PAPER

Check for updates

Toward sustainable polymer design: a molecular dynamics-informed machine learning approach for vitrimers

Yiwen Zheng, [ID] [a] Agni K. Biswal,[a] Yaqi Guo,[b] Prakash Thakolkaran,[b] Yash Kokane,[c] Vikas Varshney,[d] Siddhant Kumar [ID] [b] and Aniruddh Vashisth [ID] *[a]

Vitrimers represent an emerging class of sustainable polymers with self-healing capabilities enabled by dynamic covalent adaptive networks. However, their limited molecular diversity constrains their property space and potential applications. Recent developments in machine learning (ML) techniques accelerate polymer design by predicting properties and virtually screening candidates, yet the scarcity of available experimental vitrimer data poses challenges in training ML models. To address this, we leverage molecular dynamics (MD) data generated by our previous work to train and benchmark seven ML models covering six feature representations for glass transition temperature (T_g) prediction. By averaging predicted T_g from different models, the model ensemble approach outperforms individual models, allowing for accurate and efficient property prediction on unlabeled datasets. Two novel vitrimers are identified and synthesized, exhibiting experimentally validated higher T_g than existing bifunctional transesterification vitrimers, along with demonstrated healability. This work explores the possibility of using MD data to train ML models in the absence of sufficient experimental data, enabling the discovery of novel, synthesizable polymer chemistries with a wide range of desirable properties. The integrated MD–ML approach offers polymer chemists an efficient tool for accurate property prediction and designing polymers tailored to diverse applications.

Introduction

Polymers play a crucial role in a wide array of industries, including coating, microelectronics, automobile, and aerospace. However, their performance decays over time as molecular bonds break under conditions such as mechanical stress and temperature during utilization. In traditional thermosets and thermoplastic polymers, this damage is irreversible since the rupture of covalent bonds cannot be reversed. This leads to the formation of cracks and eventual material failure.[1] As a result, most end-of-life polymer products end up as waste after failure, posing significant challenges to sustainability. The inability to repair degraded polymers leads to product replacement and increases economic costs in plastic manufacturing. In addition, the lack of intrinsic ability to repair molecular damage

in polymers limits the recyclability of used thermoplastics and thermosets.

Since the introduction by Leibler et al.,[2–5] a class of healable polymers called vitrimers have gained significant attention as a potential solution to the plastic pollution issue. Combining the reprocessability of thermoplastics and the superior mechanical properties of thermosets, vitrimers can potentially reduce plastic production and costs in polymer recycling. Vitrimers feature dynamic covalent adaptive networks (CANs), which enable building blocks to attach to and detach from each other via bond exchange reactions, imparting healability and recyclability (Fig. 1a). Vitrimers can be classified based on the types of bond exchange reactions including transesterification, disulfide exchange, and imine exchange reactions.[4] However, the thermo-mechanical properties of currently available vitrimers are constrained by the limited variety of commercially available molecular building blocks (i.e., monomers) used in their synthesis. Therefore, to broaden the applications of vitrimers, it is essential to efficiently identify the available or synthesizable monomers that constitute vitrimers with desirable properties.

Traditionally, novel polymer discovery has been accomplished in a trial-and-error fashion.[6] For a given set of polymers, chemists characterize the properties of each one of them by experiments and select those with desired properties. Since

[a]Department of Mechanical Engineering, University of Washington, Seattle, WA, USA. E-mail: vashisth@uw.edu

[b]Department of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, Netherlands

[c]Department of Materials Engineering, Indian Institute of Technology Gandhinagar, Palaj, Gandhinagar, India

[d]Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH, USA
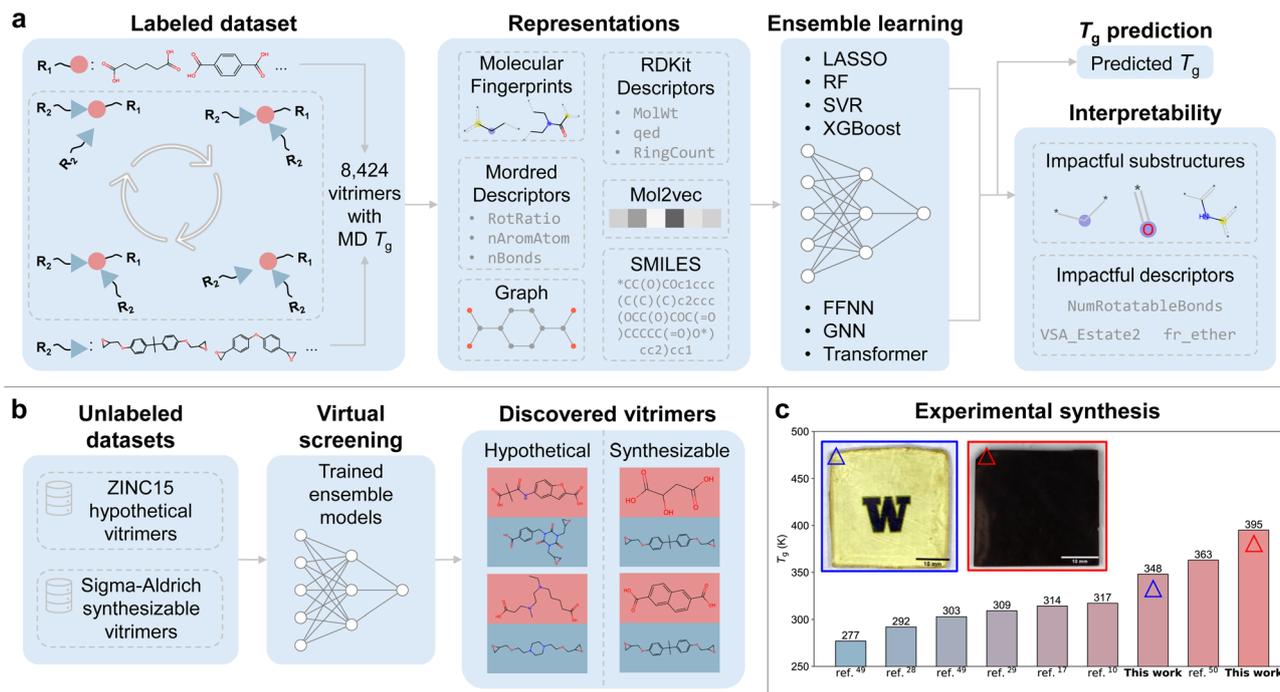
Fig. 1 Overview of this work. (a) Leveraging a previously developed vitrimer dataset, the performance of seven ML models covering six feature representation types is benchmarked. The ensemble model, which averages predictions from different models, achieves the best performance. We also provide interpretability for the ML models and identify the most impactful descriptors and substructures. (b) The trained ensemble model is applied to predict T_g of unlabeled datasets and novel vitrimers are discovered. (c) Two screened vitrimers are selected for experimental synthesis and characterization, demonstrating high T_g compared with bifunctional transesterification vitrimers present in the literature.

experimental synthesis of novel polymer chemistries is costly, molecular dynamics (MD) simulations have been favored as an alternative method to characterize polymers. MD is a simulation technique that bridges quantum mechanics and classical mechanics, and it has been extensively utilized to accelerate the discovery process.[7] MD simulations have provided valuable insights into how polymer molecular structures influence mechanical properties,[8] glass transition temperature,[9] and self-healing performance.[10] Despite these advancements, large-scale computational screening using MD or other simulation techniques remains computationally expensive even with the progress in high-performance computing.[11] Consequently, the exploration of polymer design space is typically constrained and fails to cover the vast design space of molecular compositions.

Recent developments in machine learning (ML) methods have opened a new path to accelerate polymer discovery by orders of magnitude. ML models are trained using available data to predict polymer properties (labels) from molecular structures (features). The trained models are employed to predict the properties of unlabeled polymers (whose properties are unknown) and the polymers with promising predicted properties are selected for further validation. This process is also known as virtual screening and has been applied to design polymers with various properties including glass transition temperature (T_g),[12] thermal conductivity,[13] bandgap,[14] and gas uptake properties.[15] Generally, there are three considerations when applying machine learning to polymer property prediction:

● Data: high-quality datasets are necessary for training accurate ML models. The largest polymer database PolyInfo[16] contains ~18 000 polymers with experimentally measured properties but lacks sufficient data on specific properties. For example, there are only 173 data points for thermal conductivity. As an experimental database, PolyInfo also lacks recently synthesized polymers. To address these challenges, high throughput MD simulations allow for creating a large and consistent dataset to train ML models. Researchers have established and curated polymer datasets of various properties including glass transition temperature,[17] free volume,[18] thermal conductivity,[13] and ionic conductivity.[19] ML models have been trained using these datasets for virtual screening of polymers and novel insights into structure–property relationships have been identified.

● Representations: monomers or polymer repeating units consist of atoms and bonds, which, in the original form, are not interpretable to ML models. To achieve accurate predictions, it is essential to select an appropriate way to convert molecular structures into machine-readable numerical feature representations. Molecular descriptors, which are theoretically derived values incorporating physicochemical information of molecules, serve as a commonly used representation.[20,21] Another prevalent representation is molecular fingerprints, which are vectors with substructure occurrences as entries.[22] Molecular structures can also be described using string notations such as the simplified molecular-input line-entry system (SMILES).[23] Graph-based representations have drawn increasing attention

due to the inherent graph nature of molecules, where atoms or substructures serve as nodes and bonds as edges. The selection of the best representation type is typically case-specific and requires empirical trials.

● ML models: the choice of ML models is partially dependent on the selected representation of polymers. For example, graph neural networks (GNNs) are specialized for graphs while transformers are more suitable for sequence-based input such as SMILES.[24,25] Molecular descriptors and fingerprints are more universal and accommodate both shallow and deep ML models such as linear regression, random forest, and feed-forward neural networks. Similar to representations, the choice of best ML models is also empirical.

Despite the growing applications of ML for polymers, efforts to design vitrimers using ML remain limited. Yan *et al.*[26] employed a variational autoencoder (VAE)-based virtual screening approach to identify vitrimers with desirable glass transition temperature (T_g) and recyclability. In their method, vitrimers were encoded as latent vectors using a pretrained VAE, while a feedforward neural network (FFNN) predicted T_g from these latent vectors. However, although the VAE was pretrained on ~420 000 drug-like molecules, the property predictor was trained on a small experimental dataset (389 data points for T_g) manually curated from the literature, resulting in limited predictive accuracy. In our previous work,[17] we addressed this limitation by conducting large-scale molecular dynamics (MD) simulations to calculate the T_g of 8424 hypothetical vitrimers and calibrating calculated T_g against available experimental data. This dataset, along with an unlabeled set of ~1 million hypothetical vitrimers, was used to train a VAE for the inverse design of vitrimers with targeted T_g. While the generative model successfully learned the distribution of valid vitrimers and generated novel chemistries, it did not explicitly account for synthesizability, making the proposed vitrimers difficult to synthesize. In contrast, the virtual screening approach ensures the synthesizability of designed vitrimers by restricting the search space to a predefined unlabeled dataset consisting of vitrimers derived from commercially available monomers.

This study focuses on establishing a ML model informed by MD data for accurate and efficient prediction of vitrimer T_g. The trained model is subsequently employed to identify novel bifunctional transesterification vitrimers composed of carboxylic acids and epoxides in an equimolar ratio, which is the widely adopted ratio used to synthesize transesterification vitrimers in the literature.[27-31] Leveraging open source MD-generated data from our previous study,[17] we systematically evaluate the performance of ML models in predicting vitrimer glass transition temperature (T_g). We benchmark seven ML models (linear regression, random forest, support vector regression, gradient boosting, feedforward neural network, graph neural network, and transformer) across six molecular representations, including molecular fingerprints, RDKit[20] descriptors, Mordred[21] descriptors, Mol2vec embeddings,[32] SMILES, and graphs. Our findings indicate that an ensemble learning approach, which averages predictions from multiple models, outperforms individual models (Fig. 1a). Our analysis of feature importance further provides interpretability by identifying key factors that influence the T_g of vitrimers. Using the trained ensemble model, we screen an unlabeled vitrimer dataset containing approximately 1 million hypothetical vitrimers, identifying candidates with extreme (high and low) T_g values beyond the training domain, which are further validated through MD simulations (Fig. 1b). To ensure synthesizability, we conduct an additional screening of vitrimers composed of commercially available carboxylic acids and epoxides, selecting promising candidates for experimental evaluation. Two novel vitrimers are synthesized and characterized, demonstrating experimentally higher T_g than any bifunctional transesterification vitrimer reported in the literature (Fig. 1c). The simulated and experimentally characterized T_g values validate the ability of the model to expand the property space of available vitrimers. This study demonstrates the feasibility of using MD-generated data to train ML models in the absence of available experimental data and successfully applies this approach to the design and validation of novel vitrimers with desirable properties through experimental synthesis and characterization. The proposed framework serves as an effective tool for polymer chemists to design synthesizable polymers with tailored properties for diverse applications.

Methods

Datasets

We utilize the vitrimer T_g dataset curated in our previous work,[17] which comprises 1 million hypothetical vitrimers randomly sampled from 2.5 billion possible combinations of 50 000 carboxylic acid molecules and 50 000 epoxide molecules derived from the ZINC15 database.[33] Of these, 8424 vitrimers have T_g values calculated *via* molecular dynamics (MD) simulations. A Gaussian process (GP) model is employed to calibrate the MD-calculated T_g values against experimental data from the literature. For this study, the subset of 8424 vitrimers with T_g labels serves as the training and test datasets for ML models and is referred to as the labeled dataset. The labeled dataset is randomly split into two datasets: 90% and 10% of the data are used for training and testing, respectively. The remaining vitrimers within the 1 million set are treated as the hypothetical unlabeled dataset. Further details on the datasets and MD simulations can be found in the SI.

In this work, we also curate a synthesizable dataset to facilitate the experimental synthesis of screened vitrimers. This dataset is constructed by selecting 37 commercially available bifunctional carboxylic acids and 7 bifunctional epoxides from the Sigma-Aldrich website,[34] yielding a total of 259 vitrimers

Table 1 Three vitrimer datasets used in this work

Datasets	Size	With T_g	Source
Labeled dataset	8424	Yes	ZINC15[33]
Hypothetical unlabeled dataset	991 576	No	ZINC15[33]
Synthesizable unlabeled dataset	259	No	Sigma-Aldrich[34]

that can be synthesized using off-the-shelf chemicals. The sizes and sources of all three datasets are provided in Table 1.

Feature engineering

We consider six different feature representations as inputs for machine learning models: molecular fingerprints, RDKit descriptors,[20] Mordred descriptors,[21] Mol2vec embeddings,[32] SMILES,[23] and graph-based representations. For each vitrimer in the dataset, the repeating unit is obtained by opening the epoxide rings and reacting them with carboxyl groups (see Fig. S1, SI). These repeating units are then converted into the specified feature representations, which serve as direct inputs for the ML models. Further details are discussed below.

• For molecular fingerprints, the repeating units are transformed into 2048 bit Morgan fingerprints[22,35] with a radius of 3 using the RDKit package.[20] Each bit in the fingerprint vector represents the frequency of a specific chemical substructure within the vitrimer, a method known as frequency-based molecular fingerprints. This approach has been shown to outperform traditional binary fingerprints, where each bit only indicates the presence (1) or absence (0) of a substructure.[36] We select 200 substructures that appear most frequently across the labeled dataset. Substructures with zero variance (i.e., those with identical occurrence counts across all labeled vitrimers) are removed. The final fingerprint vector consists of 195 entries, each corresponding to a specific substructure present in the vitrimers.

• Molecular descriptors for each vitrimer are computed using the RDKit[20] and Mordred[21] packages. Since certain descriptors are not well defined for some vitrimer chemistries, only those that are valid across both the labeled and hypothetical unlabeled datasets are retained. Additionally, descriptors with zero variance are removed. After filtering, 166 RDKit descriptors and 620 Mordred descriptors are selected.

• Mol2vec is an unsupervised learning-based approach that represents molecules as fixed-length numerical vectors by learning embeddings of molecular substructures.[32] In this work, we utilize a pretrained Mol2vec model to encode vitrimers into 300-dimensional vector representations, capturing their structural and chemical characteristics.

• SMILES: each vitrimer repeating unit is converted into canonical SMILES using the RDKit package,[20] where the asterisk symbol (*) denotes connection sites for the repeating unit.

• Following the approach described by Hickey et al.,[37] vitrimers are represented as molecular graphs, where heavy atoms serve as nodes and bonds as edges. Each node is characterized by atomic features including atom type, number of hydrogen atoms, valency, neighbor degree and aromaticity. Edge features include the bond type, as well as indicators for whether the bond is conjugated or part of a ring. An adjacency matrix is also extracted with binary indicators of connectivity between each pair of nodes.

Machine learning models

With vitrimer repeating units encoded into various feature representations, we proceed with training and evaluating

different ML models for T_g prediction. The schematic overview of all ML models is presented in Fig. S2 (SI). We train least absolute shrinkage and selection operator (LASSO), random forest (RF), support vector regression (SVR), extreme gradient boosting (XGBoost), and feedforward neural network (FFNN) models using molecular fingerprints, Mol2vec embeddings, RDKit descriptors, and Mordred descriptors. Graph Neural Networks (GNNs) and transformer models are trained using graph-based representations and SMILES strings, respectively. A total number of 22 models are trained and benchmarked in this work. A brief introduction to each ML model is provided in the SI. The LASSO, RF, and SVR models are implemented using the scikit-learn library,[38] while the XGBoost model is built using the XGBoost package.[39] The FFNN, GNN, and transformer models are implemented in PyTorch,[40] with the GNN and transformer architectures adapted from previous implementations by Hickey et al.[37] and Xu et al.,[41] respectively. More details on model architectures, training procedures, and hyperparameter tuning are provided in the SI. To explore possible enhancement in prediction accuracy, we also employ an ensemble learning approach by averaging the T_g values predicted by different models. This strategy has been used in previous studies for polymer property prediction.[42–44] We evaluate various combinations of all 22 ML models and identify the ensemble that yields the highest prediction performance on the test set.

The trained ML models not only enable efficient T_g prediction but also provide molecular insights into the key factors influencing vitrimer T_g. To quantify feature importance, we perform Shapley additive explanations (SHAP) analysis[45] on the trained LASSO, RF, XGBoost, and FFNN models using molecular fingerprints, RDKit descriptors, and Mordred descriptors. SVR is excluded from this evaluation due to high computational cost. For molecular fingerprints, RDKit descriptors and Mordred descriptors, we compute the average SHAP values (i.e., feature importance scores) across all considered ML models and training data. The most impactful substructures and descriptors are identified based on the highest absolute average SHAP values, providing interpretability into the factors governing vitrimer T_g.

After training all models and selecting the best-performing one, we proceed with screening the unlabeled datasets. Each vitrimer in these datasets is preprocessed using the previously discussed feature representations, and its T_g is predicted using the trained model. Promising candidates are identified for further validation. For candidates from the hypothetical dataset, MD simulations are performed to compute their T_g, following the same simulation and calibration protocols used in generating the labeled dataset (details provided in the SI). For candidates from the synthesizable dataset, two vitrimers are selected for experimental synthesis and characterization based on chemical intuition, synthesizability and cost.

Experimental synthesis and characterization

Two vitrimers are synthesized by crosslinking bisphenol A diglycidyl ether (DGEBA) with D,L-malic acid (MA) and 2,6-

naphthalenedicarboxylic acid (NA), both purchased from Sigma-Aldrich.[34] The crosslinking process involves stirring a mixture of the acids with the epoxide and 5 mol% of the catalyst triazabicyclodecene (TBD) in the presence of 5 mL of acetone. The mixture is stirred on a hot plate at temperatures ranging from 140 to 170 °C. Subsequently, the slurry is left in the beaker for 5 to 10 minutes to remove the acetone. The degassed resin is then transferred into a mold and covered with a top mold for further crosslinking in a hot press. The MA-based resin is cured at 140 °C for 5 hours, while the NA-based resin was cured at 280 °C for 6 hours under a constant pressure of 1 MPa. The glass transition temperature (T_g) of the cured samples is determined using differential scanning calorimetry (DSC) at a heating rate of 5 °C per minute.

Results

Machine learning T_g prediction

The performance of five ML models across four feature representations for T_g prediction on the test set is summarized in Fig. 2. XGBoost achieves the highest accuracy using molecular fingerprints, while LASSO outperforms other models for the remaining three representations. Among 20 (out of a total of 22) models evaluated, LASSO with Mordred descriptors (denoted as LASSO_Mordred) demonstrates the highest predictive accuracy, achieving a coefficient of determination (R^2) of 0.76 and a root mean square error (RMSE) of 15.72 K. The superior performance of LASSO_Mordred can be attributed to the high dimensionality of Mordred descriptors, which include 620 features, more than any other feature representation (195, 300, and 166 for molecular fingerprints, Mol2vec embeddings and RDKit descriptors, respectively). Additionally, LASSO incorporates an L1 regularization term, enforcing sparsity in model coefficients, thereby reducing overfitting and improving the accuracy of T_g predictions.

The performance of LASSO_Mordred is compared with that of GNN and transformer models in Fig. 3a, where it also demonstrates superior prediction accuracy over the deep learning models. To further improve T_g prediction, we implement a model ensemble approach by averaging the predicted T_g
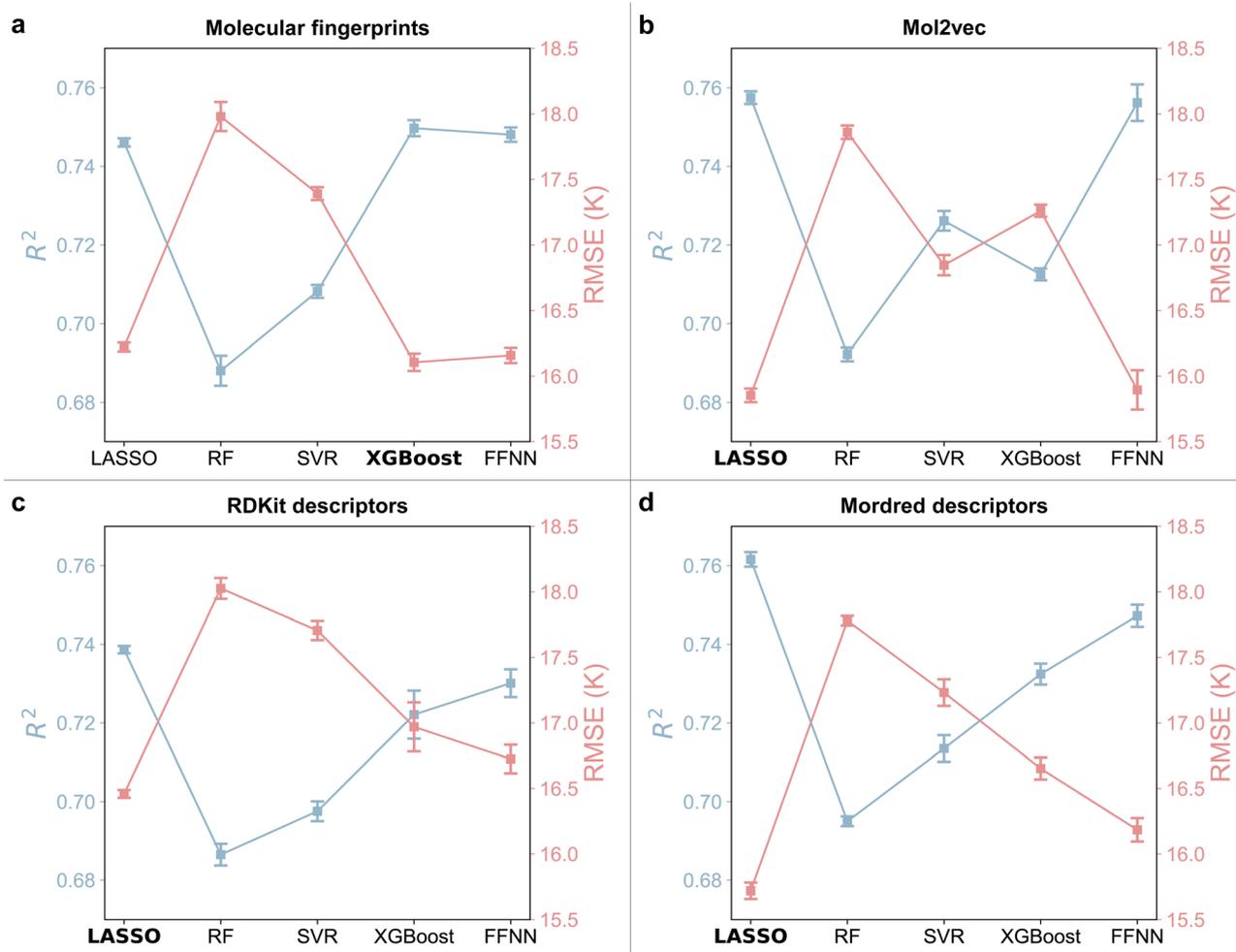

Fig. 2 Coefficient of determination (R^2) and root mean square error (RMSE) between the ground truth T_g and predicted T_g of LASSO, RF, SVR, XGBoost and FFNN models using (a) molecular fingerprints, (b) Mol2vec embeddings, (c) RDKit descriptors and (d) Mordred descriptors. The error bars indicate the standard deviation in the metrics across five-fold cross-validation. The best-performing models are highlighted in bold.
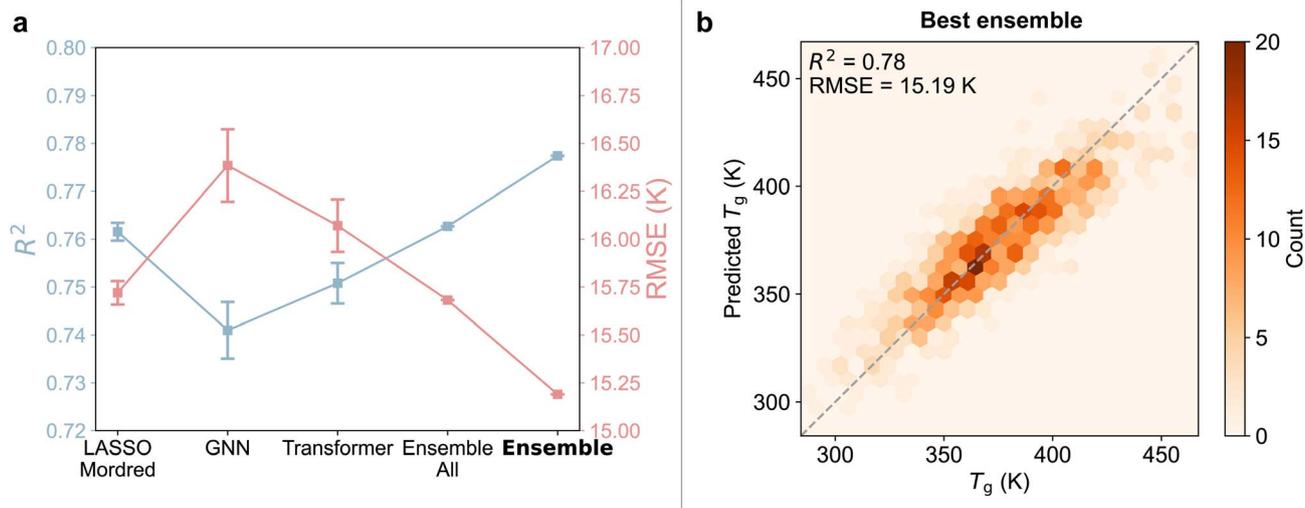
Fig. 3 Performance of other models and the ensemble model in T_g prediction. (a) R^2 and RMSE between the ground truth T_g and predicted T_g of LASSO_Mordred, GNN, transformer, the ensemble of all models and the best ensemble model. (b) Predicted T_g vs. ground truth T_g of the best ensemble model. The best-performing model is highlighted in bold.

values from multiple models. We find that an ensemble consisting of XGBoost with fingerprints (XGBoost_fp), XGBoost with Mordred descriptors (XGBoost_Mordred), GNN, and transformer models achieves the highest correlation with ground truth values in the test set, yielding an R^2 value of 0.78 and a RMSE of 15.19 K (Fig. 3b). In contrast, an ensemble incorporating all 22 models performs worse due to redundancy in overlapping features. Notably, the four models in the best-performing ensemble encompass four major types of feature representations: molecular fingerprints, molecular descriptors, graphs, and SMILES strings. These results suggest that averaging predictions from models with diverse feature representations leads to higher accuracy than relying on any single model alone, demonstrating the effectiveness of the ensemble learning approach.

Interpretability offered by machine learning models

We perform SHAP analysis on the trained models (LASSO, RF, XGBoost, and FFNN) to gain physical insights into the structure–T_g relationships. For molecular fingerprints and molecular descriptors, we compute SHAP values for each feature (substructure or descriptor), which quantify the impact of individual features on the T_g of vitrimers. The ten most impactful substructures are presented in Fig. 4, ranked by their relative importance. Substructure 190, identified as the most impactful, negatively influences T_g due to its aliphatic nature, which makes polymer chains more flexible. The positively impactful substructures can be classified into two groups. The first group, comprising substructures 160, 179, and 147, increases T_g due to their aromaticity and increased chain rigidity. Note that while the complete aromatic rings are not depicted in Fig. 4c due to the molecular fingerprint radius of 3, the atoms highlighted in yellow are confirmed as part of aromatic rings based on RDKit visualization.[20] The second

group, including substructures 194, 160, 147, 149, and 188, contains oxygen and nitrogen atoms, which are involved in hydrogen bonding within the vitrimer network, leading to an increase in T_g. These trends are consistent with previous experimental studies.[46,47] Furthermore, our findings align with the work of Tao et al.,[12] which identified similar substructures using a LASSO model trained on an experimental T_g dataset. This consistency with literature confirms the validity of our results.

The most impactful RDKit descriptors are shown in Fig. 4b, with their definitions provided in Table S1 (SI). The descriptor with the most negative impact, the number of rotatable bonds, decreases vitrimer chain rigidity, leading to a negative correlation with T_g. The most positively correlated descriptor, VSA_Estate2, is a hybrid descriptor that captures the combined effect of the van der Waals surface area and the electro-topological states of atoms. Another notable descriptor, FractionCSP3, represents the fraction of carbon atoms that are sp^3-hybridized and negatively impacts T_g due to a more aliphatic and flexible molecular structure. Several identified descriptors have also been reported in a previous study based on an experimental dataset,[48] further validating the effectiveness of our analysis. The most important Mordred descriptors, along with their descriptions, are presented in Fig. S3 and Table S2 (SI). Similar to the RDKit descriptors, the most significant Mordred descriptor is the number of rotatable bonds, demonstrating consistency across different models and providing cross-validation within this study.

The insights derived from molecular fingerprints and descriptors highlight two primary factors influencing vitrimer T_g across different scales: chain mobility and molecular interactions. At the polymer chain scale, increased chain flexibility, facilitated by rotatable bonds and aliphatic structures, enhances conformational freedom and segmental motion, resulting in lower T_g. In contrast, the presence of aromatic rings
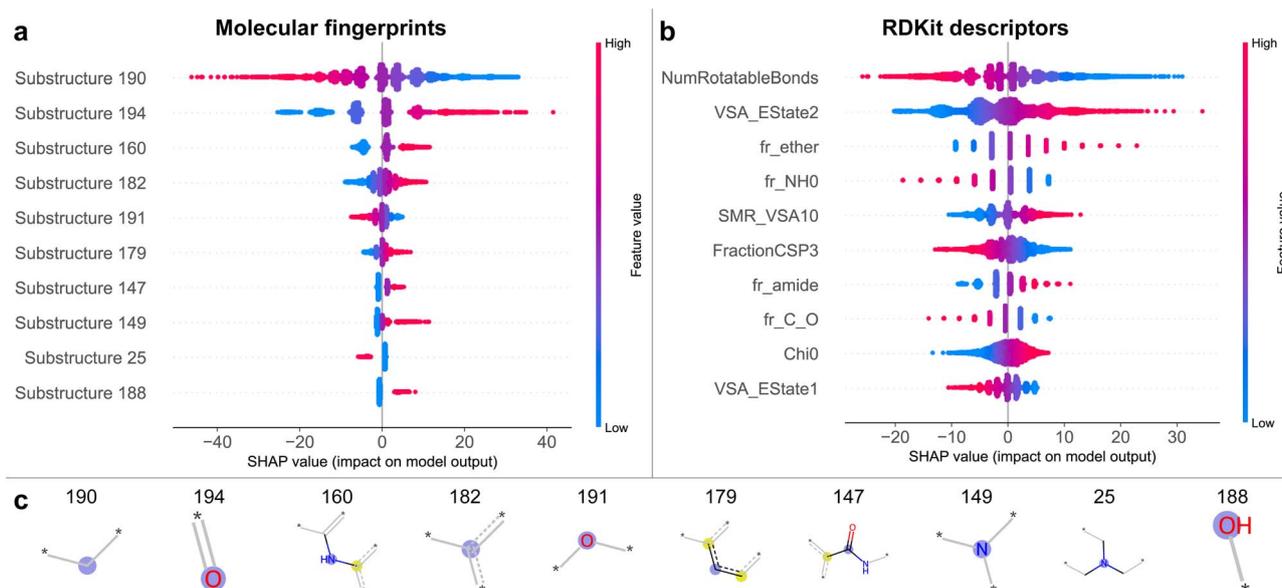
Fig. 4 Average SHAP values (feature importance scores) of ten most impactful (a) substructures and (b) RDKit descriptors calculated from the trained LASSO, RF, XGBoost and FFNN models. (c) Schematics of the substructures. The central atom of each substructure is highlighted in blue, while aromatic atoms are marked in yellow. The connectivity between atoms is represented in light gray.

and bulky side groups contributes to greater chain rigidity, restricting mobility and leading to higher T_g. At the atomistic scale, molecular interactions, such as hydrogen bonding and van der Waals forces, strengthen local interactions within the vitrimer network, increasing energy barriers to segmental motion and thereby elevating T_g. These findings offer valuable molecular design guidelines that can inform the discovery of vitrimers with tailored thermal properties.

Discovering novel hypothetical vitrimers by virtual screening

After identifying the ensemble model with the best performance on the test set, we proceed with predicting T_g for all vitrimers in the hypothetical unlabeled dataset (see Table 1). Approximately 100 vitrimers with the highest and lowest predicted T_g values are selected for validation through MD simulations and calibration. The simulation and calibration protocols remain consistent with those used in generating the labeled dataset (see the SI for details). The discovered high-T_g and low-T_g vitrimers successfully expand the property space of the training set (Fig. 5a), potentially broadening the applicability of vitrimers for use in more extreme environments. To further analyze the distribution of the discovered vitrimers relative to the training data, we apply principal component analysis (PCA) to reduce the molecular fingerprint representations to two-dimensional vectors. As shown in Fig. 5b, the novel vitrimers extend the chemical space beyond the training regime. The 10 vitrimers with the highest and lowest T_g are shown in Fig. 5 where all presented T_g values are obtained from MD simulations and GP calibration. The novel low-T_g vitrimers exhibit a range from 218 K to 237 K, while the high-T_g vitrimers range from 510 K to 549 K (Fig. 5c). These results demonstrate that the ensemble model facilitates efficient identification of novel vitrimers with a wide

T_g range, expanding the property space and potential applications of vitrimers.

Note that some vitrimers presented in Fig. 5c contain secondary amine groups, which are also potential reaction sites with epoxides. However, the secondary amine–epoxide reactions are less favorable than carboxyl–epoxide reactions in the presence of catalysts such as TBD during actual synthesis. Therefore, to maintain bifunctionality in the resulting vitrimers and reduce the complexity of MD simulations for T_g evaluation, we make the assumption that only carboxylic acids can react with epoxides while the reactions involving amines are ignored. While this simplification does not fully capture the actual experiments, undesired amine–epoxide reactions could be prevented by replacing nitrogen atoms with alternative heavy atoms. This strategy preserves the predicted T_g and minimizes side reactions, as demonstrated in our previous work.[17]

The most impactful substructure frequencies and descriptors (see Fig. 4) of the discovered vitrimers are presented in Fig. S4 (SI). The occurrence of positively correlated substructures and descriptors is higher in high-T_g vitrimers, while negatively correlated features are more prevalent in low-T_g vitrimers. This consistency cross-validates our feature importance analysis, demonstrating that the key factors identified in the training data also apply to novel vitrimers outside the training dataset. These findings further support the universality and robustness of the interpretable insights offered by our models.

The discovered vitrimers presented in Fig. 5c are currently hypothetical vitrimer candidates exhibiting desirable T_g and their self-healing capabilities remain to be evaluated. Future research could focus on quantifying the healability of these vitrimers through more complicated MD simulations, such as reactive MD. This would also allow for an investigation into the interplay between healability, T_g, and other relevant properties.
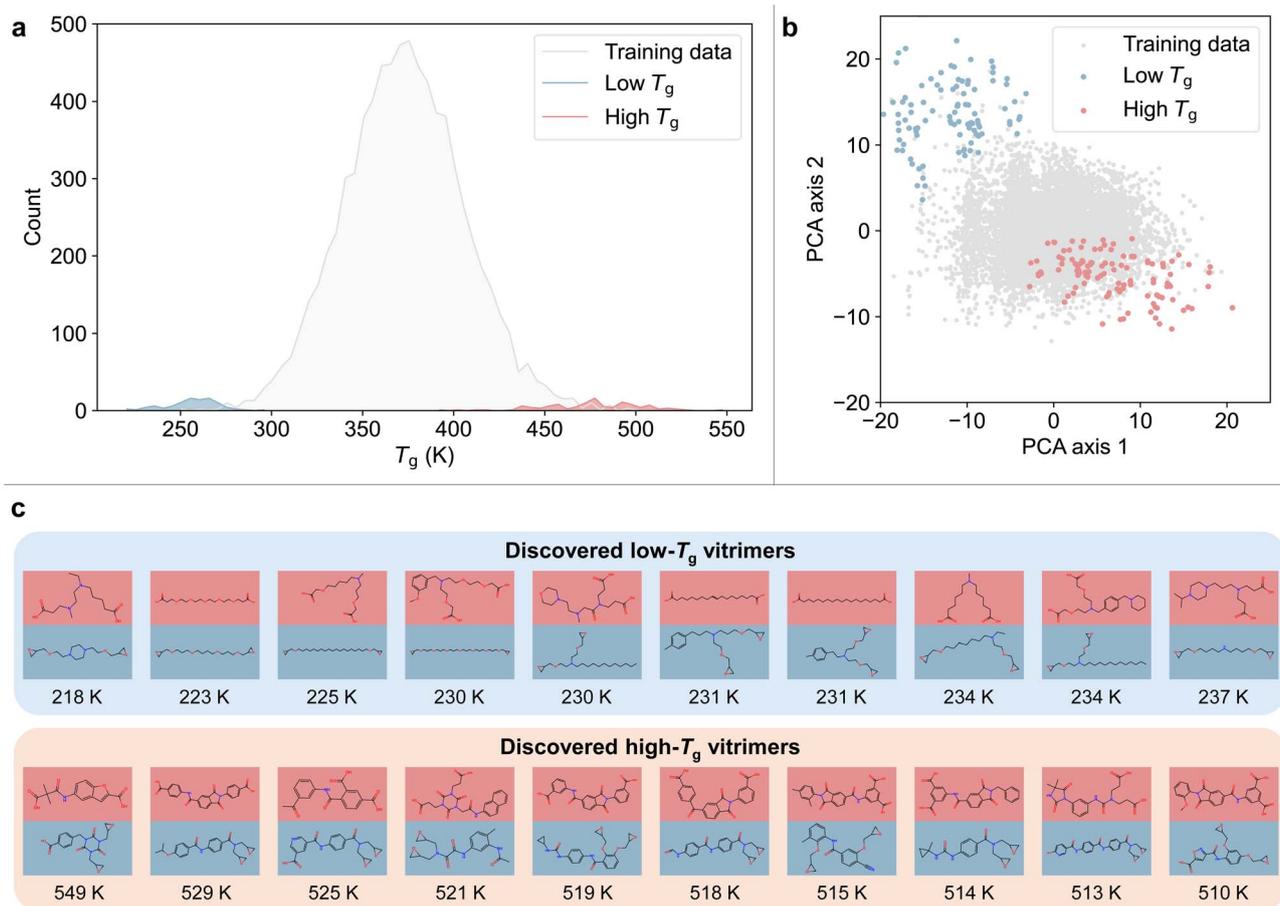
Fig. 5 Virtual screening of the hypothetical unlabeled dataset (see Table 1) and examples of discovered vitrimers whose T_g is validated by MD simulations. (a) T_g distribution of the training data and discovered vitrimers. (b) Chemical space of the training data and discovered vitrimers visualized by PCA of molecular fingerprints. (c) Molecular structures of discovered vitrimers with high and low T_g. All presented T_g values are obtained from MD simulations.

Experimental synthesis and characterization of discovered vitrimers

To validate the predictive capability and practical relevance of our trained ensemble ML model, we transition from virtual screening to experimental synthesis. The ensemble model is used to efficiently evaluate the T_g of 259 vitrimer candidates composed of commercially available acids and epoxides from the Sigma-Aldrich database.[34] Guided by the model's T_g predictions, we identify promising candidates exhibiting high thermal performance. These predictions are further filtered based on chemical intuition, ease of synthesis and purchasing, and material cost, ultimately leading to the selection of two vitrimer systems for experimental validation (Fig. 6a).

Among the screened candidates, vitrimers based on D,L-malic acid (MA) and 2,6-naphthalenedicarboxylic acid (NA) paired with diglycidyl ether (DGEBA) are selected for synthesis, as they exhibit desired T_g values from the ML ensemble model. These acids are associated with structural features—such as aromaticity and limited rotatable bonds in NA and the hydroxyl-rich backbone in MA—that are identified in our SHAP analysis as positively contributing to T_g, reinforcing the rationale for

their selection. The epoxide and acids are cured in the presence of the catalyst TBD for vitrimer synthesis (Fig. 6b). While malic acid contains hydroxyl groups in addition to carboxyl groups, these hydroxyl groups are less reactive towards esterification with epoxides due to the presence of acidic hydrogen.[51,52] Therefore, both MA-based and NA-based vitrimers are processed with an equimolar ratio of carboxyl groups to epoxides during curing.

Thermal analysis through DSC measurements reveals T_g values of 348 K (75 °C) and 395 K (122 °C) for the MA-based and NA-based vitrimers, respectively (Fig. 6c). Notably, the NA-based vitrimer exhibits a higher T_g than any bifunctional trans-esterification vitrimer reported in the literature (see Table S3, SI) to the best of our knowledge,[10,17,28,29,49,50] as shown in Fig. 6d. The MA-based vitrimer also demonstrates a relatively higher T_g compared to most existing bifunctional transesterification vitrimers. The experimental T_g of both vitrimers agrees well with the ML-predicted T_g (364 and 390 K), indicating the accuracy of our ML model in predicting the actual T_g of vitrimers. To evaluate the healability of the synthesized vitrimers, pristine specimens are cut and subsequently healed at elevated temperatures. Microscopic examination of the pristine, cut, and
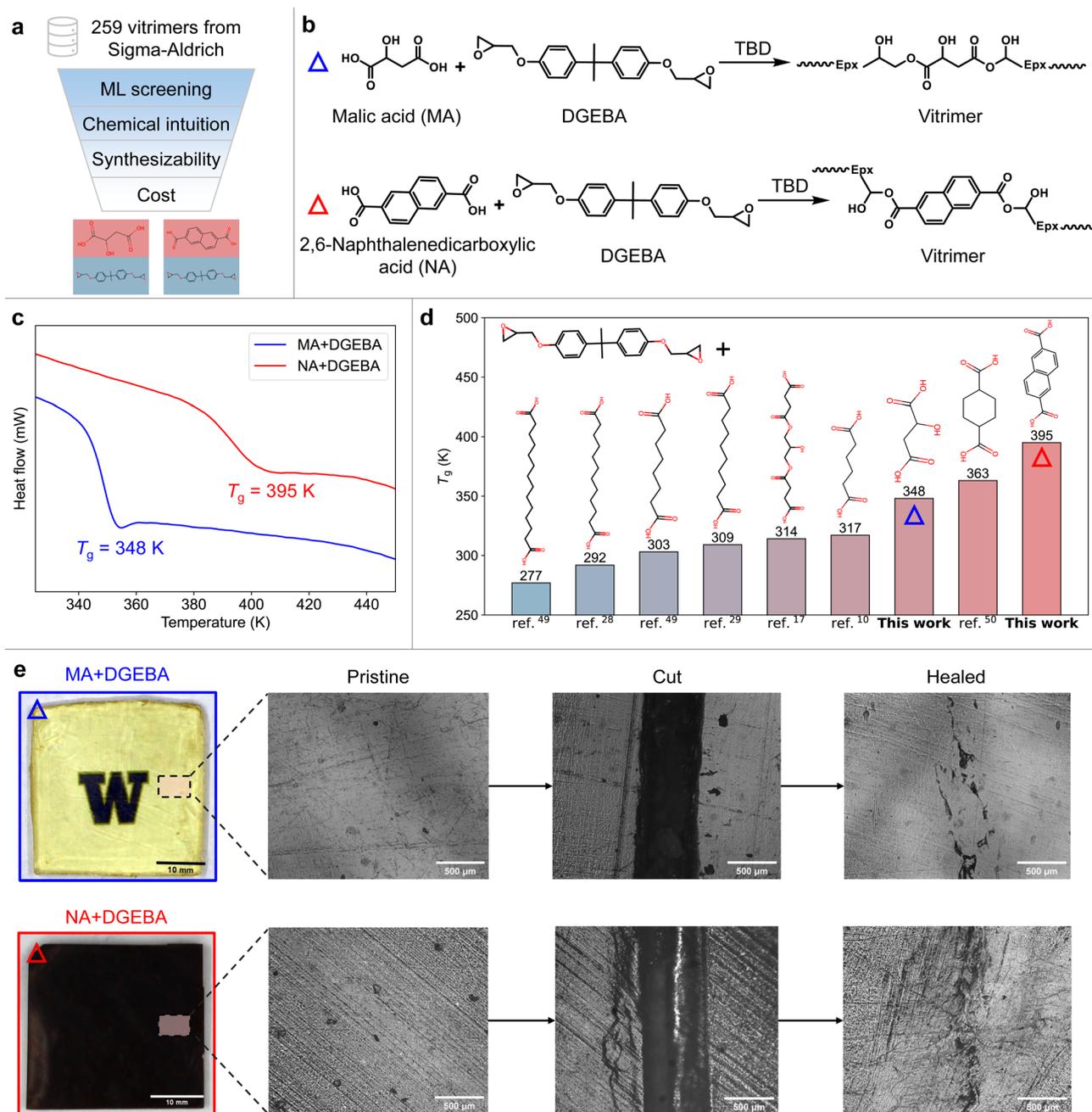
Fig. 6 Experimental synthesis and characterization of the discovered novel vitrimers. (a) The trained ensemble model is employed to virtually screen 259 vitrimers sourced from the Sigma–Aldrich database.[34] By integrating chemical intuition, we evaluate their synthesizability and cost, ultimately selecting two vitrimers for experimental synthesis and characterization. (b) Synthesis schemes of the vitrimers. (c) DSC results to measure the T_g of the synthesized vitrimers. (d) T_g of the vitrimers synthesized in this work and bifunctional transesterification vitrimers present in the literature.[10,17,28,29,49,50] (e) Images of pristine, cut and healed vitrimer specimens, demonstrating the healability of the synthesized vitrimers in this work.

healed samples (Fig. 6e) demonstrates complete damage recovery, confirming the healability of these novel vitrimer chemistries. The combination of high T_g and self-healing capability validates the effectiveness of our framework in identifying vitrimers with desirable properties, thereby broadening the scope of potential vitrimer applications.

Conclusions

This study establishes an MD-informed ML model for accurate and efficient vitrimer T_g prediction, which facilitates discovering novel bifunctional transesterification vitrimers with desirable properties. Leveraging MD-generated data from our previous study,[17] we systematically evaluate the predictive

performance of ML models for vitrimer T_g. Seven ML models are benchmarked across six molecular representations. Our findings demonstrate that an ensemble learning approach, which averages predictions from multiple models, outperforms individual models in T_g prediction. Furthermore, our feature importance analysis provides interpretability, identifying key structural and chemical factors that govern vitrimer T_g. Using the trained ensemble model, we screen an unlabeled dataset of approximately 1 million hypothetical vitrimers and discover candidates with extreme T_g values beyond the training domain. We perform an additional screening of vitrimers composed of commercially available carboxylic acids and epoxides, identifying promising candidates for experimental evaluation. Two novel vitrimers are synthesized and characterized, confirming their desirable properties. Both vitrimers exhibit relatively high T_g, with one surpassing the T_g of any bifunctional transesterification vitrimer reported in the literature. The presented T_g values confirm the capability of the model in accurate property prediction and expansion of the current vitrimer property space. This study highlights the feasibility of using MD-generated data to train ML models when experimental data are limited and successfully demonstrates the application of this approach by designing and validating novel vitrimers in experiments.

Besides T_g, self-healing efficiency is crucial for the practical applications of vitrimers. One key metric of efficient self-healing can be associated with chain mobility, suggesting that there should be sufficient difference between T_g and the topology freezing temperature (T_v) that indicates the onset of bond exchange reactions for efficient self-healing behavior. Future research could focus on designing vitrimers that exhibit both desirable T_g and reasonable healability characterized by a sufficient $T_v–T_g$ difference. For instance, with sufficient data on T_v for vitrimers, the ML framework developed in this study could be trained to predict self-healing performance. In addition, especially for high T_g vitrimers where there may not exist sufficient separation between T_g, T_v, and decomposition temperature (T_d), incorporating further chemical intuition and T_d prediction ML models could facilitate the design of vitrimers balancing T_g and T_d requirements with self-healing performance.

Although this work focuses on T_g prediction of vitrimers in a limited design space (bifunctional transesterification vitrimers with an equal molar ratio), the presented approach can be readily adapted to other types of vitrimers with higher functionality and various molar ratios, incorporating the effect of macromolecular structures with greater specificity and accuracy, provided sufficient data for training are available. The proposed framework provides an effective and efficient tool for polymer chemists to design synthesizable polymers with tailored properties, expanding the potential applications of vitrimers.

Author contributions

Yiwen Zheng: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization,

writing – original draft, and writing – review and editing. Agni K. Biswal: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing – original draft, and writing – review and editing. Yaqi Guo: methodology, software, and writing – review and editing. Prakash Thakolkaran: methodology, software, and writing – review and editing. Yash Kokane: methodology, software, and writing – review and editing. Vikas Varshney: methodology, supervision, and writing – review and editing. Siddhant Kumar: methodology, supervision, and writing – review and editing. Aniruddh Vashisth: conceptualization, funding acquisition, methodology, project administration, resources, supervision, and writing – review and editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data and code that support the findings of this study are openly available at **https://github.com/vashisth-lab/VitrimerScreening** and have been archived at Zenodo (DOI: **https://doi.org/10.5281/zenodo.16391076**). Part of the data and the Gaussian process calibration codes are sourced from Zheng *et al.*[17] (available at **https://github.com/vashisth-lab/VitrimerVAE**). The GNN and transformer models are based on implementations by Hickey *et al.*[37] (available at **https://data.mendeley.com/datasets/ydbv9t8fzr/1**) and Xu *et al.*[41] (available at **https://github.com/ChangwenXu98/TransPolymer**), respectively.

Supplementary information includes details of vitrimer datasets and machine learning models as well as additional results. See DOI: **https://doi.org/10.1039/d5dd00239g**.

References

1 R. J. Young and P. A. Lovell, *Introduction to Polymers*, CRC press, 2011.
2 D. Montarnal, M. Capelot, F. Tournilhac and L. Leibler, *Science*, 2011, **334**, 965–968.
3 M. Capelot, M. M. Unterlass, F. Tournilhac and L. Leibler, *ACS Macro Lett.*, 2012, **1**, 789–792.
4 Y. Jin, Z. Lei, P. Taynton, S. Huang and W. Zhang, *Matter*, 2019, **1**, 1456–1493.
5 B. Krishnakumar, R. P. Sanka, W. H. Binder, V. Parthasarthy, S. Rana and N. Karak, *Chem. Eng. J.*, 2020, **385**, 123820.

6 E. Hoogeboom, V. G. Satorras, C. Vignac and M. Welling, *International Conference on Machine Learning*, 2022, pp. 8867–8887.

7 T. Hansson, C. Oostenbrink and W. van Gunsteren, *Curr. Opin. Struct. Biol.*, 2002, **12**, 190–196.

8 A. Vashisth, C. Ashraf, W. Zhang, C. E. Bakis and A. C. Van Duin, *J. Phys. Chem.*, 2018, **122**, 6633–6642.

9 K.-q. Yu, Z.-s. Li and J. Sun, *Macromol. Theory Simul.*, 2001, **10**, 624–633.

10 M. Kamble, A. Vashisth, H. Yang, S. Pranompont, C. R. Picu, D. Wang and N. Koratkar, *Carbon*, 2022, **187**, 108–114.

11 J. M. Kranenburg, C. A. Tweedie, K. J. van Vliet and U. S. Schubert, *Adv. Mater.*, 2009, **21**, 3551–3561.

12 L. Tao, G. Chen and Y. Li, *Patterns*, 2021, **2**, 100225.

13 P. Thakolkaran, Y. Zheng, Y. Guo, A. Vashisth and S. Kumar, *arXiv*, 2024, preprint, arXiv:2409.06457, DOI: **10.48550/arXiv.2409.06457**.

14 R. Batra, H. Dai, T. D. Huan, L. Chen, C. Kim, W. R. Gutekunst, L. Song and R. Ramprasad, *Chem. Mater.*, 2020, **32**, 10489–10500.

15 J. Yang, L. Tao, J. He, J. R. McCutcheon and Y. Li, *Sci. Adv.*, 2022, **8**, eabn9545.

16 S. Otsuka, I. Kuwajima, J. Hosoya, Y. Xu and M. Yamazaki, *2011 International Conference on Emerging Intelligent Data and Web Technologies*, 2011, pp. 22–29.

17 Y. Zheng, P. Thakolkaran, A. K. Biswal, J. A. Smith, Z. Lu, S. Zheng, B. H. Nguyen, S. Kumar and A. Vashisth, *Adv. Sci.*, 2025, **12**, 2411385.

18 L. Tao, J. He, T. Arbaugh, J. R. McCutcheon and Y. Li, *J. Membr. Sci.*, 2023, **665**, 121131.

19 T. Xie, H.-K. Kwon, D. Schweigert, S. Gong, A. France-Lanord, A. Khajeh, E. Crabb, M. Puzon, C. Fajardo, W. Powelson, *et al.*, *arXiv*, 2022, preprint, arXiv:2208.01692, DOI: **10.48550/arXiv.2208.01692**.

20 RDKit, *RDKit: Open-source cheminformatics software*, **https://www.rdkit.org**.

21 H. Moriwaki, Y.-S. Tian, N. Kawashita and T. Takagi, *J. Cheminf.*, 2018, **10**, 1–14.

22 D. Rogers and M. Hahn, *J. Chem. Inf. Model.*, 2010, **50**, 742–754.

23 D. Weininger, *J. Chem. Inf. Comput. Sci.*, 1988, **28**, 31–36.

24 Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and S. Y. Philip, *IEEE Transact. Neural Networks Learn. Syst.*, 2020, **32**, 4–24.

25 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, *Adv. Neural Inf. Process. Syst.*, 2017, 30.

26 C. Yan, X. Feng, J. Konlan, P. Mensah and G. Li, *Phys. Chem. Chem. Phys.*, 2023, **25**, 30049–30065.

27 F. I. Altuna, C. E. Hoppe and R. J. J. Williams, *RSC Adv.*, 2016, **6**, 88647–88655.

28 S. Kaiser, P. Novak, M. Giebler, M. Gschwandl, P. Novak, G. Pilz, M. Morak and S. Schlögl, *Polymer*, 2020, **204**, 122804.

29 A. M. Hubbard, Y. Ren, D. Konkolewicz, A. Sarvestani, C. R. Picu, G. S. Kedziora, A. Roy, V. Varshney and D. Nepal, *ACS Appl. Polym. Mater.*, 2021, **3**, 1756–1766.

30 Y. Ran, L.-J. Zheng and J.-B. Zeng, *Materials*, 2021, **14**, 919.

31 A. M. Hubbard, Y. Ren, C. R. Picu, A. Sarvestani, D. Konkolewicz, A. K. Roy, V. Varshney and D. Nepal, *ACS Appl. Polym. Mater.*, 2022, **4**, 4254–4263.

32 S. Jaeger, S. Fulle and S. Turk, *J. Chem. Inf. Model.*, 2018, **58**, 27–35.

33 T. Sterling and J. J. Irwin, *J. Chem. Inf. Model.*, 2015, **55**, 2324–2337.

34 Sigma-Aldrich, **https://www.sigmaaldrich.com/US/en**, accessed, February 2025.

35 H. L. Morgan, *J. Chem. Doc.*, 1965, **5**, 107–113.

36 L. Tao, V. Varshney and Y. Li, *J. Chem. Inf. Model.*, 2021, **61**, 5395–5413.

37 K. Hickey, J. Feinstein, G. Sivaraman, M. MacDonell, E. Yan, C. Matherson, S. Coia, J. Xu and K. Picel, *Comput. Mater. Sci.*, 2024, **238**, 112933.

38 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

39 T. Chen and C. Guestrin, *Proceedings of the 22nd Acm sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

40 A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.*, *Adv. Neural Inf. Process. Syst.*, 2019, 32.

41 C. Xu, Y. Wang and A. Barati Farimani, *npj Comput. Mater.*, 2023, **9**, 64.

42 Y. S. AlFaraj, S. Mohapatra, P. Shieh, K. E. Husted, D. G. Ivanoff, E. M. Lloyd, J. C. Cooper, Y. Dai, A. P. Singhal, J. S. Moore, *et al.*, *ACS Cent. Sci.*, 2023, **9**, 1810–1819.

43 H. Esmaeili and R. Rizvi, *Comput. Mater. Sci.*, 2023, **229**, 112432.

44 W. Gao, H. Wang, Y. Xu, Y. Yang, Q. Gu, X. Duan, B. Wang and X. Zhou, *ACS Appl. Polym. Mater.*, 2024, **6**, 4501–4508.

45 S. M. Lundberg and S.-I. Lee, *Adv. Neural Inf. Process. Syst.*, 2017, 30.

46 K. Naito and A. Miura, *J. Phys. Chem.*, 1993, **97**, 6240–6248.

47 P. C. Painter, J. F. Graf and M. M. Coleman, *Macromolecules*, 1991, **24**, 5630–5638.

48 A. Babbar, S. Ragunathan, D. Mitra, A. Dutta and T. K. Patra, *J. Polym. Sci.*, 2024, **62**, 1175–1186.

49 S. Shen, V. K. Thakur and A. A. Skordos, *J. Appl. Polym. Sci.*, 2024, **141**, e56028.

50 K. Tangthana-Umrung, Q. A. Poutrel and M. Gresil, *Macromolecules*, 2021, **54**, 8393–8406.

51 S. Japon, L. Boogh, Y. Leterrier and J.-A. Månson, *Polymer*, 2000, **41**, 5809–5818.

52 S. Pesetskii, V. Shevchenko and V. Dubrovsky, *Multifunctionality of Polymer Composites: Challenges and new solutions*, 2015, pp. 302–337.