npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

https://doi.org/10.1038/s41524-025-01671-w

Discovery of liquid crystalline polymers with high thermal conductivity using machine learning

Hayato Maeda[1,10], Stephen Wu[2,3,10], Rika Marui[1], Erina Yoshida[1], Kan Hatakeyama-Sato[1], Yuta Nabae[1], Shiori Nakagawa[1], Meguya Ryu[1,4], Ryohei Ishige[1], Yoh Noguchi[2,5], Yoshihiro Hayashi[2,3], Masashi Ishii[6], Isao Kuwajima[6], Felix Jiang[7], Xuan Thang Vu[7], Sven Ingebrandt[7], Masatoshi Tokita[1], Junko Morikawa[1,8] ✉, Ryo Yoshida[2,3,6,9] ✉ & Teruaki Hayakawa[1] ✉

Next-generation power electronics require efficient heat dissipation management, and molecular design guidelines are needed to develop polymers with high thermal conductivity. Polymer materials have considerably lower thermal conductivity than metals and ceramics due to phonon scattering in the amorphous region. The spontaneous orientation of the molecular chains of liquid crystalline polymers could potentially give rise to high thermal conductivity, but the molecular design of such polymers remains largely empirical. In this study, we developed a machine learning model that predicts with more than 96% accuracy whether liquid crystalline states will form based on the chemical structure of the polymer. By exploring the inverse mapping of this model, we identified a comprehensive set of chemical structures for liquid crystalline polyimides. The polymers were then experimentally synthesized, and the results confirmed that they form liquid crystalline phases, with all polymers exhibiting calculated thermal conductivities within the range of 0.722–1.26 $W\,m^{-1}\,K^{-1}$.

Machine learning is revolutionizing materials research by enabling data-driven predictive science. In recent years, machine learning-enabled molecular design, with the objective of identifying new materials with specific, desired properties, has made significant strides[1,2]. The typical workflow consists of forward and inverse predictions[3]. First, a statistical model is built to predict the target properties of a given material based on its compositional and structural features. Subsequently, inverse mapping of this model is explored to predict materials with the desired properties in the reverse direction. Various machine learning techniques have been actively developed to apply this concept in polymer chemistry, including a wide range of property predictors[4–7] and inverse design methods using virtual polymer generators such as molecular generative artificial intelligence[8–10]. Notably, proof-of-concept examples include the discovery of amorphous polymers with high thermal conductivity[4], lithium-ion conducting polymers[11,12], gas-separating polymer membranes[13], and high-temperature polymer dielectrics[14]. However, few cases of polymers that were initially predicted by machine learning and subsequently verified experimentally have been reported. In this nascent field, our work advances the proof-of-concept of machine learning-driven polymer material design for real-world applications through the successful discovery of liquid crystalline polymers with high thermal conductivity.

As the demand for the miniaturization and portability of electronic devices increases, researchers have focused on the development of light-weight, highly insulating polymers with high thermal conductivity for efficient heat dissipation. However, the thermal conductivity of amorphous polymers is one to three orders of magnitude lower than those of metals and ceramics[15,16], which is a major barrier to their practical application. In general, polymers without free electrons exhibit low thermal conductivity because phonons, the dominant factor in heat transport, tend to scatter in the amorphous region. To overcome this empirical and theoretical limitation, inorganic fillers are commonly added to the polymer matrix[17]. However, increasing the filler content causes other polymer properties such as

[1]School of Materials and Chemical Technology, Institute of Science Tokyo, Meguro-ku, Tokyo, Japan. [2]The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tachikawa, Tokyo, Japan. [3]The Graduate University for Advanced Studies, SOKENDAI, Tachikawa, Tokyo, Japan. [4]National Metrology Institute of Japan, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan. [5]School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, Hachioji, Tokyo, Japan. [6]National Institute for Materials Science, Tsukuba, Ibaraki, Japan. [7]Institute of Materials in Electrical Engineering 1, RWTH Aachen University, Aachen, Germany. [8]Research Center for Autonomous Systems Materialogy (ASMat), Institute of Innovative Research, Institute of Science Tokyo, Midori-ku, Yokohama, Japan. [9]TRIP Headquarters, RIKEN, Wako, Saitama, Japan. [10]These authors contributed equally: Hayato Maeda, Stephen Wu. ✉e-mail: morikawa.j.4f50@m.isct.ac.jp; yoshidar@ism.ac.jp; hayakawa@mct.isct.ac.jp

adhesion, flowability, processability, and insulation to significantly deteriorate[18]. Additionally, if the thermal resistance of the matrix polymer is high, the thermal conductivity of the composite material reaches a ceiling; therefore, it is imperative to improve the thermal conductivity of the polymer matrix itself[19]. Therefore, attempts have been made to suppress phonon scattering by utilizing liquid crystalline phase formation to induce an ordered alignment of polymer chains[20]. However, the design of liquid crystalline polymers remains purely empirical, and relies extensively on trial and error. While some trends have been identified—for example, the tendency of phenyl benzoate backbones or certain alkyl spacer chains to promote liquid crystallinity[21–26]—polymers with complex molecular interactions, such as charge-transfer interactions in polyimides, often deviate from these trends, resulting in numerous exceptions[27].

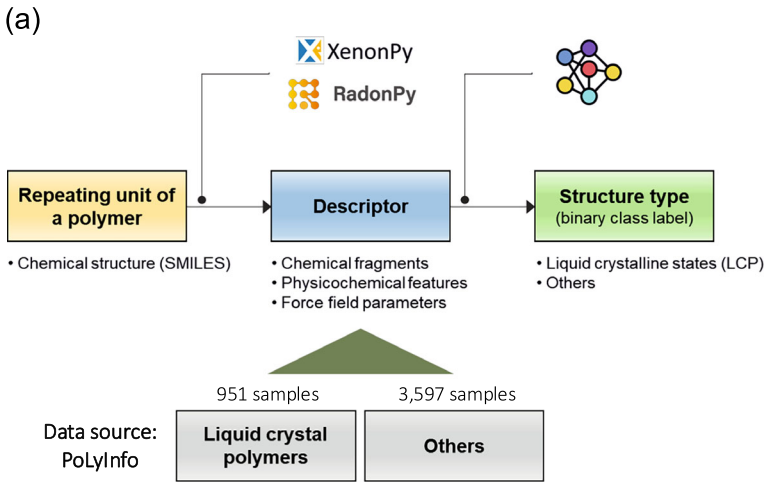
In this study, we developed a machine learning algorithm to predict the chemical structure of a polymer-repeating unit capable of forming liquid crystalline phases. This is an indirect approach that aims to enhance the thermal conductivity by addressing the challenge posed by the lack of a comprehensive dataset on the thermal conductivity of oriented polymers. We constructed a binary classifier using the compositional and structural features of the polymer-repeating unit as inputs from a labeled dataset of liquid crystalline polymers and other polymers that were previously synthesized. The prediction accuracy of the constructed model in discriminating whether a polymer exhibits a liquid-crystalline phase exceeded 96%. Using this model, we conducted high-throughput virtual screening to search for polyimides that form liquid crystalline states. In liquid crystalline polyimides, rigid mesogens composed of aromatic segments (including conjugated rings and imide groups) promote molecular alignment, while flexible spacer chains (consisting of alkyl or similar groups) enhance the molecular mobility to facilitate the formation of the liquid crystalline state[21]. In addition, six polymers were selected from a narrowed-down library of candidates for experimental verification by the de novo syntheses of their monomers followed by their polymerization reactions. As a result, all six polymers were successfully synthesized, and the resulting polymers spontaneously formed smectic liquid crystal phases. Moreover, by applying a lock-in photothermal method for measuring the in-plane thermal diffusivity using arrayed temperature sensors on suspended SiN_X membranes, their thermal conductivities were experimentally confirmed to reach 0.722−1.26 W m^{-1} K^{-1}. These are the first liquid crystalline polymers predicted and discovered via machine learning.

Results

Performance of the machine learning model

The machine learning task is formulated as a supervised learning problem aimed to classify the chemical structure of a polymer repeat unit denoted as X into two classes with the binary variable $Y \in \{-1, 1\}$ indicating a liquid crystalline state ($Y = 1$) or a non-liquid crystalline state ($Y = -1$) (Fig. 1a). The compositional and structural features of a given repeat unit are encoded into a 397-dimensional descriptor vector ($\phi(X) \in \mathbb{R}^{397}$). Here, only linear homopolymers without additives were considered for the learning and prediction tasks by focusing on their intrinsic properties and avoiding the influence of additives or copolymerization, which can affect liquid crystalline phase formation. The descriptor was formed through the concatenation of two different descriptors: a 207-dimensional vector of RDKit descriptors representing various physical, chemical, and structural features of polymer molecules (https://www.rdkit.org/), and a 190-dimensional quantitative descriptor that encodes empirical force field parameters derived using a general AMBER force field version 2 (GAFF2)[28], which is widely used in all-atom classical molecular dynamics simulations[29,30]. The calculations of these two polymer descriptors were performed using the Python libraries XenonPy[5,31] and RadonPy[29], which implement wrapper functions for RDKit and the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[32], respectively. To properly describe the periodicity of polymer repeating structures, the descriptor was calculated after linearly linking the head and tail of the repeating unit 10 times to form a decamer[30]. The descriptor calculation procedure is described in the Supplementary Information (Section 1.2; pages S2 and S3).

A machine learning classifier $Y = f\left(\phi(X)\right)$ defines a mapping from the vectorized polymer to the binary class label by indicating whether the polymer would exhibit a liquid crystalline state. As positive instances (P) for the binary classification task, we used a list of 951 liquid crystalline polymers compiled from PoLyInfo[33], a polymer property database that had been manually compiled based on a literature survey. The polymer list was partially labeled with more detailed annotations representing liquid crystalline states such as the nematic and smectic phases; however, in this study, they were merged into a single label (liquid crystal). A total of 3,597 polymers without any record of forming liquid crystalline states were extracted from the database and used as unlabeled instances (U).



(b)

		True: LCP	True: Others
Prediction	LCP	114 (9.6)	13 (7.3)
	Others	10 (3.6)	504 (11.1)

	Best	Mean (SD)
Accuracy	0.973	0.964 (0.012)
Precision	0.919	0.898 (0.044)
Recall	0.947	0.920 (0.028)
F_1 score	0.933	0.908 (0.027)

Fig. 1 | Machine learning process for predicting whether a polymer with a designed repeating unit exhibits liquid crystalline states. a Machine learning workflow. **b** Prediction accuracy (confusion matrix, precision, recall, and F_1 scores) for test datasets; standard deviations of the performance metrics obtained from 100 independent tests are shown in parentheses.

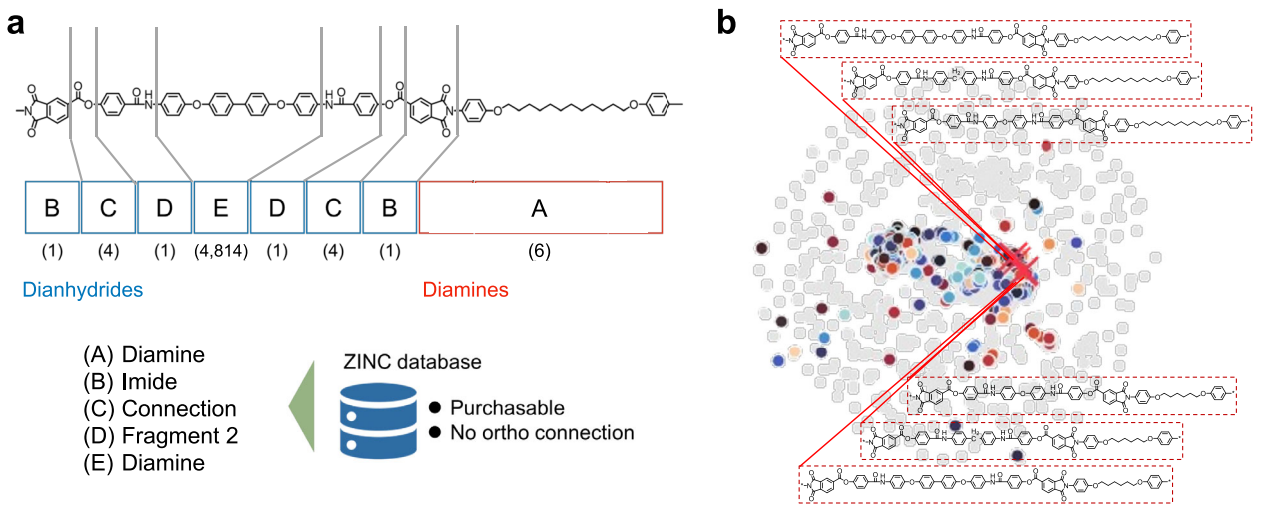


Fig. 2 | Virtual screening of liquid crystalline polyimides. a The rigid and flexible chains of polyimides were divided into five building blocks, and the corresponding fragment set for each block was extracted from the Zinc database. By combining these fragments, 115,536 virtual polyimides were computationally generated. **b** The chemical structures of the virtual polymers are visualized using the UMAP projection (in gray). The 10,825 candidates predicted to exhibit liquid crystallinity were clustered into 391 groups based on molecular similarity and visualized in different colors. From these clusters, six polyimides were selected and successfully synthesized.

Notably, these polymers are not necessarily true negative cases, and their liquid crystallinity has not been confirmed. No comprehensive database of negative cases has been constructed for liquid crystalline polymers. In machine learning, this problem is known as positive and unlabeled (PU) learning[34]. In the present study, we applied the classical PU learning algorithm proposed by Elkan et al.[35] to calibrate the classification probability. However, the effect of the PU learning calibration was insignificant. The protocols established for data preparation and PU learning are provided in Supplementary Information (Section 1.3; pages S3 and S4).

The binary classifier was modeled as a conventional multilayer perceptron neural network. A total of 85% of the entire dataset was randomly selected and used as the training set, while the remaining 15% was used as a test set for evaluating the model performance. A randomly selected 85% of the training dataset was applied for training, and the remainder was utilized for the validation set. The hyperparameters of the classification model were optimized using the black box optimization software Optuna[36], which adjusted the number and width of the hidden layers to minimize the validation F_1 score with this random split (see details in Supplementary Information (Section 1.2; page S2 and S3)). To assess the variation in the performance metrics, we independently repeated the training and testing process 100 times using different random data splits.

According to Fig. 1b, the average classification accuracy exceeds 96%. The mean values of recall and precision are 0.92 and 0.90, respectively, suggesting that the trade-off between the false and true positive rates is well balanced. Although the data are not shown herein, we confirmed that other machine learning algorithms, such as ensemble learning, achieve comparative predictive performance.

Virtual screening of liquid crystalline polyimides

Using the PU learning-calibrated classifier, we conducted an exhaustive search for polymers that undergo liquid crystallization. The search space was limited to polyimides, which are usually synthesized via the polycondensation of tetracarboxylic dianhydride and diamine monomers. To create a virtual library, we decomposed the template structure of polyimides into five building blocks labeled A-E in Fig. 2a. According to this decomposition, the acid dianhydride and diamine are composed of symmetric molecules. Fragments matching the structural pattern of each building block were searched among a set of highly available compounds in the small molecule database ZINC[37]. Considering these combinations, 115,536 virtual polyimides were computationally generated. In polyimides, the rigid chains, which consist of conjugated aromatic rings bonded to imide groups, promote molecular orientation to instill mesogenic behavior. In contrast, the flexible spacer groups tend to form an amorphous structure that can strongly absorb impact energy. Because the rigid and flexible components are arranged symmetrically, the self-assembled higher-order structures are also expected to exhibit structural equivalence (symmetry) and periodicity in the plane of the main chain. Thus, a virtual library was constructed following this design and synthesis strategy.

By applying a median liquid crystal transition probability of unity and a standard deviation of less than 0.2 as thresholds, approximately 91% of the candidate polymers were filtered out, resulting in 10,825 polyimides predicted to exhibit liquid crystallinity. Figure 2b shows a two-dimensional (2D) representation of the selected candidates using the uniform manifold approximation and projection (UMAP) algorithm[38], in which the chemical structure of each repeat unit was encoded into the descriptor following the procedure utilized during the construction of the predictive models. We applied hierarchical density-based spatial clustering of applications with noise (known as HDBSCAN)[39,40] to the projected polymers, which suggested the presence of 391 different clusters. Subsequently, we selected candidates suitable for experimental synthesis by examining representatives from the 391 clusters and those with closely related structures. Specifically, the repeating unit of a representative polyimide candidate from each cluster was reviewed by a polymer synthesis expert to assess the synthesizability of each dianhydride and diamine monomer, as well as the availability of the necessary ingredients. As a consequence, six polyimides (Fig. 3) were selected from several promising clusters and successfully synthesized; X-ray structural analysis of these polyimide bulk films revealed the formation of smectic liquid crystalline phases (Supplementary Information, Section 3.5 (pages S40–S41)). Further details of the selection procedure are provided in the Supplementary Information (Section 1.5; pages S4–S6).

Analysis of higher-order structures in spin-coated films

The higher-order structures within the spin-coated thin films were analyzed by conducting Grazing Incidence Wide-Angle X-ray Diffraction (GIWAXD) measurements. In this study, the in-plane thermal diffusivity of polyimide films—fabricated via spin coating followed by thermal imidization—was measured using SiN_X devices. The small size of these devices meant that conducting GIWAXD directly on the polyimide films formed on them was challenging. Therefore, similar films were replicated on silicon wafers under identical spin-coating conditions to facilitate GIWAXD measurements. To quantify the molecular

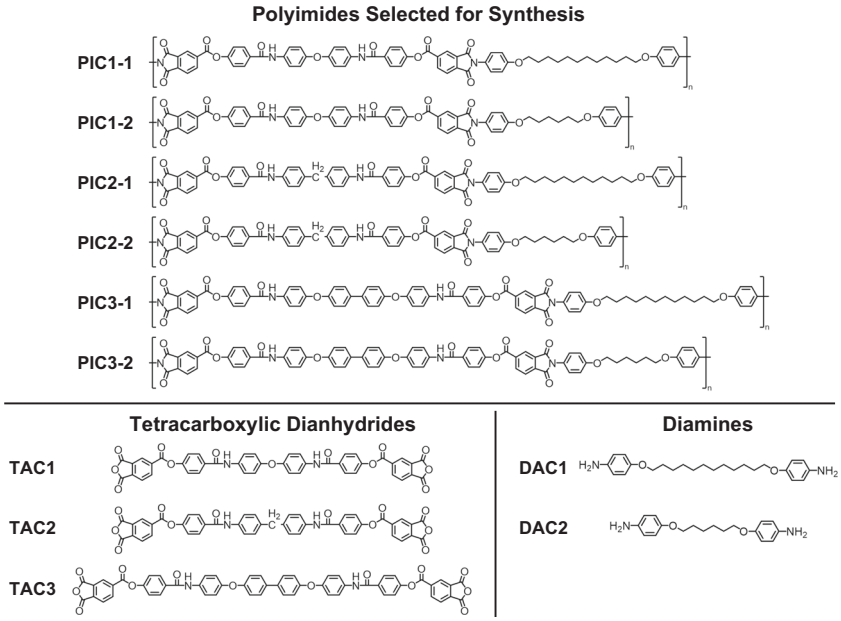


Fig. 3 | Chemical structures of the six polyimides selected for synthesis through virtual screening and their corresponding monomers. The polyimides were derived from combinations of three types of dianhydrides (TAC1, TAC2, TAC3) and two types of diamines (DAC1, DAC2).

orientation within the spin-coated films, the uniaxial orientational order parameter $S = \left(3\langle\cos^2\phi\rangle - 1\right)/2$ was calculated using the azimuthal distribution of the diffraction intensity at $2\theta \approx 20^\circ$. This calculation method follows the approach described by Ishige et al.[41].

The measured GIWAXD profiles show that the diffraction is somewhat concentrated on the meridional axis, indicating preferential in-plane orientation of mesogens in the thin films (Fig. 4a–f). The S values calculated for the films ranged from −0.33 to −0.45 (see the rightmost column of Table 1).

Mesogens of which the molecular orientation is such that the long axis is oriented perfectly parallel to the substrate surface would display a 2D GIWAXD profile with the diffraction concentrated on the meridional line parallel to the film normal, yielding an S value of −0.5. The range of S values obtained in our work suggests that the alignment of the mesogens is not entirely parallel to the substrate surface. The in-plane molecular orientation develops during the imidization process, where a polyimide film fixed onto a silicon wafer undergoes contraction. GIWAXD profiles measured immediately after spin coating and solvent removal, i.e., before the imidization process, showed azimuthally spread diffraction, indicating a lack of preferential molecular orientation (see Fig. S57 in the Supplementary Information, Section 3.5 (page 44)).

Moreover, SAXS and WAXD analyses of the bulk samples suggested that PIC1-1, 1-2, and 2-1 exhibited a smectic B-like phase, whereas PIC3-1 and 3-2 demonstrated a smectic E-like phase. PIC2-2, similar to PIC3-1 and 3-2, formed a close-packed arrangement between neighboring molecules (Fig. S54). In thin films, however, the intensity vs. q profiles shown in Fig. 4g indicated that the peak shape around $q \sim 15\ \mathrm{nm}^{-1}$ for PIC2-2, 3-1, and 3-2 differed from those observed in bulk analyses. This discrepancy suggests that in thin films, the molecular alignment over long distances may be insufficient to produce the diffraction patterns characteristic of a smectic E-like phase.

According to the S values listed in Table 1, no significant differences in the degree of in-plane orientation were observed between the PIC1-1 and PIC1-2 samples, nor between the PIC3-1 and PIC3-2 samples. However, a different trend was observed for the PIC2 series, where PIC2-1 exhibited a higher degree of in-plane orientation compared to PIC2-2. Additionally, the PIC3 series displayed a lower degree of in-plane orientation than both the PIC1 and PIC2 series. These findings suggest that differences in the structure of the mesogen significantly affect the S values, whereas variations in the spacer length do not have a substantial impact.

The presence of a methylene group at the center of the mesogen in the PIC2 series influences the higher-order structure and S values. The results of subsequent analyses of the bulk samples presented in Fig. S57 reveal the distinct higher-order structural characteristics of the PIC2 samples as compared with the other samples. This suggests that the methylene group of the mesogen in the PIC2 series plays a critical role in determining the degree of molecular orientation and structural evolution during thermal imidization.

Evaluation of thermophysical properties

The in-plane thermal conductivity (λ) and phase transition temperatures of the synthesized liquid crystalline polyimides were investigated. The results, including the in-plane thermal diffusivity (α), specific heat capacities (C_p), densities (ρ), and λ (calculated using α, C_p, ρ), as well as the phase transition temperatures, are summarized in Table 1. Measurement of the thermal conductivity and diffusivity of microscopically ordered liquid crystalline polymer thin films requires unique setups and equipment with high precision and fine spatial resolution. Several techniques, such as a thermal bridge method[42], atomic force microscopy cantilever-assisted measurements[43], and time-domain thermoreflectance[44], have been employed for measuring the microscale thermal conductivity of polymers. In this study, the thermophysical properties were evaluated by a microscale temperature wave analysis method (μ-TWA) using a device consisting of a line-shaped or spiral-shaped temperature sensor array on a 2D silicon nitride (SiN_x) membrane with a thickness of 50 or 150 nm[45] and ultrafast scanning calorimetry (FSC)[46]. The method employing the array-type sensors (μ-TWA) enabled rapid measurements of the in-plane thermal diffusivity of the aligned polyimide nanofilms to obtain thermal diffusivity distributions at different positions. The thermal conductivities of liquid crystalline polymers were obatind by measured thermal diffusivities, densities and specific heat capacities according to the following equation $\lambda = \alpha\rho C_p$. Detailed descriptions of the measurement techniques and the corresponding results are provided in the Supplementary Information (Sections 3.3 and 3.4; pages S36–S39).

Using this method, the in-plane thermal conductivity of the newly synthesized liquid crystalline aromatic polyimide nanofilm (PIC2-1) prepared via spin coating was computed from the measured thermal diffusivity, which reached $1.26\ \mathrm{W\ m^{-1}\ K^{-1}}$. This value was significantly higher than those of commercial polyimides in their nearly amorphous states[47] and of previously reported non-crosslinked liquid crystalline polyimides[48]. The PICX-2 samples exhibited higher thermal conductivities compared to the PICX-1 samples. This result aligns with previous studies showing that increased molecular rigidity promotes thermal conductivity[49]. The PICX-2 samples, with their shorter alkyl spacer chains, have more rigid molecular

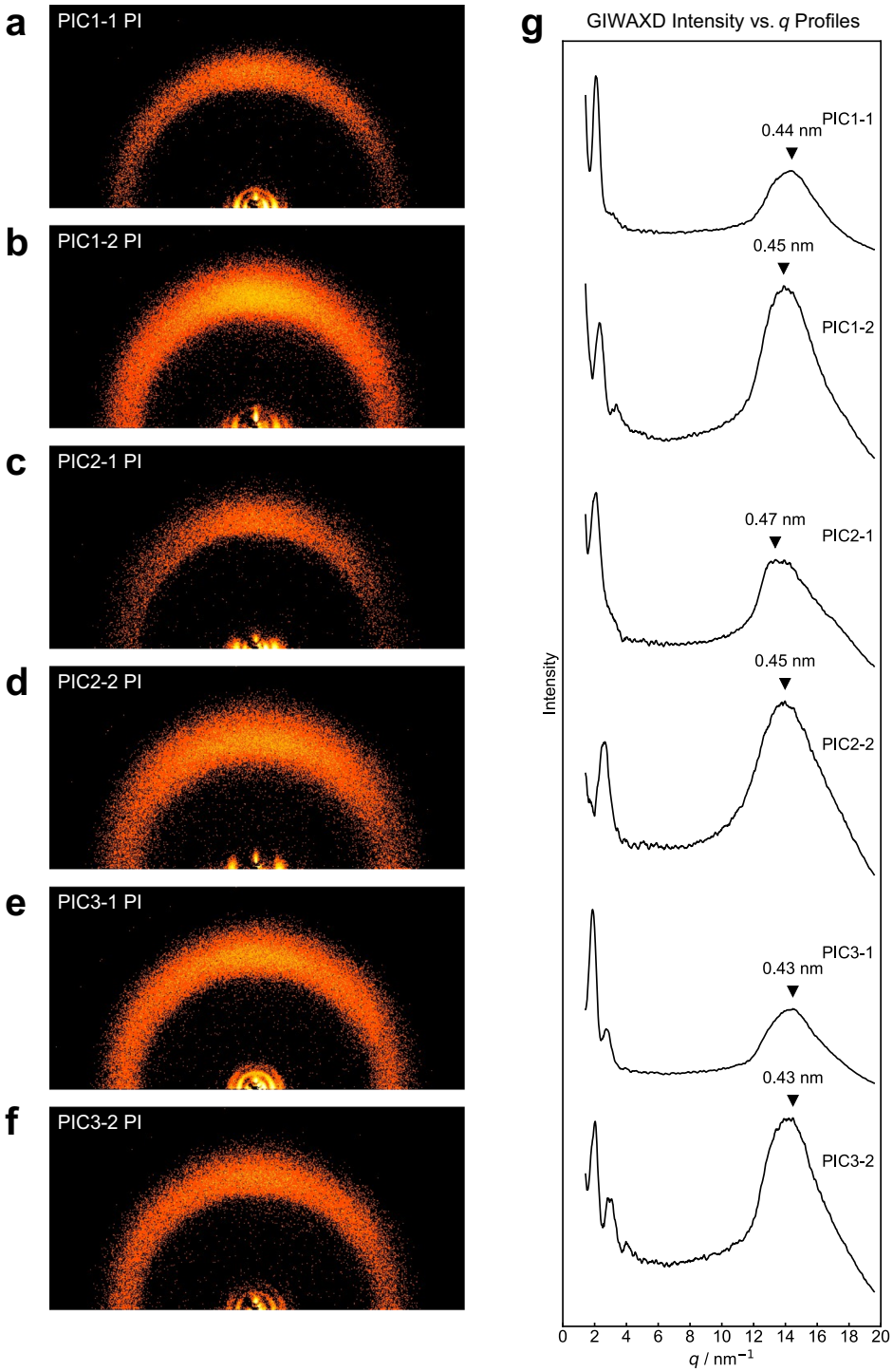


Fig. 4 | GIWAXD profiles of the spin-coated polyimide films on silicon wafers. a PIC1-1 PI, **b** PIC1-2 PI, **c** PIC2-1 PI, **d** PIC2-2 PI, **e** PIC3-1 PI, **f** PIC3-2 PI, and **g** Intensity vs. q profiles. The films were prepared under the same spin coating conditions as the samples fabricated for μ-TWA to ensure consistency in the film thickness and morphology.

structures than the PICX-1 samples, which likely contributes to their superior thermal conductivities. Furthermore, among the PICX-2 samples, those with S values closer to −0.5, indicative of a higher degree of in-plane molecular orientation, exhibited higher thermal conductivities. Specifically, PIC1-2 demonstrated the highest thermal conductivity, whereas the thermal conductivity of PIC3-2, which had the largest S value, was the lowest. These findings clearly indicate that the molecular chain orientation exerts a significant effect on the thermal conductivity. Figure 5 illustrates the relationship between the structural characteristics of the molecule (chain rigidity), degree of in-plane molecular orientation (evaluated on the basis of the order parameter S), and thermal conductivity.

The FSC measurements revealed first-order phase transitions at temperatures higher than 350 °C in the course of an ultra-fast temperature scan (10,000 K/s using FSC) during the heating and cooling processes. These observations were made possible by the extremely short residence times at high temperatures, which significantly suppressed the thermal decomposition of aromatic polyimides. To the best of our knowledge, this is the first study in which a liquid crystal transition at temperatures above 350 °C has been actually observed as a transient transition with an enthalpy change.

Discussion

This study demonstrates that machine learning can predict the formation of liquid crystalline phases for various polymers based on their chemical

Table 1 | Experimental properties of the six synthesized polyimides

Sample	T_{LC_h} (°C)	T_m (°C)	T_{LC_c} (°C)	T_C (°C)	α (×10^{-7} m^2 s^{-1})	d (nm)	C_p (J g^{-1} K^{-1})	ρ (g cm^{-3})	λ (W m^{-1} K^{-1})	S
PIC1-1	391/399	459/457	314/305	403/393	5.74 ± 0.01	600	1.18	1.32	0.894 ± 0.003	−0.45
PIC1-2	424/-	478/459	334/326	445/424	7.93 ± 0.08	500	1.16	1.37	1.26 ± 0.013	−0.44
PIC2-1	402/-	474/477	273/259	360/352	5.77 ± 0.04	470	1.2	1.31	0.904 ± 0.007	−0.40
PIC2-2	-	460/463	298/332	356/454	6.88 ± 0.17	260	1.13	1.36	1.06 ± 0.026	−0.35
PIC3-1	-	464/492	325/321	411/411	4.71 ± 0.05	540	1.16	1.32	0.722 ± 0.008	−0.33
PIC3-2	-	456/420	-/349	374/372	6.35 ± 0.36	270	1.13	1.38	0.990 ± 0.056	−0.33

T_{LC}, T_m, and T_c values measured via FSC at scan rates of 5000 and 10,000 K/s during heating and cooling; thermal diffusivities α measured by μ-TWA; thickness values d measured by the step profilometer (Alpha-Step IQ); specific heat capacities C_p determined via differential scanning calorimetry; densities ρ measured by the density gradient tube method; and thermal conductivities λ calculated from the α, C_p, and density values. The uniaxial orientational order parameter S was calculated from the GIWAXD results.

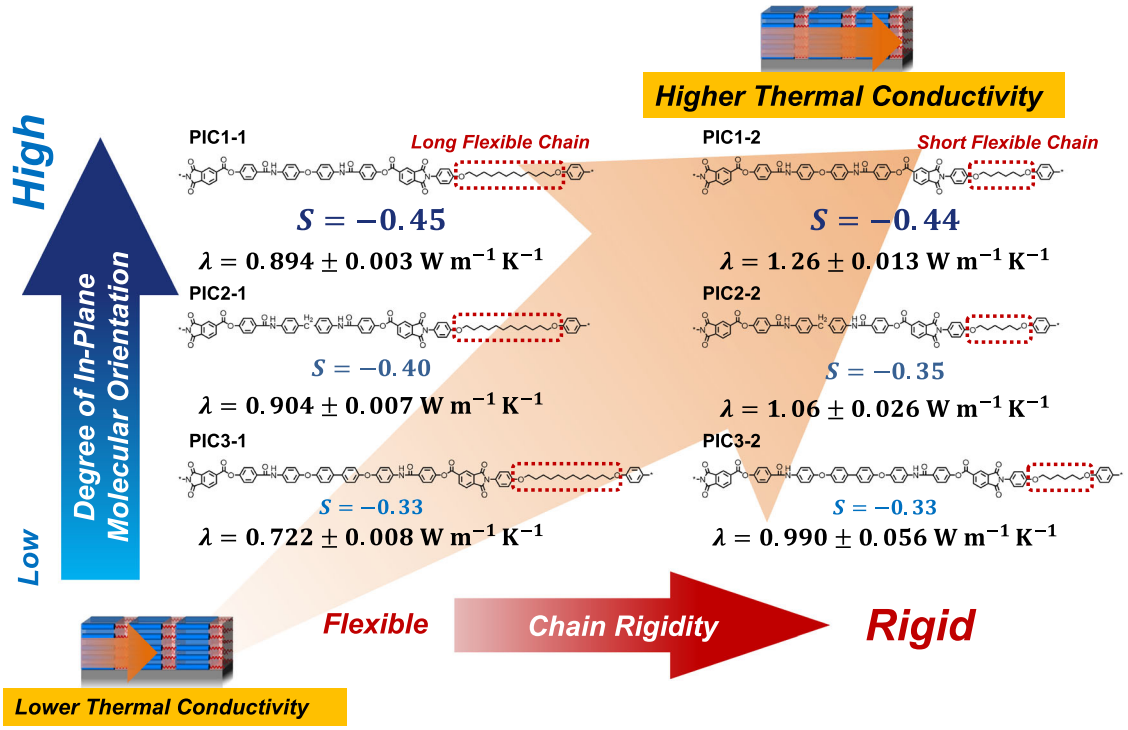


Fig. 5 | Relationship between rigidity of the molecular chain, in-plane molecular orientation, and in-plane thermal conductivity (λ).

structures. The trained predictor achieved an accuracy of over 96% within the chemical space of the tested polymer sets. Six polyimides were successfully synthesized according to the molecular design predicted via machine learning. They formed higher-order structures with smectic liquid crystalline phases. These are the first liquid crystalline polymers predicted and discovered by machine learning in the history of polymer materials research. In this proof-of-concept study, we aimed to suppress phonon scattering and enhance the thermal conductivity of the polymers by increasing the molecular orientation order through liquid crystalline phase formation. As anticipated, the thermal conductivity of the synthesized liquid crystalline polyimides exceeded 1.2 W m^{-1} K^{-1}, which is significantly higher than those of commercial polyimides in their nearly amorphous states.

Currently, fundamental methodology is not available for computationally predicting the likelihood of liquid crystalline phase formation in any given polymer, and the molecular design of these materials remains highly empirical. The proposed method offers potential as a powerful tool for facilitating the study of not only liquid crystalline polyimides but also other liquid crystalline polymers; however, several challenges persist. The reported predictive accuracy is an estimate based on a limited set of polymers, which necessitates the determination of the true generalization performance outside the current data distribution. Additionally, this study focused on binary classification to predict whether a polymer would exhibit liquid crystalline behavior, yet the prediction of specific phase types, such as nematic, smectic, and crystalline phases, should also be addressed to enable the synthesis of polymers with targeted liquid crystalline phases. Another matter is the predictability of phase transition temperatures. Furthermore, predicting the physical properties of liquid crystalline polymers represents another important challenge. Constructing a comprehensive database and establishing a methodological basis for data-driven liquid crystalline polymer chemistry are desirable goals for future studies.

Methods

Synthesis of predicted polyimides

Detailed reaction conditions and nuclear molecular resonance (NMR) and Fourier-transform infrared (FT-IR) spectra of the resulting compounds are provided in the Supplementary Information (Sections 2.1–2.5; pages S6–S34). Examples of the synthesis procedures are provided in Fig. 6.

Synthesis of tetracarboxylic dianhydrides. The three types of tetracarboxylic dianhydrides (TAC1-3) were synthesized through a three-

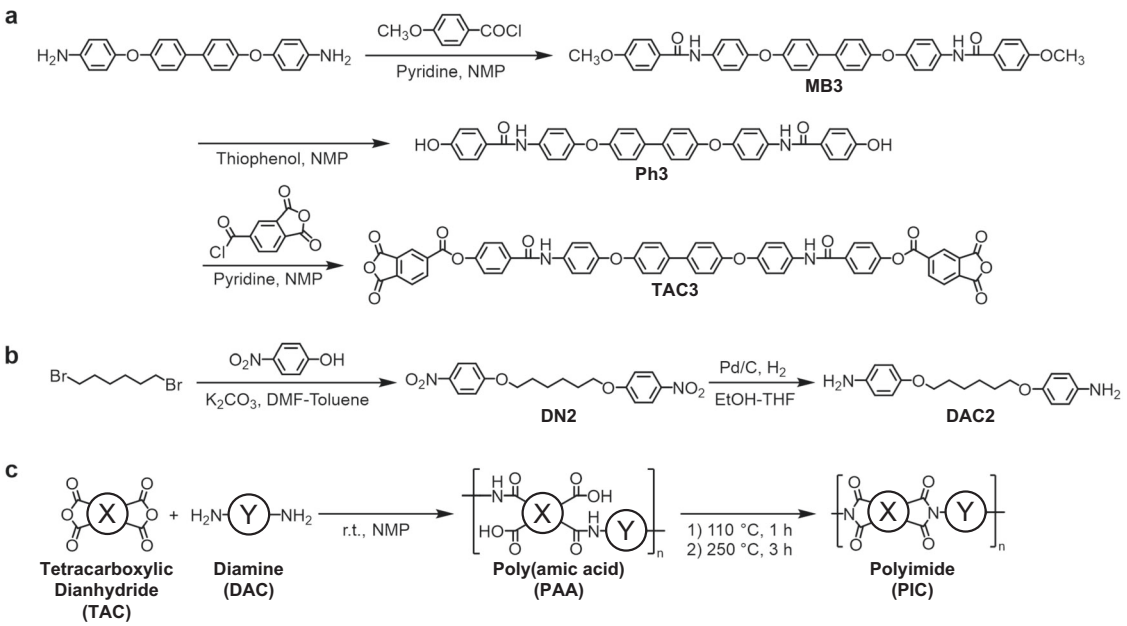


Fig. 6 | Synthesis procedure for the designed polyimides. a Synthesis of acid dianhydride TAC3; TAC1 and TAC2 were synthesized via the same procedure after replacing the starting compounds. **b** Synthesis of diamine DAC2; DAC1 was synthesized via the same procedure after replacing the starting compounds. **c** Synthesis of polyimides using a drop casting method.

step process using different aromatic diamines as starting materials. In the first step, the aromatic diamines reacted with 4-methoxybenzoic acid chloride to form amides, using *N*-methyl-2-pyrrolidone (NMP) as the solvent and pyridine as the base. The resultant compounds were purified via recrystallization from dimethylformamide. The second step involved a demethylation reaction using thiophenol, following the method developed by Chakraborti et al.[50]. The demethylated products were purified through reprecipitation in methanol. In the third step, the hydroxyl-terminated compounds were esterified with trimellitic anhydride chloride, followed by recrystallization using γ-butyrolactone.

Synthesis of diamines. The two diamines were synthesized from dibromoalkanes in a two-step process. The first step involved Williamson ether synthesis with 4-nitrophenol, followed by reduction of the nitro groups to amino groups via hydrogenation using a Pd/C catalyst. The resulting compounds were purified by recrystallization in ethanol for subsequent polymerization.

Preparation of polyimide samples. The three tetracarboxylic dianhydrides and two diamines were polymerized in NMP to prepare six poly(amic acid) solutions. Portions of these solutions were drop-cast onto glass plates and thermally imidized for wide-angle X-ray analyses. Additional samples were prepared for FSC measurements. The remaining solutions were reprecipitated, dried, and redissolved for further analysis. For thermal diffusivity measurements, a 10 wt.% solution of redissolved poly(amic acid) was spin-coated onto the sensor device and silicon wafer (2 cm × 2 cm) for GIWAXD measurements, followed by thermal imidization. Inherent viscosities were measured using an Ostwald viscometer after dissolving 0.5 g/dL of each poly(amic acid) in NMP.

Analysis of higher-order structures in spin-coated films

The molecular orientation and higher-order structures of polyimide films were analyzed by conducting GIWAXD measurements (Bruker D8 DISCOVER instrument equipped with a VANTEC-500 detector and CuKα radiation). The incident angle was consistently set to 0.25°.

Measurement of thermophysical properties

Detailed procedures for the measurement of the polymer properties are provided in the Supplementary Information (Sections 3.2–3.4; pages S36–S39).

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Information. Data and Python codes supporting the findings of this study will be made available upon reasonable request to the corresponding authors.

Received: 26 September 2024; Accepted: 18 May 2025;
Published online: 02 July 2025

References

1. Elton, D. C., Boukouvalas, Z., Fuge, M. D. & Chung, P. W. Deep learning for molecular design—a review of the state of the art. *Mol. Syst. Des. Eng.* **4**, 828–849 (2019).
2. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
3. Ikebata, H., Hongo, K., Isomura, T., Maezono, R. & Yoshida, R. Bayesian molecular design with a chemical language model. *J. Comput. Aided Mol. Des.* **31**, 379–391 (2017).
4. Wu, S. et al. Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *npj Comput. Mater.* **5**, 66 (2019).
5. Yamada, H. Predicting Materials Properties with Little Data Using Shotgun Transfer Learning. *ACS Cent. Sci.* **5**, 1717–1730 (2019).
6. Aoki, Y. et al. Multitask Machine Learning to Predict Polymer–Solvent Miscibility Using Flory–Huggins Interaction Parameters. *Macromolecules* **56**, 5446–5456 (2023).
7. Kuenneth, C. & Ramprasad, R. polyBERT: a chemical language model to enable fully machine-driven ultrafast polymer informatics. *Nat. Commun.* **14**, 4099 (2023).
8. Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R. & Jensen, K. F. Generative models for molecular discovery: Recent advances and challenges. WIREs Comput. *Mol. Sci.* **12**, e1608 (2022).

9. Ma, R. & Luo, T. PI1M: A Benchmark Database for Polymer Informatics. *J. Chem. Inf. Model.* **60**, 4684–4690 (2020).
10. Ohno, M., Hayashi, Y., Zhang, Q., Kaneko, Y. & Yoshida, R. SMiPoly: Generation of a Synthesizable Polymer Virtual Library Using Rule-Based Polymerization Reactions. *J. Chem. Inf. Model.* **63**, 5539–5548 (2023).
11. Hatakeyama-Sato, K., Tezuka, T., Nishikitani, Y., Nishide, H. & Oyaizu, K. Synthesis of Lithium-ion Conducting Polymers Designed by Machine Learning-based Prediction and Screening. *Chem. Lett.* **48**, 130–132 (2018).
12. Li, K., Wang, J., Song, Y. & Wang, Y. Machine learning-guided discovery of ionic polymer electrolytes for lithium metal batteries. *Nat. Commun.* **14**, 2789 (2023).
13. Yang, J., Tao, L., He, J., McCutcheon, J. R. & Li, Y. Machine learning enables interpretable discovery of innovative polymers for gas separation membranes. *Sci. Adv.* **8**, eabn9545 (2022).
14. Gurnani, R. et al. AI-assisted discovery of high-temperature dielectrics for energy storage. *Nat. Commun.* **15**, 6107 (2024).
15. Qian, X., Zhou, J. & Chen, G. Phonon-engineered extreme thermal conductivity materials. *Nat. Mater.* **20**, 1188–1202 (2021).
16. Xie, X. et al. High and low thermal conductivity of amorphous macromolecules. *Phys. Rev. B* **95**, 35406 (2017).
17. Du, Y.-K. et al. Recent Progress in Fabrication and Structural Design of Thermal Conductive Polymer Composites. *Chin. J. Polym. Sci.* **42**, 277–291 (2024).
18. Hwang, J. S. Filler size and content effects on the composite properties of anisotropic conductive films (ACFs) and reliability of flip chip assembly using ACFs. *Microelectron. Reliab.* **48**, 645–651 (2008).
19. Wang, Y. et al. Epoxy Composites with High Thermal Conductivity by Constructing Three-Dimensional Carbon Fiber/Carbon/Nickel Networks Using an Electroplating Method. *ACS Omega* **6**, 19238–19251 (2021).
20. Lv, G. et al. Odd-even effect on the thermal conductivity of liquid crystalline epoxy resins. *Proc. Natl. Acad. Sci.* **119**, e2211151119 (2022).
21. Kricheldorf, H. R. Liquid-Crystalline Polyimides. in *Progress in Polyimide Chemistry II* (ed. Kricheldorf, H. R.) 83–188 https://doi.org/10.1007/3-540-49814-1_3 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999).
22. Collings, P. J. & Hird, M. *Introduction to Liquid Crystals Chemistry and Physics* https://doi.org/10.1201/9781315272801 (CRC Press, 2017).
23. Donald, A. M., Windle, A. H. & Hanna, S. *Liquid Crystalline Polymers*. https://doi.org/10.1017/CBO9780511616044 (Cambridge University Press, Cambridge, 2006).
24. Wang, X.-J. & Zhou, Q.-F. *Liquid Crystalline Polymers*. https://doi.org/10.1142/5309 (WORLD SCIENTIFIC, 2004).
25. Ober, C. K., Jin, J.-I., Zhou, Q. & Lenz, R. W. Liquid crystal polymers with flexible spacers in the main chain. in Platé, N.A. (eds) Liquid Crystal Polymers I. Advances in Polymer Science, Springer, Berlin, Heidelberg, **59**, 103–146 https://doi.org/10.1007/3-540-12818-2_8 (1984).
26. Wunderlich, B. & Grebowicz, J. Thermotropic mesophases and mesophase transitions of linear, flexible macromolecules. in Platé, N.A. (eds) Liquid Crystal Polymers II/III. Advances in Polymer Science, Springer, Berlin, Heidelberg, **60-61**, 1–59 https://doi.org/10.1007/3-540-12994-4_1 (1984).
27. Tao, K. et al. Structural Design and Research Progress of Thermally Conductive Polyimide Film - A Review. *Macromol. Rapid Commun.* **44**, 2300060 (2023).
28. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
29. Hayashi, Y., Shiomi, J., Morikawa, J. & Yoshida, R. RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics. *npj Comput. Mater.* **8**, 222 (2022).
30. Kusaba, M., Hayashi, Y., Liu, C., Wakiuchi, A. & Yoshida, R. Representation of materials by kernel mean embedding. *Phys. Rev. B* **108**, 134107 (2023).
31. Wu, S., Lambard, G., Liu, C., Yamada, H. & Yoshida, R. iQSPR in XenonPy: A Bayesian Molecular Design Algorithm. *Mol. Inform.* **39**, 1900107 (2020).
32. Thompson, A. P. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **271**, 108171 (2022).
33. Masashi Ishii Takuro Ito, H. S. & Kuwajima, I. NIMS polymer database PoLyInfo (I): an overarching view of half a million data points. *Sci. Technol. Adv. Mater.: Methods* **4**, 2354649 (2024).
34. Bekker, J. & Davis, J. Learning from positive and unlabeled data: a survey. *Mach. Learn.* **109**, 719–760 (2020).
35. Elkan, C. & Noto, K. Learning Classifiers from Only Positive and Unlabeled Data. in *Proc. 14th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining* 213–220 https://doi.org/10.1145/1401890.1401920 (Association for Computing Machinery, New York, NY, USA, 2008).
36. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework. in *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining* (2019).
37. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **52**, 1757–1768 (2012).
38. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).
39. Campello, R. J. G. B., Moulavi, D. & Sander, J. Density-Based Clustering Based on Hierarchical Density Estimates. in *Advances in Knowledge Discovery and Data* Mining (eds. Pei, J., Tseng, V. S., Cao, L., Motoda, H. & Xu, G.) 160–172 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).
40. Campello, R. J. G. B., Moulavi, D., Zimek, A. & Sander, J. Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection. *ACM Trans. Knowl. Discov. Data* **10**, 1–51 (2015).
41. Ishige, R., Tanaka, K. & Ando, S. Quantitative analysis of stereoscopic molecular orientations in thermally reactive and heterogeneous noncrystalline thin films via variable-temperature infrared pMAIRS and GI-XRD. *Polym. J.* **53**, 603–617 (2021).
42. Sadeghi, M. M., Pettes, M. T. & Shi, L. Thermal transport in graphene. *Solid State Commun.* **152**, 1321–1330 (2012).
43. Shen, S., Henry, A., Tong, J., Zheng, R. & Chen, G. Polyethylene nanofibres with very high thermal conductivities. *Nat. Nanotechnol.* **5**, 251–255 (2010).
44. Cahill, D. G. Analysis of heat flow in layered structures for time-domain thermoreflectance. *Rev. Sci. Instrum.* **75**, 5119–5122 (2004).
45. Jiang, F. et al. Lock-in photothermal method for in-plane thermal diffusivity measurements using arrayed temperature sensors on suspended SiNx membranes. *Rev. Sci. Instrum.* **94**, 94903 (2023).
46. Minakov, A., Morikawa, J., Zhuravlev, E., Ryu, M. & Schick, C. Thermal contact conductance at melting and crystallization of metal micro-droplets. *Mater. Res. Express* **7**, 66524 (2020).
47. Guo, Y., Ruan, K., Shi, X., Yang, X. & Gu, J. Factors affecting thermal conductivities of the polymers and polymer composites: A review. *Compos. Sci. Technol.* **193**, 108134 (2020).
48. Hasegawa, M., Inoue, T., Miyauchi, R. & Ishii, J. Polyimides Containing Benzoxazole Units and Their Liquid-Crystalline Behavior. *Macromol. Res.* **26**, 900–912 (2018).
49. Zhang, T. & Luo, T. Role of Chain Morphology and Stiffness in Thermal Conductivity of Amorphous Polymers. *J. Phys. Chem. B* **120**, 803–812 (2016).
50. Chakraborti, A. K., Sharma, L. & Nayak, M. K. Demand-Based Thiolate Anion Generation under Virtually Neutral Conditions: Influence of Steric and Electronic Factors on Chemo- and Regioselective Cleavage of Aryl Alkyl Ethers. *J. Org. Chem.* **67**, 6406–6414 (2002).

Acknowledgements
The synchrotron radiation experiments were performed at the BL40B2 beamline of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2022B1131), with support from Dr. Noboru Ohta (JASRI) and Prof. Tomoyasu Hirai (Osaka Institute of Technology). This work was supported by the Japan Science and Technology Agency (JST) under the CREST program (Grant Number JPMJCR19I3) (J.M., T.H., R.Y., M.T., R.M.). Additionally, we acknowledge the Japan Society for the Promotion of Science (JSPS) for their support through the KAKENHI program (Grant Number 21K04828) (Y.N.) and the Ministry of Education, Culture, Sports, Science and Technology for "Program for Promoting Researches on the Supercomputer Fugaku" (project ID: hp210264) (R.Y.). HM and RM were supported by JST SPRING (Grant No. JPMJSP2106).

Author contributions
Conceptualization: R.Y., J.M., T.H.; Data Curation: H.M., S.W., J.M.; Formal Analysis: H.M., S.W., J.M.; Funding Acquisition: Y.N. (Yuta Nabae), R.Y., J.M.; Investigation: H.M., S.W., R.M., E.Y., S.N., Y.N. (Yoh Noguchi), J.M.; Methodology: M.R., R.I., F.J., X.T.V., S.I., M.T., J.M., R.Y., T.H.; Resources: M.I., I.K.; Software: S.W., Y.N. (Yoh Noguchi), Y.H.; Supervision: R.Y., J.M., T.H.; Visualization: H.M., S.W., S.N., J.M.; Writing: H.M., S.W., J.M., R.Y.; Writing—review & editing: H.M., S.W., K.H., Y.N. (Yuta Nabae), R.Y., J.M., T.H.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-025-01671-w.

Correspondence and requests for materials should be addressed to Junko Morikawa, Ryo Yoshida or Teruaki Hayakawa.