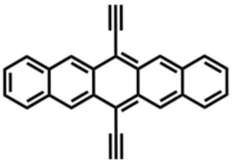
No. 250
CAS: 316384-85-7



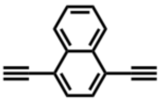
No. 251
CAS: 75610-48-9



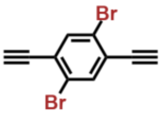
No. 252
CAS: 137000-70-5



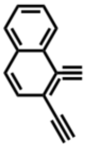
No. 253
CAS: 427879-51-4



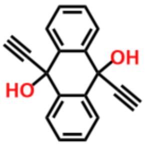
No. 254
CAS: 64856-90-2



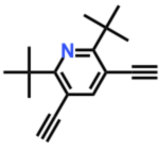
No. 255
CAS: 478190-80-6



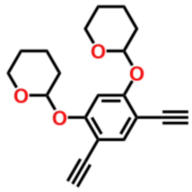
No. 256
CAS: 1600521-75-2



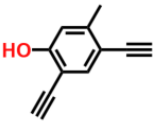
No. 257
CAS: 47031-58-3



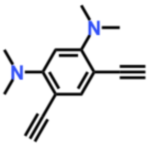
No. 258
CAS: 2417378-32-4



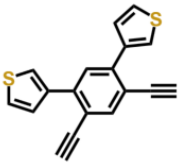
No. 259
CAS: 1224886-54-7



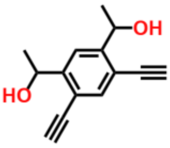
No. 260
CAS: 32229-04-2



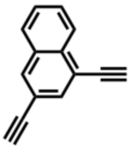
No. 261
CAS: 2417377-44-5



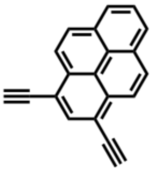
No. 262
CAS: 2191388-74-4



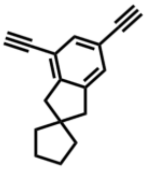
No. 263
CAS: 1231998-91-6



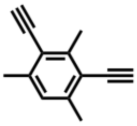
No. 264
CAS: 2417377-03-6



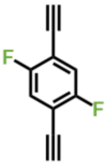
No. 265
CAS: 173678-78-9



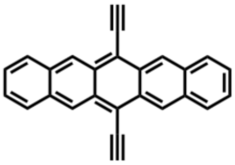
No. 266
CAS: 1615221-10-7



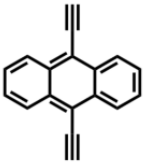
No. 267
CAS: 54407-99-7



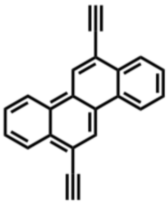
No. 268
CAS: 156016-23-8



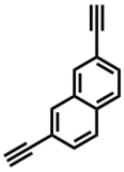
No. 269
CAS: 427879-51-4



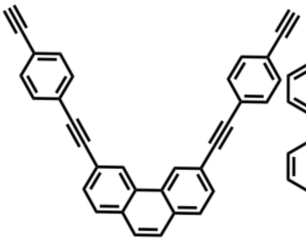
No. 270
CAS: 18512-55-5



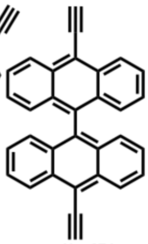
No. 271
CAS: 225225-99-0



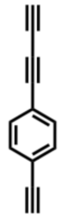
No. 272
CAS: 113705-27-4



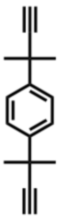
No. 273
CAS: 918778-87-7



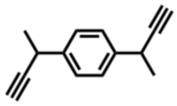
No. 274
CAS: 1429231-42-4



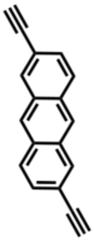
No. 275
CAS: 79109-92-5



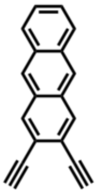
No. 276
CAS: 2417378-21-1



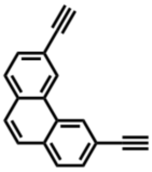
No. 277
CAS: 440121-70-0



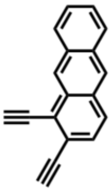
No. 278
CAS: 960223-15-8



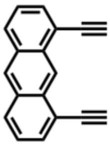
No. 279
CAS: 256439-93-7



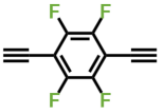
No. 280
CAS: 918778-82-2



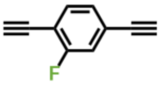
No. 281
CAS: 256439-93-7



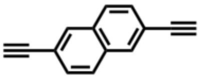
No. 282
CAS: 78053-58-4



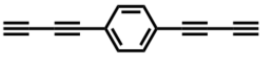
No. 283
CAS: 38002-32-3



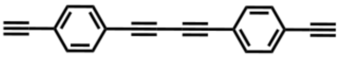
No. 284
CAS: 137000-66-9



No. 285
CAS: 507276-82-6



No. 286
CAS: 141135-35-5

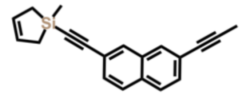
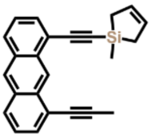
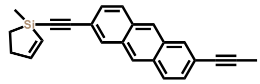
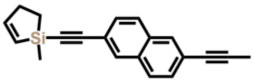
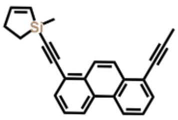
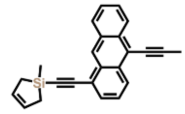
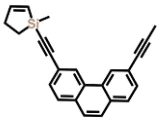
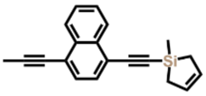
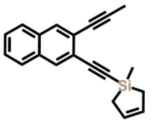


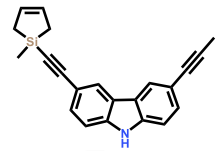
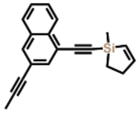
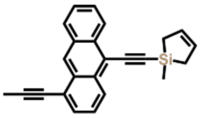
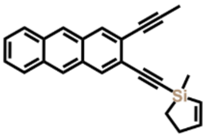
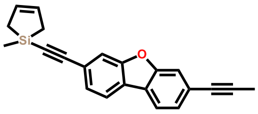
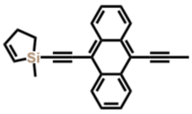
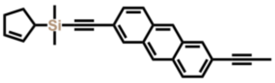
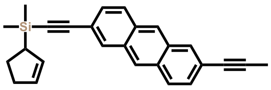
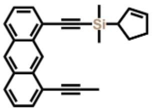
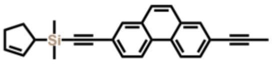
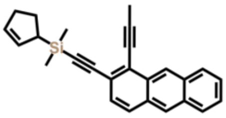
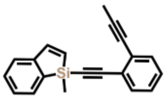
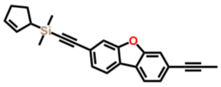
No. 287
CAS: 124417-94-3

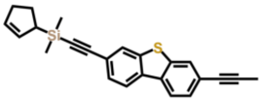
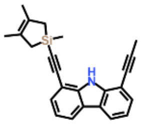
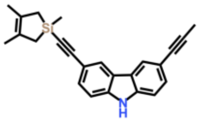
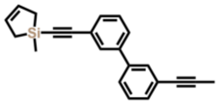
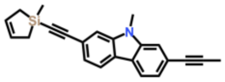
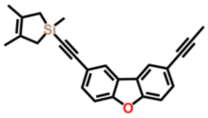
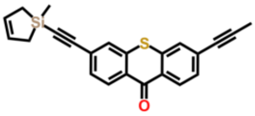
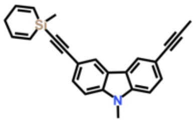
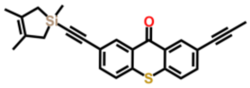
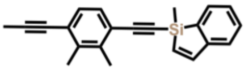
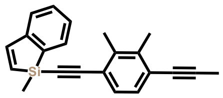
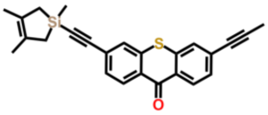
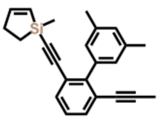
S10. The list of the promising PSAs for wave-transparent application

To provide various synthetic options for the experimentalists, the screened PSAs containing genes with significant structural differences are considered as promising PSAs. Here, the list of the total 35 promising PSA structures for wave-transparent materials is given in Table S3 (the 5 representative PSAs listed in Table 1 are included).

Table S3. The list of the promising PSAs for wave-transparent application*

ID	Structure	*ε*	tan*δ*	T_{d5} / °C	*ScoreT*	*ScoreF*
1		2.39	0.0016	665	0.18	0.19
2		2.33	0.0015	662	0.14	0.20
3		2.35	0.0020	661	0.12	0.19
4		2.40	0.0020	661	0.20	0.16
5		2.30	0.0015	659	0.20	0.18
6		2.31	0.0015	658	0.15	0.21
7		2.34	0.0018	658	0.09	0.23
8		2.35	0.0011	658	0.24	0.21
9		2.38	0.0012	658	0.20	0.21

ID	Structure	ε	tanδ	T_{d5} / °C	*ScoreT*	*ScoreF*
10		2.27	0.0021	657	0.24	0.20
11		2.39	0.0017	654	0.24	0.17
12		2.29	0.0014	652	0.16	0.21
13		2.32	0.0017	649	0.10	0.21
14		2.35	0.0016	640	0.22	0.20
15		2.30	0.0014	630	0.18	0.22
16		2.28	0.0020	629	0.14	0.24
17		2.28	0.0020	629	0.15	0.24
18		2.24	0.0018	623	0.23	0.22
19		2.30	0.0019	622	0.13	0.24
20		2.26	0.0017	621	0.19	0.24
21		2.40	0.0016	618	0.13	0.24
22		2.35	0.0021	611	0.23	0.21

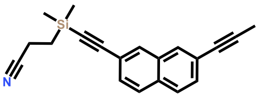
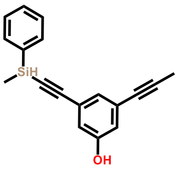
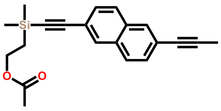
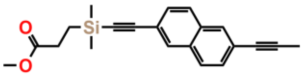
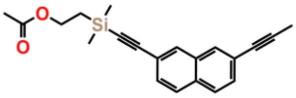
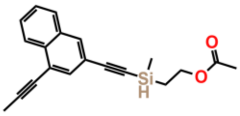
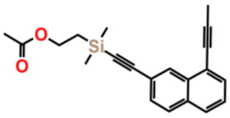
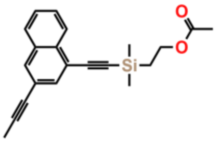
ID	Structure	ε	tanδ	T_{d5} / °C	*ScoreT*	*ScoreF*
23		2.26	0.0020	602	0.16	0.25
24		2.25	0.0016	593	0.22	0.20
25		2.31	0.0017	593	0.13	0.20
26		2.37	0.0019	593	0.18	0.24
27		2.30	0.0014	592	0.19	0.22
28		2.38	0.0013	589	0.11	0.23
29		2.40	0.0012	588	0.21	0.24
30		2.39	0.0014	578	0.16	0.22
31		2.39	0.0015	563	0.13	0.24
32		2.26	0.0007	562	0.22	0.23
33		2.27	0.0007	562	0.23	0.24
34		2.38	0.0015	561	0.13	0.24
35		2.37	0.0015	560	0.21	0.19

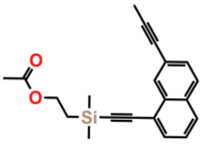
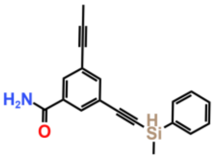
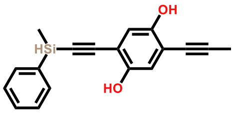
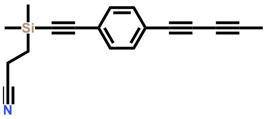
* The values of ε and tanδ given in the table were predicted at 25°C and 1kHz.

S11. The list of the promising PSAs for thin-film capacitor application

For the 36 screened PSAs with high ε, low tanδ, and high T_{d5}, we evaluated their dielectric property stabilities on temperature and frequency. Following the screening criteria of $ScoreT \leq 0.3$ and $ScoreF \leq 0.3$, 12 promising PSA structures were identified to exhibit the application potential as the resin matrix in the field of thin-film capacitors. These promising PSAs are given in Table S4 (the 5 representative PSAs listed in Table 1 are included).

Table S4. The list of the promising PSAs for thin-film capacitor application*

ID	**Structure**	ε	**tanδ**	T_{d5} / °C	$ScoreT$	$ScoreF$
1		3.74	0.0009	613	0.29	0.18
2		3.78	0.0003	595	0.29	0.27
3		3.79	0.0019	579	0.13	0.28
4		3.73	0.0016	576	0.16	0.29
5		3.78	0.0018	575	0.13	0.28
6		3.74	0.0016	572	0.18	0.29
7		3.77	0.0017	571	0.18	0.28
8		3.74	0.0016	568	0.18	0.30

ID	**Structure**	ε	**tanδ**	T_{d5} / °C	***ScoreT***	***ScoreF***
9		3.77	0.0017	567	0.18	0.29
10		3.86	0.0004	558	0.13	0.30
11		3.95	0.0014	547	0.15	0.20
12		3.80	0.0021	547	0.30	0.19

* The values of ε and tanδ given in the table were predicted at 25°C and 1kHz.

S12. Synthesis and characterizations of PSNP-MV resin

To validate the reliability of the structural design method for heat-resistant PSAs with enhanced dielectric performance, we conducted experimental verifications. Taking into account the synthesis accessibility, we chose one kind of PSA resin containing 2,7 diethynylnaphthalene and dichloromethylvinylsilane, named PSNP-MV. The predicted ε and tanδ of the PSNP-MV resin under the condition of 25°C and 1kHz are 2.97 and 0.00075, ranking at the top 2.5% and top 50% of the low ε and tanδ for the PSA candidates, respectively. In addition, the corresponding T_{d5} predicted by the ML model is 673 °C, ranking at the top 0.02% of the high heat resistance for the candidates. Regarding the dielectric property stability, the predicted values of *ScoreT* and *ScoreF* for the PSNP-MV resin are 0.22 and 0.28, which satisfies the screening criteria of *ScoreT* $\leq$ 0.3 and *ScoreF* $\leq$ 0.3.

The 2,7-diethynylnaphthalene monomer was synthesized by the Sonoga-Shira coupling method.[S6] Then, the PSNP-MV was prepared through a Grignard reaction between dichloromethylvinylsilane and 2,7-diethynylnaphthalene, as shown in Figure S7a. For example, ethyl magnesium bromide (37.5 mL, 37.5 mmol) was added to the three-necked flask with the stirring turned on, and then 2,7-diethynyl naphthalene (3.0 g, 17.04 mmol) in THF (30 mL) was slowly dripped through the funnel in an ice/water bath. The mixture was heated to 70 °C and stirred for 2 hours to produce 2,7-diethynylnaphthalene Grignard reagents. The reaction system was cooled to 20 °C, and then a solution of dichloromethylvinylsilane (1.60 g, 11.36 mmol) in THF (30 mL) was added. The reaction was then refluxed at 70 °C for 2 hours. After cooling down, 8% aqueous hydrochloric acid (50 mL) and methyl tert-butyl ether (100 mL) were slowly added to the flask. After that, the mixture was washed to neutral with deionized water, and then the solution was dried with anhydrous sodium sulfate. Finally, the solvent was distilled off to get a liquid resin (*i.e.*, PSNP-MV) with a yield of 92% (3.47 g).

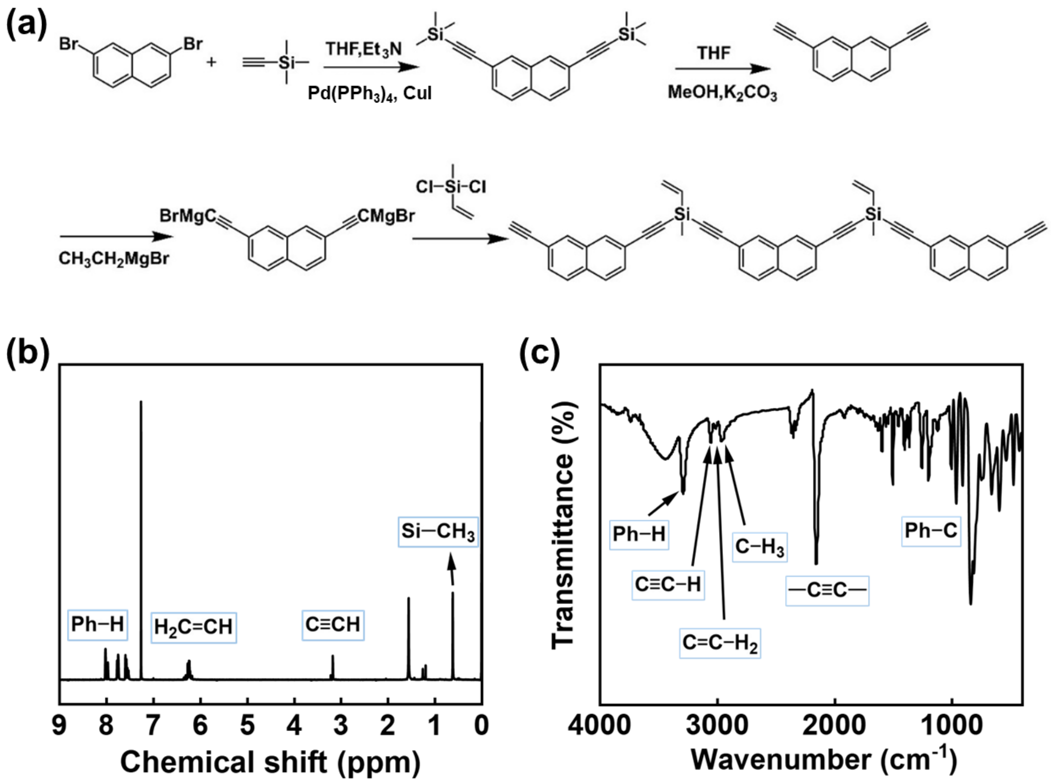


Figure S7. (a) Synthesis route of PSNP-MV. The PSNP-MV is prepared through a Grignard reaction between dichloromethylvinylsilane and 2,7-diethynylnaphthalene. (b) ^{1}H NMR ($CDCl_3$, ppm) spectra and (c) FTIR spectra of PSNP-MV.

The chemical structures of PSNP-MV were characterized by ^{1}H NMR and FTIR spectra. Figure S7b shows the ^{1}H NMR ($CDCl_3$, ppm) spectra of PSNP-MV. The resonance peaks at 7.57–8.02 ppm are assigned to the aromatic protons (Ph-H), and the peak at 3.15 ppm is caused by the terminal acetylene groups (C≡C-H). The vinyl protons on silicon and the silicon methyl protons resonate at 6.20–6.28 ppm and 0.71 ppm, respectively. The ratio of the relative peak integral area is close to the design value, indicating that the designed structure was synthesized. In the present work, the number of repeating units of PSNP-MV is 2.

Figure S7c shows the FTIR spectra of PSNP-MV. The absorption at 3292 cm^{-1} is attributed to the vibration of the C≡C-H group. There is a characteristic absorption peak at 3014 cm^{-1} for the stretching vibration of the C=C-H_2 bond. The characteristic peak at 2155 cm^{-1} belongs to the stretching vibration of the -C≡C- bond. The bands of 1593 cm^{-1} and 1502 cm^{-1} are stretching vibrations of the naphthalene ring. The above results indicate that the spectra meet the designed structure of PSNP-MV.

The thermal curing process of PSNP-MV was carried out in air conditions, involving 150 °C for 4 hours, 180 °C for 2 hours, 200 °C for 2 hours, and 230 °C for 2 hours. The thermal stability of the cured PSNP-MV resin was examined by a thermal gravimetric analyzer under a nitrogen atmosphere. Figure S8 shows the TGA curve of the cured PSNP-MV resin. The PSNP-MV thermoset has a T_{d5} of 671 °C, exhibiting excellent thermal stability. The experimental T_{d5} value of 671 °C is close to the predicted value of 673 °C, indicating the reliability of the T_{d5} prediction model.

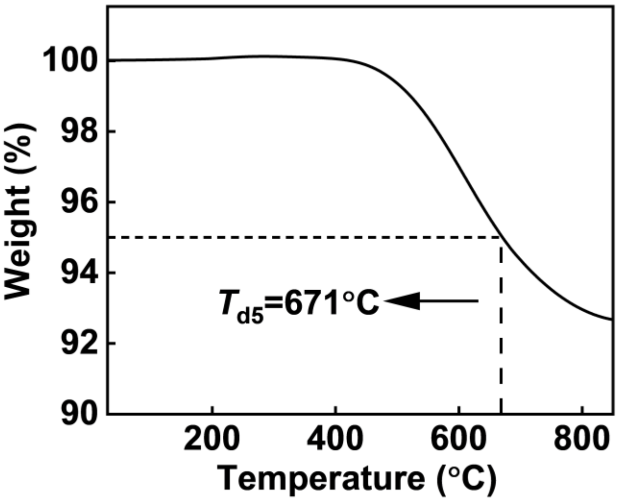


Figure S8. Thermo-gravimetric analysis curve of the cured PSNP-MV resin at a heating rate of 10 °C·min^{-1} under a nitrogen atmosphere.

S13. Additional experimental validation of PSA-V resin

Here, we chose another kind of PSA resin and carried out the additional experimental validation. Poly(silicon-containing arylacetylene vinyl)s (PSA-V) resin was synthesized from m-diacetylene benzene and dichloromethylvinylsilane, and its 5% thermal decomposition temperature (T_{d5}) can reach 630 °C (see Figure S9). By utilizing the established ML models, the predicted T_{d5} value of PSA-V is 642 °C, which is consistent with the experimental value. In addition, the PSA-V resin is predicted to exhibit a relatively higher dielectric constant and higher dielectric loss than the PSNP-MV resin. We examined the dielectric properties of the cured PSA-V at various temperatures and frequencies. Figure S10 shows the comparison between experimental and predicted values of the dielectric properties, where the orange symbols correspond to ε and the green symbols represent tanδ. It can be seen that at various test conditions, the experimental values of ε and tanδ are in the ranges of 3.7~4.0 and 0.005~0.01, respectively, which agrees with the predicted values.

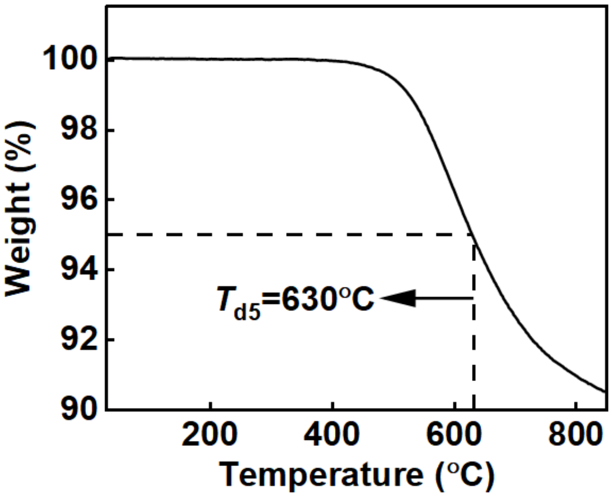


Figure S9. Thermo-gravimetric analysis curve of the cured PSA-V resin at a heating rate of 10 °C· min^{-1} under a nitrogen atmosphere.

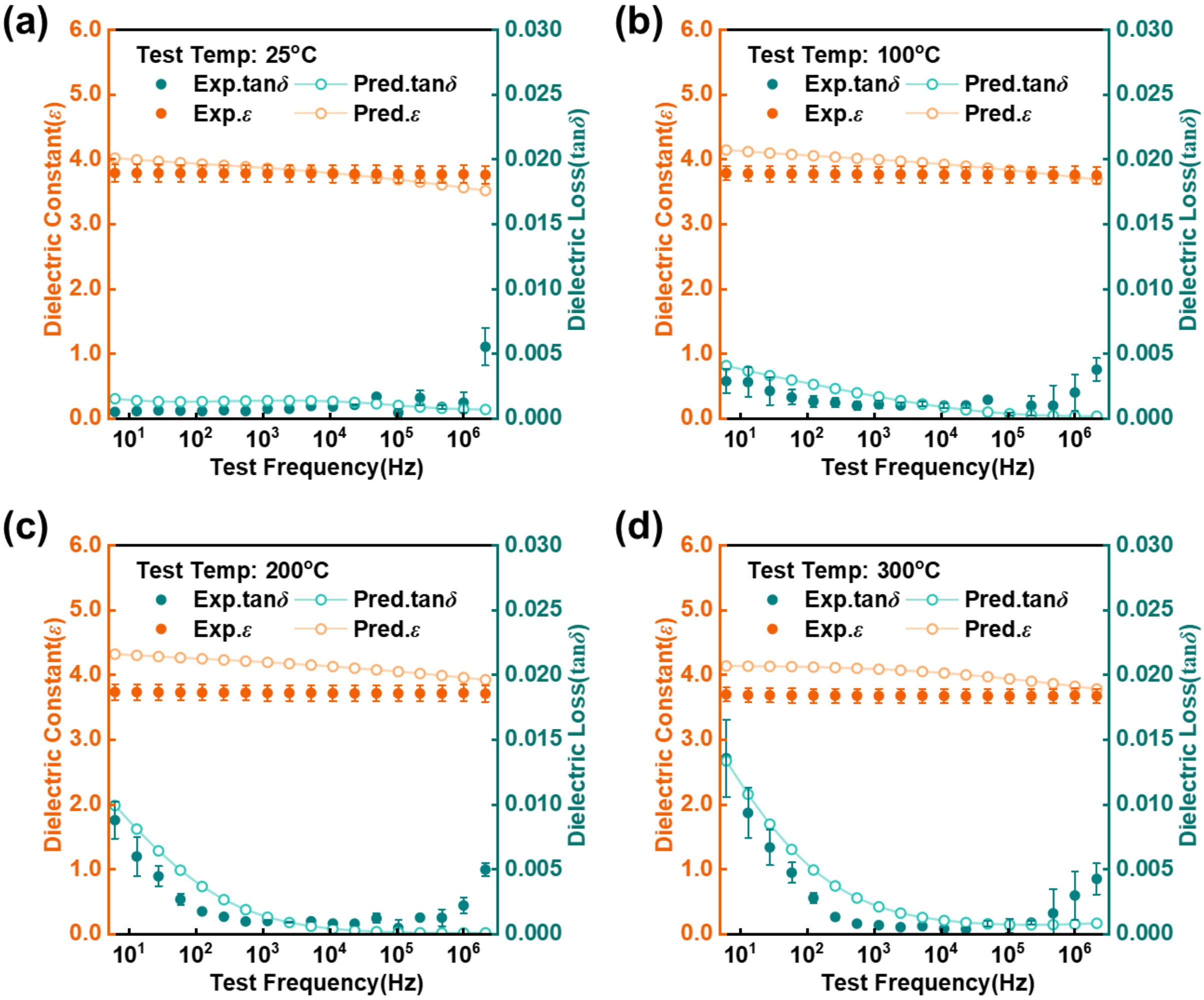


Figure S10. The variations of dielectric constant and loss with frequency at various test temperatures (a) 25 °C, (b) 100 °C, (c) 200 °C, and (d) 300 °C for the cured PSA-V resins.

References

[S1] H. Moriwaki, Y. S. Tian, N. Kawashita and T. Takagi, *J. Cheminformatics*, 2018, **10**, 4.

[S2] RDKit: Open Source Toolkit for Cheminformatics, http://www.rdkit.org/, (accessed November 2023).

[S3] R. Gurnani, D. Kamal, H. Tran, H. Sahu, K. Scharm, U. Ashraf and R. Ramprasad, *Chem. Mater.*, 2021, **33**, 7008.

[S4] H. Moriwaki, Y. S. Tian, N. Kawashita and T. Takagi, J. *Cheminformatics*, 2018, **10**, 4.

[S5] PubChem, https://pubchem.ncbi.nlm.nih.gov/, (accessed March 2023).

[S6] S. Zhang, S. Du, L. Wang, J. Lin, L. Du, X. Xu and L. Gao, *Chem. Eng. J.*, 2022, **448**, 137643.