npj | computational materials

Article

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

https://doi.org/10.1038/s41524-025-01696-1

A machine learning approach to designing and understanding tough, degradable polyamides

Yoshifumi Amamoto [1,2,3,4] ✉, Chie Koganemaru[1], Ken Kojio[1,5], Atsushi Takahara[1,5], Sayoko Yamamoto[6], Kazuki Okazawa[1], Yuta Tsuji[1,6], Toshimitsu Aritake[7] & Kei Terayama[8,9,10] ✉

The development of environmentally friendly plastics has received renewed attention for a sustainable society. Although the trade-off between toughness and degradability is a common challenge in biodegradable polymers, the design of biodegradable polymers to overcome these issues is often difficult. In this study, we demonstrated that machine learning techniques can contribute to the development of multiblock polyamides composed of Nylon6 and α-amino acid segments that are mechanically tough and degradable. Multi-objective optimization based on Gaussian process regression for the degradation rate, strain at break, and Young's modulus (the last two parameters correspond to toughness) suggested appropriate α-amino acid sequences for polyamides endowed with both properties. Ridge regression revealed that the physical factors associated with the sequences, as well as the higher-order multiblock-derived structures (such as the crystal lattice structure, melting points, and hydrogen bonding), were essential for endowing these polymers with satisfactory properties among the multimodal measurement/calculation data. Our method provides a useful approach for designing and understanding environment-friendly plastics and other materials with multiple properties based on machine learning techniques.

A sustainable society represents the ultimate goal of preserving the environment and advancing human development, and the use of environment-friendly plastics is necessary to achieve such a society[1,2]. Plastics released into the natural environment persist for a long time and can affect ecosystems and human health[3–5]. Recently, segregated microplastics formed in the sea have become an issue of concern as an ocean plastic problem[6,7], and the use of biodegradable polymers represents an option for overcoming this problem[8–10]. The functional groups in a polymer dissociate during plastic degradation; however, biodegradable polymers are not typically used in everyday products because of issues associated with the trade-off relationship between the mechanical toughness and degradability of the polymer, where high degradability is often associated with low toughness, and vice versa.

Machine learning techniques have enabled more sophisticated human development of material and chemical science[11,12]. In terms of polymer design, machine learning techniques have been used to determine the optimal monomer sequences in polymers/oligomers[13–15]. For example, optimized α-amino acids in luminescent proteins have been explored using Bayesian optimization (BO)[16]. Furthermore, for the estimation of peptide structures, optimized sequences of self-assembled peptides were explored using Monte Carlo tree search, along with using random forest methods. This resulted in the development of a novel sequence with exceptional performance by eliminating human bias[17]. It should be noted that these polymer-design targets have single properties. Recently, multiobjective optimization, which targets multiple properties, has been applied to the design of materials based on BO and generic algorithms[18,19]. However, to the best of our knowledge, machine learning-assisted designs of biodegradable polymers that are both tough and degradable have not yet been developed.

Machine learning techniques have also contributed to the understanding of important factors in materials, including biodegradable

[1]Institute for Materials Chemistry and Engineering, Kyushu University, Fukuoka, Japan. [2]Graduate School of Artificial Intelligence and Science, Rikkyo University, Toshima-ku, Tokyo, Japan. [3]Graduate School of Social Data Science, Hitotsubashi University, Kunitachi, Tokyo, Japan. [4]Japan Synchrotron Radiation Research Institute, Sayo-gun, Hyogo, Japan. [5]Research Center for Negative Emission Technology, Kyushu University, Fukuoka, Japan. [6]Faculty of Engineering Sciences, Kyushu University, Kasuga, Fukuoka, Japan. [7]Hitotsubashi Institute for Advanced Study, Hitotsubashi University, Kunitachi, Tokyo, Japan. [8]Graduate School of Medical Life Science, Yokohama City University, Yokohama, Kanagawa, Japan. [9]RIKEN Center for Advanced Intelligence Project, Tokyo, Japan. [10]MDX Research Center for Element Strategy, Institute of Science Tokyo, Yokohama, Kanagawa, Japan. ✉e-mail: y.amamoto@r.hit-u.ac.jp; terayama@yokohama-cu.ac.jp

polymers[20–27]. In the case of polymers, not only the chemical structure but also the complicated structures related to the surface and higher-order structures are essential for this property. Because a wide range of aspects are utilized, using multiple measurement and calculation methods, humans often struggle to recognize the complicated relationships among multiscale, multimodal, and multivariate data on many samples[11,27]. Recently, machine learning techniques, including explainable artificial intelligence (XAI), have contributed to the specification of important physical factors or features from the measurement data of biodegradable polymers. For instance, key features have been extracted from the molecular descriptor and measurement data based on feature selection[26]. Furthermore, the X-ray scattering images were analyzed using XAI techniques such as "gradient-weighted class activation mapping" and "Shapley additive explanations," in which the diffraction peaks and small-angle regions were recognized as significant regions[21]. Currently, the research focus has primarily been restricted to single measurement techniques or monomodal data. Therefore, the remaining challenge involves establishing a methodology to evaluate the essential multiscale and/or multimodal information factors for integration analysis.

Herein, we report the design and understanding of multiblock polyamides composed of Nylon6 and α-amino acid segments in terms of both toughness and degradability using machine learning techniques (Fig. 1). BOs have been used to suggest optimized α-amino acid sequences in alternating multiblock copolymers that satisfy multiple properties associated with trade-off relationships. Furthermore, the essential physical factors for these properties were extracted from multimodal data based on ridge regression.

Results

Polyamide preparation

Tough and degradable polymers have been designed using polyamides[28–34]. The amide bonds in polyamides provide strong intermolecular interactions that endow them with toughness, high degradation selectivity, and high thermal stability; consequently, various methods for synthesizing multiblock polyamides have been established[32,33]. In addition, amino acid-based fibers have been used in sportswear and bulletproof vests[35]. Furthermore, a wide range of monomers, including α-amino acids, can be utilized, and the monomer-sequence combinations and regularities are easily controlled. Prior to polymer synthesis, the length of the α-amino acid (αAA) sequence required for degradability was investigated using model reactions involving oligopeptides in aqueous solutions. Oligopeptides with different αAA sequence lengths positioned between 6-aminohexanoic acid (AHA) as the monomer units of Nylon6 were enzymatically degraded using Proteinase K in Tris buffer solution (Fig. S1a). Minimal enzymatic degradation was observed for Met (Fig. S1b). The degradation rates of the oligopeptides ($\rho_{\mathrm{enzyme,\ oligo}}$) increased with increasing αAA content. The three αAAs dramatically enhanced oligopeptide degradation, even at low enzyme concentrations (Ala-Met-Ala, Fig. S1b, oligopeptide: 1 mM, Proteinase K: 0.1 μM). Liquid chromatography–mass spectrometry (LC-MS) revealed that the C-terminus of Met is the most degraded part of the oligopeptide. On the other hand, the type of αAA sequence was also important for enhancing the oligopeptide degradation rate. For example, the degradation rate of AHA-Ala-AA^2-Ala-AHA clearly depended on the central αAA sequence (Fig. S1c), with AHA-Ala-AA^2-Ala-AHA being more suitable for enzymatic degradation than random sequences (Fig. S1c, d). It is worth mentioning that AHA-Ala-AHA-Ala-AHA was barely degraded under the current conditions (AHA, Fig. S1c). We conclude that a sequence of three continuous αAAs is required for sufficient enzymatic degradation.

Polyamides were synthesized by reacting oligopeptides with coupling agents (Fig. 2a). Oligopeptides with different numbers of AHA units and fixed Ala-Met-Ala segments were reacted to determine the suitable length of the AHA segment. In gel permeation chromatography (GPC), large molecules corresponding to high molecular weight compounds elute earlier. A peak in the higher molecular weight region was observed in the GPC curves after purification; the position of this peak effectively corresponded to that of commercially available Nylon6 (Fig. 2b). Mass spectrometry revealed multiple peaks with periodic widths corresponding to dehydrated oligopeptides (Fig. S2). These results indicated that a polyamide with an alternating multiblock structure was successfully formed. Polyamide films with different numbers of AHA units were enzymatically degraded to determine

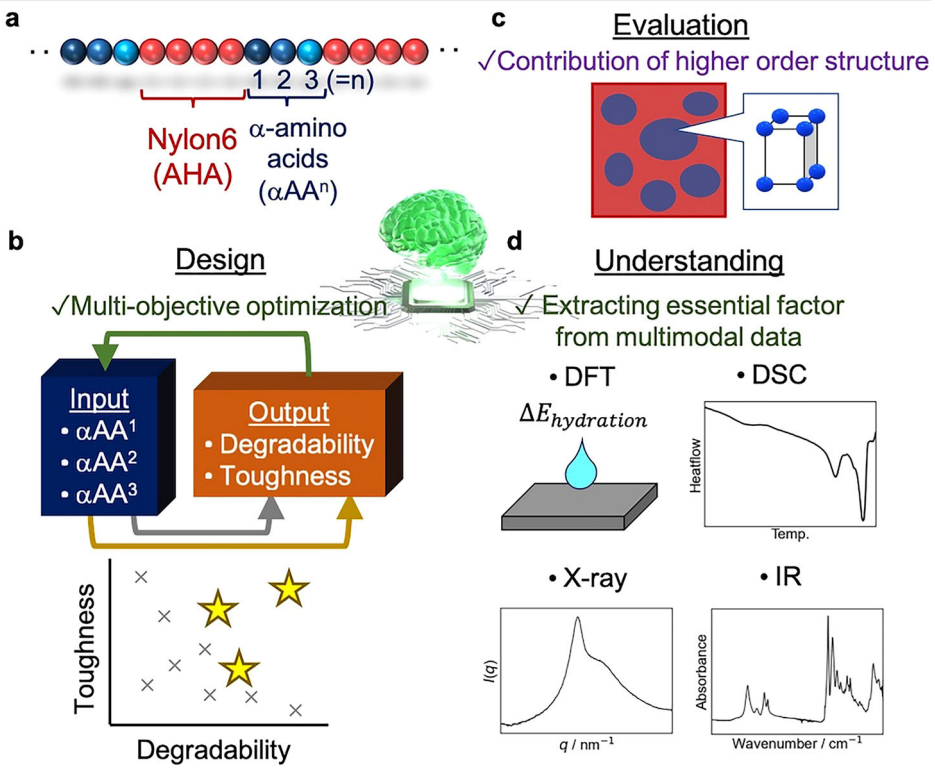

Fig. 1 | Schematic representation of this work. **a** Multiblock polyamide consisting of AHA and αAA segments. **b** Multi-objective polyamide optimization for degradability and toughness. **c** Evaluating the material properties of higher-order polyamide structures. **d** Extracting important physical factors for property from experimental/calculational data.

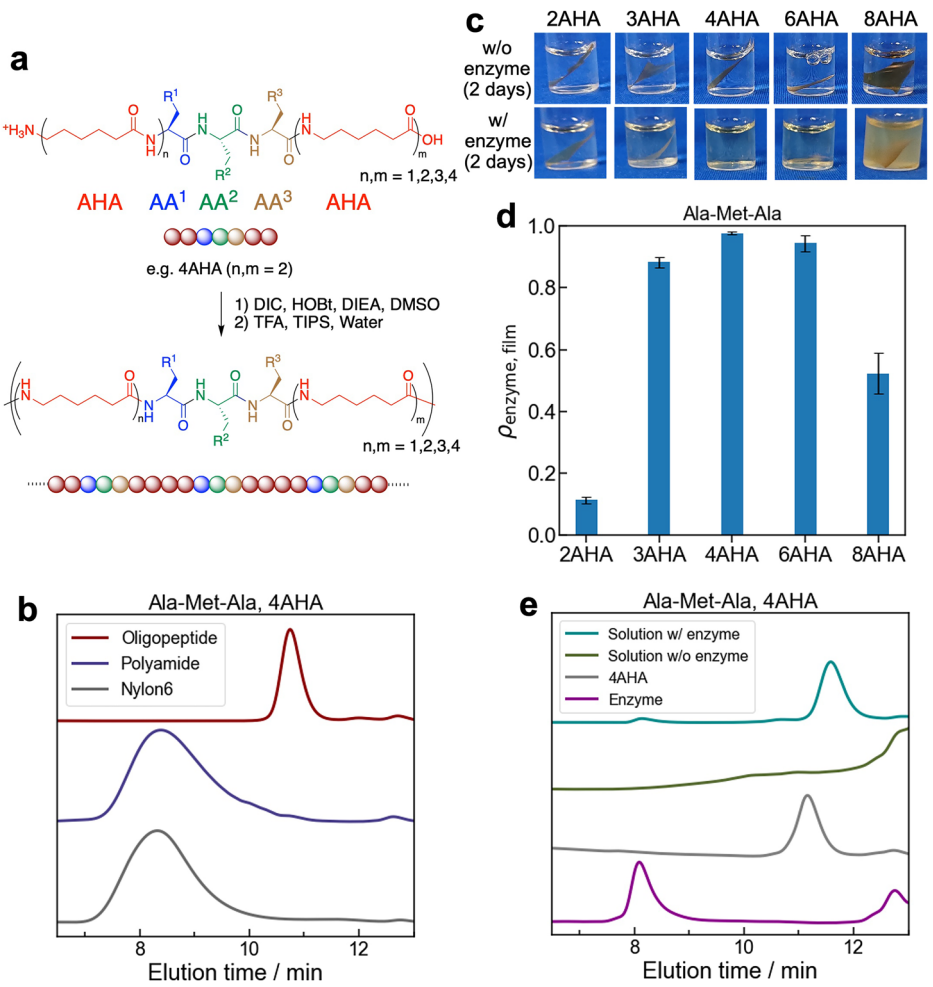

Fig. 2 | Polyamide preparation and enzymatic degradability. a Multiblock polyamide synthesis by coupling an oligopeptide composed of AHA and αAA units. **b** GPC traces for Nylon6 and the polyamide before and after coupling. **c** Photographic images. The films with Ala-Met-Ala units and varying numbers of AHA units in buffer solutions were observed after 2 days, both without (w/o) and with (w/) the enzyme. **d** Enzymatic degradation rates of polyamide films ($\rho_{enzyme,\ film}$) after 2 days in a buffer solution of Proteinase K. The degradation rates were estimated using control samples immersed in buffer without the enzyme; these rates reflect the enzymatic degradational process, primarily occurring at the surface. **e** GPC traces for reaction products with/without enzymes, 4AHA, and enzymes.

the optimal number of AHA units. While 2AHA was hardly degraded (Fig. 2c, d), the degradation rate increased as the number of AHA units increased to four, after which it decreased (Fig. 2c, d). The polyamide films with 4AHA and 6AHA disappeared after 2 days (Fig. 2c). Hence, we concluded that 4AHA has a suitable number of segments for enzymatic degradation. The mechanisms underlying these differences are discussed below. AHA and αAA sequences that were four and three units long, respectively, were used hereafter.

The degradation products of the polyamide films composed of 4AHA and Ala-Met-Ala were evaluated by GPC after degradation testing with and without the enzyme. The major product had a slightly lower molecular weight than the original oligopeptide (Fig. 2b, dark red) in the presence of the enzyme (Fig. 2e), whereas no corresponding peaks were observed in the absence of the enzyme. Since cyclic compounds were formed, the peak could have appeared later in the elution process. As mentioned above, amide bonds involving AHA units hardly degrade under the current conditions; therefore, the 4AHA oligopeptide was observed after enzymatic degradation of the αAAs in the polyamide. Nylon 6 itself is reportedly poorly biodegradable. Although amide bonds in AHA units were scarcely degraded by the currently used enzyme, Proteinase K, it has been reported that Nylon6 oligomers can be metabolized in the natural environment using different enzymes[36]. Hence, multiblock structures offer a strategy for the biodegrading of Nylon6 derivatives.

Multi-objective polyamide optimization

Prior to the multi-objective optimization study, we investigated the material properties of polyamides with various αAA sequences. Water-soluble αAAs, such as lysine and aspartic acid, were excluded because the polyamide films were dissolved in a buffer solution without enzymes. Therefore, fourteen αAAs among twenty essential ones for humans were utilized, which led to $14 \times 14 \times 14 = 2744$ candidate combinations. We prepared polyamides (4AHA) with eight predetermined sequences (Ala-AA^2-Ala) and 17 random sequences (AA^1-AA^2-AA^3). The mechanical properties were evaluated by the uniaxial elongation of the polyamide films at room temperature. The stress–strain curve clearly depended on the αAA sequence. For example, some films were brittle-like glasses (blue, Fig. 3a), whereas some stress–strain curves showed yield points similar to those observed for crystalline polymers (green, Fig. 3a). A small number of polyamide films exhibited elastomer-like stress–strain curves with low Young's moduli and high strains at break (orange, Fig. 3a). Polyamide films with different sequences were also subjected to enzymatic degradation. Enzymatic degradation tests were chosen since they are time efficient and facilitate the comparison across different time periods. The degradation rate of films ($\rho_{enzyme,\ film}$) depended on the αAA sequence (Fig. 3b, Table S1). Several films almost disappeared after 2 days, whereas the Nylon6 film hardly degraded under the current conditions (Fig. S3a). Polyamides with high degradation rates contain Ala in AA1 and/or AA3 and specific amino acids, such as Met, Glu, and Leu in AA2. This tendency was also confirmed by the degradation tests of oligopeptides (Fig. S1c, d). Other sequences exhibited low or moderate degradation rates. Therefore, degradation rates in the range of 0.5–0.7 were not observed. Furthermore, the polyamide films were degraded using a different enzyme, pepsin, and the degradation patterns differed because of the substrate specificity of the enzymes (Table S1). Other representative biodegradable polymers such as poly(L-lactic acid) (PLLA), polybutylene succinate (PBS), and polybutylene succinate-*co*-adipate (PBSA) were less degradable (Fig. S3a). This suggests that amide bonds with

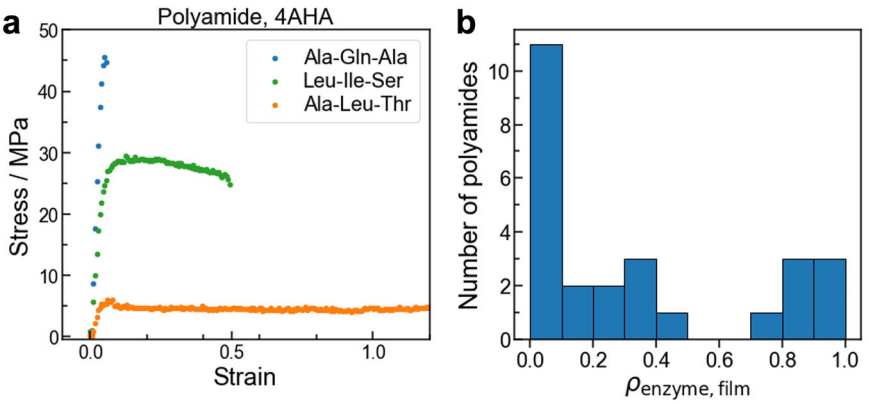

Fig. 3 | Properties of polyamide films with several α-amino acid sequences. a Stress–strain curves of representative polyamide films at room temperature. **b** Enzymatic degradation rates of polyamide films ($\rho_{enzyme,\ film}$) with several αAA sequences.

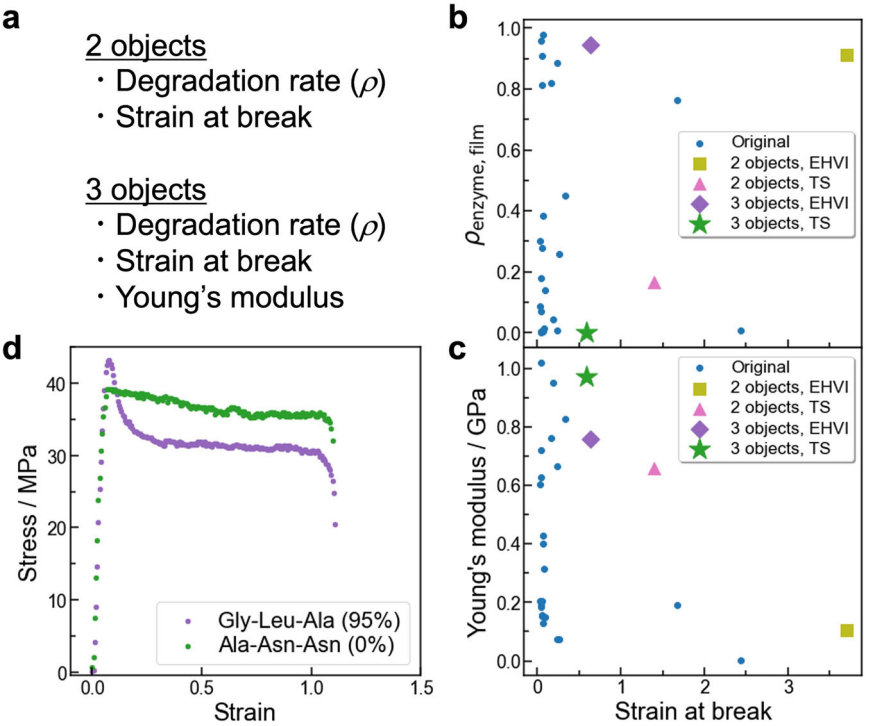

Fig. 4 | Multi-objective polyamide optimization for toughness and degradability using BO. **a** Target properties for the multi-objective optimization. **b**, **c** Multi-objective optimization results for polyamide films in terms of $\rho_{enzyme,\ film}$, Young' modulus, and strain at break. **d** Stress–strain curves for polyamide films with sequences suggested by BO; $\rho_{enzyme,\ film}$ values are shown in parentheses.

appropriate sequences degrade faster because of the substrate specificity of Proteinase K, although the enzyme is capable of breaking both amide and ester bonds. The polyamide films did not always show a degradation behavior consistent with that of the oligopeptides (Fig. S7a), which will be discussed later. These results revealed that the αAA sequence significantly affected the toughness and enzymatic degradability of the polyamide film.

Next, we subjected the AA sequence in the polyamide to a multi-objective optimization for toughness and degradability. T-scales were used as the αAA descriptors[37]. Bayesian optimization (BO) based on Gaussian process regression was used, where expected hyper-volume improvement (EHVI)[38] and Thompson sampling (TS)[39] were the objective functions of the multi-objective BO. The number of iterations of the BO was set to one in all cases. We first attempted to optimize the enzymatic degradation rate and strain at break, which revealed trade-off relationships in all samples except one (Fig. 4a, b). The predicted values moderately agree with the actual values, indicating that the current BO approach is effective (Fig. S4a, b). The EHVI exhibited a remarkable improvement in the Pareto solution (ocher, Fig. 4b); therefore, we used multi-objective optimization in the current approach. However, this sample exhibited elastomer-like stress–strain curves with a low Young's modulus (ocher, Fig. S4c). Therefore, we performed a multi-objective optimization for three parameters (degradation rate, strain at break, and Young's modulus). The Gly-Leu-Ala containing polyamide exhibited Pareto regions in two scatter plots (purple, Fig. 4b, c), as suggested by the EHVI. This behavior was also confirmed by the increase in the hypervolume of the Pareto points through EHVI (Fig. S4d). Furthermore, the stress–strain curve of the polyamide revealed behavior similar to that of a crystalline polymer (Fig. 4d). The strain at break of the polyamide was higher than that of Nylon6 prepared by solvent casting, but lower than that of Nylon6 molded by hot pressing (Fig. S3b). The current single iteration provides sufficient improvement in the Pareto regions. Owing to the relatively high experimental costs, further progress using BO will be explored in the near future. Films with the suggested αAA sequences exhibited low enzymatic degradation rates for both two and three objects when TS was used as the objective function. Nevertheless, polyamides with superior enzymatic degradability and high mechanical performance have been obtained via BO-based multi-objective optimization.

Phase separation in a multiblock polyamide

A multiblock polyamide is expected to form an aggregated structure on the nanoscale because AHA and αAA segments are repeatedly located in one

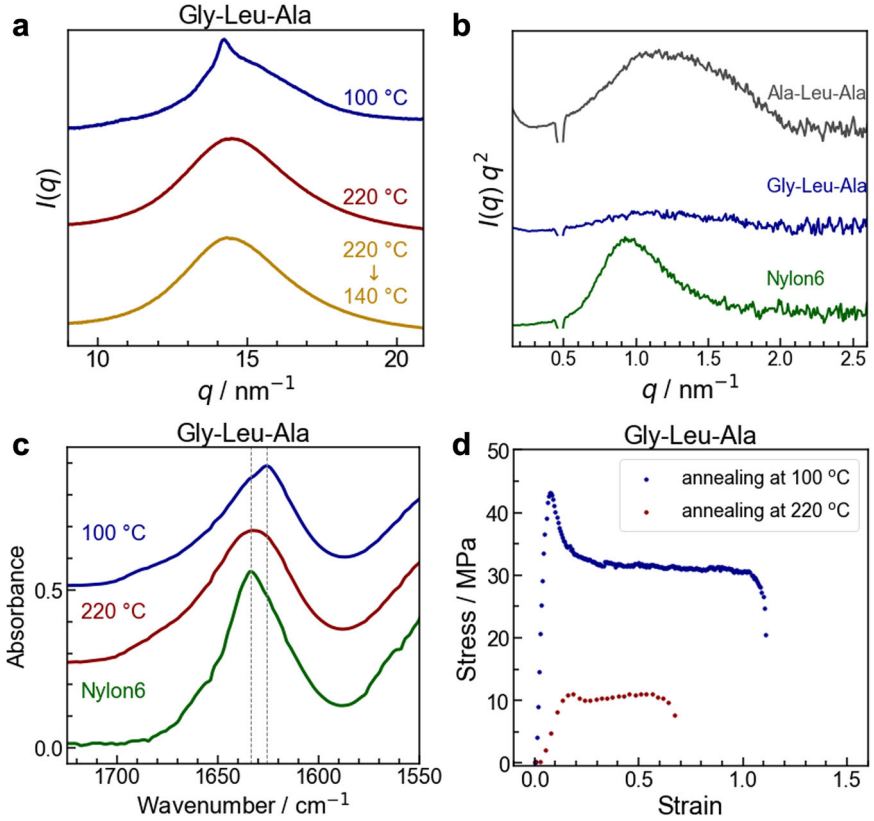

Fig. 5 | Multiblock polyamide phase separation and its effect on mechanical properties. a WAXS profiles of a thermally treated polyamide containing Gly-Leu-Ala sequences and 4AHA units, with annealing temperatures listed. **b** Kratky plots of SAXS data for Nylon6 and polyamides with different sequences. **c** IR spectra of Nylon6 and a thermally treated polyamide containing Gly-Leu-Ala sequences and 4AHA units, with annealing temperatures listed. **d** Stress–strain curves for a Gly-Leu-Ala-containing polyamide thermally treated at 100 and 220 °C. Tensile testing was carried out at room temperature.

chain. Therefore, the thermal properties and nanoscale structures of the polyamides were evaluated using differential scanning calorimetry (DSC), wide-angle X-ray scattering (WAXS), and small-angle X-ray scattering (SAXS) experiments. The DSC curve of the Gly-Leu-Ala-containing polyamide exhibited a baseline shift at 50 °C during first heating (Fig. S5a), which is close to the glass transition temperature (T_g) of Nylon6. Two overlapping melting peaks were observed at 190 °C; these peaks were separately observed at 190 and 240 °C in the case of Ala-Met-Ala (blue, Fig. 6b). The melting point of poly(αAA) is reportedly higher than that of Nylon6[40]; hence, the two melting peaks are derived from the AHA-rich and αAA-rich phases. Diffraction peaks at q~14 nm^{-1} and an amorphous halo, whose ratios depended on the αAA sequence, were observed by WAXS (Fig. 5a, black, and Fig. S5b), while SAXS revealed a scattering peak at q~1.2 nm^{-1} for the Ala-Leu-Ala-containing polyamide, which is higher than that of a long-range structure that corresponds to the lamella thickness of Nylon6 (Fig. 5b, green). This peak was less intense for some samples (e.g., Gly-Leu-Ala, blue, Fig. 5b). Furthermore, the Ala-Met-Ala film was transparent, and no structures were observed by polarization microscopy (Fig. S5c). These results show that these multiblock copolymers form phase-separated structures composed of AHA and αAA segments on the several-nanometer scale without the formation of spherulites, and that the clarity of the structure depends on the sequences.

The phase-separated structure was altered by heating, as evidenced by changes in the DSC curves observed during the second heating process, in which T_g peaks became more intense and melting peaks disappeared (Fig. S5a). WAXS and IR techniques were used to confirm heating-related changes in crystal structure. A diffraction peak was observed for Ala-Met-Ala up to 210 °C; this peak disappeared with further heating above 250 °C for Ala-Met-Ala and 220 °C for Gly-Leu-Ala, which is above the second melting peak in the DSC curve (Figs. 5a and S5d), indicating that the diffraction peaks at q~14 nm^{-1} are derived from the crystal structures of αAA-rich phases[29]. Hence, αAA-rich phases crystallize in the phase-separated polyamide films. The crystal peaks in the WAXS profiles were not regenerated by maintaining the film at 140 °C for 10 min after being heating above their melting points (T_m), which is consistent with the absence of any melting peak during the second DSC heating process (Figs. 5a and S5a). Furthermore, the two overlapping peaks that correspond to the stretching vibrations of hydrogen-bonded C=O groups in the AHA-rich and αAA-rich phases were observed as a single peak in the IR spectrum after the polyamide films had been heated above T_m (Figs. 5c and S5e). In terms of T_g, the baseline for the Gly-Leu-Ala-containing polyamide appeared to shift at ~80 °C, which is between the T_g regions of the AHA-rich and αAA-rich phases (green, Fig. S5a). Hence, the phase-separated structure disappeared when heated above the melting point of the αAA-rich phases to form a non-crystalline miscible system. We conclude that a phase-separated structure was formed during the solvent-casting process given that the two segments are miscible when heated[41].

We speculated that both the αAA sequence and the higher-order structure of a polyamide contribute to both of the abovementioned properties, as heating a polyamide film above its melting point alters its phase-separated and crystal structure. Therefore, we subjected films thermally treated at 100 and 220 °C, which are below and above their melting points, respectively, to tensile testing. The polyamide films treated at 100 °C still showed yield points in their stress–strain curves when heated, although less stress was observed (Fig. S5f, blue). The mechanical properties of the polyamide were sufficiently maintained because the crystal structure is maintained up to the melting point. On the other hand, the polyamide films treated at 220 °C (which are not crystalline) are brittle at room temperature and become quite soft when heated (Figs. 5d and S5f, red). Because the thermally treated films have glass transition temperatures of about 80 °C, they transform from their glass states to melt/rubber states when heated during tensile testing. Hence, a phase-separated crystal structure endows the multiblock polyamide with high mechanical performance, especially when heated.

As mentioned in the previous section, enzymatic degradation of polyamides with different numbers of AHA units and fixed Ala-Met-Ala

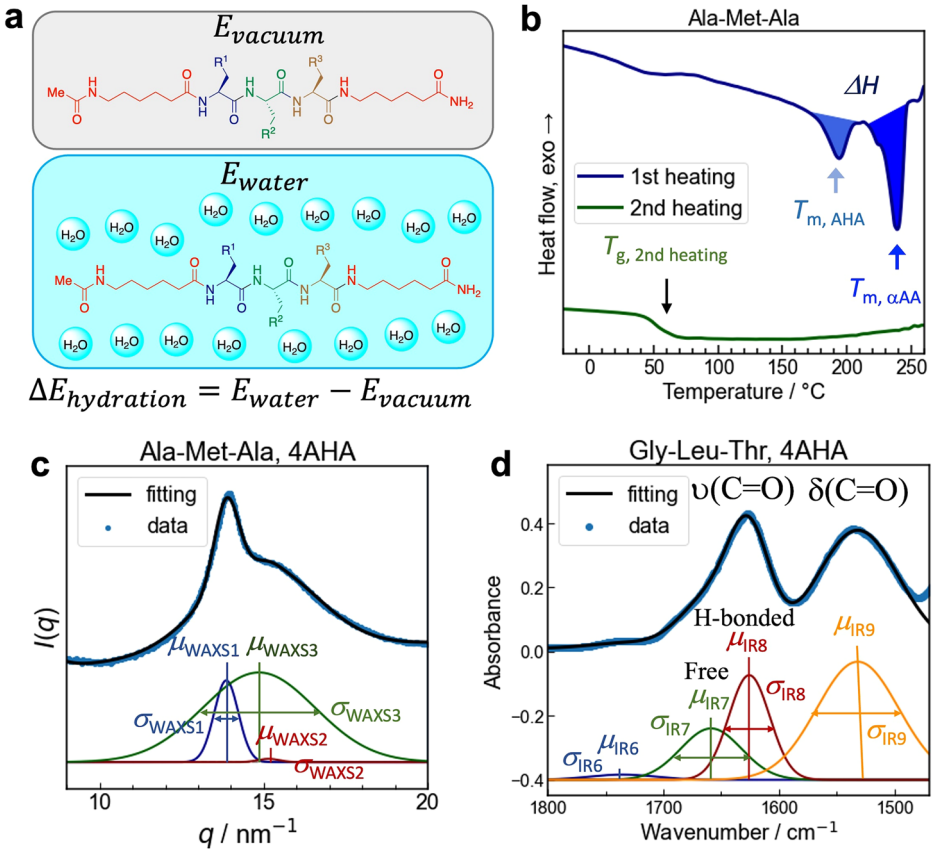

Fig. 6 | Extraction of physical values of higher-order structures from multimodal data. **a** Evaluating hydration energies based on quantum-chemical calculations. Extracting physical values from **b** DSC, **c** WAXS, and **d** IR data. Gaussian distributions were used to peak-fit the IR and WAXS data.

Table 1 | Higher-order structures essential for degradability

Positive		Negative	
Physical value	**Coefficient**	**Physical value**	**Coefficient**
$\rho_{enzyme, oligo}$	0.098 ± 0.015	$T_{m, AHA}$	−0.109 ± 0.019
μ_{WAXS1}	0.095 ± 0.015	σ_{IR8}	−0.103 ± 0.014
μ_{WAXS3}	0.063 ± 0.011	σ_{IR7}	−0.077 ± 0.015
σ_{WAXS1}	0.055 ± 0.014	$T_{g, 2nd heating}$	−0.077 ± 0.012
ΔH_{cp}	0.054 ± 0.016	$Area_{IR8}$	−0.075 ± 0.010
$Area_{IR7}$	0.050 ± 0.008	$\Delta E_{hydration}$	−0.055 ± 0.012
σ_{WAXS2}	0.048 ± 0.013	$Area_{IR5}$	−0.052 ± 0.012

Top seven average values of the ridge regression coefficients for different leave-one-out cross-validation (LOOCV) test samples and various physical values

sequences revealed that 4AHA was optimal (Fig. 2d). This tendency is contrary to the expectation that a low AHA ratio results in a high degradation rate owing to the high volume fraction of αAA segments during enzyme dissociation. Therefore, we investigated the effects of higher-order structures on the degradability. The intensity of the hydrogen-bonded C=O peak (1621 cm^{-1}) in the IR spectrum was observed to decrease as the number of AHA units was increased, while the peak corresponding to Nylon6 became less intense (Fig. S6a); furthermore, the wide-angle X-ray scattering (WAXS) diffraction peaks at q~14 nm^{-1} also became less intense (Fig. S6b). Differential scanning calorimetry (DSC) measurements revealed that the melting points of the αAA-rich phases (220–270 °C) decreased as the number of AHA units was increased, despite the AHA-rich phase maintaining a melting point of ~190 °C (Fig. S6c). These observations suggest that an increase in the volume fraction of AHA segments reduces the crystallite size or order of the αAA-rich phases. It is well known that enzymes preferentially degrade amorphous regions of crystalline polymers. In other words, the multiblock structure destroyed the crystal structures of the αAA segments and enhanced enzymatic degradation. These decreases in the volume fraction and crystal order of the αAA-rich phases contributed to the reversal of the observed degradability trend; hence, the 4AHA-containing polyamide showed maximum enzymatic degradation.

Multifactor analysis of physical factors on properties from multimodal data

The enzymatic degradation ratios of some films deviated from those of the oligopeptides, indicating that other factors also influenced degradability (Fig. S7a). For instance, oligopeptide samples that degrade poorly in solution are highly degradable in the film state. These differences can be attributed to the presence of surfaces or higher-order structures in the films. Although the differences could be evaluated by conventional human analytical methods, the complicated relationships and multiple measurement techniques are troublesome. Therefore, machine learning techniques were adopted for the integral analysis.

Machine learning-based analysis was investigated by extracting more than 40 physical values from a number of experimental and quantum chemical studies and by specifying their importance (Figs. 6a–d and S7b). For example, the widths of the WAXS diffraction peaks (σ_{WAXS}) were used as indicators of the crystallite size of the αAA segments. The affinities of the αAA segments for water were calculated based on quantum-chemically calculated hydration energies ($\Delta E_{hydration}$). How these factors impact film degradability ($\rho_{enzyme, film}$) was evaluated based on ridge regression, LASSO regression, and multiple linear regression with sequential feature selectors (Fig. S8), with mean-squared errors of 0.135, 0.202, and 0.292 determined, respectively, for the test data; consequently, we discuss the ridge regression results. In addition to oligopeptide degradation, several other important features were also identified (Table 1). For example, σ_{WAXS1} and μ_{WAXS1} were positively correlated with $\rho_{enzyme, film}$, indicating that polyamide films with smaller or more disordered αAA-rich phase crystals are more

degradable. The melting temperatures of the AHA-rich phases ($T_{m,AHA}$) were negatively correlated with $\rho_{enzyme,\ film}$, indicating that segment disorder increased degradability. The relative area of peak7 in the infrared (IR) spectrum ($Area_{IR7}$) was positively correlated with $\rho_{enzyme,\ film}$, revealing that more free C=O bonds or miscible phases enhanced degradability. Furthermore, the hydration energy ($\Delta E_{hydration}$) negatively affected $\rho_{enzyme,\ film}$, suggesting that the strong affinity of the polyamide for water enhances degradability. The feature importances were also confirmed by removing the top k features (Fig. S9). As the number of removed features increased, the value of the coefficient of determination decreased. After removing the top three features, the coefficients of determination remained almost identical, indicating that these three features were significantly effective in predicting the degradation rates. Consequently, we revealed that in addition to the degradability of the αAA sequence itself, chain properties (e.g., hydrophilicity) and higher structures (e.g., crystal structures) significantly affect enzymatic degradability.

Polyamide degradation in the natural environment

Although degradation experiments using enzymes have been conducted in vitro, the degradation of polyamides in natural environments is desirable. Therefore, polyamides were degraded in muddy water; here, polyamide films were immersed at room temperature in ten-fold concentrated muddy water collected from paddy fields around Kyushu University. The degradation rates were calculated using control samples in sterilized muddy water, which excluded abiotic processes. Several samples were confirmed to have lost weight after 5 days, although Nylon6 was hardly degraded (Fig. S10a, b). Notably, the Gly-Leu-Ala containing polyamide, which exhibited high mechanical performance (as discussed above), was one of the degraded samples (Fig. 4). Thus, we conclude that the current polyamide design promotes degradation in the natural environment. Degradation under other natural conditions and associated degradation mechanisms will be addressed in future studies.

Discussion

We demonstrate tough and degradable multiblock polyamides designed using machine learning techniques. Alternating multiblock copolymers composed of Nylon6 and α-amino acid segments were prepared by coupling oligopeptides. A sequence of three continuous α-amino acids endowed the polymer with versatile material properties and enhanced enzyme degradability. Multi-objective optimization based on BO suggested appropriate α-amino acid sequences that were both degradable and mechanically tough (as indicated by properties such as the Young's modulus and strain at break), which is difficult to achieve using the polymer design skills of humans. In addition, smaller αAA-rich phase crystals or a lower crystal order associated with the inclusion of AHA segments enhanced the enzymatic degradability. The ridge regression revealed the essential factors between the experimental and calculated data. In addition, some materials are degradable in natural environments or muddy water. It is worth mentioning that the Nylon6 oligomer generated by degradation can be metabolized in the natural environment.

The key points to consider when breaking down the trade-off relationship observed for biodegradable polymers in this study include adaptation of the multiblock structure and the utilization of machine learning techniques. The current study demonstrates the applicability of BO in copolymer design derived from many monomer candidates with multiple properties (as each monomer in the copolymers plays a distinct role, compensating for the weaknesses of the others). The Nylon6 (AHA) and α-amino acid segments in the multiblock copolymers play different roles in determining toughness and degradability. The phase-separated two-segment structure enables the formation of a moderate nanometer-scale crystal structure, which endows toughness even when heated owing to the high melting temperatures of these materials. In addition, smaller αAA-rich-phase crystals or a lower crystal order associated with the inclusion of AHA segments enhances enzymatic degradability. On the other hand, machine learning techniques have contributed to the design of monomer sequences and mechanistic understanding. Generally, humans struggle to predict the material properties of many possible three-amino acid sequences, especially multiple properties, and identifying the essential factors of higher-order structures using multimodal experimental and simulation data is difficult. Furthermore, this framework for the monomer selection in sequence-controlled polymers is applicable to different types of polymers and to different properties having trade-off relationships. Notably, each monomer plays a specific role in each property, helping to overcome trade-off relationships. The current technique for fabricating polyamides with both high degradability and mechanical properties is applicable to protein-based fibers using the sequence control of amino acids. Our approach, which uses machine learning techniques, is useful for designing and understanding environmentally friendly plastics and other materials that require multiple properties.

Methods

Materials

Fmoc-αAA-OH, *N*-Hydroxysuccinimide (NHS), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC·HCl), 6-aminohexanoic acids (AHA), 1-hydroxybenzotriazole (HOBt), diisopropylcarbodiimide (DIC), piperidine, 3-[bis(dimethylamino)methyliumyl]-3H-benzotriazol-1-oxide hexafluorophosphate (HBTU), and dimethyl formamide (DMF) were purchased from Watanabe Chem. Ind., LTD. Trifluoro acetic acid (TFA), diisopropyl ethyl amine (DIEA), 2,2,2-trifluoroethanol (TFEt), and tetraethylene glycol monomethyl ether (TEG) were obtained from Tokyo Chemical Industry Co., Ltd. Triisopropyl silane (TIPS) was purchased from FUJIFILM Wako Pure Chemical Corporation. 2-Chlorotrityl chloride resin was obtained from GL Biochem Ltd. Nylon6 was purchased from Sigma-Aldrich Co. LLC. Fmoc-AHA-OH was synthesized by a coupling reaction of Fmoc-OSu and AHA as previously reported[42].

Synthesis of Fmoc-AHA-AHA-OH

Fmoc-AHA-OH (43.8 g, 124 mmol), NHS (21.4 g, 186 mmol), and chloroform (620 mL) were poured into in 1 L round-bottom flask after which EDC·HCl (35.7 g, 186 mmol) was added portion wise over 5 min. The mixture was stirred at room temperature for 3 h, washed twice with water (270 mL) containing brine (30 mL), and three times with aqueous 0.1 N HCl (270 mL) containing brine (30 mL). The organic layer was dried using magnesium sulfate for 30 min, filtered, and the filtrate was evaporated and dried under vacuum. AHA (17.9 g, 136 mmol), water (300 mL), acetone (300 mL), and sodium hydrogen carbonate (20.8 g, 248 mmol) were then added and the mixture was stirred at room temperature for 20 h. The acetone was removed by evaporation, and chloroform (300 mL) and 2 N aqueous HCl (170 mL) were added to neutralize the solution, which was then extracted twice with chloroform (2 × 150 mL). The combined organic layers were dried using magnesium sulfate, after which the residue was dissolved in hexane (300 mL) and the solution was stored in a freezer for 20 h. The solid was collected by filtration and dried under vacuum. This precipitation process was repeated three times to afford the product as a white powder (44.7 g, 95.8 mmol (77% yield)). ^{1}H-NMR (DMSO-d6): δ/ppm 1.25 (m, 4H, CH_2), 1.38 (m, 4H, CH_2), 1.49 (m, 4H, CH2), 2.03 (t, J = 7 Hz, 2H, CH_2), 2.19 (t, J = 7 Hz, 2H, CH_2), 3.00 (m, 4H, CH2), 4.21 (t, J = 7 Hz, 1H, CH), 4.29 (d, J = 7 Hz, 2H, CH2), 7.33-7.90 (m, 8H, aromatic). ^{13}C-NMR (DMSO-d6): δ 24.7, 25.5, 26.4, 26.5, 29.4, 29.6, 34.1, 35.9, 38.7, 47.3, 65.6, 120.5, 125.6, 127.5, 128.0, 141.2, 144.4, 156.5, 172.3, 174.9. HRMS exact mass calculated for $[M + 1]^+$ $C_{27}H_{35}N_2O_5$ 467.2540, found 467.2546.

Synthesis of Fmoc-AHA-AHA-AHA-OH

Fmoc-AHA-AHA-AHA-OH was synthesized in the same manner as Fmoc-AHA-AHA-OH using Fmoc-AHA-AHA-OH (21.0 g, 45.0 mmol), NHS (7.77 g, 67.5 mmol), EDC·HCl (12.9 g, 67.5 mmol), AHA (6.49 g, 49.5 mmol), and sodium hydrogen carbonate (7.56 g, 90 mmol). The product was obtained as a white powder (22.2 g, 38.2 mmol (85% yield)). ^{1}H-NMR (DMSO-d6): δ/ppm 1.22 (m, 6H, CH_2), 1.37 (m, 6H, CH_2), 1.47

(m, 6H, CH_2), 2.03 (t, J = 7 Hz, 4H, CH_2), 2.19 (t, J = 7 Hz, 2H, CH_2), 3.00 (m, 6H, CH_2), 4.21 (t, J = 6 Hz, 1H, CH), 4.29 (d, J = 6 Hz, 2H, CH_2), 7.31-7.90 (m, 8H, aromatic). ^{13}C-NMR (DMSO-d6): δ 24.7, 25.5, 26.4, 26.6, 29.4, 29.6, 34.1, 35.9, 38.7, 38.8, 39.4, 47.3, 65.6, 120.5, 125.6, 127.5, 128.0, 141.2, 144.4, 156.5, 172.3, 174.9. HRMS exact mass calculated for $[M+1]^+$ $C_{33}H_{46}N_3O_6$: 580.3381, found: 580.3387.

Synthesis of Fmoc-AHA-AHA-AHA-AHA-OH

Fmoc-AHA-AHA-AHA-AHA-OH was synthesized in the same manner as Fmoc-AHA-AHA-OH using Fmoc-AHA-AHA-AHA-OH (11.6 g, 20.0 mmol), NHS (3.45 g, 30.0 mmol), EDC·HCl (5.75 g, 30.0 mmol), AHA (2.89 g, 22.0 mmol), and sodium hydrogen carbonate (3.36 g, 40.0 mmol). The product was obtained as a white powder (11.3 g, 16.3 mmol (82% yield)). ^{1}H-NMR (DMSO-d6): δ/ppm 1.22 (m, 8H, CH_2), 1.37 (m, 8H, CH_2), 1.47 (m, 8H, CH_2), 2.03 (t, J = 7 Hz, 6H, CH_2), 2.19 (t, J = 7 Hz, 2H, CH_2), 3.00 (m, 8H, CH_2), 4.21 (t, J = 6 Hz, 1H, CH), 4.29 (d, J = 6 Hz, 2H, CH_2), 7.31-7.90 (m, 8H, aromatic). ^{13}C-NMR (DMSO-d6): δ 24.7, 25.5, 26.4, 26.6, 29.4, 29.6, 34.1, 35.9, 38.7, 38.8, 47.3, 65.6, 120.5, 125.6, 127.5, 128.0, 141.2, 144.4, 156.5, 172.3, 174.8. HRMS exact mass calculated for $[M+1]^+$ $C_{39}H_{57}N_4O_7$: 693.4222, found: 693.4227.

Preparing oligopeptides

Oligopeptides composed of AHA and α-amino acids were prepared by solid-phase peptide synthesis. In a typical run, Fmoc-AHA-AHA-OH (303 mg, 649 μmol) was coupled to 2-chlorotrityl chloride resin (500 mg, loading: 1.18 mmol/g) using dichloromethane (9 mL) and DIEA (200 mL, 1.15 mmol) in a reaction tube. After stirring for 40 min at room temperature, the resin was washed three times with dichloromethane containing DIEA (2.5% v/v) and methanol (2.5% v/v) and three times with DMF. The Fmoc group was deprotected by treatment with a mixture of DMF and piperidine (80/20 (v/v)) twice. After 10 min, the resin was washed seven times with DMF. Fmoc-Ala-OH (606 mg, 1.94 mmol) was coupled with DIC (302 μL, 1.95 mmol) and HOBt (298 mg, 1.95 mmol) in DMF for 2 h at room temperature. The resin was washed seven times with DMF. This coupling cycle was repeated using Fmoc-Phe-OH (754 mg, 1.95 mmol) and Fmoc-Ala-OH (606 mg, 1.95 mmol), after which Fmoc-AHA-AHA-OH (908 mg, 1.95 mmol) was coupled with HBTU (671 mg, 1.77 mmol) and DIEA (514 μL, 2.95 mmol) at room temperature for 2.5 h. After washing and deprotection, the resin was treated with methanol three times and dried under vacuum. The resin was removed using TFA (400 μL) and TIPS (0.6 mL) in dichloromethane (12 mL) at room temperature for 40 min. The reaction mixture was filtered, evaporated, and precipitated with DMF (5 mL) in chloroform and hexane (1/2 (v/v)). The product (H_2N-AHA-AHA-Ala-Phe-Ala-AHA-AHA-OH·TFA) was obtained as white powder (396 mg, 453 μmol, 70% yield). ^{1}H-NMR (DMSO-d6): δ/ppm 1.11-1.55 (m, 24H, CH_2), 2.01-2.09 (m, 6H, CH_2), 2.18 (t, J = 7H, CH_2), 2.74-2.78 (m, 2H, CH_2), 2.79-2.85 (m, 1H, CH_2), 2.98-3.06 (m, 7H, CH_2), 4.19 (q, J = 7 Hz, 2H, CH), 4.46 (m, 1H, CH), 7.16-7.26 (m, 5H, aromatic), 7.68-7.97 (m, 6H, NH). MALDI-TOF-MS analysis calculated for $[M+1]^+$ $C_{39}H_{66}N_7O_8$: 761.0, found: 760.5.

Polyamide synthesis

In a typical run, HOBt (175 mg, 1.14 mmol), DIC (179 μL, 1.15 mmol), and DIEA (331 mL, 1.90 mmol) were added to H_2N-AHA-AHA-Ala-Phe-Ala-AHA-AHA-OH·TFA (332 mg, 0.38 mmol) in DMSO (1.21 mL) at room temperature. After 1 h, the reaction mixture was heated to 60 °C and maintained at this temperature for 2 days. The solid was dissolved in TFEt (4 mL) and precipitated twice from a mixture of chloroform and hexane (4/1 v/v)). The polyamide was treated with TFA, TIPS, and water (= 95/2.5/2.5 v/v/v) for 2 h when amino acids with protected side-chain were involved, and precipitated in a mixture of diethyl ether and hexane (= 9/1 (v/v). The product was obtained as a white powder (154 mg, 47% yield). GPC (TFEt with sodium trifluoroacetate (5 mM)): M_n = 4800, M_w/M_n = 5.3. MALD-TOF-MS analysis calculated for [3 M (cyclic)+Na]$^+$ $C_{117}H_{189}N_{21}NaO_{21}$: 2247.5, found: 2249.1.

Film moulding

The polyamide was moulded by solvent casting, while Nylon6 was moulded by solvent casting and hot pressing. Polyamide or Nylon6 (180 mg) was dissolved in TFEt (3.4 mL) and poured into a PTFE mould (4.5 × 3.5 cm) and kept horizontal overnight. The dried sample was heated at 80 °C for 3 h and at 100 °C for 2 h under vacuum, to afford an ~100-μm-thick polyamide film. In the case of hot pressing, Nylon6 was melted at 240 °C in a hot-press machine under vacuum for 10 min at 10 MPa using a silicon wafer modified with *n*-octadecyltrimethoxysilane. The sample was then crystallized in another hot-press machine at 142 °C for 20 min at 5 MPa. Melted samples were treated in ice water without crystallization to produce quenched samples.

Tensile testing

Tensile testing was performed by uniaxially elongating polyamide films. Samples were cut using a 2 × 12 mm dumbbell-shaped mould. Uniaxial elongation was carried out using a tensile-testing machine (Imoto Machinery CO., Ltd) at 10 mm/min and room temperature or 100 °C. Young's moduli were calculated based on the initial slopes of the stress–strain curves. Strains at break were evaluated as the strains at which samples separated. Both values are reported as the averages of three tests.

Enzymatically degrading oligopeptides

Oligopeptides (AHA-AA^1-AA^2-AA^3-AHA) were prepared as described above. The N-terminus of each peptide consisted of an amino group for sequences with AA^1 and AA^3 = Ala, and an acetyl group for random sequences. The C-terminus was modified with an amide group using a Rink-amide resin. In a typical run, Proteinase K (0.1 μM) with TEG (100 μM) as an internal standard, and an oligopeptide (1 mM) were added to a Tris-buffered saline solution and allowed to react in a heated bath at 37 °C. After 30 min, the reaction solution was diluted 10 times with water containing formic acid (0.1%, v/v) and allowed to deactivate at 98 °C for 15 min. The amount of oligopeptide was determined by liquid chromatography–mass spectrometry using an LCMS-2020 and Nexera X2 system (Shimadzu Corp.). The eluent, which consisted of aqueous formic acid and methanol (98/2 to 0/100 (v/v)), was flowed at 1.5 mL/min. The reaction mixture was separated using an ODS column (TSKgel ODS120-H (TOSOH Corp.)) at 40 °C. Compounds were ionized using the ESI method and detected using a quadrupole mass spectrometer in positive mode. The areas of the peaks in the mass spectra of the oligopeptides before and after reaction (A_{before} and A_{after}) were standardized against those of TEG. The enzymatic degradation rate was calculated as $\rho_{enzyme,\ oligo} = 1 - A_{after}/A_{before}$ using the average values from three runs.

Enzymatically degrading polyamides

Polyamides were enzymatically degraded by immersing films in Proteinase K solution. Polyamide films (~100-μm thick) were cut to rectangular shapes that were ~5 mg in weight. Each film was immersed in a solution of Proteinase K in Tris-buffered saline (0.5 mg/mL, 0.8 mL, pH 7.4) for 2 days at 37 °C. The ratio of the film remaining following enzyme treatment (r_{enzyme}) was calculated from the weights of the film before and after degradation ($w_{before,\ enzyme}$ and $w_{after,\ enzyme}$, respectively) as: $r_{enzyme} = w_{after,\ enzyme}/w_{before,\ enzyme}$. Films were concurrently treated with Tris buffer devoid of enzyme as a control, and the remaining ratio in the absence of enzyme (r_{buffer}) was evaluated as described above. The enzymatic degradation rate was calculated as: $\rho_{enzyme,\ film} = 1 - r_{enzyme}/r_{buffer}$ using the average values from three experiments.

Degrading polyamides in muddy water

Polyamides were degraded under natural conditions by immersing films in muddy water. Polyamide films (~100-μm thick) were cut into rectangular shapes ~1.5 mg in weight. Muddy water was collected from a paddy field around Kyushu University, sonicated for 10 s, and concentrated 10 times under vacuum. Films were immersed in the concentrated muddy water (1 mL) at room temperature for 5 days. The ratio of each film remaining in

the muddy water (r_{muddy}) was determined by weighting the film before and after degradation ($w_{before,\ muddy}$ and $w_{after,\ muddy}$, respectively) as: $r_{muddy} = w_{after,\ muddy}/w_{before,\ muddy}$. Films were also concurrently treated in muddy water deactivated by heating at 98 °C for 10 min as a control, and the remaining ratio ($r_{control}$) was evaluated using the same equation. The degradation rate of the polyamide in muddy water was calculated as: $\rho_{muddy,\ poly} = 1 - r_{muddy}/r_{control}$. Average values and standard deviations determined from four of five test samples were used, with one outlier excluded.

X-ray scattering experiments

WAXS and SAXS measurements were performed on the BL05XU beamline of the SPring-8 facility. Each film was exposed to X-rays with a wavelength of 0.1 nm for 1 s at room temperature, and detected using SOPHIAS and PILATUS detectors with sample-to-detector distances of 159 mm and 1378 mm for WAXS and SAXS, respectively. The sample-to-detector distance and beam center were determined using CeO_2 and AgBe as standards, respectively. Two-dimensional images were converted into one-dimensional profiles using FIT2D software. Profiles were subtracted from background samples without the use of any coefficients. The scattering vector is defined as: $q = 4\pi \sin\theta/\lambda$, where λ is the X-ray wavelength, and 2θ is the scattering angle.

Differential scanning calorimetry (DSC)

DSC measurements were performed using a NEXTA DSC200 instrument (Hitachi High-Tech Corp.) in temperature-modulated mode. Each film (~3 mg) was placed in an aluminum pan and sealed, and a film-free aluminum pan was prepared as a reference sample. The samples were cooled to −40 °C at 10 °C/min, and then heated to 240 °C at 2 °C/min with a temperature amplitude of 0.5 °C and a frequency of 0.012 Hz. This protocol was repeated under the same conditions for a second cycle. The acquired curves (total) were separated into reversing and non-reversing components using NEXTA software, and the reverse components were plotted in this study.

Gel permeation chromatography (GPC)

GPC measurements were conducted using a Prominence HPLC system (Shimadzu Corp.) equipped with LC-20AD, CTO-20AC, SPD-20A, and RID-10A components. Sodium trifluoroacetate (5 mM) in TFEt was used as the eluent at 3 mL/min. The polyamide was dissolved in TFEt (5 mg/mL, 20 μL) and injected onto and separated by TSKgel SuperAW4000 and SuperAW3000 columns at 40 °C, and detected by an RI detector. Five polystyrene standard samples (M_p = 1.1×10^5, 5.3×10^4, 2.1×10^4, 4.9×10^3, and 1.9×10^3) were used to construct the calibration curve from which M_n and M_w/M_n values were evaluated.

IR spectroscopy

Fourier-transform infrared (FT-IR) spectroscopy was carried out using an INVENIO X (Bruker) spectrometer in attenuated total reflection mode. Each polyamide film was placed in the Gladi ATR unit (PIKE Tech.) and fixed using a pin. Spectra were acquired at room temperature in the 4000–500 cm^{-1} wavenumber range with a resolution of 2 cm^{-1} and sixteen scans. IR spectra were obtained by subtracting the sample-free background spectra.

MALDI-TOF MS analysis

MALDI-TOF MS analysis was performed using an Autoflex spectrometer (Bruker). The polyamide, dissolved in dimethyl sulfoxide, was mixed with a solution of α-cyano-4-hydroxycinnamic acid as the matrix on the sample plate. After the solvent was dried, mass spectra were acquired in positive ion mode over an m/z range of 2000–6000.

Quantum chemical calculations

Oligopeptides with Ac-AHA-AA^1-AA^2-AA^3-AHA-NH_2 sequences were subjected to quantum-chemical calculations using Gaussian 16 software. We first used Balloon[43], a conformer-searching tool implemented in Winmostar[44] that uses a multi-objective genetic algorithm, to search for stable oligopeptide conformations using the MMFF94 force field[43]. Structures were optimized at the B3LYP/6-31 G** level[45–48] in Gaussian 16[49]. The effect of water as the solvent was included using the polarizable continuum model (ε = 78.3553)[50]. The hydration energy of each polyamide ($\Delta E_{hydration}$) was obtained as the difference between the energies in water (E_{water}) and in a vacuum (E_{vacuum}).

Multi-objective optimization using BO

Multiple properties were subjected to Bayesian optimizations using Gaussian process regression (GPR) in PHYSBO library[51]. The T-scale was adopted as the α-amino acid descriptor[37], in which the T1, T2, and T3 values of the three α-amino acids (a total of nine values) were used as explanatory variables. Young' modulus, strain at break, and the enzymatic degradability of each polyamide film thermally treated at 80 °C were utilized as objective variants. Eight Ala-AA^2-Ala samples and 17 samples of random sequences (AA^1-AA^2-AA^3) were used to construct GPR models. Multi-objective optimizations were performed using hypervolume-based probability of improvement (HVPI)[38], EHVI[38], and TS[39] as objective functions. First, the model was defined using the "search.discrete_multi.policy" function for optimization. The recommended sequences were then suggested using the policy_m.bayes_search function with the following hyperparameters: max_num_probes = 100, num_search_each_probe = 1, and interval = 2. The suggested sequence of three α-amino acids was obtained by excluding of water-soluble ones from 14 (AA^1) × 14 (AA^2) × 14 (AA^3) – 25 (original) = 2719 combinations, which were synthesized and re-evaluated. In the current study, the sequences suggested by HVPI and EHVI were the same in most cases.

Evaluating important physical values

Important physical values for enzymatic degradability were evaluated by ridge regression[52], LASSO regression[53], and linear regression with sequential feature selectors (SFS)[54] using the scikit-learn library. Information, such as T_g, T_m, ΔH_{melt}, peak position/width/relative area after peak fitting, and film crystallinity were extracted from DSC, WAXS, and IR data. The hydrophilicities of the αAA sequences were evaluated using quantum-chemical calculations. Consequently, more than 40 values were extracted as explanatory variables. Nested leave-one-out cross-validation (LOOCV) was adopted, in which one test sample was eliminated to evaluate the mean-squared error, while the others were used to determine regularity parameters and coefficients; this treatment was repeated by changing the test sample. Therefore, many more coefficients than the sample size (26) were calculated. New explanatory variables were sequentially added using improved ones (compared with other explanatory variables) based on the mean squared error (MSE) of the validation data when the SFS linear regression method was used. Explanatory variables were then added until the MSE improved. The important feature values for each combination were evaluated using the average values of the regression coefficients of the various models.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability

The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Received: 8 October 2024; Accepted: 4 June 2025;
Published online: 01 July 2025

References

1. Manker, L. P. et al. Sustainable polyesters via direct functionalization of lignocellulosic sugars. *Nat. Chem.* **14**, 976–984 (2022).

2. Li, X. L., Clarke, R. W., Jiang, J. Y., Xu, T. Q. & Chen, E. Y. X. A circular polyester platform based on simple gem-disubstituted valerolactones. *Nat. Chem.* **15**, 278–285 (2023).
3. Gross, R. A. & Kalra, B. Biodegradable polymers for the environment. *Science* **297**, 803–807 (2002).
4. Haider, T. P., Volker, C., Kramm, J., Landfester, K. & Wurm, F. R. Plastics of the future? The impact of biodegradable polymers on the environment and on society. *Angew. Chem. Int. Edit.* **58**, 50–62 (2019).
5. Fagnani, D. E. et al. 100th anniversary of macromolecular science viewpoint: redefining sustainable polymers. *ACS Macro Lett.* **10**, 41–53 (2021).
6. Isobe, A., Iwasaki, S., Uchida, K. & Tokai, T. Abundance of non-conservative microplastics in the upper ocean from 1957 to 2066. *Nat. Commun.* **10**, 417 (2019).
7. Shruti, V. C. & Kutralam-Muniasamy, G. Bioplastics: missing link in the era of microplastics. *Sci. Total Environ.* **697**, 134139 (2019).
8. Nair, L. S. & Laurencin, C. T. Biodegradable polymers as biomaterials. *Prog. Polym. Sci.* **32**, 762–798 (2007).
9. Tokiwa, Y., Calabia, B. P., Ugwu, C. U. & Aiba, S. Biodegradability of plastics. *Int. J. Mol. Sci.* **10**, 3722–3742 (2009).
10. Delre, C. et al. Near-complete depolymerization of polyesters with nano-dispersed enzymes. *Nature* **592**, 558–563 (2021).
11. Amamoto, Y. Data-driven approaches for structure-property relationships in polymer science for prediction and understanding. *Polym. J.* **54**, 957–967 (2022).
12. Okazawa, K. et al. Exploring the optimal alloy for nitrogen activation by combining Bayesian optimization with density functional theory calculations. *ACS Omega* **7**, 45403–45408 (2022).
13. Kosuri, S. et al. Machine-assisted discovery of chondroitinase ABC complexes toward sustained neural regeneration. *Adv. Healthc. Mater.* **11**, 2102101 (2022).
14. Tamasi, M. J. et al. Machine learning on a robotic platform for the design of polymer-protein hybrids. *Adv. Mater.* **34**, 2201809 (2022).
15. Suh, J. S., Suh, B. C., Bae, J. H. & Kim, Y. M. Machine learning-based design of biodegradable Mg alloys for load-bearing implants. *Mater. Design* **225**, 111442 (2023).
16. Saito, Y. et al. Machine-learning-guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synth. Biol.* **7**, 2014–2022 (2018).
17. Batra, R. et al. Machine learning overcomes human bias in the discovery of self-assembling peptides. *Nat. Chem.* **14**, 1427–1435 (2022).
18. Terayama, K., Sumita, M., Tamura, R. & Tsuda, K. Black-box optimization for automated discovery. *Accounts Chem. Res.* **54**, 1334–1346 (2021).
19. Zamengo, M., Wu, S. P., Yoshida, R. & Morikawa, J. Multi-objective optimization for assisting the design of fixed-type packed bed reactors for chemical heat storage. *Appl. Therm. Eng.* **218**, 119327 (2023).
20. Min, K., Cuiffi, J. D. & Mathers, R. T. Ranking environmental degradation trends of plastic marine debris based on physical properties and molecular structure. *Nat. Commun.* **11**, 727 (2020).
21. Amamoto, Y., Kikutake, H., Kojio, K., Takahara, A. & Terayama, K. Visualization of judgment regions in convolutional neural networks for X-ray diffraction and scattering images of aliphatic polyesters. *Polym. J.* **53**, 1269–1279 (2021).
22. Takamura, A., Tsukamoto, K., Sakata, K. & Kikuchi, J. Integrative measurement analysis via machine learning descriptor selection for investigating physical properties of biopolymers in hairs. *Sci. Rep.* **11**, 24359 (2021).
23. Wang, J. et al. Estimating the relative crystallinity of biodegradable polylactic acid and polyglycolide polymer composites by machine learning methodologies. *Polymers* **14**, 527 (2022).
24. Fransen, K. A. et al. High-throughput experimentation for discovery of biodegradable polyesters. *Proc. Natl. Acad. Sci. USA* **120**, e2220021120 (2023).
25. Yuan, W. L. et al. Revealing factors influencing polymer degradation with rank-based machine learning. *Patterns* **4**, 100846 (2023).
26. Okada, M., Amamoto, Y. & Kikuchi, J. Designing sustainable hydrophilic interfaces via feature selection from molecular descriptors and time-domain nuclear magnetic resonance relaxation curves. *Polymers* **16**, 824 (2024).
27. Schuett, T. et al. Application of digital methods in polymer science and engineering. *Adv. Funct. Mater.* **34**, 2309844 (2024).
28. Ryadnov, M. G. & Woolfson, D. N. Engineering the morphology of a self-assembling protein fibre. *Nat. Mater.* **2**, 329–332 (2003).
29. Tsuchiya, K., Ishii, T., Masunaga, H. & Numata, K. Spider dragline silk composite films doped with linear and telechelic polyalanine: Effect of polyalanine on the structure and mechanical properties. *Sci. Rep.* **8**, 3654 (2018).
30. Wu, J. H. et al. Rationally designed synthetic protein hydrogels with predictable mechanical properties. *Nat. Commun.* **9**, 620 (2018).
31. Gudeangadi, P. G. et al. Poly(alanine-nylon-alanine) as a bioplastic: chemoenzymatic synthesis, thermal properties and biological degradation effects. *Polym. Chem.* **11**, 4920–4927 (2020).
32. Tsuchiya, K. & Numata, K. Facile terminal functionalization of peptides by protease-catalyzed chemoenzymatic polymerization toward synthesis of polymeric architectures consisting of peptides. *Polym. Chem.* **11**, 560–567 (2020).
33. Koga, T., Morishita, T., Harumoto, Y., Nishimura, S. & Higashi, N. Spider silk-inspired peptide multiblock hybrid copolymers for self-healable thin film materials. *Mater. Adv.* **2**, 7851–7860 (2021).
34. Hu, J. Y. et al. Design of synthetic collagens that assemble into supramolecular banded fibers as a functional biomaterial testbed. *Nat. Commun.* **13**, 6761 (2022).
35. Numata, K. & Kaplan, D.L. Silk proteins: designs from nature with multipurpose utility and infinite future possibilities. *Adv. Mater*. https://doi.org/10.1002/adma.202411256.
36. Kinoshita, S. et al. Purification and characterization of 6-aminohexanoic acid oligomer hydrolase of flavobacterium Sp-Ki72. *Eur. J. Biochem.* **116**, 547–551 (1981).
37. Tian, F. F., Zhou, P. & Li, Z. L. T-scale as a novel vector of topological descriptors for amino acids and its application in QSARs of peptides. *J. Mol. Struct.* **830**, 106–115 (2007).
38. Couckuyt, I., Deschrijver, D. & Dhaene, T. Fast calculation of multiobjective probability of improvement and expected improvement criteria for Pareto optimization. *J. Glob. Optim.* **60**, 575–594 (2014).
39. Yahyaa, S. & Manderick, B. Thompson sampling for multi-objective multi-armed bandits problem. In *Proc. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 47–52 (2015).
40. Aharoni, S. M. *n-Nylons: Their Synthesis, Structure, and Properties* (John Wiley & Sons Ltd, 1997).
41. Peinemann, K. V., Abetz, V. & Simon, P. F. Asymmetric superstructure formed in a block copolymer via phase separation. *Nat. Mater.* **6**, 992–996 (2007).
42. Sakamoto, T. et al. Evaluation of dynamic features of Escherichia coli 16S ribosomal RNA in homogeneous physiological solution. *Biophys. J.* **89**, 4122–4128 (2005).
43. Vainio, M. J. & Johnson, M. S. Generating conformer ensembles using a multiobjective genetic algorithm. *J. Chem. Inf. Model* **47**, 2462–2474 (2007).
44. Winmostar V11, X-Ability Co. Ltd, *Tokyo, Japan*, *2023*.
45. Hariharan, P. C. & Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **28**, 213–222, https://doi.org/10.1007/BF00533485 (1973).

46. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
47. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
48. Hehre, W. J., Ditchfield, R. & Pople, J. A. Self—consistent molecular orbital methods. XII. Further extensions of Gaussian—type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **56**, 2257–2261 (2003).
49. Gaussian 16 Rev. C.01 (Wallingford, CT, 2016).
50. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999–3093 (2005).
51. Motoyama, Y. et al. Bayesian optimization package: PHYSBO. *Comput. Phys. Commun.* **278**, 108405 (2022).
52. Hoerl, A. E. & Kennard, R. W. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* **42**, 80–86 (2000).
53. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. B* **58**, 267–288 (1996).
54. Ferri, F. J., Pudil, P., Hatef, M. & Kittler, J. in *Machine Intelligence and Pattern Recognition* Vol. 16 (eds Gelsema, E. & Kanal, S.) 403–413 (Springer, 1994).

Acknowledgements

This work was supported by the Cabinet Office, Government of Japan, Cross-ministerial Strategic Innovation Promotion Program (SIP), and "Technologies for Smart Bio-industry and Agriculture" (funding agency: Bio-oriented Technology Research Advancement Institution, NARO). This study is based on the results obtained from project JPNP18016, commissioned by the New Energy and Industrial Technology Development Organization (NEDO). This work was also supported by JSPS Grant-in-Aid for Scientific Research on Innovative Areas, Discrete Geometric Analysis for Materials Design: 20H04644, Grant-in-Aid for Scientific Research (B): 20H02800, and Data Creation and Utilization Type Material Research and Development Project Grant Numbers JPMXP1122683430 and JPMXP1122714694, and by Institute of Mathematics for Industry, Joint Usage/Research Center in Kyushu University (Workshop (II), Reference No. 2023a011 and 2024a011). Y.A. and K.T. acknowledge the financial support from the Grant-in-Aid for the RIKEN-Kyushu University Science and Technology Hub Collaborative Research Program. Synchrotron radiation experiments were performed at the BL40XU and BL05XU beamlines of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal Nos. 2020A1525, 2021B1476, and 2022B1029).

Author contributions

Y.A. and K.T. conceived the study. Y.A. and C.K. performed the experiments. Y.A., C.K. and T.A. analyzed the data. S.Y., K.O. and Y.T. carried out the quantum chemical calculation. K.K. and A.T. bear responsibility for the measurements. K.T. developed the method of multi-objective optimization. Y.A. drafted the manuscript, and the others reviewed it.

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/s41524-025-01696-1.

Correspondence and requests for materials should be addressed to Yoshifumi Amamoto or Kei Terayama.