Structure-Property Relationships for Biodegradability in Copolyesters

Authors: Katharina A. Fransen[1], Julia Casey[1], Gabrielle Godbille-Cardona[1], Natalie Mamrol[2], Jiale Shi[1], Alex Zappi[1], Jignesh S. Mahajan[1], Jiarui Lu[3], Debra J. Audus[4], and Bradley D. Olsen[1]*
*Corresponding Author

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]Department of Computer Science, Wellesley College, Wellesley, MA 02481, USA

[4]Materials Science and Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

*Corresponding Author
Bradley D. Olsen
Massachusetts Institute of Technology, Room 66-558A
77 Massachusetts Ave.
Cambridge, MA 02139
Phone: 617-715-4548
Email: bdolsen@mit.edu

Abstract: With growing concerns about increasing plastic pollution, interest in biodegradable polymers, particularly polyesters, continues to grow. Copolymerization is an important molecular handle to tune properties to achieve material performance and biodegradation simultaneously. To better understand the structure-property relationships that govern biodegradability, a 300-member copolymer library was synthesized and tested using a high-throughput clear-zone biodegradation assay; testing shows over 80 % of the copolymer library is biodegradable even though only 50 % of the homopolymers from which they are derived are degradable. Repeat units with longer carbon backbones decreased biodegradability, though copolymers showed biodegradability at all average repeat unit lengths examined. For both homopolymers and copolymers, oxygen substitution of backbone carbons was established as a lever to improve biodegradability. A novel chemical similarity-informed embedding that considers polymer chemical structure and composition was developed and implemented for the simultaneous

quantitative structure-property relationship modeling of homopolymers and copolymers. Random forest models could simultaneously capture homopolymer and copolymer behavior with 78 % and 95 % accuracy, respectively; however, models trained only on homopolymer data could not predict copolymer biodegradability. Unlike random forest models, linear models were not able to capture both homopolymer and copolymer biodegradability.

Main Text:

Introduction:

Biodegradable polyesters present a promising end-of-life solution for materials that often suffer accidental environmental release, are targets for industrial degradation,[1-4] or shed microplastics into the environment through daily use and wear.[2] Polyesters are a promising material class for the development of biodegradable polymers as they contain ester bonds capable of undergoing both chemical degradation and biodegradation.[5-6] Additionally, polyesters such as poly(3-hydroxybutyrate) and poly(3-hydroxyvalerate) have been synthesized and leveraged as energy sources by microbes, providing naturally occurring microbes capable of biodegrading polyesters.[7-8] Studies on the biodegradability of polyesters have focused on small (< 10) sets of polymers,[9-12] examining the effect of isolated chemical structure effects such as side chains or repeat unit length,[13] or developing biodegradable polymers for a specific application such as packaging,[10] fibers,[14] or medicine.[15-16] As the need for sustainable polymer development grows,[17] an understanding of how material structure affects biodegradation and the development of predictive biodegradability models becomes increasingly critical to address the accumulation of polymers and microplastics in the environment.[2, 18-19]

At the same time, different intended applications require materials with varying mechanical, optical, thermal, and transport properties.[20-21] Copolymerization presents a demonstrated route to achieve tuned material properties without sacrificing all of the benefits from the original material selection. For example, terephthalic acid, ethylene glycol, and cyclohexanedimethanol are often copolymerized to obtain polyethylene terephthalate glycol

(PETG) for applications such as 3D printing.[22-23] Moving towards greener materials, lactone copolymers, including copolymerization of lactide and its stereoisomers, have been used to tune physical properties beyond those available among homopolymers.[11, 24-25] In addition, terephthalic acid, adipic acid, and butanediol are copolymerized to form polybutylene adipate terephthalate (PBAT), a quickly growing competitor to polylactic acid with improved flexibility and toughness, which presents a promising tunable and biodegradable route towards polymer waste management.[12, 26-27]

Despite the importance of biodegradability for the development of these sustainable and green materials, the systematic design of biodegradable polymers remains a major challenge. Criteria for biodegradability under various degradation conditions have been set forth by agencies such as the American Society for Testing and Materials (ASTM) and the International Organization for Standardization (ISO), including standards for anerobic environments, aerobic environments accelerated by high temperatures, marine environments, and industrial composting.[28-32] The standardized tests provide valuable information for end-of-life fate determination in their targeted environments, but their required testing timeframes (30-180 days)[28-33] and variability in the microbial consortia between individual tests and test types present challenges for the use of these tests when comparing the biodegradability and compostability of polymers. While biodegradable and compostable are often used interchangeably for polymers,[34] in this work biodegradability is used to refer to the ability of materials to degrade under ambient microbial action[34] and compostability is interpreted as defined in ASTM D6400 and establishes a material as certifiable by the Biodegradable Products Institute in the United States.[33]

Accelerated biodegradation testing methods have been identified as a key challenge in biodegradable polymer development[35] as timeframes[28-29, 31-32] associated with standard test methods hinder their use as primary screening tools. Several accelerated biodegradation strategies have been developed, which enable testing for polymer biodegradation within 30 days.[36-39] Recently, the accelerated clear-zone biodegradation test[39] was adapted for high-

throughput. This test leveraged a *Paucimonas lemoignei (P. lemoignei)* colony to degrade polymer particulates, providing control over the microbial action governing the degradation and guaranteeing replicability. *P. lemoignei* contains 7 known polyester depolymerases which are excreted extracellularly.[7, 40-41] Automation of the visual degradation data collection and processing supported the simultaneous testing of 100-200 polymers in two weeks. The high-throughput accelerated clear-zone biodegradation test has enabled the testing of a polyester and polycarbonate library with over 600 members, the identification of important trends for biodegradability in the tested polymers, and the development of polyester biodegradation structure-property relationships using machine learning.[42]

Quantitative structure-property relationship (QSPR) modeling to predict polymer properties promises to help researchers navigate increasingly large design spaces to find optimal polymer designs. Polymer property prediction using machine learning has been demonstrated[43-46] but continues to face challenges across data generation and curation, polymer featurization and fingerprinting, and model selection and training.[47] Dataset sizes, especially experimental datasets, are expensive to generate and limited in size even when produced through high-throughput experimentation.[47] Polymers also present unique challenges for representation as they are comprised of stochastic chemical structures which most chemical representations, developed for small molecules,[46, 48] are not designed to capture. Although strides have been made in capturing the stochasticity of polymers in a singular representation such as BigSMILES,[49-50] fingerprint embedding methods for polymer machine learning remain challenging due to not only the stochastic nature of the polymer but topological variability and variable monomer compositions in copolymers.[47] Advances in applying fingerprinting strategies remain limited to polymers with one or two monomers[51] or rely on additional machine learning models to generate embeddings.[52] As a result, monomers, repeat units, or ring repeat units[53-54] are often used as a basis for small molecule chemical fingerprinting algorithms to generate the polymer representations. Copolymers face additional representation challenges due to their variable local monomer sequences which

are not captured by monomer or repeat unit-based representation strategies. One proposed solution to address this challenge has been summing monomer fingerprint vectors in proportion to the monomer composition,[51] but such strategies do not capture functional group connections between monomers or their relative abundances in the polymer, and a significant need for additional polymer representations remains.

This work applied high-throughput experimentation and QSPR modeling in order to develop a deeper understanding of the impact of copolymerization on polymer biodegradability. A 300-member copolymer library was synthesized and evaluated for biodegradability to augment an existing polyester and polycarbonate biodegradability dataset and identify the impact of copolymerization of more than two monomers in AA/BB polycondensation. A novel chemical similarity-informed embedding incorporating polymer chemical structure and composition was introduced and leveraged to develop predictive random forest and linear machine learning models for polyester biodegradability, probing the transferability of homopolymer biodegradation data in predicting copolymer biodegradability.

Experimental Section:

Polymer Synthesis and Characterization

Materials: Monomers and catalysts were purchased from Sigma Aldrich, TCI America, and AK Scientific and used as received. Information on the purity and source of each monomer is available in the *Supplementary Dataset*. Deuterated solvents were purchased from Cambridge Isotope Laboratories and Sigma Aldrich. Gel Permeation Chromatography (GPC) solvents were purchased from Sigma Aldrich. Polymers were synthesized using standard melt condensation techniques. Details for the polymerization methods for each specific polymer are provided in the *Supplemental Dataset* along with polymer characterization, including molecular mass (both number average and mass average), dispersity, biodegradability, and final monomer incorporation ratios. Polymer molecular mass and dispersity were determined via GPC for polymers soluble in tetrahydrofuran (THF); for products insoluble in THF, physical state changes,

increased viscosity after reaction, insolubility in THF and dichloromethane (DCM) (when monomers were soluble), and peak broadening in the ^{1}H Nuclear Magnetic Resonance Spectroscopy (NMR) were used as evidence of polymerization. Polymers with a degree of polymerization less than 5 were resynthesized at higher temperatures, under higher vacuum, and with longer reaction times to achieve higher molecular weights. Under the resynthesis conditions, 41 of 58 attempted monomer combinations yielded polymers with higher molecular weights; these polymers are included in the Supplementary Dataset and denoted with a "_RE"; the lower molecular weight materials with the same monomers are also reported. The remaining 17 polymers were charred under the resynthesis conditions. The molecular weight of polymers that were resynthesized was characterized via size exclusion chromatography in either THF or 1,1,1,3,3,3-hexafluoroisopropanol (HFIP), depending on the resulting material's solubility.

Melt Polycondensation Procedure (General): To a 20mL vial with a Teflon stir bar, diol 1 (1.05 equivalents), diol 2 (1 equivalent), and dicarboxylic acid (2 equivalents) were added. The vial was fitted with a septum cap and purged with nitrogen before injecting 0.02 mL of titanium(IV) tert-butoxide. The vial was set to stir at 250 rpm (1 rpm = 2π per 60 $^{s-1}$) throughout the procedure. The vial was heated to 140 °C for 2 hours under 1 atm of nitrogen, followed by heating to 160 °C for 3 hours before being placed under vacuum (~10 torr) and heated to 175 °C for 2 hours. The vacuum level was increased to 75 mtorr for 12 hours before the vial was refilled with nitrogen (1 atm) and removed from heat before characterization. Polymerization procedure details for each polymer are provided in the *Supplementary Dataset*. Information for the customized parallel synthesis reactor can be found in *Figure S1, Figure S2,* and *Table S5.*

Molecular Weight and Dispersity Determination: For polymers soluble in THF, a polystyrene-equivalent molecular weight was determined using size exclusion chromatography on an Agilent 1260 Infinity liquid chromatography system with a Wyatt Technology Optilab T-rEX Refractive Index Detector and Wyatt Technology Heleos Dawn 8+ Light Scattering detector at a flow rate of 0.35 mL/min and 30 °C or a Tosoh EcoSEC Elite HLC-8420GPC equipped with a refractive index

detector at a flow rate of 0.45 mL/min and 40 °C. For both THF GPC systems, a Tosoh TSKGel guard column and two TSKGel Alpha-M columns were used for the separation and samples were prepared at 10 mg/mL. For resynthesized polymers only soluble in HFIP, poly(methyl methacrylate) equivalent molecular weight was determined using size exclusion chromatography on an Agilent 1260 Infinity liquid chromatography system equipped with a refractive index detector at a flow rate of 0.3 mL/min and 30 °C using a Agilent PL HFIPgel 4.6 mm diameter column and guard column.

Polymer Composition Determination: A Brucker Avance Neo600 or Bruker Avance Neo500 nuclear magnetic resonance spectrometer was used to measure ^{1}H NMR spectra of all polymers; 32 scans and a 3 second pulse delay were used for all samples. Molar ratios of the incorporated monomers were determined from peak integration and polycondensation reaction stoichiometry. Data was acquired in deuterated chloroform when possible, and otherwise in deuterated dimethyl sulfoxide or deuterated 1,1,1,3,3,3-hexafluoroisopropanol.

Biodegradation Testing: Biodegradation testing was carried out following procedures from Fransen, Av-Ron, et al.[42] Briefly, agar growth media containing suspended polymer particles was prepared by homogenizing 0.8 mL of polymer dissolved in dichloromethane at 40 mg/mL into 40 mL of American Type Culture Collection (ATCC) growth medium 179 containing 20 g/L agar using a Scilogexe D500 homogenizer with a with S20F/ER20 probe head. Pouring of the dissolved polymer solution in DCM into the agar medium resulted in precipitation of the polymer to form small particulates for testing. For polymers insoluble in DCM, a suspension was generated by sonicating the polymer in DCM in a water bath sonicator for 2-3 minutes. DCM was evaporated during the homogenization process. Homogenization was carried out in a fume hood for safety. Either 0.8 or 1 mL of the polymer containing ATCC growth medium with agar was added into 8 of 12 wells in a plate. The remaining 4 wells were filled with 0.8 or 1 mL of ATCC growth medium containing no polymer. *Paucimonas lemoignei (P. lemoignei)* was incubated in liquid ATCC growth medium until reaching an absorbance at 600 nm (OD600) of 0.5-0.8 at 30 °C and 180

rpm. A drop of *P. lemoignei* liquid culture (1 µL) was added to the center of 6 of the polymer-containing wells and 2 of the wells containing no polymer. The remaining 4 wells were not inoculated with bacterial culture. Plates were incubated for 13 days at 30 °C and monitored via imaging once every 24 hours using an Opentrons 1st generation robot or custom 2-D Gantry system[55] with a 2.8-12mm Varifocal USB Webcam Mini Camera and an 8x11 in Kaiser Slimlite Plano light tablet to provide background light. Further details are available in the *Supplementary Dataset.*

Biodegradation Analysis: Optical density curves showing the progression of degradation in each well were extracted using methods and software developed by Av-Ron.[56] Biodegradability was determined through a comparison of the optical density (OD) over time of the sample wells to the control wells. A sample was considered biodegradable when the control rate of change in OD (no polymer) fell outside of one standard deviation of the sample rate of change of OD (polymer samples inoculated with bacteria). Additional details on the biodegradation analysis can be found in the *Supplementary Information*. Three different types of biodegradation behavior could be observed visually. Briefly, clear zone behavior showed circular areas of optical clarity in the initially opaque plate, indicative of a diffusion limitation to the observed degradation. For distributed behavior, the opacity of the plate decreased uniformly during testing, and color interference showed the appearance of colored degradation products. Polymers for which the agar wells showed no colony growth were considered non-biodegradable as there was no microbial action which could be contributing to the polymer degradation; the inhibition of colony growth indicated toxicity of the sample to the bacteria. The biodegradability of the polymer and the biodegradation behavior (clear zone, distributed, or color interference as defined in Fransen, Av-Ron, et al.[42]) is provided in the *Supplemental Dataset*, and further information on the designations is provided in the *Supplementary Information*.

Libraries

In addition to the copolymer library synthesized in this work, polymers and data from Fransen, Av-Ron, et al.[42] were used for biodegradation trends analyses and machine learning. The libraries are described in **Table 1.**

Table 1: Description of libraries analyzed in this work.

Name	Description	Data Availability
Copolymer Library	A 300-member copolymer library synthesized from combinations of diols and dicarboxylic acids; each copolymer contains two repeat units.	Polymers were synthesized and published as part of this work. Details for specific polymers are available in the *Supplementary Dataset.*
Original Library	A 642-member library of polyesters and polycarbonates. Polyesters include homopolymers (synthesized from a diol and dicarboxylic acid derivative) and copolymers (synthesized from two lactone monomers).	Polymers from ref. 42. Data is available under a CC BY-NC-ND license. Copyright 2023 Fransen, Av-Ron, et al.[42]
Homopolymer Library	A 236-member subset of the original library comprised of homopolymers synthesized from the diols and dicarboxylic acids used in this study.	Polymers from ref. 42. Data is available under a CC BY-NC-ND license. Copyright 2023 Fransen, Av-Ron, et al.[42]

Machine Learning of Polymer Biodegradability

Polymer Representation

Previously, homopolymers have been represented for machine learning using SMILES (Simplified Molecular Input Line Entry System) representations[57] of several monomer unit repeats joined into a ring structure,[53] or as just the monomer repeat unit.[45] However, such approaches cannot be directly extended to copolymers due to the presence of multiple repeat units. An alternative approach is to use a weighted average of monomer representations;[51] though, this approach is inconsistent if the most prevalent chemical fingerprints, bit vector fingerprints, are used. To overcome this challenge, polymers were embedded for machine learning as a vector of similarity scores; the n^{th} score in the vector was the pairwise similarity between the polymer being

represented and the nth polymer in a landmark set of polymers known as the landscape.[58] The algorithm from Shi et al.[59] was used to compute each of these similarity scores.

Specifically, the polymers *P* and *Q* were each represented as an ensemble of monomers with each monomer (dicarboxylic acid or diol) embedded via an RDKit Fingerprint[60] with a radius of 5 and 512 bits. Polymer *P* with *m* monomers where p_i is monomer *i* and w_{pi} is the mole fraction of monomer p_i in polymer *P* can be denoted as

$$P = \{(p_1, w_{p,1}), (p_2, w_{p,2}), \ldots, (p_m, w_{p,m})\}$$

If polymer Q has n monomers q_j with mole fraction $w_{q,j}$ then the Tantimoto similarity score[61] between fragments p_i and q_j can be defined as $s_{i,j}$ and the distance between the fragments as $d_{i,j}$=1-$s_{i,j}$. The earth mover's distance metric,[62] EMD*(P,Q)*, between *P* and *Q* can then be obtained from:

$$EMD_{P,Q} = min_F \sum_{\mathrm{i}=1}^{m} \sum_{j=1}^{n} f_{i,j} \cdot d_{i,j}$$

subject to $f_{i,j} \geq 0, \sum_{j=1}^{n} f_{i,j} = w_{p,i}$ and $\sum_{i=1}^{m} f_{i,j} = w_{q,j}$ for all i and j that do not appear in summations. To generate a vector embedding representation for polymer *P* using landscape ***Q***={$Q_1, Q_2, \ldots, Q_k, \ldots Q_z$} containing *z* number of polymers, the EMD(P,Q_l) can be calculated for all *z* polymers in ***Q*** and mapped into an embedding vector of EMD similarity scores (***S***) for polymer *P*:

S=[EMD(P,Q_1), EMD(P,Q_2),…, EMD(P,Q_k),…, EMD(P,Q_z)]

The mapping for the embedding vector for one generic polymer compared to a generic 5-polymer landscape is shown in *Figure S6.*

Landscape Optimization

For both the original library and the copolymer library, the effect of landscape size was explored. Ten different landscapes were explored. The first 5 landscapes were made of polymers from *i* splits (where *i* is 1 through 5 to obtain 5 landscapes) of the original library data. The second 5

landscapes were made similarly, but using *i* splits of both the original library data and the copolymer data.

Data Splits

For all libraries, the data was split into 6 groups. One split was used as the test set; the remaining 5 splits were used as training data. The chemically representative data splits of the original library were taken from Fransen, Av-Ron et al.[42] The splitting strategy used to ensure splits representative of chemical structures in the copolymer library is described as follows. For both the homopolymer library and the copolymer library, each polymer was first categorized as terephthalic/isophthalic acid-containing and biodegradable, terephthalic/isophthalic acid-containing and non-biodegradable, aliphatic dicarboxylic acid-containing and biodegradable, and aliphatic dicarboxylic acid-containing and non-biodegradable. For these two libraries, one sixth of each category was randomly selected and combined to generate each split of the homopolymer library and the copolymer library. For test sets, the 'Combined Polymers' test set was the combination of the splits from the original library and copolymer library used as the test sets.

Model Optimization

To determine optimized model hyperparameters, 5-fold cross-validation was carried out using 5 data splits. The models were then retrained using the full training set and evaluated using the test set (the remaining 6th data split). For random forest classifier models,[63] the number of estimators (trees), maximum tree depth, and maximum leaf nodes were optimized; the minimum sample split and minimum samples per leaf were fixed at 2 and 1, respectively. Linear models using stochastic gradient descent learning[63] were implemented and the model type (linear state vector machine (SVM) classifier, logistic classifier, or linear perceptron classifier) and regularization term multiplier were optimized. All models were optimized unless stated otherwise.

Landscape Optimization

For both the original library and the copolymer library, the effect of landscape size was explored. Ten different landscapes were used for the original library. The first 5 landscapes were made of

polymers from *i* data splits (where *i* is 1 through 5 to obtain 5 landscapes) of the original library data. The second 5 landscapes were made similarly, but using *i* splits of both the original library data and the copolymer data.

Model Error Evaluation

For hyperparameter optimization, the first 5 splits of each library were used as training data for cross-validation and the 6th split was used as a test set. For evaluation of optimized models with different amounts of data or different landscapes, cross-validation with all six data splits was performed.

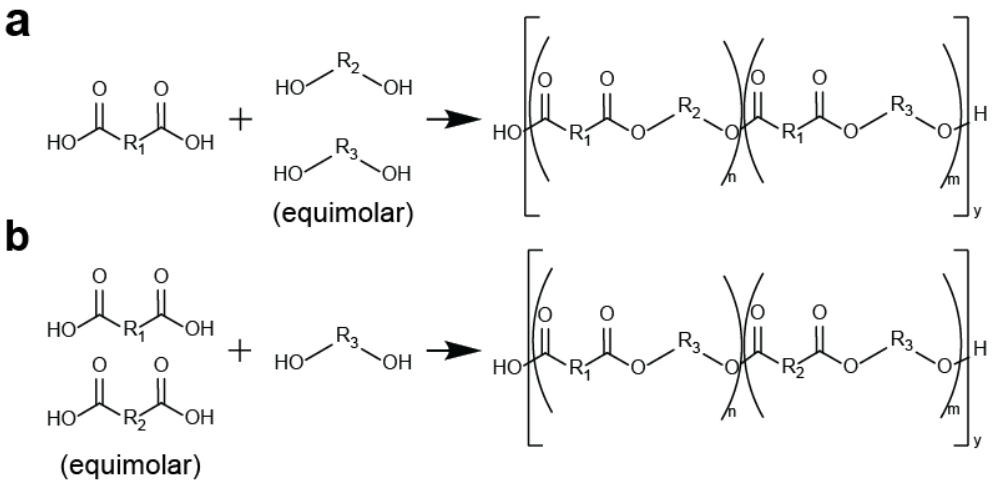
Results and Discussion:

Library Design and Synthesis

A library of copolyesters prepared by polycondensation was designed by selecting 24 diols and 12 diacids from the monomers of a polyester and polycarbonate library (original library)[42] based on their polymerizability and diversity of chemical functionality. These monomers spanned a range of different chemical substructures, such as ethers, thioethers, aromatic rings, and non-aromatic rings, with the intention of emphasizing chemical diversity among substructures in commercially available monomers. A summary of the carboxylic acids and derivatives, their chemical structures, their number of occurrences in the library, their SMILES,[57] and the three-letter abbreviation code used in polymer naming can be found in *Table S1*; the same information is available for the diols in *Table S2.* To synthesize each copolymer, two diols and one diacid (200 polymers) or two diacids and one diol (100 polymers) were selected at random from the monomer set.

Polymers were synthesized with equimolar ratios of the monomer (diol or dicarboxylic acid) with two variants as depicted in **Scheme 1**. A slight stoichiometric excess of the diol, or lower boiling point diol when two diols were present, was added to ensure conversion to high molecular mass. A high-throughput reactor was customized to enable the simultaneous synthesis of 11 copolymers under high-vacuum (100 mtorr) conditions; images of the customized parts and

the assembled reactor are available in *Figure S1*, and schematics of the custom-designed parts are available in *Figure S2*. The polymers combined in each synthesis batch were selected based on the boiling points of the diol (or lightest diol), and polymers with the same composition (two diols and one dicarboxylic acid vs. two carboxylic acids and one diol) were preferentially placed in the same synthesis set. Of the 300 polymers, 283 reached degrees of polymerization > 5; the remaining materials could not be synthesized to higher molecular weights due to a combination of factors such as monomer purity, degradation of the synthesized materials at polymerization temperatures, limited reactivity of the monomers, or loss of a monomer under reaction conditions leading to charred samples. Two polymer compositions, one containing all aliphatic monomers and one containing an aromatic monomer were synthesized in triplicate to demonstrate the replicability of the synthesis and biodegradation testing. The resulting polymers had consistent monomer composition, M_n values within 10% error, and yielded the sample biodegradation result, demonstrating the reproducibility of the synthesis under the same synthetic conditions and consistency of the resulting biodegradation determination (see the *Supplementary Dataset* for polymer details).



Scheme 1: Synthetic schemes used in the synthesis of the copolymer library. **A:** 200 polymers were synthesized as a combination of one random dicarboxylic acid and two random and unique diols. **B:** 100 polymers were synthesized as a combination of two random and unique dicarboxylic acids and one random diol.

Trends in Copolymer Biodegradability

Copolymers were biodegradable for 82% of the tested library. Compared to polymers from Fransen, Av-Ron, *et al.*[42] (referred to as the original library) which spanned a > 90 monomers space and were biodegradable for 49 % of the tested polymers,[42] the copolymers showed an increase in biodegradability. Compared only to homopolymers based on the subset of 36 monomers used in this study (homopolymer library), the copolymers were still more biodegradable: 50% of the homopolymers containing the subset of 36 monomers in this study were degradable (*Figure S3*). **Figure 1** shows the fraction of biodegradable copolymers containing each combination of dicarboxylic acid and diol (shown on the axes), across the tertiary component of the polymers. An overlay of the homopolymer library biodegradability[42] for each dicarboxylic acid and diol combination highlights the positive impact of copolymer structure on biodegradability. The observed increase in biodegradability aligns with previous copolymer studies characterizing the impact of monomer ratios on biodegradability in carbon dioxide-propylene oxide-ε-caprolactone polymers and poly(glycolide-co-L-lactide-co-ϵ-caprolactone), which demonstrated increased biodegradability with increasing comonomer fractions.[64-65]

The magnitude of the increase in biodegradability for the copolymer library over the homopolymer library was 9 % larger than expected based on the homopolymer library biodegradabilities. Based on the homopolymer library data, 73 % of the copolymer library was expected to biodegrade, but 82 % of the library showed degradation. For each copolymer, the three monomers used could also be combined in pairs to yield two unique, or constituent, homopolymers, each containing one diol and one dicarboxylic acid, referred to as the constituent homopolymers in the text. **Table 2** partitions the biodegradable and non-biodegradable copolymers by how many of their two constituent homopolymers were biodegradable in the homopolymer dataset. Copolymers with one or two biodegradable constituent homopolymers (73 % of the copolymer library) were expected to biodegrade since they contain linkages shown to be biodegradable among the homopolymer library. However, only 90 % of the copolymers expected

to be biodegradable based on their constituent homopolymers were biodegradable. The greater-than-expected overall biodegradability was achieved due to the 57 % of copolymers without biodegradable constituent homopolymers that were biodegradable. For these polymers, it was hypothesized that the random copolymer structure increased the accessibility of the cleavable ester bonds, enabling the biodegradation of these formulations.

Growth inhibition of the *Paucimonas lemoignei* (*P. lemoignei*) colony contributed significantly to whether the copolymers were biodegradable; colony growth was considered necessary for biodegradability. It is hypothesized that for polymers which did not show growth, the polymer or early-time degradation products are toxic to the bacteria. Of the 48 non-biodegradable copolymers, 20, or 42 %, had no colony growth (**Figure 2a**). The inhibition of colony growth explains the deviation of copolymer biodegradability from what is predicted based on constituent homopolymer biodegradability; 5 of the 16 non-biodegradable copolymers with expected biodegradability were non-biodegradable due to colony growth inhibition.

The toxicity of the polymer or early-time degradation products contributed to their non-biodegradability for copolymers with and without biodegradable constituent homopolymers. Identifying monomers shared among the copolymers which prevented colony growth reveals that 81% had backbone aromatic groups from bisphenol A (BPA), terephthalate, or isophthalate (**Figure 2b**), chemicals with studied toxic effects on human health,[66-67] aquatic species,[66-67] and select microbes.[67-68] For the subsets of the copolymer library which contained monomers with aromatic units (BPA, terephthalate, isophthalate, and guaiacol glycerol ether), similar biodegradability was observed compared to the equivalent subset in the homopolymer library: 41 % of the aromatic containing copolymer library subset was biodegradable compared to 39 % of the aromatic homopolymer library subset, likely due to the toxicity of their biodegradation byproducts. The remaining polymers that showed inhibited colony growth contained 1,5-pentanediol, a chemical with known antimicrobial activity.[69] Similarly to the effects of the polymers with aromatic backbone groups, early time degradation and release of 1,5-pentanediol may be

inhibiting the necessary colony growth for biodegradation. Additionally, for BPA, terephthalate, and 1,5-pentanediol, minimum inhibitory concentration testing for *P. lemoignei* demonstrated toxicity of the monomers, supporting the hypothesis that early time degradation products are able to inhibit colony growth during biodegradation testing. Details for the toxicity testing method and results of the testing can be found in the *Supplementary Information* and *Figure S4*.

Increased degradability from the copolymer structure may be contributing to the release of toxic bioproducts and subsequent inhibition of colony growth: the copolymers with no growth and known constituent homopolymer biodegradability had either an aromatic monomer or 1,5-pentanediol as part of a non-biodegradable constituent homopolymer. For copolymers with biodegradable constituent homopolymers, the biodegradation byproducts or their release rate likely differed from that observed during the homopolymer biodegradation. For one copolymer (GLA50_HND25_GGE25), both constituent homopolymers were biodegradable, but combined into a copolymer, they inhibited growth. It is hypothesized that the increased biodegradability from the copolymer structure leads to a more rapid release of toxic bioproducts, which leads to inhibited colony growth.

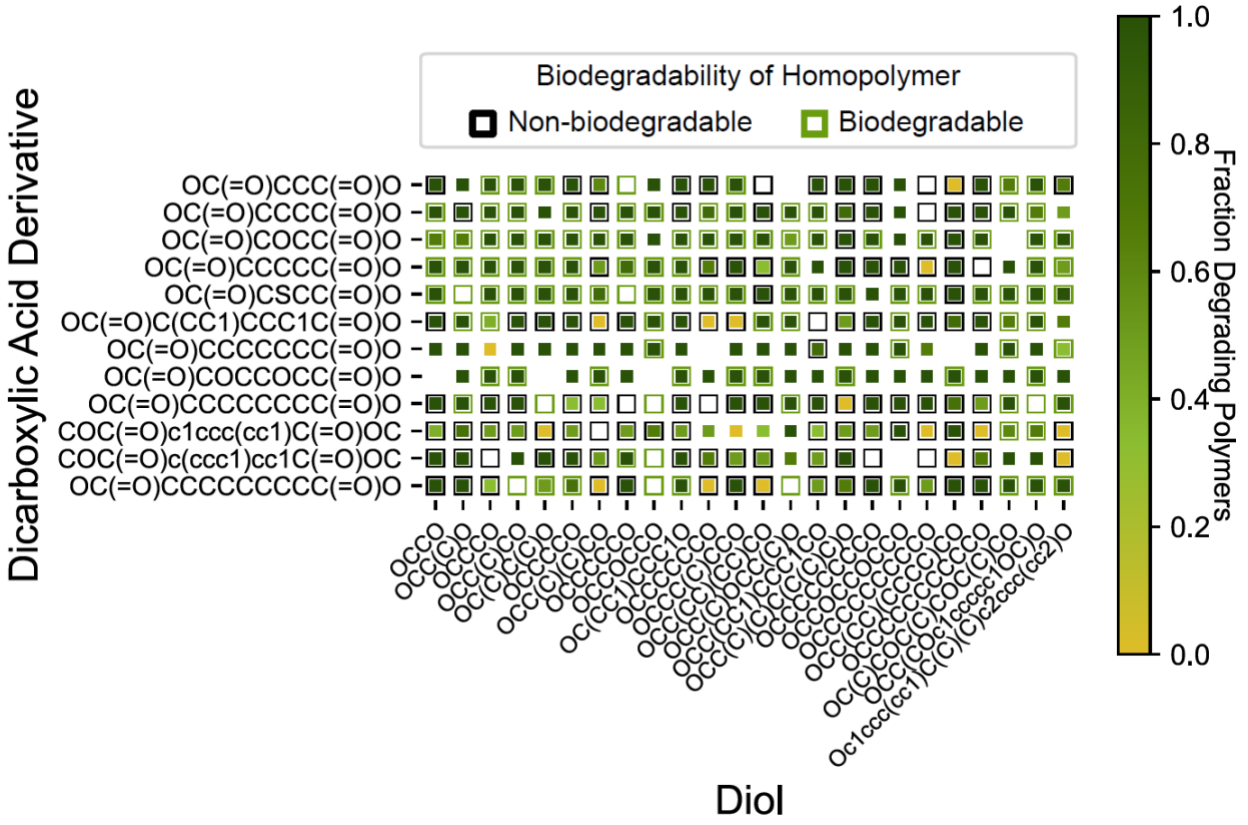


***Figure 1**:* Overview of the biodegradability of polymers containing each dicarboxylic acid derivative and diol combination. Each inner square represents the set of copolymers containing the indicated diol and dicarboxylic acid derivative as 2 of the 3 monomers used to synthesize the polymer; the inner square color indicates the fraction of copolymers in that set that showed biodegradability. The outer squares show the biodegradability of the homopolymer (from the homopolymer library) synthesized from the indicated two monomers or equivalent monomers. Homopolymer library data was reproduced from ref 42. Available under a CC BY-NC-ND license. Copyright 2023 Fransen, Av-Ron, et al.[42]

Table 2: Copolymer biodegradability breakdown with respect to constituent homopolymers. The table describes the number of polymers in each combination of categories.

	Two Biodegradable Constituent Homopolymers	**One Biodegradable Constituent Homopolymers**	**Zero Biodegradable Constituent Homopolymers**	**Missing Data for Constituent Homopolymers**	**Totals**
Biodegradable Copolymers	60	82	32	78	252
Non-Biodegradable Copolymers	2	14	26	6	48
Totals	62	96	58	84	300

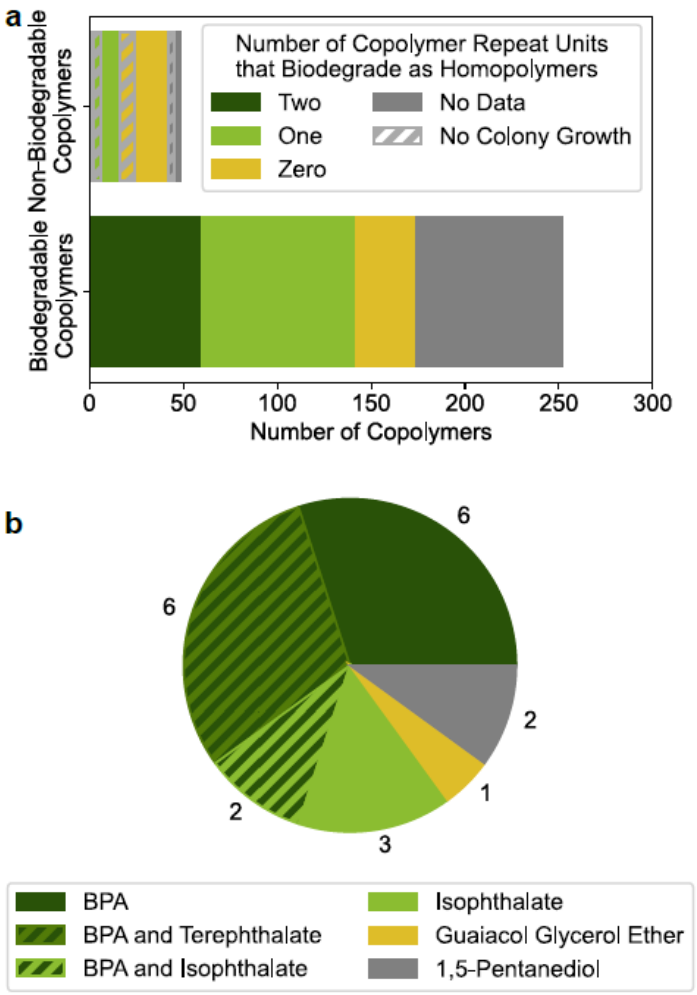


***Figure 2**:* **a:** For each copolymer, there are two constituent homopolymers that contain unique pairs of the copolymer's monomers. For copolymers where both constituent homopolymers have data in the homopolymer library, the number of biodegradable and non-biodegradable copolymers is depicted. Copolymers missing constituent homopolymer data are included in grey. The number of copolymers that inhibited colony growth is overlaid with a diagonal grey hatch for non-biodegradable copolymers. Note that there are two non-biodegradable copolymers with two biodegradable constituent homopolymers which are not visible due to the image scale; these polymers did not show colony growth. Colony growth was determined by visual observation after two weeks of incubation. **b:** Breakdown of monomers present in copolymers, which showed no colony growth. A total of 20 copolymers showed no colony growth; the number represented in each slice is shown adjacent to the relevant piece of the chart. Homopolymer data was reproduced from ref 42. Available under a CC BY-NC-ND license. Copyright 2023 Fransen, Av-Ron, et al.[42]

No singular rationalization was evident for the 11 non-biodegradable copolymers predicted to biodegrade which showed colony growth. Four of the copolymers contained terephthalate, and another 3 contained a long-chain dicarboxylic acid, which have both been correlated with inhibited biodegradation.[42] One of the terephthalate-containing polymers also

contained 1,5-pentanediol, indicating that some of these polymers may have some low-level toxicity which doesn't fully inhibit colony growth but may contribute to the inhibition of biodegradation. Three additional polymers contained 1,4-cyclohexanedicarboxylic acid, 1,3-propanediol, and a third diol unique to each of the polymers. While the homopolymer from 1,3-propanediol and 1,4-cyclohexanedicarboxylic acid was biodegradable, the copolymer from 1,4-cyclohexanedicarboxylic acid and the other diol was not biodegradable in each case. This lends evidence to the hypothesis that for certain generally biodegradation-inhibiting structures like long dicarboxylic acid chains or bulky ring-containing diacids, the copolymer structure may not be sufficient to overcome the effects limiting or inhibiting the biodegradability of the substituent homopolymer and this effect is then carried to the copolymer.

Smaller repeat unit backbone lengths and oxygen substitution of backbone carbons improved biodegradability for homopolymers and copolymers; however, the tested copolymer repeat unit lengths did not show a cutoff repeat unit length above which no copolymers were biodegradable. **Figure 3a** demonstrates the decreasing number of biodegradable polymers with increasing carbon length of the repeat unit in the backbone of the polymers for both copolymers and homopolymers. No biodegradable homopolymers were observed for the homopolymer library with repeat unit carbon lengths of 15 or greater; 50 % of the copolymers remain biodegradable for these longer chain lengths (average repeat unit lengths of 15-17).

Oxygen substitution of backbone carbons (**Figure 3b**) improved biodegradability for the copolymers, original library, and homopolymer library; however, for the copolymer library, improvements were only observed with more than one oxygen substitution. For the original library and homopolymer library, the fraction of biodegradable polymers increased by at least 0.18 for all pairs while the copolymers saw no change for single oxygen substitutions compared to increases of 0.12 or 0.17 (the maximum possible as all polymers became degradable) for oxygen substitution of two carbons. It should be noted that the perceived difference in impact for oxygen substitutions in the copolymers vs. homopolymers and the original library may be driven by the

higher number of biodegradable polymers among the copolymers, which resulted in few polymers that can become biodegradable with oxygen substitutions. A sulfur substitution in glutaric acid shows similar positive impacts on biodegradability as oxygen substitution (*Figure S5*).

For both the original library and the copolymer library, the physical state was correlated to the biodegradability of the polymers. Copolymers had fewer degradable polymers among crystalline polymers (64 % biodegradable) compared to solid, non-crystalline (84 % biodegradable) and viscous polymers (91 % biodegradable) (*Figure S6*). This trend was mirrored in the original library,[42] with both datasets supporting that disruptions to material chain packing and crystallinity provide an increase in enzymatic biodegradation rate.[10] Further, the copolymers also had a smaller fraction of crystalline polymers (26 % of solid polymers were crystalline) than the original library (40 % of solid polymers were crystalline).

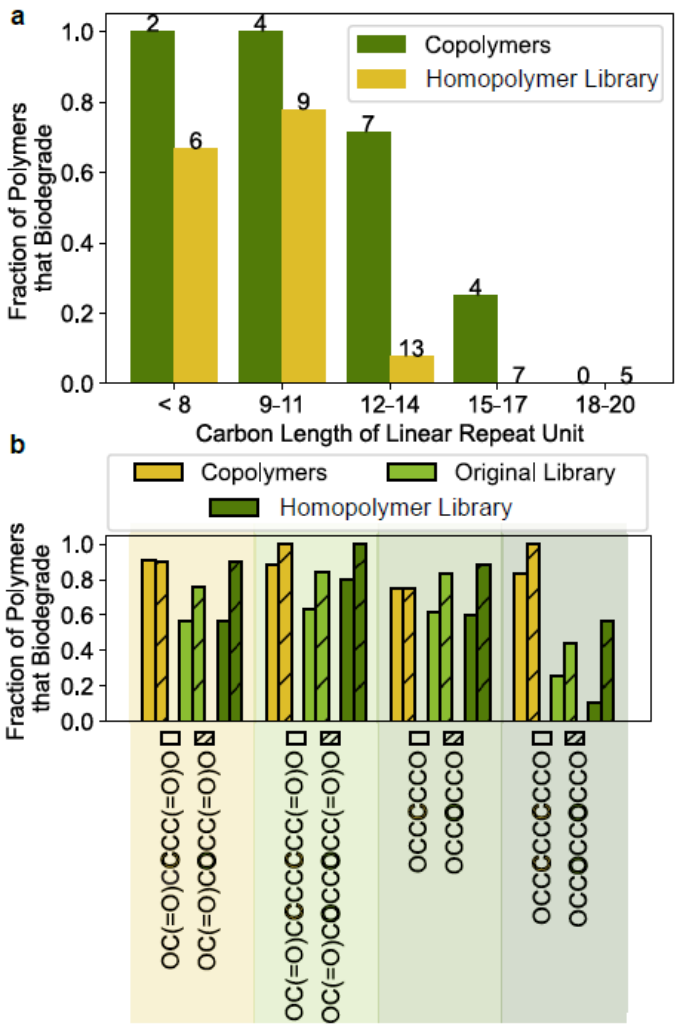


Figure 3: Original library and homopolymer library data was reproduced from ref 42. Available under a CC BY-NC-ND license. Copyright 2023 Fransen, Av-Ron, et al.[42] **a:** Fraction of biodegradable polymers in the copolymer and homopolymer library containing only linear aliphatic monomers. For the homopolymer library, the carbon length of the linear repeat unit was the sum of the carbons in the constituent diol and dicarboxylic acid. For the copolymers, the length was calculated using a molar weighted average (determined by ^{1}H NMR) of constituent monomer carbon lengths. The total number of polymers represented in each bar grouping is listed on top of each bar. **b:** Effect of oxygen substitution on biodegradability. For the copolymers, only polymers in which both repeat units had the oxygen substitution are depicted. Oxygen-containing monomers are paired with their carbon-only counterparts, and the oxygen atoms and the carbons that they replace in their monomer pair are colored green and yellow, respectively.

Polymers that were resynthesized to higher molecular weights showed the same biodegradability as their lower molecular weight counterparts in 36 of 41 cases; however, the type of observed biodegradation behavior shifted with the increase in molecular weight for 8 of 36 cases where biodegradation was the same. The change in biodegradation behavior from clear-zone to distributed is reflective of a retardation of the observable degradation rate; high molecular

weights require an increase in the number of broken bonds before the polymer particles are small enough to no longer scatter light and distributed degradation has previously been correlated to slower degradation rates.[42]

Of the five cases where biodegradation behavior changed with increasing molar mass, two (SCA50-DEG25-BPA25-RE and SUA50-TPD25-PTD25-RE) transitioned from no growth to biodegradation. In the lower molecular weight samples, it is hypothesized that there was fast enough release of toxic BPA and 1,5-pentanediol to the bacteria to inhibit growth. At high molecular weight, the release of monomers was slower, allowing for colony growth. Sustaining a colony lead to eventual biodegradation. The remaining polymers which showed a change from distributed degradation to no degradation (ADA25-DMT25-DPD50, DGA25-DMT25-ETG50, and DMT50-DPD25-HND25) all contained terephthalate. This was consistent with the low fraction of biodegradable terephthalate-containing polymers previously observed,[42] and is likely due to inaccessibility of the ester bonds for degradation with chain packing in the increased molecular weight materials.

Predicting Biodegradability Using Machine Learning

The novel landmark-based polymer embedding (*Figure S7*) was able to represent the original library, copolymer library, and both libraries simultaneously. Comparable maximum test set accuracies were achieved for modeling the original library using the similarity score-based polymer embedding (81 %, see *Table S3*) and RDKit fingerprinting (82 %)[42] with random forest models. Each polymer embedding was a set of similarity scores calculated against a landscape composed of the entire training set leveraging the similarity scoring method from Shi et al.[59] A visual summary of the embedding calculation process is shown in *Figure S7,* and the novel chemical embedding method for polymer representation enabled the simultaneous training and evaluation of predictive machine learning algorithms for both original library and copolymer biodegradation.

In order to assess the impact of the landscape on the biodegradability prediction ability of machine learning models, different landscapes were used to train a random forest model on the full copolymer and original library training datasets. The ability to use a small, chemically representative landscape, rather than the entirety of the training dataset is valuable as it prevents the need to calculate new embeddings for all polymers with any addition of data to the dataset. In order to find the minimum size landscape needed to achieve high-accuracy models, increasing landscape sizes were tested for landscapes that included only original library polymers (**Figure 4a**) and both original library polymers and copolymers (**Figure 4b**). The test set accuracies were robust to the number of polymer splits in the landscape, which was hypothesized to be due to the balanced distribution of chemical features across both the data splits used to build the landscapes and the test data. However, using only polymers from the original library in the landscape provided slightly higher accuracies than using both the original library and copolymer library polymers in the landscape across each test dataset (**Figure 4a** and **Figure 4b**). Likely, the narrower chemical range of the monomers included in the copolymer library and the narrowed distribution of similarity scores in the embeddings when including copolymers in the landscape (*Figure S8*) increased the error in the machine learning predictions. Therefore, the first split of the original library was used as a small, chemically representative landscape for all subsequent modeling unless otherwise stated.

Original library machine learning models did not show transferability to the copolymer library; these models failed to accurately predict copolymer biodegradability for both linear classifiers and random forest classifiers. Random forest classifiers and linear classifiers were investigated for their relatively limited number of trainable parameters and robustness to small dataset sizes.[63] When models were trained using only the original library or homopolymer library dataset, they failed to predict the copolymer test data accurately, achieving accuracies of between

42 % and 49 % across the two training datasets and model types (*Table S3*), the statistical equivalent of a naïve guess of biodegradability.

Unlike linear models, random forest models could be trained to capture both original library and copolymer biodegradation behavior. This optimization (*Table S4*) yielded a model that had 95 % accuracy for the copolymer test set and was able to maintain a test set accuracy of 78 % for the original library, an accuracy comparable to the 81 % achieved by a model optimized for only the original library (*Table S3*) within the model error estimated by cross-validation. When training the optimized models from the original library with increasing amounts of copolymer data, linear algorithms were unable to reliably and accurately capture both types of data (**Figure 4c**), but random forest models (**Figure 4d**) were able to maintain accuracy for the original library test set while achieving 95 % copolymer test set accuracy. Despite better maintaining the original library data accuracy, the random forest models did not show definitive plateaus in accuracy with increasing fractions of copolymers in the training data (**Figure 4d**); the dataset size used for training remained limiting to the accuracy of the predictive machine learning models over the chemical space. Being data-limited, the optimized random forest model accuracy remained sensitive to the amount of copolymer training data and original library training data used for both the original library test set (**Figure 4e**) and the copolymer test set (**Figure 4f**). The robustness of the random forest model was further probed through extrapolation studies where specific chemical functionalities (ethers, sulfurs, aromatic groups, and aliphatic rings) were withheld from the training data and used as the testing data (*Figure S9*). Only the biodegradability of sulfur-containing copolymers was predicted with accuracy (92 %) comparable to the full model accuracies (95 % for the copolymer test set). The original library aromatic polymers prediction accuracy was 47 %, worse than a naïve random guess of polymer biodegradability, emphasizing the need for a chemically diverse and representative dataset to accurately predict polymer biodegradability. Overall, these results highlight the difficulty in training a model for two differing chemical spaces with limited experimental data, as sparse data in a broad space can strongly

impact individual predictions and fluctuate the accuracy of the model, and no *a priori* method exists that can establish the amount of data required for an accurate predictive model based on experimental data.

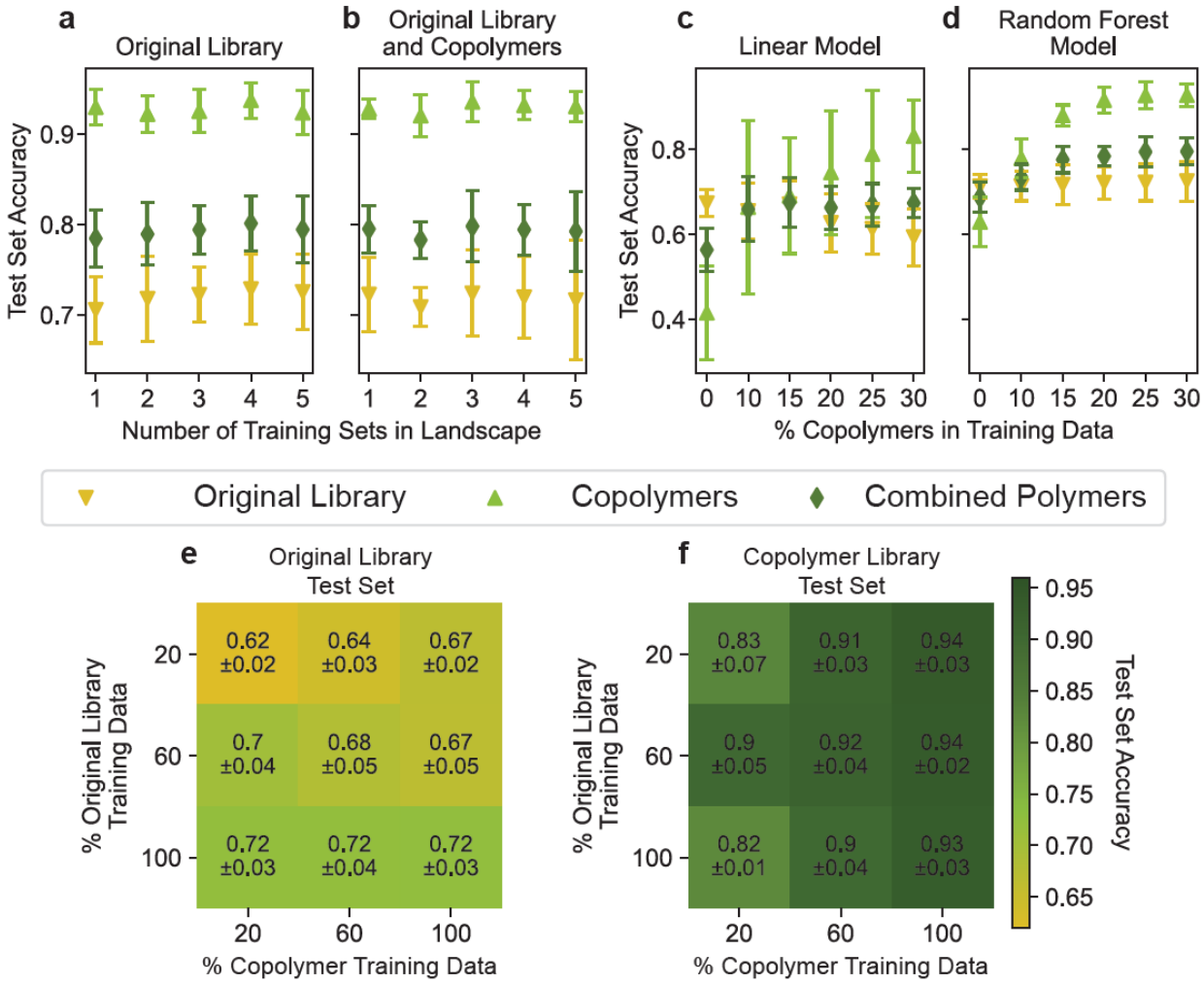


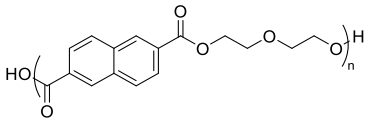
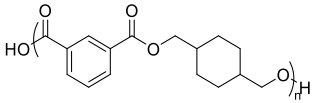
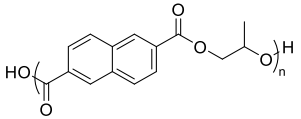
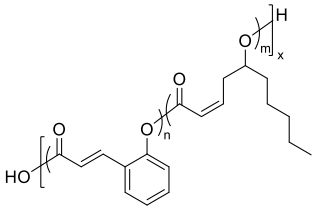
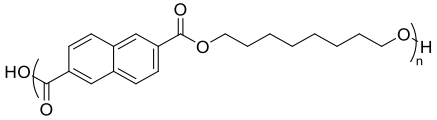
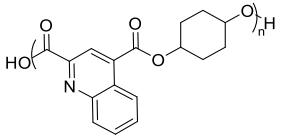
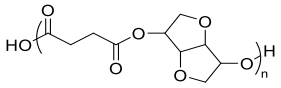
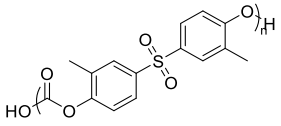
Figure 4*:* Subfigures a-d have common markers indicating the appropriate test sets. Test sets used to evaluate the models were one split of data held back during training from the original library, copolymer, and the combination of these two test sets (combined polymers). Subfigures e and f share a color bar to their right. **a and b:** Models used hyperparameters from optimization for original library polymers (Table S3). Models were random forest models. **a:** Number of splits (*i*) of original library polymers used in the landscape was varied. **b:** Number of splits of original library polymers and copolymers used in the landscape was varied. For *i* number of training sets, the polymers in *i* splits of the original training data and *i* splits of the copolymer training data were used as the landscape. **c and d:** Models used optimized hyperparameters from the original library data (Table S3). Models with the optimized hyperparameters were trained with all of the original library training data and increasing amounts of copolymer data. The first original library split was used as the landscape. **c:** The optimized linear model (original library data) was a logistic regression classifier with a 0.01 regularization constant multiplier**. d:** Random forest model with hyperparameters of 16 Trees, a maximum tree depth of 16 and 128 maximum leaf nodes **e and f:** Random forest model with 32 trees, 64 maximum leaf nodes, a maximum tree depth of 32, a minimum sample split of 2, and one sample per leaf trained with varying amounts of original library and copolymer training data. Model hyperparameters were taken from optimization of random forest model for the original library and copolymer library (*Table S4)*, and the first split of the original library was used as the landscape. The models were evaluated using original library test sets (**e**) and the copolymer library test sets (**f**).

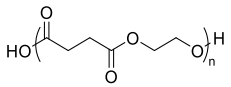
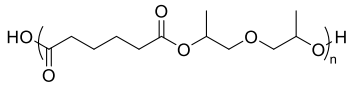
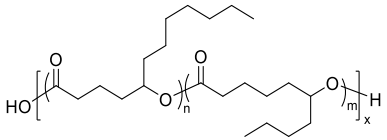
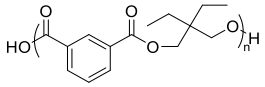
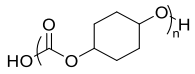
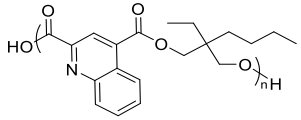
Despite data limitations, the joint models (*Table S4*) provide utility in identifying biodegradable polymer candidates. Examining the uncertainty of the model's predictions revealed that true positive predictions had higher probabilities associated with them compared to false positive results (*Figure S10*). Thus, polymers with high predicted probabilities, as opposed to simply greater than 50 %, could be directly used to bias chemistry selection to the most promising monomers, enabling polymer design. The critical evaluation of the uncertainty of the predicted biodegradability of novel candidate materials permits an informed application of the available models within their current constraints and demonstrates the utility of even data-limited modeling results.

Feature analysis revealed the chemical structures most relied upon by the model for biodegradability prediction; 9 of the 15 highest-weighted similarity-score features were to aromatic polymers. Feature analysis was conducted on the optimized random forest model trained on all the original library and copolymer data; the BigSMILEs[49-50] representation of the 15 most important polymers identified in the feature analysis and their importance score can be found in **Table 3.** Other chemical structures identified in these polymers were ethers, aliphatic rings, side chains, and long linear aliphatic monomers; however, no indication is provided through the feature analysis on whether these chemical structures show positive or negative correlation with the biodegradability of the polymer. The identified chemical features aligned with the chemical structures identified in the trends analysis of this work, as well as in previous analysis of the original library;[42] for example, aromatic containing polymers have been found to be less biodegradable than their aliphatic counterparts, and ether containing polymers show improved biodegradability over their carbon-containing counterparts. Combining these two approaches enables insight into both which chemical structures show importance for biodegradability determination and if they positively or negatively impact the biodegradability of the polymers; this

highlights the value of both approaches in designing and assessing future biodegradable polyesters.

Table 3: Results of feature analysis for the random forest model with 32 Trees, a maximum tree depth of 32 and 64 maximum leaf nodes trained on the original and copolymer libraries with the first split of the original library used as the landscape.

Original Library Polymer BigSMILES[49-50]	Chemical Structure	Feature Importance Weight
{[][<]C(=O)c(ccc1c2)cc1ccc2C(=O)[<], [>]OCCOCCO[>][]}		0.02374
{[][<]C(=O)c(ccc1)cc1C(=O)[<], [>]OCC(CC1)CCC1CO[>][]}		0.02126
{[][<]C(=O)c(ccc1c2)cc1ccc2C(=O)[<], [>]OCC(C)O[>][]}		0.01977
{[][<]C(=O)C=Cc(cccc1)c1O[>], [<]C(=O)C=CCC(CCCCC)O[>][]}		0.01811
{[][<]C(=O)c(ccc1c2)cc1ccc2C(=O)[<], [>]OCCCCCCCCCO[>][]}		0.01732
{[][<]C(=O)c1nc2ccccc2c(c1)C(=O)[<], [>]OC(CC1)CCC1O[>][]}		0.01605
{[][<]C(=O)CCC(=O)[<], [>]OC(CO1)C2C1C(CO2)O[>][]}		0.01571
{[][<]C(=O)[<], [>]Oc1c(C)cc(cc1)S(=O)(=O)c2ccc(c(C)c2)O[>][]		0.01567

{[][<]C(=O)CCC(=O)[<], [>]OCCO[>][]}		0.01540
{[][<]C(=O)CCCCC(=O)[<], [>]OC(C)COCC(C)O[>][]}		0.01533
{[][<]C(=O)CCCC(CCCCCCC)O[>], [<]C(=O)CCCCC(CCCC)O[>][]		0.01511
{[][<]C(=O)c(ccc1)cc1C(=O)[<], [>]OCC(CC)(CC)CO[>][]}		0.01499
{[][<]C(=O)[<], [>]OC(CC1)CCC1O[>][]}		0.01491
{[][<]C(=O)c1nc2ccccc2c(c1)C(=O)[<], [>]OCC(CC)(CCCC)CO[>][]}		0.01379

Conclusions

The utility of copolymerization in tuning the physical properties of polymers motivates its use when developing sustainable alternatives to tradition polymers. However, limited information is available on the impact of the copolymer structure on biodegradability; the synthesis and biodegradation testing of a 300-member copolymer library using the automated high-throughput clear zone assay enabled an understanding of biodegradation trends and the development of predictive biodegradability models for polyester copolymers.

Copolymers were biodegradable for over 80 % of the samples, demonstrating increased biodegradability over the 49 % biodegradable original library and increased biodegradability beyond predictions based on their constituent homopolymer biodegradability due to the copolymer structure. The increased biodegradability of the copolymers revealed evidence of degradation product toxicity to *P. lemoignei*; 42 % of the non-biodegradable copolymers inhibited bacterial growth. These polymers contained aromatic monomers with established toxicity such as BPA,[68] or monomers with antimicrobial properties[69] and highlight an interesting concern when developing biodegradable materials: highly degradable materials with potentially toxic degradation products may inhibit biodegradation and impact microbial consortia growth near the material. Trends observed in correlating chemical structure to biodegradability were shared between homopolymers and copolymers: biodegradability decreased with increasing linear carbon repeat unit length and increased with oxygen substitution of backbone carbons. However, these trends were less pronounced in copolymers, likely due to their already high inherent biodegradability and the limited available room for improvement. The presence of crystallinity correlated to reduced biodegradability for both homopolymers and copolymers, although no physical state information predicted biodegradability for either dataset.

Quantitative structure-property relationships for simultaneous biodegradability prediction of both original library polymers and copolymers were achieved by leveraging novel chemical similarity-informed embedding and random forest classifier models. The similarity score-based

embedding considered chemical substructures found within the polymer and its composition, allowing for direct comparison between homopolymers and copolymers. Linear models failed to capture original library and copolymer data simultaneously, whereas random forest models were able to match their accuracy in predicting original library data while simultaneously predicting copolymer library biodegradability. However, just like the linear models tested, random forest models trained only on original library data could not project to copolymer biodegradability, highlighting the need for experimental data spanning the relevant chemical and structural space when building models. Although the developed models remained data-limited, when paired with an understanding of trends in biodegradability for materials of interest, these models can be used to concentrate expensive experimental efforts for developing biodegradable materials to identified promising chemical and structural spaces. Here, the copolymer chemical architecture was established as a valuable tool for increasing polymer biodegradability in addition to the accepted tunability of physical properties for copolymers.

Data Availability: All data and code used for library generation, data processing, machine learning model optimization, training, and evaluation, and figure generation is available at https://github.com/olsenlabmit/Terpolymer_Modeling. Randomly generated data splits for each library used for training and test sets are included as part of the provided data. Software developed for extraction of biodegradation curves and biodegradability determination[42, 56] is available at DOIs 10.5281/zenodo.8187737 and 10.5281/zenodo.15236575. A summary of the 300-member copolymer library information, low molecular weight polymers which were resynthesized and replicated polymers, specific synthetic methods, actual masses used during synthesis, and monomer details including sourcing and purity are found in the *Supplemental Dataset*. The extracted optical density curves used for biodegradation determination are available at https://github.com/olsenlabmit/Terpolymer_Modeling/Biodegradability-OD-Plots. Integrated and annotated ^{1}H NMR spectra, including the original fid data are available at

https://github.com/olsenlabmit/Terpolymer_Modeling/NMR. The GPC/SEC datafiles and relevant calibration curves are available at https://github.com/olsenlabmit/Terpolymer_Modeling/GPC.

Notes: Certain equipment, instruments, software, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement of any product or service by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

Supporting Information: The following supporting information is provided.

Filename or Repository Information	Description
Supplementary Information	Additional Methods details, supplemental figures and schematics, monomer tables, supplemental tables
Supplementary Dataset	Polymer characterization data summaries including biodegradability, synthetic methods and reagent masses, monomer details
https://github.com/olsenlabmit/Terpolymer_Modeling/GPC	GPC/SEC data files for polymers
https://github.com/olsenlabmit/Terpolymer_Modeling/NMR	^{1}H NMR data files for polymers
https://github.com/olsenlabmit/Terpolymer_Modeling/Biodegradability-OD-Plots	Extracted OD plots used for biodegradability determination
https://doi.org/10.5281/zenodo.8187737 https://doi.org/10.5281/zenodo.15236575	Biodegradation robot details and processing software
https://github.com/olsenlabmit/Terpolymer_Modeling	Code used for machine learning of biodegradability, trained biodegradability models

Acknowledgements: This work was supported by Abdul Latif Jameel Water and Food Systems (J-WAFS) and the MIT Climate and Sustainability Consortium as well as the Gordon and Betty

Moore Foundation. K.A.F. was supported in part by the National Science Foundation Graduate Research Fellowship under Grant No. 2141064. We would like to thank the MIT Department of Chemistry Instrumentation Facility (DCIF) for use of their NMR Instruments and Dr. Walter Massefski for his assistance in obtaining NMR data.

Competing Interests: The authors have no competing interests to report.

Materials & Correspondence: Correspondence to Bradley D. Olsen at bdolsen@mit.edu

Author Contributions: B.D.O conceptualized the project and acquired funding. D.J.A assisted in polymer embedding conceptualization. K.A.F designed the material scope, designed and constructed synthesis equipment, carried out material synthesis, testing, and characterization, and performed data analysis and machine learning. J.C. carried out material synthesis, testing, and characterization. J.S. developed polymer similarity embedding. N.D.M. designed synthesis equipment, assisted in synthesis equipment construction and in material testing, and contributed to data interpretation discussions. G.G.C. carried out material testing and contributed to data interpretation discussions. J.L. assisted in polymer similarity embedding development. A.M.Z. assisted in material testing and carried out toxicity testing. J.S.M. assisted in material characterization.

References:

1. Satti, S. M.; Shah, A. A., Polyester-based biodegradable plastics: an approach towards sustainable development. *Lett Appl Microbiol* **2020,** *70* (6), 413-430.
2. Agarwal, S., Biodegradable Polymers: Present Opportunities and Challenges in Providing a Microplastic-Free Environment. *Macromolecular Chemistry and Physics* **2020,** *221* (6), 2000017.
3. Wu, F.; Misra, M.; Mohanty, A. K., Challenges and new opportunities on barrier performance of biodegradable polymers for sustainable packaging. *Progress in Polymer Science* **2021,** *117*, 101395.
4. Haider, T. P.; Völker, C.; Kramm, J.; Landfester, K.; Wurm, F. R., Plastics of the Future? The Impact of Biodegradable Polymers on the Environment and on Society. *Angewandte Chemie International Edition* **2019,** *58* (1), 50-62.
5. Aarsen, C. V.; Liguori, A.; Mattsson, R.; Sipponen, M. H.; Hakkarainen, M., Designed to Degrade: Tailoring Polyesters for Circularity. *Chemical Reviews* **2024,** *124* (13), 8473-8515.
6. Woodard, L. N.; Grunlan, M. A., Hydrolytic Degradation and Erosion of Polyester Biomaterials. *ACS Macro Letters* **2018,** *7* (8), 976-982.
7. Schöber, U.; Thiel, C.; Jendrossek, D., Poly(3-Hydroxyvalerate) Depolymerase ofPseudomonas lemoignei. *Applied and Environmental Microbiology* **2000,** *66* (4), 1385-1392.
8. Lenz, R. W.; Marchessault, R. H., Bacterial Polyesters: Biosynthesis, Biodegradable Plastics and Biotechnology. *Biomacromolecules* **2005,** *6* (1), 1-8.
9. Witt, U.; Müller, R.-J.; Augusta, J.; Widdecke, H.; Deckwer, W.-D., Synthesis, properties and biodegradability of polyesters based on 1,3-propanediol. *Macromolecular Chemistry and Physics* **1994,** *195* (2), 793-802.
10. Paek, K. H.; Im, S. G., Biodegradable Aromatic–Aliphatic Copolyesters Derived from Bis(2-Hydroxyethyl) Terephthalate for Sustainable Flexible Packaging Applications. *ACS Applied Polymer Materials* **2022,** *4* (8), 5298-5307.
11. Fernández, J.; Etxeberria, A.; Sarasua, J.-R., Synthesis and properties of ω-pentadecalactone-co-δ-hexalactone copolymers: a biodegradable thermoplastic elastomer as an alternative to poly(ε-caprolactone). *RSC Advances* **2016,** *6* (4), 3137-3149.
12. Fu, Y.; Wu, G.; Bian, X.; Zeng, J.; Weng, Y., Biodegradation behavior of poly (butylene adipate-co-terephthalate)(PBAT), poly (lactic acid)(PLA), and their blend in freshwater with sediment. *Molecules* **2020,** *25* (17), 3946.
13. Larrañaga, A.; Lizundia, E., A review on the thermomechanical properties and biodegradation behaviour of polyesters. *European Polymer Journal* **2019,** *121*, 109296.
14. Sousa, A. F.; Patrício, R.; Terzopoulou, Z.; Bikiaris, D. N.; Stern, T.; Wenger, J.; Loos, K.; Lotti, N.; Siracusa, V.; Szymczyk, A.; Paszkiewicz, S.; Triantafyllidis, K. S.; Zamboulis, A.; Nikolic, M. S.; Spasojevic, P.; Thiyagarajan, S.; van Es, D. S.; Guigo, N., Recommendations for replacing PET on packaging, fiber, and film materials with biobased counterparts. *Green Chemistry* **2021,** *23* (22), 8795-8820.
15. Singhvi, M. S.; Zinjarde, S. S.; Gokhale, D. V., Polylactic acid: synthesis and biomedical applications. *Journal of Applied Microbiology* **2019,** *127* (6), 1612-1626.
16. Ulery, B. D.; Nair, L. S.; Laurencin, C. T., Biomedical applications of biodegradable polymers. *Journal of Polymer Science Part B: Polymer Physics* **2011,** *49* (12), 832-864.
17. Haque, F. M.; Ishibashi, J. S. A.; Lidston, C. A. L.; Shao, H.; Bates, F. S.; Chang, A. B.; Coates, G. W.; Cramer, C. J.; Dauenhauer, P. J.; Dichtel, W. R.; Ellison, C. J.; Gormong, E. A.; Hamachi, L. S.; Hoye, T. R.; Jin, M.; Kalow, J. A.; Kim, H. J.; Kumar, G.; LaSalle, C. J.; Liffland, S.; Lipinski, B. M.; Pang, Y.; Parveen, R.; Peng, X.; Popowski, Y.; Prebihalo, E. A.; Reddi, Y.; Reineke, T. M.; Sheppard, D. T.; Swartz, J. L.; Tolman, W. B.; Vlaisavljevich, B.; Wissinger, J.;

Xu, S.; Hillmyer, M. A., Defining the Macromolecules of Tomorrow through Synergistic Sustainable Polymer Research. *Chemical Reviews* **2022,** *122* (6), 6322-6373.
18. Koelmans, A. A.; Redondo-Hasselerharm, P. E.; Nor, N. H. M.; de Ruijter, V. N.; Mintenig, S. M.; Kooi, M., Risk assessment of microplastic particles. *Nature Reviews Materials* **2022,** *7* (2), 138-152.
19. da Costa, J. P.; Santos, P. S. M.; Duarte, A. C.; Rocha-Santos, T., (Nano)plastics in the environment – Sources, fates and effects. *Science of The Total Environment* **2016,** *566-567*, 15-26.
20. Campo, E. A., *Selection of Polymeric Materials*. William Andrew Publishing: 2008.
21. van Krevelen, D. W., *Properties of Polymers*. Elsevier Science: 2012.
22. Hassan, M. H.; Omar, A. M.; Daskalakis, E.; Liu, F.; Bartolo, P. In *Preliminary studies on the suitability of PETG for 4D printing applications*, MATEC Web of Conferences, EDP Sciences: 2020; p 01010.
23. Balou, S.; Ahmed, I.; Priye, A., From Waste to Filament: Development of Biomass-Derived Activated Carbon-Reinforced PETG Composites for Sustainable 3D Printing. *ACS Sustainable Chemistry & Engineering* **2023,** *11* (34), 12667-12676.
24. Altay, E.; Jang, Y.-J.; Kua, X. Q.; Hillmyer, M. A., Synthesis, Microstructure, and Properties of High-Molar-Mass Polyglycolide Copolymers with Isolated Methyl Defects. *Biomacromolecules* **2021,** *22* (6), 2532-2543.
25. Jem, K. J.; Tan, B., The development and challenges of poly (lactic acid) and poly (glycolic acid). *Advanced Industrial and Engineering Polymer Research* **2020,** *3* (2), 60-70.
26. Kanwal, A.; Zhang, M.; Sharaf, F.; Li, C., Polymer pollution and its solutions with special emphasis on Poly (butylene adipate terephthalate (PBAT)). *Polymer Bulletin* **2022,** *79* (11), 9303-9330.
27. Burford, T.; Rieg, W.; Madbouly, S., Biodegradable poly(butylene adipate-co-terephthalate) (PBAT). *Physical Sciences Reviews* **2023,** *8* (8), 1127-1156.
28. ASTM, Standard Test Method for Determining Anaerobic Biodegradation of Plastic Materials Under High-Solids Anaerobic-Digestion Conditions. ASTM International: West Conshohocken, PA, 2018; Vol. D5511-18.
29. ASTM, Standard Test Method for Determining Aerobic Biodegradation of Plastic Materials Under Controlled Composting Conditions, Incorporating Thermophilic Temperatures. ASTM International: West Conshohocken, PA, 2021; Vol. D5338-15(2021).
30. ASTM, Standard Test Method for Determining Aerobic Biodegradation of Plastic Materials in the Marine Environment by a Defined Microbial Consortium or Natural Sea Water Inoculum. ASTM International: West Conshohocken, PA, 2018; Vol. D6691-17.
31. ASTM, Standard Test Method for Determining Aerobic Biodegradation of Plastic Materials in Soil. ASTM International: West Conshohocken, PA, 2018; Vol. D5988-18.
32. ASTM, Standard Guide for Exposing and Testing Plastics that Degrade in the Environment by a Combination of Oxidation and Biodegradation. ASTM International: West Conshohocken, PA, 2018; Vol. D6954-18.
33. ASTM, Standard Specification for Labeling of Plastics Designed to be Aerobically Composted in Municipal or Industrial Facilities. ASTM International: West Conshohocken, PA, 2022; Vol. D6400-22.
34. Goel, V.; Luthra, P.; Kapur, G. S.; Ramakumar, S. S. V., Biodegradable/Bio-plastics: Myths and Realities. *Journal of Polymers and the Environment* **2021,** *29* (10), 3079-3104.
35. Albright, V. C.; Chai, Y., Knowledge Gaps in Polymer Biodegradation Research. *Environmental Science & Technology* **2021,** *55* (17), 11476-11488.
36. Jandas, P. J.; Prabakaran, K.; Mohanty, S.; Nayak, S. K., Evaluation of biodegradability of disposable product prepared from poly (lactic acid) under accelerated conditions. *Polymer Degradation and Stability* **2019,** *164*, 46-54.

37. da Silva, S. A.; Hinkel, E. W.; Lisboa, T. C.; Selistre, V. V.; da Silva, A. J.; da Silva, L. O. F.; Faccin, D. J. L.; Cardozo, N. S. M., A biostimulation-based accelerated method for evaluating the biodegradability of polymers. *Polymer Testing* **2020,** *91*, 106732.
38. Šerá, J.; Serbruyns, L.; De Wilde, B.; Koutný, M., Accelerated biodegradation testing of slowly degradable polyesters in soil. *Polymer Degradation and Stability* **2020,** *171*, 109031.
39. Augusta, J.; Müller, R.-J.; Widdecke, H., A rapid evaluation plate-test for the biodegradability of plastics. *Applied Microbiology and Biotechnology* **1993,** *39* (4), 673-678.
40. Kanesawa, Y.; Tanahashi, N.; Doi, Y.; Saito, T., Enzymatic degradation of microbial poly(3-hydroxyalkanoates). *Polymer Degradation and Stability* **1994,** *45* (2), 179-185.
41. Reinhardt, S.; Handrick, R.; Jendrossek, D., The "PHB Depolymerase Inhibitor" of Paucimonas lemoignei Is a PHB Depolymerase. *Biomacromolecules* **2002,** *3* (4), 823-827.
42. Fransen, K. A.; Av-Ron, S. H. M.; Buchanan, T. R.; Walsh, D. J.; Rota, D. T.; Van Note, L.; Olsen, B. D., High-throughput experimentation for discovery of biodegradable polyesters. *Proceedings of the National Academy of Sciences* **2023,** *120* (23), e2220021120.
43. Schustik, S. A.; Cravero, F.; Ponzoni, I.; Díaz, M. F., Polymer informatics: Expert-in-the-loop in QSPR modeling of refractive index. *Computational Materials Science* **2021,** *194*, 110460.
44. Hasnaoui, H.; Krea, M.; Roizard, D., Neural networks for the prediction of polymer permeability to gases. *Journal of Membrane Science* **2017,** *541*, 541-549.
45. Barnett, J. W.; Bilchak, C. R.; Wang, Y.; Benicewicz, B. C.; Murdock, L. A.; Bereau, T.; Kumar, S. K., Designing exceptional gas-separation polymer membranes using machine learning. *Science advances* **2020,** *6* (20), eaaz4301.
46. Tao, L.; Varshney, V.; Li, Y., Benchmarking Machine Learning Models for Polymer Informatics: An Example of Glass Transition Temperature. *Journal of Chemical Information and Modeling* **2021**.
47. Martin, T. B.; Audus, D. J., Emerging Trends in Machine Learning: A Polymer Perspective. *ACS Polymers Au* **2023,** *3* (3), 239-258.
48. Ma, R.; Liu, Z.; Zhang, Q.; Liu, Z.; Luo, T., Evaluating Polymer Representations via Quantifying Structure–Property Relationships. *Journal of Chemical Information and Modeling* **2019,** *59* (7), 3110-3119.
49. Lin, T.-S.; Coley, C. W.; Mochigase, H.; Beech, H. K.; Wang, W.; Wang, Z.; Woods, E.; Craig, S. L.; Johnson, J. A.; Kalow, J. A.; Jensen, K. F.; Olsen, B. D., BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules. *ACS Central Science* **2019,** *5* (9), 1523-1531.
50. Lin, T.-S.; Rebello, N. J.; Lee, G.-H.; Morris, M. A.; Olsen, B. D., Canonicalizing BigSMILES for Polymers with Defined Backbones. *ACS Polymers Au* **2022,** *2* (6), 486-500.
51. Kuenneth, C.; Schertzer, W.; Ramprasad, R., Copolymer Informatics with Multitask Deep Neural Networks. *Macromolecules* **2021,** *54* (13), 5957-5961.
52. Patel, R. A.; Borca, C. H.; Webb, M. A., Featurization strategies for polymer sequence or composition design by machine learning. *Molecular Systems Design & Engineering* **2022,** *7* (6), 661-676.
53. Arora, A.; Lin, T.-S.; Rebello, N. J.; Av-Ron, S. H. M.; Mochigase, H.; Olsen, B. D., Random Forest Predictor for Diblock Copolymer Phase Behavior. *ACS Macro Letters* **2021,** *10* (11), 1339-1345.
54. Yu, M.; Shi, Y.; Jia, Q.; Wang, Q.; Luo, Z.-H.; Yan, F.; Zhou, Y.-N., Ring Repeating Unit: An Upgraded Structure Representation of Linear Condensation Polymers for Property Prediction. *Journal of Chemical Information and Modeling* **2023,** *63* (4), 1177-1187.
55. Fransen, N. S. G.-C., Gabrielle F.; Fransen, Katharina A.; Mamrol, Natalie D., High-Throughput Biodegradation Test Gantry Robot. Zenodo, 2025.
56. Av-Ron, S. H. M. *High-Throughput Clear-Zone*, Zenodo: 2022.

57. Weininger, D., SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **1988,** *28* (1), 31-36.
58. Blay, V.; Radivojevic, T.; Allen, J. E.; Hudson, C. M.; Garcia Martin, H., MACAW: An Accessible Tool for Molecular Embedding and Inverse Molecular Design. *Journal of Chemical Information and Modeling* **2022,** *62* (15), 3551-3564.
59. Shi, J.; Walsh, D.; Zou, W.; Rebello, N. J.; Deagen, M. E.; Fransen, K. A.; Gao, X.; Olsen, B. D.; Audus, D. J., Calculating Pairwise Similarity of Polymer Ensembles via Earth Mover's Distance. *ACS Polymers Au* **2024,** *4* (1), 66-76.
60. RDKit: Open-source cheminformatics. https://www.rdkit.org (accessed 08/28/2024).
61. Rogers, D. J.; Tanimoto, T. T., A Computer Program for Classifying Plants. *Science* **1960,** *132* (3434), 1115-1118.
62. Flamary, R.; Courty, N.; Gramfort, A.; Alaya, M. Z.; Boisbunon, A.; Chambon, S.; Chapel, L.; Corenflos, A.; Fatras, K.; Fournier, N.; Gautheron, L.; Gayraud, N. T. H.; Janati, H.; Rakotomamonjy, A.; Redko, I.; Rolet, A.; Schutz, A.; Seguy, V.; Sutherland, D. J.; Tavenard, R.; Tong, A.; Vayer, T., POT: Python optimal transport. *J. Mach. Learn. Res.* **2021,** *22* (1), Article 78.
63. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É., Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011,** *12* (null), 2825–2830.
64. Liu, S.; Xiao, H.; Huang, K.; Lu, L.; Huang, Q., Terpolymerization of Carbon Dioxide with Propylene Oxide and ε-Caprolactone: Synthesis, Characterization and Biodegradability. *Polymer Bulletin* **2006,** *56* (1), 53-62.
65. Jung, Y.; Lee, S.-H.; Kim, S.-H.; Lim, J. C.; Kim, S. H., Synthesis and characterization of the biodegradable and elastic terpolymer poly(glycolide-co-L-lactide-co-ϵ-caprolactone) for mechano-active tissue engineering. *Journal of Biomaterials Science, Polymer Edition* **2013,** *24* (4), 386-397.
66. Mathieu-Denoncourt, J.; Wallace, S. J.; de Solla, S. R.; Langlois, V. S., Plasticizer endocrine disruption: Highlighting developmental and reproductive effects in mammals and non-mammalian aquatic species. *General and Comparative Endocrinology* **2015,** *219*, 74-88.
67. Djapovic, M.; Milivojevic, D.; Ilic-Tomic, T.; Lješević, M.; Nikolaivits, E.; Topakas, E.; Maslak, V.; Nikodinovic-Runic, J., Synthesis and characterization of polyethylene terephthalate (PET) precursors and potential degradation products: Toxicity study and application in discovery of novel PETases. *Chemosphere* **2021,** *275*, 130005.
68. Zhao, J.; Chen, X.; Lin, F.; Yang, N.; Huang, H.; Zhao, J., Mechanism of toxicity formation and spatial distribution in activated sludge treating synthetic effluent containing bisphenol A (BPA). *Chemical Engineering Journal* **2014,** *250*, 91-98.
69. Faergemann, J.; Hedner, T.; Larsson, P., The in vitro activity of pentane-1,5-diol against aerobic bacteria. A new antimicrobial agent for topical usage? *Acta Dermato-Venereologica* **2005,** *85* (3), 203-205.