Machine-Assisted Discovery of Chondroitinase ABC Complexes toward Sustained Neural Regeneration

Shashank Kosuri, Carlos H. Borca, Heloise Mugnier, Matthew Tamasi, Roshan A. Patel, Isabel Perez, Suneel Kumar, Zachary Finkel, Rene Schloss, Li Cai, Martin L. Yarmush, Michael A. Webb,* and Adam J. Gormley*

Among the many molecules that contribute to glial scarring, chondroitin sulfate proteoglycans (CSPGs) are known to be potent inhibitors of neuronal regeneration. Chondroitinase ABC (ChABC), a bacterial lyase, degrades the glycosaminoglycan (GAG) side chains of CSPGs and promotes tissue regeneration. However, ChABC is thermally unstable and loses all activity within a few hours at 37 °C under dilute conditions. To overcome this limitation, the discovery of a diverse set of tailor-made random copolymers that complex and stabilize ChABC at physiological temperature is reported. The copolymer designs, which are based on chain length and composition of the copolymers, are identified using an active machine learning paradigm, which involves iterative copolymer synthesis, testing for ChABC thermostability upon copolymer complexation, Gaussian process regression modeling, and Bayesian optimization. Copolymers are synthesized by automated PET-RAFT and thermostability of ChABC is assessed by retained enzyme activity (REA) after 24 h at 37 °C. Significant improvements in REA in three iterations of active learning are demonstrated while identifying exceptionally high-performing copolymers. Most remarkably, one designed copolymer promotes residual ChABC activity near 30%, even after one week and notably outperforms other common stabilization methods for ChABC. Together, these results highlight a promising pathway toward sustained tissue regeneration.

1. Introduction

Central Nervous System (CNS) injuries have devastating, long-term physical, psychological, and socio-economic consequences for patients and families. Although mortality rates have substantially improved in patients with CNS injuries, improvement in functional outcomes remains elusive. When it comes to functional recovery, many studies point to a general common theme: CNS neurons attempt to regenerate after a traumatic injury, but the post-injury environment is highly inhibitory and results in abortive regeneration.[1] This is mainly due to the complex pathophysiology of the CNS, which undergoes enormous biochemical and physical changes post-injury. Initial trauma results in extensive tissue damage, compromise of the blood-brain/spinal cord vasculature, and necrotic cell death. Compromised vasculature allows an influx of inflammatory cells that begin secreting pro-inflammatory cytokines and vasoactive peptides that potentiate further damage resulting in edema, excitotoxicity, altered gene expression, and enhanced cell signaling. Reactive astrocytes surround the site of injury and secrete a wide range of pro-inflammatory factors including chondroitin sulfate proteoglycans (CSPGs) resulting in a dense network of scar tissue that acts as a mechanical and chemical barrier to tissue regeneration.[2] Although glial scar and CSPGs within the scar play an important role in isolating injury site from healthy tissue, thereby minimizing uncontrolled tissue damage, their highly inhibitory nature sacrifices long-distance functional regeneration. CSPGs are known to be potent inhibitors of neurite outgrowth, and their inhibitory nature has been well documented in vitro and in vivo.[3] CSPG family members share two common features: a protein core that varies in structure, and glycosaminoglycan (GAG) side chains that vary in number and sulfation pattern.[4] Extensive research has shown that high levels of CSPGs in CNS injury models in vivo correlates well with presence of dystrophic growth cones that are hallmarks of failed attempts at neuronal regeneration.[5] The growth suppressive nature of CSPGs is mainly attributed to their GAG side chains and sulfation pattern and degrading these inhibitory molecules has become a viable therapeutic strategy for promoting tissue regeneration.[3a]

S. Kosuri, H. Mugnier, M. Tamasi, I. Perez, S. Kumar, Z. Finkel, R. Schloss, L. Cai, M. L. Yarmush, A. J. Gormley
Biomedical Engineering
Rutgers, The State University of New Jersey
Piscataway, NJ 08854, USA
E-mail: adam.gormley@rutgers.edu
C. H. Borca, R. A. Patel, M. A. Webb
Chemical and Biological Engineering
Princeton University
Princeton, NJ 08544, USA
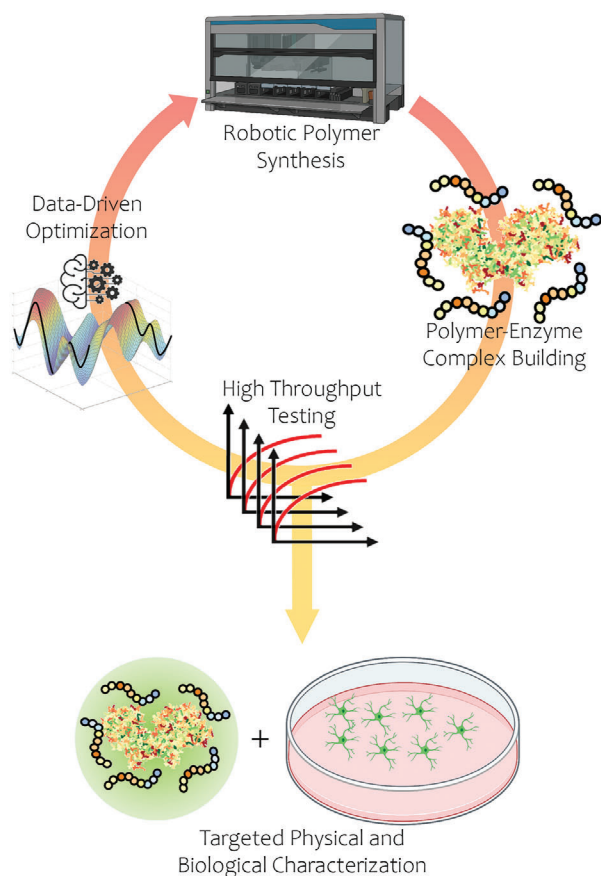E-mail: mawebb@princeton.edu

Figure 1. Workflow schematic. Automated polymer synthesis was conducted using a Hamilton Microlab Starlet. Initially, a library with over 500 unique copolymers were quickly evaluated for their ability to retain ChABC activity at 37 °C for 24 h. Quantified retained enzyme activity (REA) information along with polymer composition and chain length were used to train a Gaussian Process Regression model (GPR) that predicts REA based on copolymer chemistry and chain length. Using active learning, 24 new polymers were proposed based on GPR modeling, and these polymers were subsequently synthesized and tested to provide additional data for GPR. After three of these active learning iterations, a few of the most promising designs were selected for further characterization.

Chondroitinase ABC (ChABC), a 115 kDa bacterial lyase that degrades the GAG side chains of CSPGs, has proven highly effective in promoting axonal sprouting and neuronal regeneration in various animal models.[6] However, its use as a therapeutic is limited by its thermal instability as it completely loses activity within a few hours at 37 °C under dilute conditions.[7] To maintain therapeutic efficacy, continuous intrathecal administration every two weeks for up to six weeks is required. Therefore, there is an immediate need to develop highly stable, nanodispersed ChABC for glial scar degradation. Several approaches have been utilized for stabilizing ChABC, such as stabilization in high-concentration sugar solutions (1 M trehalose, 2.5 M sucrose), immobilization onto porous scaffolds, and protein structure modification.[8]

Herein, we report the discovery of new polymer-enzyme complexes (PECs) of thermostabilized ChABC enabled by robotic synthesis, high-throughput testing, and data-driven optimization (**Figure 1**). PECs contain enzyme wrapped inside a syn-

thetic chaperone-like polymeric shell that safeguards it from surrounding harsh microenvironments.[9] However, rational identification of complementary copolymer composition on a case-by-case basis is highly time consuming and labor intensive. Meanwhile, data-driven design of copolymers has been pursued in silico.[10] Recent advances in the field of oxygen tolerant polymer chemistries now enable synthesis of complex polymers in bench top well plates. To increase throughput, the Gormley lab has programmed liquid-handling robotics to perform autonomous polymer chemistry and post-polymerization functionalization in well plates, which increases synthetic efficiency while maintaining excellent reproducibility.[11] This allows for automated and facile generation of datasets containing polymer chemistry and retained enzyme activity (REA) that can be effectively used to train machine learning (ML) models to represent the complex structure-activity relationship between polymer-enzyme interactions and REA. Using active learning—a data acquisition paradigm wherein new data points for inclusion in ML model training are selected by interactive query—three iterations of 24 copolymer candidates were proposed for experimental synthesis and testing. From these 72 new copolymers, the most promising were selected to conduct preliminary long-term studies, and the best of these was selected for further biological and biophysical characterization.

2. Results and Discussion

2.1. Data-Driven Design of PECs

We used active machine learning to identify copolymer designs with high likelihood to stabilize ChABC following thermal exposure based on iterative synthesis, testing, modeling and optimization (Figure 1, top); all copolymers were restricted to have four or fewer distinct monomers and chain lengths with degree of polymerization (DP) between 50 and 200 in increments of 25. For this study, we selected eight methacrylate or methacrylamide-based monomers based on their compatibility with PET-RAFT polymerization, their good solubility in DMSO, and their chemical diversity. In particular, these monomers are subdivided into three groups: hydrophobic, hydrophilic and ionic based on their logP and pKa values. Hydrophilic monomers were selected to improve polymer solubility in buffers and also to interact with the enzyme through hydrogen bonding. Hydrophobic monomers were selected to interact with the non-polar residues present on the protein surface while ionic monomers were selected to interact with charged amino acid residues. The active learning process was seeded by a dataset of 504 copolymers with systematic variation in monomer composition and chain length (see Seed Library Design) and accompanying data on their ability to stabilize ChABC, as quantified by REA at 37 °C for 24 h.

Figure 2A shows the distribution of REAs of all samples tested, including the seed database and each set of candidates produced by the active learning scheme. The three iterations (1–3) of active learning resulted in remarkable improvement of REA relative to the seed data, with many designs yielding *enhanced* activity above 100%. This is reflected by an overall upward trend of the distribution of data in terms of quartile positions, median, mean, and maximum. In the original seed dataset, we observed an average REA of 27.1% for 504 samples with one sample exhibiting 125%
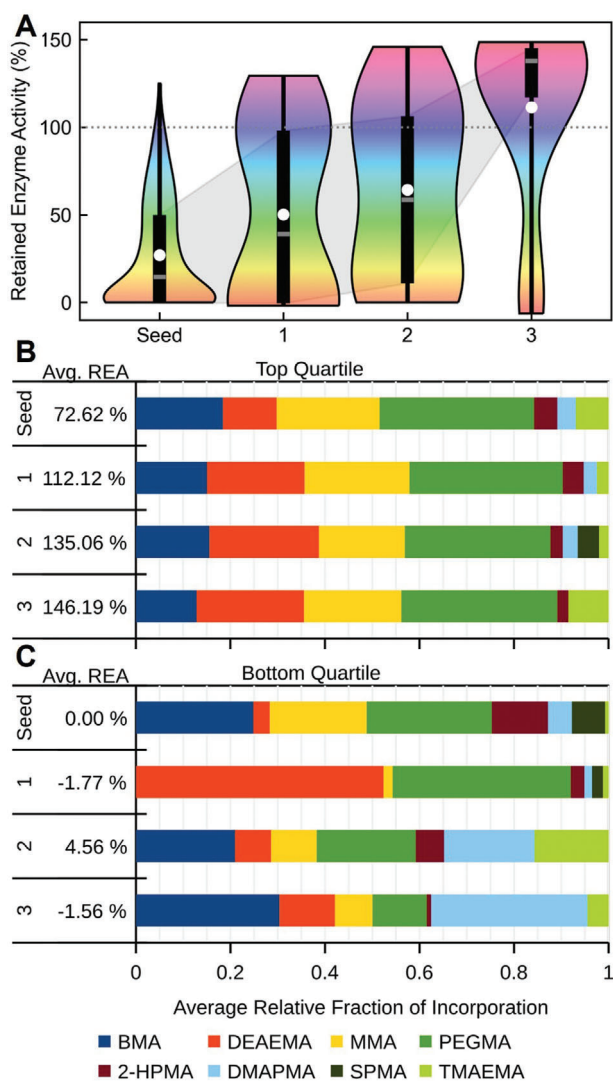
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
HEALTHCARE
MATERIALS**

www.advhealthmat.de

Figure 2. Generational improvement for PECs. A) Comparison of distributions of REA for ChABC-PECs for the seed database and three iterations of active learning. The distributions are represented as violins colored by REA as well as black candlesticks delimiting the second and third quartiles of the data, meaning that 50% of all data points in each iteration is included in the candle. The gray dash indicates the median REA, and the white circle indicates the average. The transparent gray shading is a guide to the eye to follow changes in summary statistics. B) Comparison of average compositions of copolymers in the top quartile for the seed dataset and polymers proposed in the three iterations of active learning. C) Comparison of average compositions of copolymers in the top quartile for the seed dataset and polymers proposed in the three iterations of active learning. In (B) and (C), the width of horizontal bars indicate the fraction of incorporation for each monomer; bars are organized in decreasing order of octanol–water partition coefficient, such that more hydrophobic monomers are at the left and more hydrophilic monomers are at the right.

of the native enzyme activity at the end of 24 h. The average REA of samples across different iterations improved with global average REAs of 50.3%, 64.2%, and 111.4% for iterations 1, 2, and 3, respectively. The highest performing samples for iterations 1, 2, and 3 exhibited REAs of 129.4%, 146%, and 148%, respectively. Details of the copolymers along with REA for iterations 1, 2, and

3 are given in the supporting information (Tables S1–S3, Supporting Information). Native enzyme activity at $t = 0$ served as positive control for all experiments and residual enzyme activity after heating at 37 °C for 24 h served as negative control. Because native enzyme retained some amount of residual activity (<1–2%) depending on the batch used, some of the PECs destabilized the enzyme resulting in negative REA values (Tables S1–S3, Supporting Information).

Figure 2B presents an aggregated analysis of the average REA and average polymer composition for ChABC-PECs in the top quartile for REA, and Figure 2C shows an identical analysis for those in the bottom quartile. Variations of average composition of the top quartiles of the different sets of polymers are small, especially when compared to those of the bottom quartiles. Nevertheless, the average REA increases upon successive iterations of active learning. This indicates a delicate balance of polymer chemistry is required to achieve high-performing polymers and highlights the underlying challenge of the polymer design task. Comparing the second and third iterations, the complete removal of DMAPMA and SPMA increases average REA by more than 10%. This is likely related to a notable imbalance between the fractions of more hydrophobic versus more hydrophilic monomers. In this scenario, the active learning process identifies that a preponderance of hydrophilic monomers is more likely associated with high REA. However, trying to rationalize design choices from active learning is not always straightforward because apparently similar compositions can yield drastically different results. By comparison, the bottom quartile ChABC-PECs exhibit substantial chemical diversity with no obvious trend. In general, this seems sensible as there are likely many more copolymers that do not effectively stabilize ChABC compared to those that do. We note that the degree of polymerization did not seem to influence the GPR model as much as the variations in polymeric composition, suggesting it is a relatively less important design feature for ChABC, at least over the range studied here. These observations highlight the importance of implementing an iterative active learning approach: it can capture fine details that could easily escape a rational design approach, and it has a higher probability of success compared to a random or systematic search.

2.2. Long-Term Stability Study in Artificial Cerebrospinal Fluid (aCSF)

To assess the long-term stability of designed ChABC-PECs, we selected the top five copolymers identified during active learning and performed an initial evaluation of their propensity to sustain ChABC activity over the course of five days (Figure S1, Supporting Information). The concentration of polymers used for this study was 12.5×10^{-6} M and the enzyme concentration was maintained at 1 ng μL^{-1} (8.8×10^{-9} M). Of these, we selected the PEC with the highest REA retention of ChABC after five days for additional long-term study and evaluation. The composition of the best performing copolymer was DEAMA (23.1 mol%): BMA (33.2 mol%): PEGMA (31.7 mol%): TMAEMA (12.0 mol%) and had a DP of 75; the molecular weight of the copolymer is 33 kDa and has a dispersity of $Đ = 1.5$.

Five copolymer concentrations were tested for the long-term stability assessment to study the effect of copolymer concentra-

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
HEALTHCARE
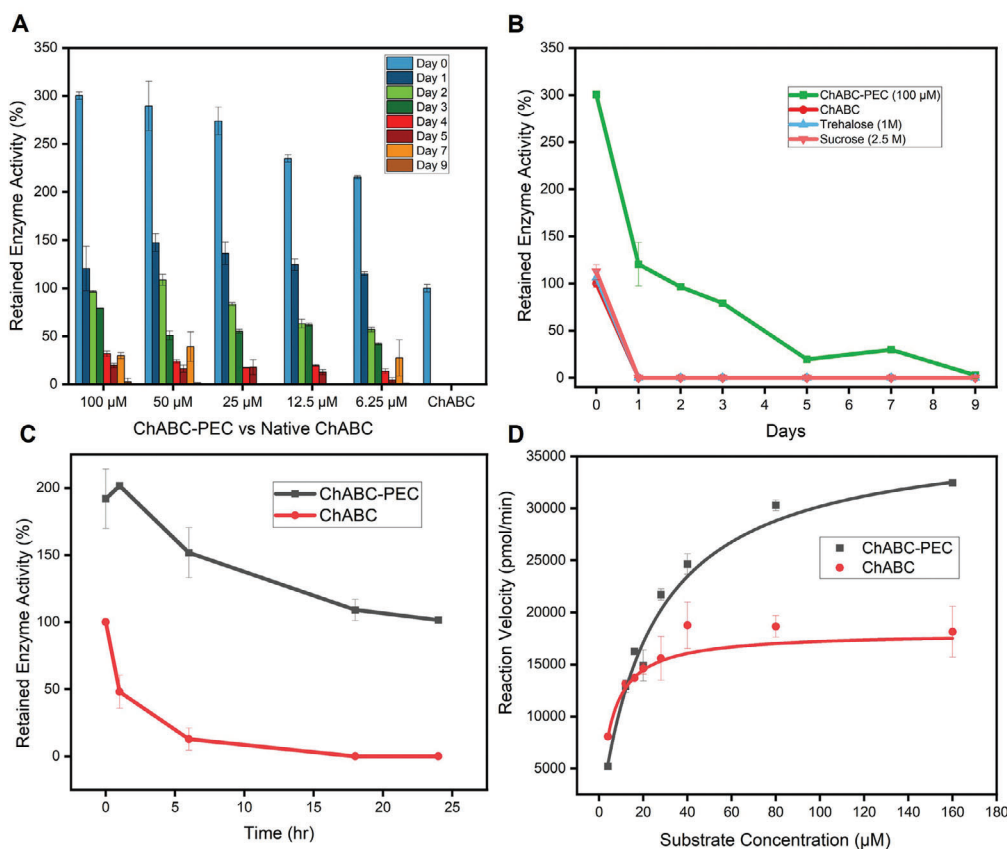MATERIALS**

www.advhealthmat.de

Figure 3. Retained enzyme activity (REA). A) REA of ChABC in the presence of polymer at different concentrations at 37 °C. PECs at different concentrations increased activity of enzyme at $t = 0$ at least 2–3 fold and maintained high levels of enzyme for the initial few days. B) Comparison of PEC with common enzyme stabilizers Trehalose and Sucrose. C) PEC retained >100% activity while native ChABC lost all activity within 24 h. D) Activity of native ChABC and ChABC-PEC (12.5×10^{-6} M) at varying substrate concentrations. Data represented here as mean $\pm$ SD, $n = 3$ for all experiments.

tion on enzyme activity retention. Enzyme concentration was maintained at 2 ng μL^{-1} (17.63×10^{-9} M) for long-term stability experiments and copolymer concentrations were varied from 6.25–100×10^{-6} M. Remarkably, copolymer at all concentrations increased the initial activity of the enzyme at $t = 0$. The initial activity of the enzyme increased by threefold for 100 and 50×10^{-6} M copolymer concentrations while 2–2.5-fold increases were observed for 25, 12.5, and 6.25×10^{-6} M concentrations (**Figure 3**A). At the end of 24 h, PECs at all five concentrations retained around 110–150% native enzyme activity within the same period. The native enzyme did not retain any activity at the end of 24 h. By contrast, ChABC-PECs continued to retain high levels of activity (>60%) at the end of 72 h. By the end of day 7, samples at 25 and 12.5×10^{-6} M lost all activity but 100, 50, and 6.25×10^{-6} M samples continued to have REAs of 29%, 39%, and 28%, respectively.

The efficiency of PECs to stabilize ChABC was compared with common enzyme stabilizers trehalose and sucrose at an enzyme concentration of 2 ng μL^{-1} (17.63×10^{-9} M). ChABC stabilized with 1 M trehalose and 2.5 M sucrose lost all enzyme activity within 24 h while PEC retained 100% activity within the same time period (Figure 3B,C). While trehalose and sucrose have been shown to stabilize ChABC in various studies, we suspect that the

inability of these sugars to stabilize in our study is due to the low concentration of enzyme that was used in this study.[8a,c]

We also investigated the kinetic parameters of ChABC in the presence of our copolymer and observed a significant increase in the maximum velocity (V_{max}), while the affinity to the substrate decreased, as can be observed from K_m (**Table 1**, Figure 3D). Deviation from Michaelis–Menten kinetics is commonly associated with substrate inhibition, which is mainly attributed to simultaneous binding of two or more substrate molecules to the active site of the enzyme. The fact that the activity of the enzyme in the presence of polymers is unaffected by substrate concentration suggests that the polymers prevent interactions between the substrate and the enzyme that may disrupt the active site. We therefore hypothesize that the lower activity of the native enzyme compared to ChABC-PEC could be because

Table 1. Kinetic parameters of ChABC with and without copolymer.

Sample	V_{max} [pmol μg^{-1} min^{-1}]	K_m [$\times 10^{-6}$ M]	K_{cat} [s^{-1}]
ChABC	18 070	5.01	284.7
ChABC-PEC	37 315	23.62	587.93

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
HEALTHCARE
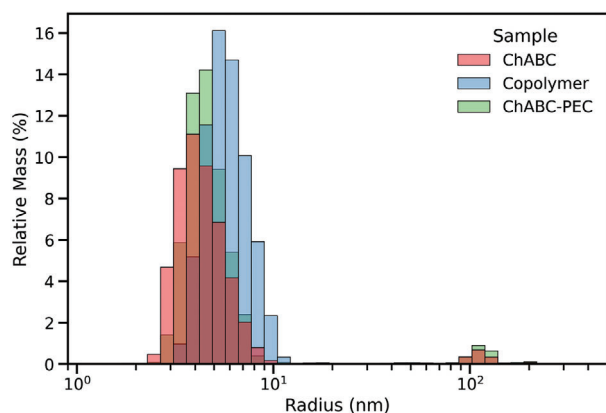MATERIALS**

www.advhealthmat.de

Figure 4. Dynamic light scattering. Biophysical characterization of ChABC, copolymer, and ChABC-PEC using DLS.

of substrate inhibition, which is prevented in the presence of the copolymer resulting in nanodispersed PEC (Figure 3D). Future experiments are needed to confirm this hypothesis.

2.3. Hydrodynamic Size of PECs Using Dynamic Light Scattering

To characterize the size of PECs in solution, we measured the hydrodynamic size of copolymer, native enzyme, and PEC using dynamic light scattering (DLS). The scattering data suggest that the copolymer is swollen in solution but complexes around nanodispersed ChABC without forming multi-molecular species (**Figure 4**).

2.4. Polymer Solutions Cytotoxicity and Inflammation on Astrocytes

To evaluate adverse cytotoxic side effects of copolymers, we performed testing on astrocyte cultures. Astrocyte cultures have been established as previously described.[12] Briefly, astrocytes seeded in 24 well plates were treated with different concentrations of copolymer on the day of the experiment. After 24 h, supernatants were collected, and cell viability was assessed using standard Live/Dead assay. No cytotoxicity was observed even at the highest concentration of 400×10^{-6} M (7.2 mg L^{-1}) (**Figure 5**A). A wide range of concentrations from 400 to 6.25×10^{-6} M were tested for possible cytotoxicity and no adverse effects were observed at all concentrations (data shown until 100×10^{-6} M).

We wanted to evaluate any inflammatory properties of our copolymers that could stimulate the secretion of Tumor Necrosis Factor-alpha (TNF-α) in astrocytes. Lipopolysaccharide (LPS) is a well-known inflammatory agent that upregulates the secretion of TNF-α in astrocytes and has been well characterized. Astrocytes treated with LPS had high levels of TNF-α secretion while astrocytes treated with copolymers had no significant differences compared to control group (no LPS) (Figure 5B). We also evaluated the biocompatibility of PECs on astrocytes, and no cytotoxic effects were observed at tested concentrations (Figure 5C). The polymer concentrations for the two groups were 12.5 and 6.25×10^{-6} M, respectively, while the enzyme concentration was maintained at 1 ng µL^{-1}.
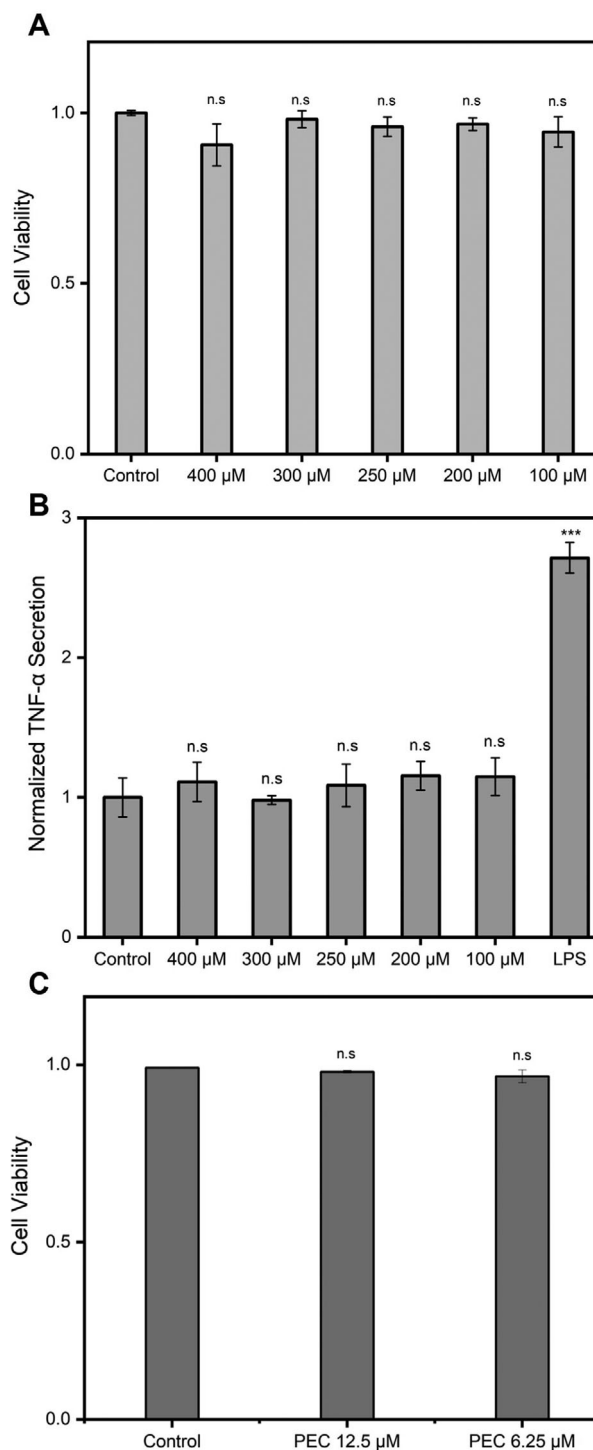

Figure 5. Toxicity and activity in vitro. A) No cytotoxity was observed when astrocytes were treated with heteropolymer at various concentrations. B) TNF-α secretion by astrocytes treated with heteropolymers. No inflammation was observed in the presence of our copolymer constructs compared to control group. C) No cytotoxicity was observed when PECs were treated with astrocytes at two different concentrations. Data presented as mean ± SD, $n = 6$–9, p-values calculated using paired comparison plot with Tukey analysis.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
HEALTHCARE
MATERIALS**

www.advhealthmat.de

3. Conclusions

ChABC suffers from poor thermal stability under physiological conditions, and this severely restricts its use as a viable therapeutic option for treating CNS injuries. While various techniques have been reported for ChABC stabilization, we report the novel use of copolymers for the very first time to stabilize ChABC in aCSF. In particular, by coupling automated polymer chemistry with active learning, we discovered several polymers that stabilize ChABC at physiological temperature in aCSF. In general, polymers identified by active learning exhibited significantly better REA for ChABC compared to a large, systematic screen, and designs tended to improve with the acquisition of more data. We further evaluated the long-term stability for one of the designs, which remarkably stabilized ChABC for over 7 days at a very low concentration. While polymer chemistry has been historically low throughput, recent advancements in the field of automated polymer synthesis using oxygen tolerant chemistries enable higher throughput screening of combinatorial polymer libraries with relative ease. While systematic exploration-based design yielded modest results, coupling active learning with high-throughput experimentation enabled the efficient identification of unique copolymers with propensity to stabilize ChABC. Importantly, the use of these designed copolymers for thermostabilization of ChABC was shown to be competitive with or superior to other existing stabilization strategies at very low concentrations of ChABC. These encouraging results motivate future preclinical studies to test the ability of these designs to facilitate glial scar degradation and promote neural tissue regeneration after injury.

4. Experimental Section

Materials: Monomers 2-Diethylamino ethyl methacrylate (DEAEMA), Hydroxypropyl methacrylate (2-HPMA), [2-(Methacryloyloxy)ethyl]trimethylammonium chloride solution (TMAEMA), *N*-[3-(Dimethylamino)propyl]methacrylamide (DMAPMA) were purchased from Sigma-Aldrich, Methyl methacrylate (MMA) and 3-Sulfopropyl methacrylate potassium salt (SPMA) from VWR, Butyl methacrylate (BMA) from Alfa Aesar and Poly(ethyleneglycol) (*n*) monomethyl ether monomethacrylate (PEGMA, $M_n \approx 400$) from Polysciences. Monomers were deinhibited prior to use by passing over mono-methyl ether hydroxyquinone (MEHQ) inhibitor removal resin. Chemicals Ethyl 2-(phenylcarbonothioylthio)-2-phenylacetate, Sodium Chloride, Magnesium Chloride, Calcium Chloride, Sodium Phosphate (mono and di basic), Lithium Bromide were purchased from Sigma-Aldrich, Zinc Tetraphenyl Porphyrin, Potassium Chloride, Dimethyl Sulfoxide (DMSO) from Fisher Scientific, Potassium phosphate (mono and dibasic) and Sodium Acetate anhydrous from VWR. Chondroitin sulfate sodium salt from bovine cartilage was purchased from Sigma-Aldrich. Recombinant *Proteus vulgaris* Chondroitinase ABC Protein (20 µg, 0.809 mg mL^{-1} stock concentration), was purchased from R&D Systems (6877-GH-020).

Automated PET-RAFT Synthesis: Polymer libraries were prepared as previously described.[11a] Briefly, Hamilton MLSTARlet sequences and processes were generated from Python with sample concentration, reagent volumes, and well position. Files containing reaction information were transferred to the Hamilton MLSTARlet to prime the robotic transfers. Stock solutions of monomer (2 M), ethyl 2-(phenylcarbonothioylthio)-2-phenylacetate (RAFT agent, 100 or 50 × 10^{-3} M) and ZnTPP (4 or 2 × 10^{-3} M) were prepared in DMSO and pipetted into 1 mL aliquots. Aliquots were loaded into a Hamilton MLSTARlet liquid handling robot and automatically pipetted into 96-well clear flat-bottom polypropylene

well plates (Greiner bio-one). Monomer/CTA ratio was varied from 100–400 while ZnTPP/CTA remained at 0.01. Polymer mixtures were dispensed to a total volume of 200 µL and a final monomer concentration of 1 M. They were then covered with a well-plate sealing tape and radiated under 560 nm LED light (5 mW cm^{-2}, TCP 12-Watt Yellow LED BR30 bulb) for 16 h.

Seed Library Design: Polymers that make up the seed library incorporated combinations of eight monomer chemistries. As previously described, these monomers were classified into three groups based on their logP and pKa values: hydrophobic, hydrophilic, and ionic. This allows for an extremely large number of physiochemically distinct copolymers. The initial seed library was therefore designed in a way such that each copolymer composition had a maximum of four distinct monomers. The compositions of the 504 polymers were varied systematically, and the fraction of incorporation of any one particular monomer was limited to be no more than 70% in a single polymer composition. This ensured that the copolymer compositions in the seed library were chemically diverse. Another factor that was considered was the solubility of copolymers in aqueous buffers. For example, a copolymer with 80% BMA, 10% MMA and 10% 2-HPMA will be insoluble in aqueous buffer solutions because of the extreme hydrophobic nature of the copolymer resulting in unreliable results. Therefore, the seed library only allowed up to 70% of hydrophobic content. The list of compositions constituting the seed library is provided in Table S5 (Supporting Information).

Machine Learning: The Python libraries Scikit-learn v0.24.1 and Hyperopt v0.2.5 were used to optimize Gaussian Process Regression (GPR) models to predict REA from a feature vector describing a copolymer.[13] The input feature vector for each copolymer was a size-explicit composition vector, which here is a nine-dimensional vector with the chain length divided by 200 (the maximum prescribed chain length) in the first dimension and the fractional incorporation of each of the eight possible monomers in the remaining dimensions.[14] Nested *k*-fold cross-validation was used to construct the GPR models prior to each iteration of proposed polymers. In particular, the dataset was first split into five outer folds. For each outer fold, a set of optimal hyperparameters (i.e., those present in the radial basis function kernel and white noise kernel) were obtained by 20-fold cross-validation over the remaining outer folds. This yielded five sets of hyperparameters, which were averaged to obtain a final set of hyperparameters. Using these hyperparameters, a final GPR model was trained with all experimental data acquired to that point. In all cases, models were optimized by minimizing the average mean squared error of a power-transform of the REA.

The final GPR model was used in tandem with a Bayesian optimization scheme to identify 200 new copolymers that maximize a modified expected improvement (EI) acquisition function of the form:

$$EI\ (\mathbf{x}) = (\mu_\mathbf{x} - \tau_* - \xi)\,\Phi\left(\frac{\mu_\mathbf{x} - \tau_* - \xi}{\sigma_\mathbf{x}}\right) + \sigma_\mathbf{x}\phi\left(\frac{\mu_\mathbf{x} - \tau_* - \xi}{\sigma_\mathbf{x}}\right) + \pi\ (\mathbf{x})$$

$$(1)$$

In Equation (1), $\mathbf{x}$ denotes the feature vector for the copolymer, $\mu_\mathbf{x}$ and $\sigma_\mathbf{x}$ are the corresponding predicted mean and standard deviation from the GPR model, τ_* indicates the maximum predicted mean from the GPR model amongst proposed or synthesized copolymers. In addition, $\Phi(z)$ and $\varphi(z)$ respectively denote the standard normal cumulative and probability density distribution functions associated with a random variable z, and ξ is a hyperparameter that controls the balance between exploration and exploitation. Finally, $\pi(x)$ is a penalty function associated with design x, which is computed as

$$\pi\ (x) = 100H\left(d - 0.05\sqrt{2}\right)\left[1 - \tanh\left(\frac{-d}{1 - 0.05\sqrt{2}}\right)\right] \qquad (2)$$

where $H(\cdot)$ denotes the Heaviside function and $d = \min_{x_j \in A} |x - x_j|$ yields the minimum Euclidean distance between the polymer x and all previously tested or proposed polymers, forming the set A. Each of the 200 polymers

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
HEALTHCARE
MATERIALS**

www.advhealthmat.de

produced optimize Equation (1) with a different value of ξ. In particular, 200 ξ values that are base-10 logarithmically spaced, starting at 0.001 and ending at 30 are considered.

To further promote diversity in the polymers proposed for synthesis, the initial 200 candidates were downselected to 24 polymers using an unsupervised learning approach based on the DBSCAN and k-Means clustering algorithms.[13b,15] For DBSCAN, the maximum allowable distance for two samples to be in the same neighborhood was set as $0.05\sqrt{2}$, and the minimum number of samples to be considered a cluster was three. In prior discussion, the appearance of the quantity $0.05\sqrt{2}$ originates from the notion that polymers that differ by 0.05 in terms of their fraction of incorporation for one component are reliably distinct in the automated synthesis. Representative samples from the clusters were subsequently identified based on their proximity to the cluster centroid. These representative samples, as well as any noise samples, were further considered as potential candidates. In the event that DBSCAN produced greater than 24 candidate polymers, these were then downselected using k-Means clustering with $k = 24$. As before, representative samples were based on proximity to the cluster centroid. The net result of this filtering was identification of 24 chemically distinct candidate polymers. Following experimental synthesis and evaluation of these 24 proposed polymers, the newly acquired data would be added to the dataset, and the process, beginning with the training of a new GPR model was repeated. Three iterations of active learning with this protocol were conducted. The first iteration of active learning used a GPR model trained on the seed library.

Thermal Stability Assay: Activity of PECs was evaluated by their ability to digest chondroitin sulfate substrate resulting in unsaturated disaccharides. Briefly, polymers were synthesized and diluted in DMSO before further dilution into assay buffer (aCSF: 149×10^{-3} M NaCl, 3×10^{-3} M KCl, 0.8×10^{-3} M MgCl$_2$, 1.4×10^{-3} M CaCl$_2$ 1.5×10^{-3} M Na$_2$HPO$_4$, 0.2×10^{-3} M NaH$_2$PO$_4$, pH 7.4) to a final concentration of 25×10^{-6} M (<1% DMSO). Copolymers that either gelled in DMSO due to high hydrophobicity or precipitated in aCSF buffer were excluded from further experiments. While some gelled polymers exhibited enzyme protection capability after manually redissolving the gel in a more suitable solvent, they were excluded from the dataset used to train the GPR models due to inconsistent sample handling. All iterative stabilization experiments were performed for 24 h at 37 °C and at a fixed polymer concentration of 12.5×10^{-6} M. 15 µL of 25×10^{-6} M polymer samples were mixed with 15 µL of 2 ng µL^{-1} ChABC (8.69×10^{-9} M) in UV-star polystyrene 384 well plates that were thermally sealed with a plate sealing film before being thermally challenged in an incubator at 37 °C for 24 h. Substrate solution was prepared by diluting 10 mg mL^{-1} of chondroitin sulfate in DI water to a final concentration of 4 mg mL^{-1} in assay buffer. 30 uL of 4 mg mL^{-1} substrate was added to 30 µL of polymer-enzyme complexes and an increase in absorbance was measured in kinetic mode for 60 min at 232 nm with 20 s intervals at 25 °C. The initial rate of change of absorbance was used to calculate the specific enzyme activity using the following equation.

Specific Activity

$$= \frac{\text{Adjusted } V_{max} \left(\frac{OD}{min}\right) \times \text{well volume (L)} \times 10^{12} \left(\frac{pmol}{mol}\right)}{\text{ext coeff } (M^{-1}\ cm^{-1}) \times \text{path correction} \times \text{ amount of enzyme } (\mu g)}$$

(3)

Ext Coeff for CS substrate was 3800 M^{-1} cm^{-1} and path correction was 0.95 cm.

Polymer Characterization: The molecular weight (M_w and M_n) and dispersity ($Đ$) of selected polymers were measured by gel permeation chromatography (GPC) using an Agilent 1260 Infinity II. Polymer samples were eluted through a Phenomenex 5.0 µm guard column (50 × 7.5 mm) preceded by two Phenomenex Phenogel columns (10E4 and 10E3 Å). GPC calibration was completed with Agilent PMMA standards. Polymers were prepared at 50:1 eluent/polymer ratio in DMF and filtered with a 0.45 µm PTFE filter. Polymer conversion was calculated by obtaining ^{1}H NMR spec-

tra using a Varian VNMRS 500 MHz spectrometer with mesitylene as an internal standard and processed using Mestrenova 11.0.4.

Dynamic Light Scattering (DLS): DLS of polymers and polymer-enzyme mixtures were performed on a DynaPro DLS Plate Reader III, Wyatt Technologies. Concentration of ChABC for DLS experiments was maintained at 0.2 mg mL^{-1} while polymer concentration was at 2 mg mL^{-1}. The data was collected using a wavelength of 830 nm and a scattering angle of 173°. Fifteen acquisitions were collected for each sample with an acquisition time of 5 s per acquisition using auto attenuation. Regularization analysis was performed using Rayleigh spheres model for hydrodynamic size measurement.

E18 Rat Cortical Astrocyte Isolation: For astrocyte cultures, cortical tissues from Sprague Dawley rat embryos were purchased from BrainBits. Prior to isolation, T-75 flasks were coated with 10 mL of PDL for at least 4 h at room temperature. On the day the tissue was received, it was immediately dissociated for E18 cortical astrocyte isolation. Briefly, the media in which the tissue was preserved was removed and stored for a later step. The tissue was then incubated in 2 mL of cell dissociation media composed of Hibernate-E without Calcium (BrainBits) media containing papain (BrainBits, 2 mg mL^{-1}) for 10 min at 37 °C. The cell dissociation media was removed, and the preservation media was re-added to the tissue. The tissue was then mechanically dissociated with a fire-polished Pasteur pipette. The cell suspension was spun down (200 RCF for 1 min), the supernatant was removed, and the cells were re-suspended in 3 mL of "astrocyte maintenance media" composed of low glucose Dulbecco's Modified Eagle Medium (DMEM, Gibco) containing sodium pyruvate and L-glutamine and supplemented with 10% fetal bovine serum (FBS, Gibco) + 1% penicillin-streptomycin (PS, 10000 U mL^{-1}, Gibco). The isolated astrocytes were then seeded on to the pre-coated T-75 flasks at a density of 1 million cells per flask. The media was changed 24 h after plating and subsequently changed every 3–4 days until 70% confluency was reached. Astrocytes were used anywhere between P1 and P3.

Cytotoxicity: Astrocytes were seeded onto 24 well-plates at a density of 50 000 cells per well and were allowed to attach overnight. On the day of the experiment, the media was changed and 100 µL of polymer at varying concentrations were added for final concentrations of $6.25–400 \times 10^{-6}$ M. After 24 h, supernatants were collected, and viability was assessed via LIVE/DEAD (Invitrogen) assay according to manufacturer instructions. Briefly, cells were incubated with 1.5 µg mL^{-1} calcein-AM, 1.3 µg mL^{-1} ethidium-homodimer-1, and 2 µg mL^{-1} Hoechst-33342 (Invitrogen) for 10–15 min, followed by three washes with media. The cells were subsequently fixed with 200 µL of 4% paraformaldehyde (PFA, Sigma Aldrich) per well. After 15 min of PFA incubation, 600 µL of phosphate-buffered saline (PBS) was added to each well to dilute the PFA to 1%. The plate was then stored at 4 °C for any future analysis. The supernatants were later analyzed for TNF-α levels via an enzyme linked immunosorbent assay (ELISA, BioLegend) according to manufacturer instructions.

Astrocyte Inflammation in Presence of Copolymers: Astrocytes were seeded onto 24 well-plates at a density of 50 000 cells per well and were allowed to attach overnight. On the day of the experiment, the media was switched to low FBS media (low glucose DMEM containing sodium pyruvate and L-glutamine and supplemented with 1% FBS + 1% PS) and 100 µL of polymer at varying concentrations were added for final concentrations of $6.25–400 \times 10^{-6}$ M. Control groups for the study included astrocytes exposed to ultrapure lipopolysaccharide (LPS, Invivogen, 4 µg mL^{-1}) and without LPS, a well-known inflammatory stimulant. After 24 h, cell supernatants were collected, and cells were fixed with PFA as described above. Activation was determined by evaluating secretion of TNF-α via ELISA.

Statistical Analysis: All enzyme-activity measurements for active learning iterations 1–3 were performed in triplicates. Cytotoxicity and inflammation experiments were measured with a sample size of $n = 6–9$, and the statistical significance was assessed using Tukey's paired comparison test.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

[1] a) J. Silver, J. H. Miller, *Nat. Rev. Neurosci.* **2004**, *5*, 146; b) M. B. Leal-Filho, *Surg. Neurol. Int.* **2011**, *2*, 112; c) E. J. Bradbury, E. R. Burnside, *Nat. Commun.* **2019**, *10*, 3879.

[2] a) R. J. Mckeon, R. C. Schreiber, J. S. Rudge, J. Silver, *J. Neurosci.* **1991**, *11*, 3398; b) M. V. Sofroniew, *Trends Neurosci.* **2009**, *32*, 638.

[3] a) H. Wang, Y. Katagiri, T. E. McCann, E. Unsworth, P. Goldsmith, Z. X. Yu, F. Tan, L. Santiago, E. M. Mills, Y. Wang, A. J. Symes, H. M. Geller, *J. Cell Sci.* **2008**, *121*, 3083; b) J. R. Siebert, A. C. Steencken, D. J. Osterhout, *Biomed Res. Int.* **2014**, *2014*, 845323.

[4] B. Caterson, *Int. J. Exp. Pathol.* **2012**, *93*, 1.

[5] a) R. A. Asher, D. A. Morgenstern, M. C. Shearer, K. H. Adcock, P. Pesheva, J. W. Fawcett, *J. Neurosci.* **2002**, *22*, 2225; b) L. L. Jones, R. U. Margolis, M. H. Tuszynski, *Exp. Neurol.* **2003**, *182*, 399.

[6] a) E. J. Bradbury, L. D. Moon, R. J. Popat, V. R. King, G. S. Bennett, P. N. Patel, J. W. Fawcett, S. B. McMahon, *Nature* **2002**, *416*, 636; b) A. W. Barritt, M. Davies, F. Marchand, R. Hartley, J. Grist, P. Yip, S. B. McMahon, E. J. Bradbury, *J. Neurosci.* **2006**, *26*, 10856; c) A. O. Caggiano, M. P. Zimber, A. Ganguly, A. R. Blight, E. A. Gruskin, *J. Neurotrauma* **2005**, *22*, 226; d) C. H. Chau, D. K. Shum, H. Li, J. Pei, Y. Y. Lui, L. Wirthlin, Y. S. Chan, X. M. Xu, *FASEB J.* **2004**, *18*, 194; e) E. S. Rosenzweig, E. A. Salegio, J. J. Liang, J. L. Weber, C. A. Weinholtz, J. H. Brock, R. Moseanko, S. Hawbecker, R. Pender, C. L. Cruzen, J. F. Iaci, A. O. Caggiano, A. R. Blight, B. Haenzi, J. R. Huie, L. A. Havton, Y. S. Nout-Lomas, J. W. Fawcett, A. R. Ferguson, M. S. Beattie, J. C. Bresnahan, M. H. Tuszynski, *Nat. Neurosci.* **2019**, *22*, 1269.

[7] N. J. Tester, A. H. Plaas, D. R. Howland, *J. Neurosci. Res.* **2007**, *85*, 1110.

[8] a) H. Lee, R. J. McKeon, R. V. Bellamkonda, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 3340; b) M. E. Shahaboddin, K. Khajeh, M. Maleki, A. Golestani, *Enzyme Microb. Technol.* **2017**, *105*, 38; c) A. Raspa, E. Bolla, C. Cuscona, F. Gelain, *CNS Neurosci. Ther.* **2019**, *25*, 86; d) M. Nazari-Robati, K. Khajeh, M. Aminian, M. Fathi-Roudsari, A. Golestani, *Int. J. Biol. Macromol.* **2012**, *50*, 487; e) A. Raspa, L. Carminati, R. Pugliese, F. Fontana, F. Gelain, *J. Controlled Release* **2021**, *330*, 1208; f) H. Z. Hu, N. Granger, S. B. Pai, R. V. Bellamkonda, N. D. Jeffery, *Brain* **2018**, *141*, 1017.

[9] a) R. Chapman, M. H. Stenzel, *J. Am. Chem. Soc.* **2019**, *141*, 2754; b) C. DelRe, Y. Jiang, P. Kang, J. Kwon, A. Hall, I. Jayapurna, Z. Ruan, L. Ma, K. Zolkin, T. Li, C. D. Scown, R. O. Ritchie, T. P. Russell, T. Xu, *Nature* **2021**, *592*, 558; c) B. Panganiban, B. Qiao, T. Jiang, C. DelRe, M. M. Obadia, T. D. Nguyen, A. A. A. Smith, A. Hall, I. Sit, M. G. Crosby, P. B. Dennis, E. Drockenmuller, M. Olvera de la Cruz, T. Xu, *Science* **2018**, *359*, 1239.

[10] a) M. A. Webb, N. E. Jackson, P. S. Gil, J. J. de Pablo, *Sci. Adv.* **2020**, *6*, 6216; b) A. J. Gormley, M. A. Webb, *Nat. Rev. Mater.* **2021**, *6*, 642.

[11] a) M. Tamasi, S. Kosuri, J. D. Stefano, R. Chapman, A. J. Gormley, *Adv. Intell. Syst.* **2020**, *2*, 1900126; b) R. Upadhya, S. Kosuri, M. Tamasi, T. A. Meyer, S. Atta, M. A. Webb, A. J. Gormley, *Adv. Drug Delivery Rev.* **2021**, *171*, 1; c) R. Chapman, A. J. Gormley, K. L. Herpoldt, M. M. Stevens, *Macromolecules* **2014**, *47*, 8541; d) R. Chapman, A. J. Gormley, M. H. Stenzel, M. M. Stevens, *Angew. Chem.* **2016**, *55*, 4500; e) S. Oliver, L. Zhao, A. J. Gormley, R. Chapman, C. Boyer, *Macromolecules* **2018**, *52*, 3; f) J. Yeow, R. Chapman, A. J. Gormley, C. Boyer, *Chem. Soc. Rev.* **2018**, *47*, 4357; g) J. Xu, K. Jung, A. Atme, S. Shanmugam, C. Boyer, *J. Am. Chem. Soc.* **2014**, *136*, 5508; h) J. Yeow, R. Chapman, J. T. Xu, C. Boyer, *Polym. Chem.* **2017**, *8*, 5012; i) Z. Li, S. Kosuri, H. Foster, J. Cohen, C. Jumeaux, M. M. Stevens, R. Chapman, A. J. Gormley, *J. Am. Chem. Soc.* **2019**, *141*, 19823; j) R. Upadhya, N. S. Murthy, C. L. Hoop, S. Kosuri, V. Nanda, J. Kohn, J. Baum, A. J. Gormley, *Macromolecules* **2019**, *52*, 8295.

[12] E. C. Stucky, J. Erndt-Marino, R. S. Schloss, M. L. Yarmush, D. I. Shreiber, *Nano LIFE* **2017**, *7*, 1750005.

[13] a) G. V. Fabian Pedregosa, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É.d. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825; b) J. Bergstra, D. Yamins, D. Cox, in Int. Conf. on Machine Learning 2013, Atlanta, USA, p. 115; c) C. K. Williams, C. E. Rasmussen, Cambridge, MA, MIT Press, 2006, Vol. 2, p. 4.

[14] R. A. Patel, C. H. Borca, M. A. Webb, *ChemRxiv*, **2021**, https://doi.org/10.33774/chemrxiv-2021-m74c8.

[15] a) J. Močkus, in Optimization Techniques IFIP Technical Conference, Novosibirsk, Russia, 1975, p. 400; b) J. Mockus, in IFIP Congress 1977 (August), Toronto, Canada, p. 195; c) J. Mockus, In System Modeling and Optimization 1982, p. 473; d) M. Ester, H. P. Kriegel, J. Sander, X. Xu, In KDD 1996, August, 96, 226; e) H. Steinhaus, *Bull. Acad. Pol. Sci.* **1957**, *4*, 801; f) J. MacQueen, *Proc. 5th Berkeley Symp. Math. Stat. Probab.*, Univ. Calif., Los Angeles 1965/66 **1967**, p. 281.