A High-Throughput Platform for Efficient Exploration of Functional Polypeptides Chemical Space

Guangqi Wu [a], Haisen Zhou [a], Jun Zhang [b], Zi-You Tian [a], Xingyi Liu [a], Shuo Wang [a], Connor W. Coley [c], Hua Lu [a]*

[a] Beijing National Laboratory for Molecular Sciences, Center for Soft Matter Science and Engineering, Key Laboratory of Polymer Chemistry and Physics of Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, People's Republic of China.

[b] Changping Laboratory, Yard 28, Science Park Road, Changping District, Beijing 102206, People's Republic of China.

[c] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 United States

* Email: chemhualu@pku.edu.cn

Abstract

Rapid and in-depth exploration of the chemical space of high molecular weight synthetic polypeptides via the ring-opening polymerization (ROP) of *N*-carboxyanhydride (NCA) is a viable approach towards protein mimics and functional biomaterials. The traditional synthetic workflow, however, has been labour-intensive and with limited throughput. Here, we develop an efficient chemistry for the high throughput diversification of polypeptides based on a click-like reaction between selenolate and various electrophiles in aqueous solutions. Importantly, the platform is amenable to automation, which allows rapid generation of up to 1200 homopolypeptides or random heteropolypeptides (RHP) within one day. With the assistance of machine learning, iterative exploration of the RHP library efficiently and effectively identifies candidates with improved glutathione peroxidase-like activity from complex chemical space of which we have little prior knowledge. This automated and high-throughput platform provides potential solutions to unmet challenges such as *de novo* design of artificial enzyme, biomacromolecule delivery, and understanding of intrinsically disordered proteins.

Introduction

Proteins are natural biopolymers with vast chemical space and sophisticated functions such as binding, catalysis, transportation and signaling. For decades, an overarching goal of polymer science is to create protein-like functional polymeric materials for not only fundamental understanding of proteins but also solving real-world challenges.[1-5] To this end, peptides made by the solid phase peptide synthesis (SPPS) have been widely explored. However, SPPS is generally limited by the small scale and short length of its products. Recent studies have shown that synthetic random heteropolymers with statistically controlled side chain compositions are able to exhibit protein-like functions even without the peptide backbone.[6-9] To this end, synthetic polypeptides prepared by the ring-opening polymerization (ROP) of *N*-carboxyanhydrides (NCA) have emerged as promising protein mimics with the potential to synergize the advantages of both peptides and synthetic polymers.[10-15] Specifically, polypeptides possess the same backbone and secondary conformations as protein and can be

efficiently produced at up to kilogram scales and a high number-averaged molecular weight (M_n).[16] Nevertheless, similar to other polymers, polypeptides are subjected to the curse of dimensionality, i.e., the combination of just a few residues can lead to a chemical space that is too large to be fully explored.[17,18] To reach functional protein-mimicking polypeptides from the enormous chemical space, one needs to (1) facilely produce polypeptides from the design space with high fidelity, and (2) establish an efficient strategy for effective exploration of the space at affordable labour and time cost.

The development of automated and high throughput synthesis (HTS) have greatly facilitated small-molecule drug discovery and biomacromolecule evolution.[19-23] Furthermore, recent applications of machine learning (ML) in chemistry have further accelerated the function mining processes and the interpretation of the experimental results.[24-30] Although the employment of HTS and/or ML for making functional polymers can date back to early 1990s,[31-35] progress in this field has been largely lagged behind compared to those for small molecules and biomacromolecules.[18,36] Currently, HTS of polymers is primarily accomplished by parallel step-growth polymerizations or chain-growth radical polymerizations.[31,35,37] Langer *et al*. applied Michael addition to synthesize poly(β-amino ester)s for non-viral gene delivery.[38,39] Boyer *et al.* used photo-induced electron transfer-reversible addition-fragmentation chain transfer (PET-RAFT) for the HTS of poly(acryl amide)s to explore polymers with protein binding[40] and antimicrobial activity[41]. More recently, Leibfarth and Gormley *et al.* applied controlled radical polymerization in combination with automated synthesis and ML to identify polymers for enhanced magnetic resonance signals[18] and protein preservation[42,43], respectively. As a complementary approach to the aforementioned parallel polymerization, post-polymerization modifications (PPM)[44,45] is also common for library generation.[46] For this, highly efficient Huisgen cycloaddition[47], activated ester-amine conjugation[48,49] and thiol-ene reactions[50] are among the most frequently employed reactions.

To date, attempts to incorporate the NCA and polypeptide chemistry to HTS workflow have been challenging and sparse. This is partially attributable to the high sensitivity to moisture of the ROP system, poor aqueous solubility of polypeptides, and laborious workup processes. In a pioneering work, Deming applied the parallel copolymerization strategy to generate hundreds of polypeptides in two weeks[51]. However, the synthesis procedure was time-consuming with a relatively low throughput. On the other hand, PPM was also utilized for the synthesis of both homopolypeptides and random heteropolypeptides (RHP) (Figure 1A),[52-58] but again most of the early seminal works were performed in a low throughput fashion. Moreover, many click reactions often introduce bulky and redundant spacer moieties (*e.g.,* triazole) that might affect both the secondary structure and function of the resulting polypeptides.

One underlying challenge for ML-assisted polymer design is the limited availability of high-quality data.[36] While most studies exploit data from literature and virtual experiments (e.g., electronic structure calculations or simulations), an ideal platform is capable of performing new experiments to support model training. We envision that a powerful PPM chemistry would not only simplify the HTS procedure but also eliminate the difference in M_n and dispersity, which is beneficial for the generation of standard data. Herein, we report the development of a HTS platform in aqueous solutions for polypeptides based on a click-like reaction between selenolate and electrophiles. This quantitative chemistry gave accurate control of the molecular composition of RHPs and was amenable to automated synthesis, which allowed efficient generation and purification of over 1200 polymers within one day. Compared with other click-type reactions, this chemistry only introduced a selenium atom as a miniature

yet reactive linker, which pose minimum influence on the overall polymer structure and offer desirable responsiveness and functions to the materials. With the assistance of ML model-guided optimization, we were able to perform iterative exploration of the RHP chemical space for enzyme mimics and identified candidates with improved glutathione peroxidase (GPx)-like activity in a more efficient and effective way.

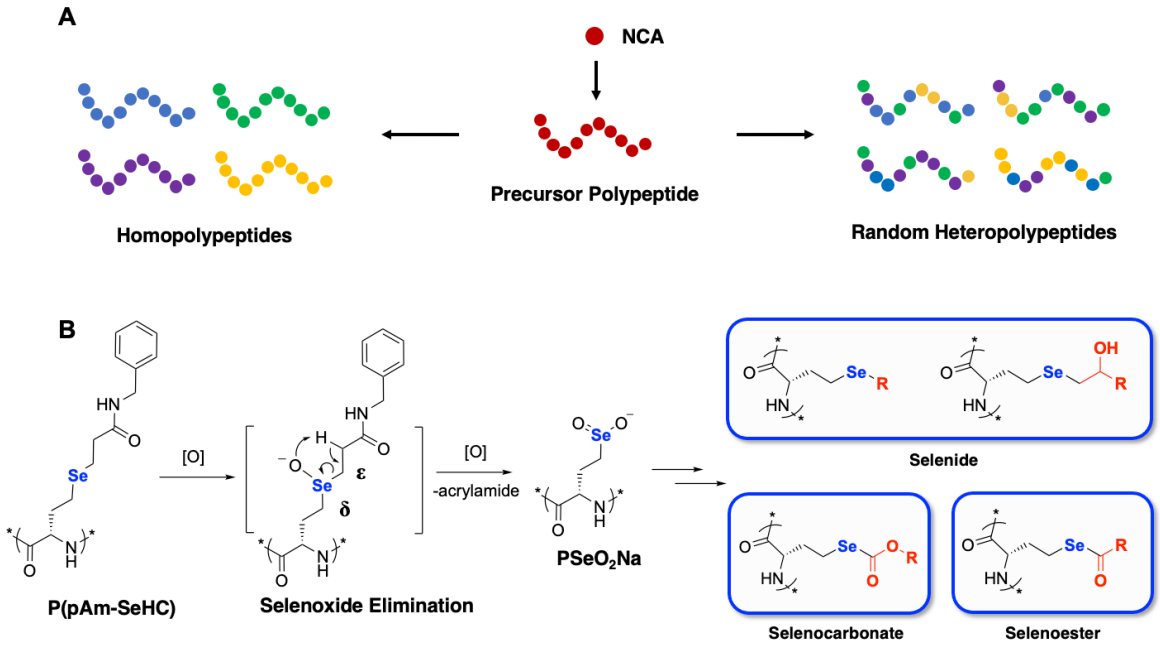


Figure 1. Post-polymerization modification (PPM) of polypeptides through highly efficient selenium chemistry. (A) Schematic illustration of the PPM strategy for making homopolypeptides (left) and random heteropolypeptides (RHP, right). (B) The selenopolypeptide $PSeO_2Na$, derived from the selenoxide elimination of P(pAm-SeHC), is utilized to generate reactive selenium species for PPM.

Results and Discussion

Design and synthesis of the precursor selenopolypeptide for PPM

Inspired by recent studies on selenocysteine-based bioconjugation[59-65], we sought to introduce a selenolate, arguably one of the most nucleophilic functional groups in aqueous solution, to the side chain for the PPM of polypeptides (Figure 1B). Among various selenol protecting strategies[66], we chose to generate the reactive selenium species after polymerization through selenoxide elimination (Figure 1B) owing to its efficiency and mildness. This type of reaction was previously documented for the introduction of double bond.[67] However, the resulting selenium species were normally considered as byproducts and discarded as mixture.[68-71] Notably, Chen *et al.* reported that if a selenium and an amide carbonyl were placed at the δ- and ζ-position of the side chain of amino acids, the elimination would exclusively take place at the Se_δ-C_ε instead of the C_β-Se_δ bond (Figure 1B).[72-74] Based on this discovery, we designed a selenopolypeptide, P(pAm-SeHC), whose pendant group is a latent selenolate. The seleno-amino acid (pAm-SeHC) was synthesized in an one-pot manner with no need of chromatography, affording the pure product at multi-gram scale (Figure 2A-B, S1-2). The monomer, namely pAm-SeHC NCA, was then obtained using a moisture-tolerant method recently developed by our group (Figure S3-4).[75] Benzyl amine-initiated ROP of pAm-SeHC

NCA at various monomer-to-initiator ([M]/[I]) ratios all gave complete monomer conversion, generating the desired selenopolypeptide P(pAm-SeHC) with a controlled M_n up to 35.8 × 10^3 g mol^{-1} and dispersity ($Đ$) of ~1.10 (Figure 2C, Figure S5, Table S1). It was also worth mentioning that the whole process did not involve any volatile selenium species, and thus avoided undesired stench which has been a common safety concern for organoselenium compounds.

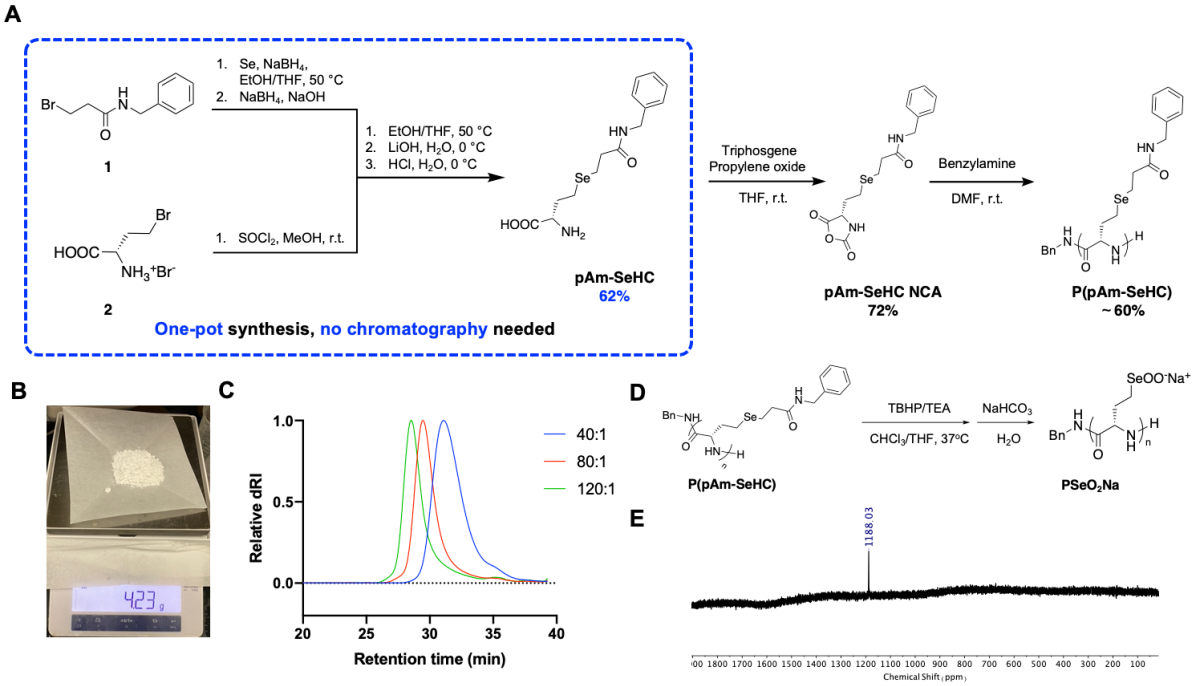
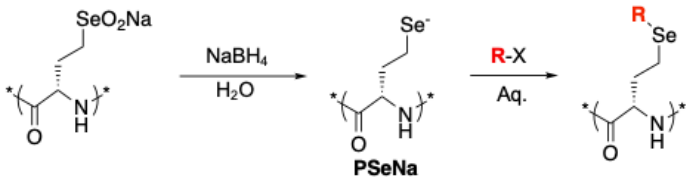


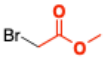
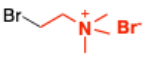
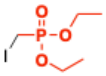
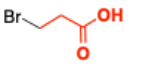
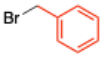
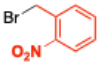
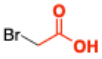
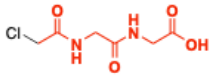
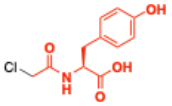
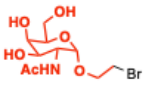
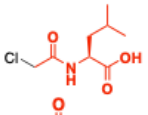
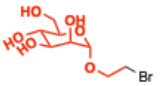
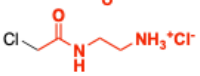
Figure 2. Synthesis of the precursor polypeptide $PSeO_2Na$. (A) Scheme for the synthesis of P(pAm-SeHC). (B) A photograph of the amino acid pAm-SeHC in gram scale. (C) Size exclusion chromatography (SEC) of P(pAm-SeHC) prepared at various [M]/[I] ratios. (D) Oxidative elimination of P(pAm-SeHC) in a biphasic system gave the precursor polypeptide $PSeO_2Na$. (E) The ^{77}Se NMR spectrum of $PSeO_2Na$.

We next performed the selenoxide elimination of P(pAm-SeHC) with *tert*-butyl hydroperoxide (TBHP). However, initial attempts all failed in most common organic solvents, possibly owing to the poor organic solubility of both the oxidized intermediate and eliminated product. This hurdle was eventually overcome using a one-pot, two-step process in a biphasic system (Figure 2D).[71] Briefly, P(pAm-SeHC) was firstly oxidized in a mixed THF/chloroform solution using TBHP, followed by the addition of sodium bicarbonate solution and trimethylamine (TEA), affording exclusively the expected $PSeO_2Na$ in the aqueous phase. The generation of $PSeO_2Na$ was confirmed by both ^1H and ^{77}Se NMR with an overall yield around 60% calculated from NCA (Figure 2E, Figure S6). Remarkably, $PSeO_2Na$ was stable in ambient atmosphere for at least three months under -20 °C, making it an ideal intermediate for scale-up synthesis and long-term storage. The block copolymer PEG-*b*-$PSeO_2Na$ was also prepared from PEG-*b*-P(pAm-SeHC) with a modest yield (Figure S7-8).

PPM for the generation of various homopolypeptides

Table 1. PPM of $PSeO_2Na$ with activated and inactivated halides.



Activated R-X		Product (Yield [a], GE [b])	Inactivated R-X		Product (Yield [a], GE [b])
Hal-1		P1 (79%, >95%)	Hal-14		P14 (60%, >95%)
Hal-2		P2 (73%, >95%)	Hal-15		P15 (73%, >95%)
Hal-3		P3 (62%, >95%)	Hal-16		P16 (68%, >95%)
Hal-4		P4 (48%, >95%)	Hal-17		P17 (75%, >95%)
Hal-5		P5 (99%, >95%) [c]	Hal-18		P18 (84%, >95%)
Hal-6		P6 (97%, >95%) [c]	Hal-19		P19 (53%, >95%)
Hal-7		P7 (75%, >95%) [c]	Hal-20		P20 (42%, >95%)
Hal-8		P8 (78%, >95%)	Hal-21		P21 (81%, >95%)
Hal-9		P9 (89%, >95%)	Hal-22		P22 (63%, >95%)
Hal-10		P10 (78%, >95%)	Hal-23		P23 (62%, >95%)
Hal-11		P11 (86%, >95%)	Hal-24		P24 (85%, >95%)
Hal-12		P12 (87%, >95%)	Hal-25		P25 (100%, >95%)
Hal-13		P13 (72%, >95%)			

[a] yield = purification yield, [b] GE: grafting efficiency based on the ^1H NMR spectra of the product, [c] Prepared from PEG-*b*-$PSeO_2Na$

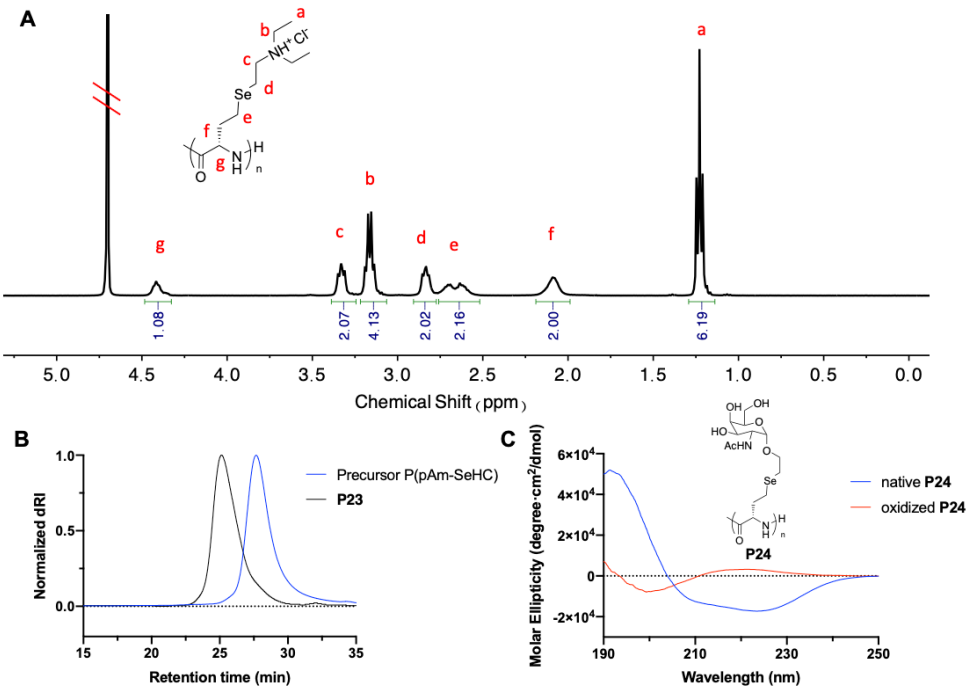


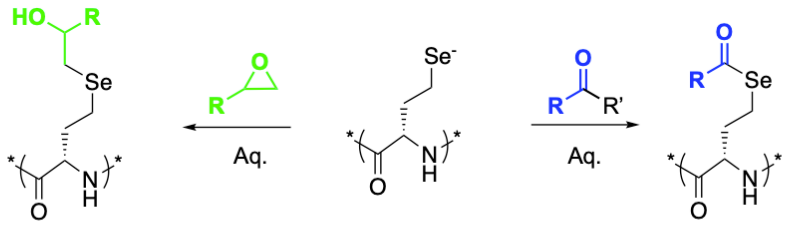
Figure 3. Representative characterizations of homopolypeptides prepared by PPM of $PSeO_2Na$. (A) The ^1H NMR spectrum of **P17** in D_2O. (B) Overlay of the SEC traces of **P23** and its precursor P(pAm-SeHC). (C) Circular dichroism spectrum of the GalNAc-grafted selenopolypeptide (**P24**) and its oxidized form.

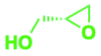
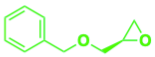
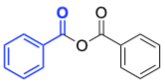
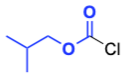
Next, $PSeO_2Na$ was reduced with $NaBH_4$ in water, affording the selenolate-bearing polypeptide (PSeNa) for further functionalization (Figure S9). This *in situ* generated PSeNa was soluble in aqueous solutions and highly reactive for substitution reactions with various activated (Table 1, **Hal-1–13**) or inactivated (Table 1, **Hal-14–25**) electrophiles. The PPM exhibited almost quantitative side chain conversion as indicated by ^1H NMR (Table 1, Figure 4A, Figure S10-S35). The SEC chromatogram of the modified polymer remained a unimodal peak after the modification, suggesting negligible backbone degradation, chain scission, or crosslink during the process (Figure 4B). Some of the obtained polypeptides (e.g. **P21**, **P23** and **P24,** Figure 4C, Figure S36) exhibited typical α-helical conformation in water, implying minimum racemization of the backbone. Oxidation-induced helix-to-coil transition (Figure 4C) was also observed, which was similar to the oligo(ethylene glycol)-grafted selenopolypeptides.[76]

The PPM of PSeNa showed remarkable tolerance to various functionalities, including those that are hard to be introduced directly through the ROP of NCA. For instance, poly(selenomethionine) (**P1**) and the hydrofluorocarbon-bearing polymer (**P20**) were smoothly synthesized in spite of their poor solubility in common solvents. Reactive functional groups such as alkenes and alkynes were also facilely grafted to the side chain (**P5** and **P6**).[77] Photo-labile *o*-nitro benzyl group was efficiently incorporated (**P8**) for potential light-responsive materials[78]. We also generated various anionic and cationic polyelectrolytes bearing diverse functionalities such as carboxylic acids, primary, secondary, tertiary amines, and quaternary ammonium (**P9-P19**). The applicability of the strategy was further demonstrated by the preparation of densely packed brush-like polymer (**P23,** Figure S34), which has long been challenging owing to the steric hindrance.[56] The synthesis of glycopolypeptides used to be

laborious; here, polypeptides modified with GalNAc (**P24**) and mannose (**P25**) were readily achieved in one step, holding great promise for biomedical applications including lysosome targeting chimeras [79,80] and immunotherapy[81]. Disappointingly, modification of PSeNa with secondary organohalides or primary organochlorides gave low grafting efficiency, likely due to their low reactivity. Of note, some of the synthesized selenopolypeptides had extremely poor solubility even in trifluoroacetic acid (TFA). To circumvent the problem, those polymers were prepared by using the more soluble block copolymer PEG-*b*-PSeO$_2$Na as the precursor of PPM (**P5**, **P6**, **P7**).

Table 2. PPM of PSeO$_2$Na with epoxide, acyl chloride, anhydride, and chloroformate



Epoxide		Product (Yield [a], GE [b])	ROR'		Product (Yield [a], GE [b])
EP-1		P26 (46%, >95%)	AC-1		P28 (88%, >95%) [c]
EP-2		P27 (79%, >95%)	AC-2		P29 (83%, >95%) [c]
			AC-3		P30 (81%, >95%) [c]

[a] yield = purification yield, [b] GE: grafting efficiency based on the ^{1}H NMR spectra of the product, [c] Prepared from PEG-*b*-PSeO$_2$Na

Other electrophilic substrates besides organohalides were also applied for PSeNa modification (Table 2). For example, modification of PSeNa with epoxides smoothly generated polypeptides with β-hydroxyl selenide (**P26** and **P27**, Figure S37-38). We also successfully fabricated for the first time two polypeptides bearing selenoester (**P28** and **P29**, Figure S39-40), a functional group incompatible with the ROP of NCA due to its vulnerability to nucleophiles[82,83]. Similarly, a polypeptide tethering selenocarbonate was prepared from chloroformate (**P30**, Figure S41). These polymers were all obtained with more than 80% separation yields and almost quantitative grafting efficiency.

PPM for making random heteropolypeptides

With the success in making homopolypeptides, we next exploited the chemistry for RHP synthesis by simultaneously treating the precursor with multiple organohalides. To facilitate future HTS and ML algorithm development, we sought to predictively control the molecular composition of the RHP through machine-readable input.[36,84] For this, we choose the feeding volume ratio of the organohalides, as it could be directly transformed to the command for automated liquid handling workstation. To achieve this goal, the condition (stock solution concentration, stoichiometry, temperature, *etc.*) for the RHP synthesis was carefully optimized on three binary systems: (1) both organohalides are inactivated (**Hal-16** and **Hal-22**); (2) both

organohalides are activated (**Hal-11** and **Hal-12**); (3) one organohalide is activated and the other is inactivated (**Hal-9** and **Hal-22**). After the trial-and-error attempts and considering the difference in relative reactivity, we fixed the concentration of the stock solutions at 1.2 and 1.0 equivalent to that of the selenolate for those inactivated and activated organohalides, respectively. Meanwhile, the total volume of organohalides was set to be equal to the volume of PSeNa solution. With this optimized condition, a good match between the input volume ratio and the actual molecular composition of the RHP was obtained for all three scenarios (Figure 4, Figure S42-44). Thanks to the high reactivity of the selenolate, it was unnecessary, sometimes even deteriorative, to use a largely excessive organohalides in the RHP synthesis (Figure S45-46). Overall, this capability to precisely control the composition of RHP laid a firm foundation for generating high quality datasets for ML and paved the way for quantitative structure-activity relationship (QSAR).

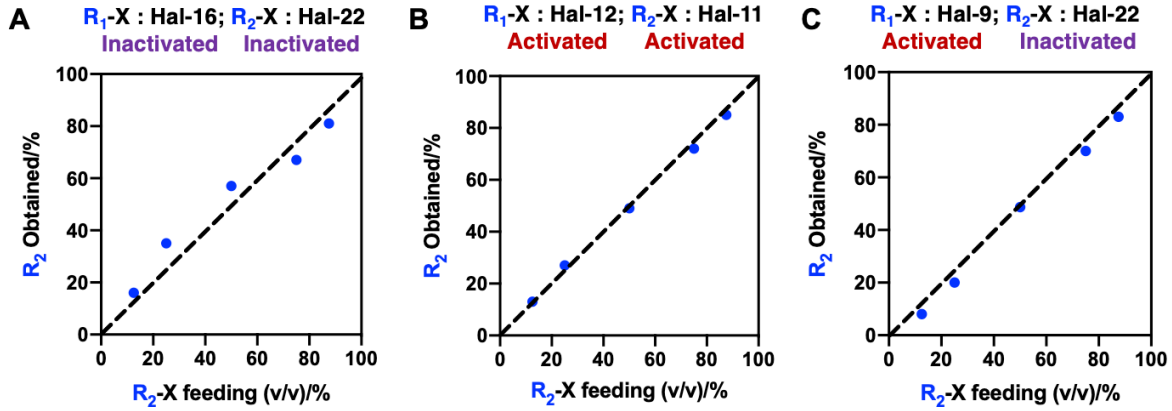


Figure 4. Control of molecular composition of RHP in binary organohalide systems. (A) Modification with two inactivated (**Hal-16** and **Hal-22**) organohalides. Concentration of stock solutions: [**Hal-16**] = [**Hal-22**] = [selenolate] ×1.2; condition: 50 °C for 6 hours. (B) Modification with two activated (**Hal-11** and **Hal-12**) organohalides. Concentration of stock solutions: [**Hal-11**] = [**Hal-12**] = [selenolate]; condition: room temperature for 4 hours. (C) Modification with an activated (**Hal-9**) and an inactivated (**Hal-22**) organohalides. Concentration of stock solutions: [**Hal-9**] = [selenolate], [**Hal-22**] = [selenolate] ×1.2; condition: incubation at 50 °C for 6 hours. The composition of the obtained RHP were determined by ^{1}H NMR spectroscopy after purification.

Function-oriented exploration of RHP with GPx-like activity with the assistance of automation and machine learning

We then explored the feasibility of transferring the synthesis from flasks (5.0 mg/mL of **PSeO$_2$Na,** 1.0 – 4.0 mL) to multi-well plates (1.0 mg/mL of **PSeO$_2$Na,** 100 – 200 μL). In model studies within an NMR tube, the PPM was found completed after 6 hours at 50 °C or 8 hours at 37 °C when an inactivated organohalide **Hal-16** was used (Figure S47). For the activated organohalide **Hal-9**, the completed modification required only 1 hour at 50 °C (Figure S48). Based on this finding, the HTS of RHP was established with the assistance of a commercialized automated workstation for dispensing stock solutions of organohalides to plates (Figure 5A). Then the freshly reduced PSeNa solution was added through pipetting and the plate was sealed and incubated in an oven at 50 °C for 6-8 hours. The resulting polymers were purified in parallel using a desalting plate. This semi-automated workflow greatly boosted the synthesis capability and enabled the parallel preparation of ~400 RHP (4 plates) in one day.

The throughput could be easily improved to ~1200 RHP (12 plates) per day if only activated organohalides were used for modification.

Next, the HTS system was coupled to functional analysis for protein-like activity. Many organoselenium compounds show activities similar to glutathione peroxidase (GPx)[85,86], a class of proteins that catalyze the reduction of peroxide by glutathione (GSH). Since GPx plays important roles in retaining cellular redox homeostasis,[87] developing GPx mimics may lead to antioxidative therapeutics for the treatment of stroke, reperfusion injury, and neurodegenerative diseases.[88-92] We previously synthesized two selenopolypeptides P(EG_x-SeHC) (x = 3 or 4) but they only exhibited weak GPx-like activities.[76] Here, we sought to improve the GPx-like activity of the RHP with the platform and provide information of structure-activity relationships (SAR) for future works.

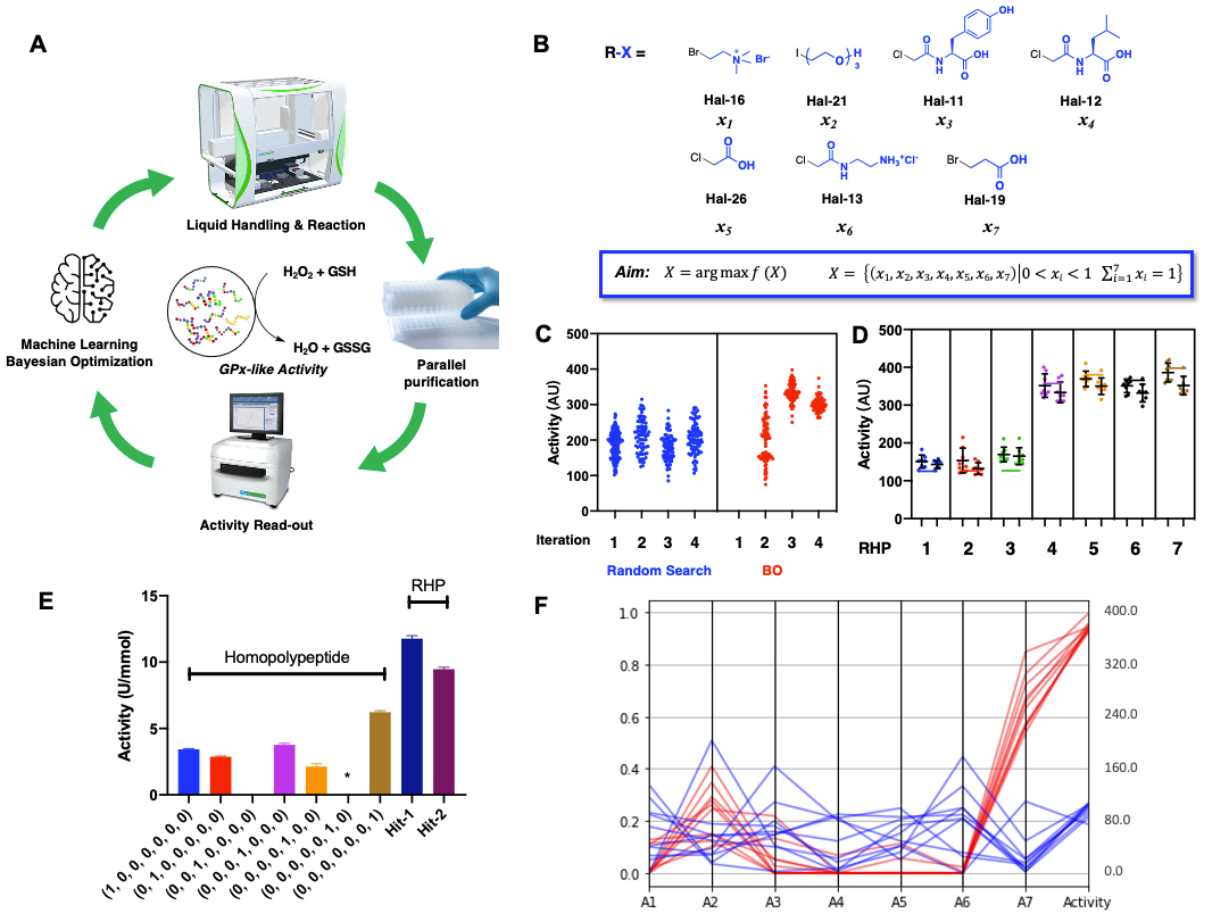


Figure 5. Closed-loop optimization of GPx activity of the RHPs via HTS and machine learning. (A) Cartoon illustration of the closed-loop workflow containing four modules, namely HTS, parallel purification, activity read-out, and Bayesian optimization. (B) Structure of the seven selected organohalides for RHP library generation and aim of optimization. The molecular composition of RHPs are descripted as seven-dimensional vectors, $\boldsymbol{x} = (x_1, \ldots, x_7)$, where x_n (n = 1 to 7) is the relative volume ratio of the organohalide. (C) GPx-like activity of RHPs in each iteration via random searching (blue) or Bayesian optimization (red). (D) Data validation within a plate (n =8) and between two different plates. RHPs with low (lane 1-3) and high (lane 4-7) GPx-like activities from the database were selected for validation. Dots on the right and left side in each lane represent the results from different plates. Black lines and error bars in each lane represent the mean and standard deviation. The colored line in each lane is the original activity of the RHP from the database (E) Comparison of the GPx-like activity

of the two RHP hits (Hit-1: with the seven homopolypeptides each modified with one individual organohalide used in HTS (n =3). **Hit-1**: (0.12, 0.12, 0, 0, 0, 0, 0.76) and **Hit-2**: (0, 0.24, 0.22, 0, 0, 0, 0.54). All polymers are synthesized in flask and then purified for GPx activity. *Activity of the homopolypeptide (0, 0, 0, 0, 0, 1, 0) cannot be measured properly owing to precipitation during testing. (F) Parallel coordinate plot of best ten (red) and worst ten (blue) performing RHPs.

Seven organohalides with good water solubility were selected for PPM (Figure 5B). These organohalides offer different structural features including charges, hydrogen bonding donor and acceptors, neutral polar groups, alky groups, and aromatic groups. Mole fractions of the components were combined as a vector representation of each RHP. A fluorescein was tethered to the N-terminus of the precursor polymer as an internal label to quantify the polymer content in each well. The GPx-like activity of each RHP was determined by normalizing the absolute readout from a modified NADPH-coupled assay93 to the fluorescent intensity of each well (see Supporting Information). Based on previous optimization studies (Figure 4 and S49-S50), we built a workflow that execute the HTS, parallel purification, and activity analysis of the RHPs within 10 hours (83 experiments/plate, 2 plates in parallel). However, even with this throughput, the space of hypothetical RHPs is too large to be fully explored. This situation motivates the use of a model-guided optimization strategy to help prioritize experiments and accelerate discovery. Bayesian optimization (BO) is a powerful tool for various design problems94 and is receiving increasing attention in the chemistry community43,95-97. Because of the data efficiency of BO relative to brute force or random screening, it is especially useful for problems where evaluation is expensive. We applied a BO framework based on BoTorch and Ax98 and established a closed-loop design-build-test-learn workflow (Figure 5A).

Within 4 days, four iterations comprising a total of ~660 experiments were performed (Figure 5C) with 166 experiments per iteration. Initially, 166 RHPs were randomly chosen analyzed from the designed space to train a Gaussian process (GP) regression model. Candidates for successive iterations were chosen by selecting compositions that optimized an expected improvement (EI) acquisition function, subject to the constraint that total mole fractions equal 1. To avoid trapping into local minimums, random search and BO were performed simultaneously in each round. Both were used to select 83 RHP compositions to synthesize, evaluate, and retrain the surrogate GP before proposing the candidates for the next iteration. It was found that while random search consistently found candidates with activities near a range of 150-200, BO efficiently found RHPs with substantially higher GPx-like activity, particularly in the 3rd and 4th iteration (Figure 5C). T-distributed stochastic neighbor embeddings (t-SNE) showed that BO quickly identified an area in the design space that achieves higher activity (Figure S49). Replicate experiments of seven randomly-selected RHPs from the database were carried out in different plates, and the results validated the reproducibility of the data (Figure 5D, Table S2-3).

To further validate the results, two hits from BO were synthesized in flask (Figure S50-51) and their GPx-like activity normalized to the amount of selenium was evaluated. **Hit-1** exhibited ~2 times higher GPx-like activity than the homopolypeptide modified with **Hal-19**, the most active homopolypeptide in the design space (Figure 5E), which meant that activity of RHPs is not merely the normalized average of the activity of each component. Analysis of the high-performing RHPs showed that RHPs with high x_7 tend to have higher activity (Figure 5F, Table S4-5), which was also indicated by retrospectively analyzing the full dataset with a linear regression surrogate model (Figure S52). While the linear model suggests maximizing x_7 to be advantageous for activity, the homopolymer with x_7 = 1 does not outperform **Hit-1** or **Hit-2**.

While detailed SAR is currently under investigation, the above results illustrate that the application of machine learning model or other optimization algorithm can facilitate materials discovery.

Conclusion

In summary, we reported a robust, quantitative, and divergent strategy for the rapid expansion of polypeptide library based on a universal precursor selenopolypeptide. After the controlled ROP and a regioselective elimination reaction, the *in situ* generated selenolate on the side chain were readily modified with a school of electrophiles, creating homopolypeptides and RHPs with broad chemical diversity. This PPM strategy avoided the laborious efforts of making a variety of NCA that are synthetically challenging. Compared with many other click reactions, the selenolate is miniature in size and does not create a bulky linkage moiety after reaction. Moreover, these polymers can be used to design materials with interesting properties and functions by harnessing the unique chemistry of selenium.

The potential of this modification chemistry was further highlighted by the establishment of HTS and ML-model-guided optimization of functional RHP. Enabled by the efficiency of the reaction, a map from the feeding volume ratio to the molecular composition was directly created. Because all polypeptides were derivatized from the same precursor, this strategy could be particularly useful to generate standardized dataset. Moreover, the HTS was performed in aqueous solutions and open air, which allowed convenient transferring of the resulting polymers to subsequent biological assays. As a proof-of-concept, we demonstrated a concise workflow enabling the rapid identification of RHPs with promising GPx-like activity. The identified RHP exhibited GPx-like activity ~2 times higher than the most active homopolypeptide. While detailed SAR is still under investigation, these results underscored the power of HTS and ML in exploring systems of which people have little knowledge. With the rich and robust chemistry, we envision that the potential of this platform is far beyond artificial enzymes and can be readily expanded to applications such as the discovery of antimicrobial agents, understanding of protein phase separation, and development of intracellular delivery systems for therapeutic biomacromolecules.

Acknowledgements

G. W. thank Samuel Goldman and Dr. Zitao Wang for the discussions and Dr. Rocío Mercado for checking with the code. This work was supported by the National Key Research and Development Program of China (2019YFA0904203), National Natural Science Foundation of China (21975004, 22105008), fellowship of China Postdoctoral Science Foundation (2020M680192) and Li Ge-Zhao Ning Life Science Youth Research Fund. G. W. thank the Boya postdoctoral fellowship of Peking University for financial support.

Methods

Selenoxide elimination of P(pAm-SeHC)

p(pAm-SeHC) obtained from ROP of 680 mg pAm-SeHC NCA was dissolved in 50 mL THF and 100 mL chloroform. To the solution was added 70% TBHP (2760 μL, 12 eq of Se) and stirred at room temperature for 2 hours before TEA (800 μL, 3 eq of Se) and 75 mL $NaHCO_3$ solution (1M) were added. The system was stirred under 37 °C for 16 hours during which a

clear phase separation was observed, which was used as an indication of the completion of the reaction. The aqueous phase was washed with 150 mL DCM, and dialyzed (MWCO 3500 Da) against 0.5 M NaCl for 12 hours, followed by water for 48 hours with changed twice every day. The remaining content was lyophilized to give the final product as a pale yellow powder (256 mg, yield: 64% from NCA).

General procedure for post polymerization modification of $PSeO_2Na$

Under the protection of nitrogen, $PSeO_2Na$ was dissolved in water (5-10 mg/mL). To the solution was added $NaBH_4$ (5 mg $NaBH_4$ for 10 mg $PSeO_2Na$) followed by the addition of TFA (1.5 μL for 5 mg $NaBH_4$, **caution: gas emission**). Precipitation was observed in ~10 minutes. After 20 minutes, another portion of TFA (1.5 μL) was added. The system was stirred at room temperature for 30 minutes and the completion of the reduction was indicated by the re-dissolving of the precipitate. Then to the polymer solution was added the modifier in THF or water (specified below). The system was stirred under indicated temperature. The selenopolypeptide was purified by dialysis and/or SEC and recovered by lyophilization. The post polymerization modification of PEG-*b*-$PSeO_2Na$ was carried out similarly.

Please see the support information for other methods.

Reference

1 Cole, J. P., Hanlon, A. M., Rodriguez, K. J. and Berda, E. B. Protein-Like Structure and Activity in Synthetic Polymers. *Journal of Polymer Science Part a-Polymer Chemistry* **55**, 191-206 (2017). https://doi.org:10.1002/pola.28378

2 Rothfuss, H., Knofel, N. D., Roesky, P. W. and Barner-Kowollik, C. Single-Chain Nanoparticles as Catalytic Nanoreactors. *J. Am. Chem. Soc.* **140**, 5875-5881 (2018). https://doi.org:10.1021/jacs.8b02135

3 Bonduelle, C. Secondary structures of synthetic polypeptide polymers. *Polym. Chem.* **9**, 1517-1529 (2018). https://doi.org:10.1039/c7py01725a

4 Varanko, A. K., Su, J. C. and Chilkoti, A. Elastin-Like Polypeptides for Biomedical Applications. *Annual Review of Biomedical Engineering, Vol 22* **22**, 343-369 (2020). https://doi.org:10.1146/annurev-bioeng-092419-061127

5 Callmann, C. E., Thompson, M. P. and Gianneschi, N. C. Poly(peptide): Synthesis, Structure, and Function of Peptide-Polymer Amphiphiles and Protein-like Polymers. *Acc. Chem. Res.* **53**, 400-413 (2020). https://doi.org:10.1021/acs.accounts.9b00518

6 Jiang, T., Hall, A., Eres, M., Hemmatian, Z., Qiao, B., Zhou, Y., Ruan, Z., Couse, A. D., Heller, W. T., Huang, H., de la Cruz, M. O., Rolandi, M. and Xu, T. Single-chain heteropolymers transport protons selectively and rapidly. *Nature* **577**, 216-220 (2020). https://doi.org:10.1038/s41586-019-1881-0

7 Panganiban, B., Qiao, B. F., Jiang, T., DelRe, C., Obadia, M. M., Nguyen, T. D., Smith, A. A. A., Hall, A., Sit, I., Crosby, M. G., Dennis, P. B., Drockenmuller, E., de la Cruz, M. O. and Xu, T. Random heteropolymers preserve protein function in foreign environments. *Science* **359**, 1239-1243 (2018). https://doi.org:10.1126/science.aao0335

8 Hilburg, S. L., Ruan, Z. Y., Xu, T. and Alexander-Katz, A. Behavior of Protein-Inspired Synthetic Random Heteropolymers. *Macromolecules* **53**, 9187-9199 (2020). https://doi.org:10.1021/acs.macromol.0c01886

9 Han, Z., Hilburg, S. L. and Alexander-Katz, A. Forced Unfolding of Protein-Inspired Single-Chain Random Heteropolymers. *Macromolecules* **55**, 1295-1309 (2022). https://doi.org:10.1021/acs.macromol.1c02411

10 Song, Z. Y., Tan, Z. Z. and Cheng, J. J. Recent Advances and Future Perspectives of Synthetic Polypeptides from N-Carboxyanhydrides. *Macromolecules* **52**, 8521-8539 (2019). https://doi.org:10.1021/acs.macromol.9b01450

11 Song, Z. Y., Han, Z. Y., Lv, S. X., Chen, C. Y., Chen, L., Yin, L. C. and Cheng, J. J. Synthetic polypeptides: from polymer design to supramolecular assembly and biomedical application. *Chem. Soc. Rev.* **46**, 6570-6599 (2017). https://doi.org:10.1039/c7cs00460e

12 Zhou, X. F. and Li, Z. B. Advances and Biomedical Applications of Polypeptide Hydrogels Derived from alpha-Amino Acid N-Carboxyanhydride (NCA) Polymerizations. *Adv. Healthcare Mater.* **7**, e1800020 (2018). https://doi.org:ARTN 1800020 10.1002/adhm.201800020

13 Deng, C., Wu, J. T., Cheng, R., Meng, F. H., Klok, H. A. and Zhong, Z. Y. Functional polypeptide and hybrid materials: Precision synthesis via alpha-amino acid N-carboxyanhydride polymerization and emerging biomedical applications. *Prog. Polym. Sci.* **39**, 330-364 (2014). https://doi.org:10.1016/j.progpolymsci.2013.10.008

14 Hou, Y. Q. and Lu, H. Protein PEPylation: A New Paradigm of Protein-Polymer Conjugation. *Bioconjugate Chem.* **30**, 1604-1616 (2019). https://doi.org:10.1021/acs.bioconjchem.9b00236

15 Deming, T. J. Synthetic polypeptides for biomedical applications. *Prog. Polym. Sci.* **32**, 858-875 (2007). https://doi.org:10.1016/j.progpolymsci.2007.05.010

16 Liu, Y., Li, D., Ding, J. X. and Chen, X. S. Controlled synthesis of polypeptides. *Chin. Chem. Lett.* **31**, 3001-3014 (2020). https://doi.org:10.1016/j.cclet.2020.04.029

17 Taylor, S. V., Walter, K. U., Kast, P. and Hilvert, D. Searching sequence space for protein catalysts. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10596-10601 (2001). https://doi.org:10.1073/pnas.191159298

18 Reis, M., Gusev, F., Taylor, N. G., Chung, S. H., Verber, M. D., Lee, Y. Z., Isayev, O. and Leibfarth, F. A. Machine-Learning-Guided Discovery of F-19 MRI Agents Enabled by Automated Copolymer Synthesis. *J. Am. Chem. Soc.* **143**, 17677-17689 (2021). https://doi.org:10.1021/jacs.1c08181

19 Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V. S., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U. and Sittampalam, G. S. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10**, 188-195 (2011). https://doi.org:10.1038/nrd3368

20 Coley, C. W., Eyke, N. S. and Jensen, K. F. Autonomous Discovery in the Chemical Sciences Part I: Progress. *Angew. Chem. Int. Ed.* **59**, 22858-22893 (2020). https://doi.org:10.1002/anie.201909987

21 Yang, L. L., Pijuan-Galito, S., Rho, H. S., Vasilevich, A. S., Eren, A. D., Ge, L., Habibovic, P., Alexander, M. R., de Boer, J., Carlier, A., van Rijn, P. and Zhou, Q. H. High-Throughput Methods in the Discovery and Study of Biomaterials and Materiobiology. *Chem. Rev.* **121**, 4561-4677 (2021). https://doi.org:10.1021/acs.chemrev.0c00752

22 Soheilmoghaddam, F., Rumble, M. and Cooper-White, J. High-Throughput Routes to Biomaterials Discovery. *Chem. Rev.* **121**, 10792-10864 (2021). https://doi.org:10.1021/acs.chemrev.0c01026

23 DeBenedictis, E. A., Chory, E. J., Gretton, D. W., Wang, B., Golas, S. and Esvelt, K. M. Systematic molecular evolution enables robust biomolecule discovery. *Nat Methods* **19**, 55-64 (2022). https://doi.org:10.1038/s41592-021-01348-4

24 Gromski, P. S., Granda, J. M. and Cronin, L. Universal Chemical Synthesis and Discovery with 'The Chemputer'. *Trends in Chemistry* **2**, 4-12 (2020). https://doi.org:10.1016/j.trechm.2019.07.004

25 Pollice, R., Gomes, G. D., Aldeghi, M., Hickman, R. J., Krenn, M., Lavigne, C., Lindner-D'Addario, M., Nigam, A., Ser, C. T., Yao, Z. P. and Aspuru-Guzik, A. Data-Driven Strategies for Accelerated Materials Design. *Acc. Chem. Res.* **54**, 849-860 (2021). https://doi.org:10.1021/acs.accounts.0c00785

26 Correa-Baena, J. P., Hippalgaonkar, K., van Duren, J., Jaffer, S., Chandrasekhar, V. R., Stevanovic, V., Wadia, C., Guha, S. and Buonassisi, T. Accelerating Materials Development via Automation, Machine Learning, and High-Performance Computing. *Joule* **2**, 1410-1420 (2018). https://doi.org:10.1016/j.joule.2018.05.009

27 Yang, K. K., Wu, Z. and Arnold, F. H. Machine-learning-guided directed evolution for protein engineering. *Nat Methods* **16**, 687-694 (2019). https://doi.org:10.1038/s41592-019-0496-6

28 Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M. and Zhao, S. R. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **18**, 463-477 (2019). https://doi.org:10.1038/s41573-019-0024-5

29 Kumar, R., Le, N., Tan, Z., Brown, M. E., Jiang, S. and Reineke, T. M. Efficient Polymer-Mediated Delivery of Gene-Editing Ribonucleoprotein Payloads through Combinatorial Design, Parallelized Experimentation, and Machine Learning. *ACS Nano* **14**, 17626-17639 (2020). https://doi.org:10.1021/acsnano.0c08549

30 Kumar, R., Le, N., Oviedo, F., Brown, M. E. and Reineke, T. M. Combinatorial Polycation Synthesis and Causal Machine Learning Reveal Divergent Polymer Design Rules for Effective pDNA and Ribonucleoprotein Delivery. *Jacs Au* **2**, 428-442 (2022). https://doi.org:10.1021/jacsau.1c00467

31 Goldberg, M., Mahon, K. and Anderson, D. Combinatorial and rational approaches to polymer synthesis for medicine. *Adv Drug Deliv Rev* **60**, 971-978 (2008). https://doi.org:10.1016/j.addr.2008.02.005

32 Baudis, S. and Behl, M. High-Throughput and Combinatorial Approaches for the Development of Multifunctional Polymers. *Macromol. Rapid Commun.* (2021). https://doi.org:ARTN 2100400 10.1002/marc.202100400

33 Holmes, P. F., Bohrer, M. and Kohn, J. Exploration of polymethacrylate structure-property correlations: Advances towards combinatorial and high-throughput methods for biomaterials discovery. *Prog. Polym. Sci.* **33**, 787-796 (2008). https://doi.org:10.1016/j.progpolymsci.2008.05.002

34 Gormley, A. J. and Webb, M. A. Machine learning in combinatorial polymer chemistry. *Nat. Rev. Mater.* **6**, 642-644 (2021). https://doi.org:10.1038/s41578-021-00282-3

35 Upadhya, R., Kosuri, S., Tamasi, M., Meyer, T. A., Atta, S., Webb, M. A. and Gormley, A. J. Automation and data-driven design of polymer therapeutics. *Adv. Drug. Deliv. Rev.* **171**, 1-28 (2021). https://doi.org:10.1016/j.addr.2020.11.009

36 Patel, R. A., Borca, C. H. and Webb, M. A. Featurization strategies for polymer sequence or composition design by machine learning. *Molecular Systems Design & Engineering* (2022). https://doi.org:10.1039/D1ME00160D

37 Oliver, S., Zhao, L., Gormley, A. J., Chapman, R. and Boyer, C. Living in the Fast Lane High Throughput Controlled/Living Radical Polymerization. *Macromolecules* **52**, 3-23 (2019). https://doi.org:10.1021/acs.macromol.8b01864

38 Lynn, D. M., Anderson, D. G., Putnam, D. and Langer, R. Accelerated discovery of synthetic transfection vectors: parallel synthesis and screening of a degradable polymer library. *J. Am. Chem. Soc.* **123**, 8155-8156 (2001). https://doi.org:10.1021/ja016288p

39 Green, J. J., Langer, R. and Anderson, D. G. A combinatorial polymer library approach yields insight into nonviral gene delivery. *Acc. Chem. Res.* **41**, 749-759 (2008). https://doi.org:DOI 10.1021/ar7002336

40 Gormley, A. J., Yeow, J., Ng, G., Conway, O., Boyer, C. and Chapman, R. An Oxygen-Tolerant PET-RAFT Polymerization for Screening Structure-Activity Relationships. *Angew. Chem. Int. Ed.* **57**, 1557-1562 (2018). https://doi.org:10.1002/anie.201711044

41 Judzewitsch, P. R., Corrigan, N., Trujillo, F., Xu, J. T., Moad, G., Hawker, C. J., Wong, E. H. H. and Boyer, C. High-Throughput Process for the Discovery of Antimicrobial Polymers and Their Upscaled Production via Flow Polymerization. *Macromolecules* **53**, 631-639 (2020). https://doi.org:10.1021/acs.macromol.9b02207

42 Kosuri, S., Borca, C. H., Mugnier, H., Tamasi, M., Patel, R. A., Perez, I., Kumar, S., Finkel, Z., Schloss, R., Cai, L., Yarmush, M. L., Webb, M. A. and Gormley, A. J. Machine-Assisted Discovery of Chondroitinase ABC Complexes toward Sustained Neural Regeneration. *Adv Healthc Mater*, e2102101 (2022). https://doi.org:10.1002/adhm.202102101

43 Tamasi, M. J., Patel, R. A., Borca, C. H., Kosuri, S., Mugnier, H., Upadhya, R., Murthy, N. S., Webb, M. A. and Gormley, A. J. Machine Learning on a Robotic Platform for the Design of Polymer-Protein Hybrids. *Adv. Mater.*, e2201809 (2022). https://doi.org:10.1002/adma.202201809

44 Gauthier, M. A., Gibson, M. I. and Klok, H. A. Synthesis of Functional Polymers by Post-Polymerization Modification. *Angew. Chem. Int. Ed.* **48**, 48-58 (2009). https://doi.org:10.1002/anie.200801951

45 Gunay, K. A., Theato, P. and Klok, H. A. Standing on the shoulders of hermann staudinger: Post-polymerization modification from past to present. *Journal of Polymer Science Part a-Polymer Chemistry* **51**, 1-28 (2013). https://doi.org:10.1002/pola.26333

46 Zhong, Y. B., Zeberl, B. J., Wang, X. and Luo, J. T. Combinatorial approaches in post-polymerization modification for rational development of therapeutic delivery systems. *Acta Biomater.* **73**, 21-37 (2018). https://doi.org:10.1016/j.actbio.2018.04.010

47 Ladmiral, V., Mantovani, G., Clarkson, G. J., Cauet, S., Irwin, J. L. and Haddleton, D. M. Synthesis of neoglycopolymers by a combination of "click chemistry" and living radical polymerization. *J. Am. Chem. Soc.* **128**, 4823-4830 (2006). https://doi.org:10.1021/ja058364k

48 Wong, S. Y., Sood, N. and Putnam, D. Combinatorial Evaluation of Cations, pH-sensitive and Hydrophobic Moieties for Polymeric Vector Design. *Mol. Ther.* **17**, 480-490 (2009). https://doi.org:10.1038/mt.2008.293

49 Pedone, E., Li, X. W., Koseva, N., Alpar, O. and Brocchini, S. An information rich biomedical polymer library. *J. Mater. Chem.* **13**, 2825-2837 (2003). https://doi.org:10.1039/b306857a

50 Yan, Y. F., Liu, L., Xiong, H., Miller, J. B., Zhou, K. J., Kos, P., Huffman, K. E., Elkassih, S., Norman, J. W., Carstens, R., Kim, J., Minna, J. D. and Siegwart, D. J. Functional polyesters enable selective siRNA delivery to lung cancer over matched normal cells. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E5702-E5710 (2016). https://doi.org:10.1073/pnas.1606886113

51 Wyrsta, M. D., Cogen, A. L. and Deming, T. J. A parallel synthetic approach for the analysis of membrane interactive copolypeptides. *J. Am. Chem. Soc.* **123**, 12919-12920 (2001). https://doi.org:10.1021/ja017307y

52 Deming, T. J. Synthesis of Side-Chain Modified Polypeptides. *Chem. Rev.* **116**, 786-808 (2016). https://doi.org:10.1021/acs.chemrev.5b00292

53 Deming, T. J. Functional Modification of Thioether Groups in Peptides, Polypeptides, and Proteins. *Bioconjugate Chem.* **28**, 691-700 (2017). https://doi.org:10.1021/acs.bioconjchem.6b00696

54 Lu, H., Bai, Y., Wang, J., Gabrielson, N. P., Wang, F., Lin, Y. and Cheng, J. Ring-Opening Polymerization of γ-(4-Vinylbenzyl)-l-glutamate N-Carboxyanhydride for the Synthesis of Functional Polypeptides. *Macromolecules* **44**, 6237-6240 (2011). https://doi.org:10.1021/ma201164n

55 Zhou, J. R., Chen, P. P., Deng, C., Meng, F. H., Cheng, R. and Zhong, Z. Y. A Simple and Versatile Synthetic Strategy to Functional Polypeptides via Vinyl Sulfone-Substituted L-Cysteine N-Carboxyanhydride. *Macromolecules* **46**, 6723-6730 (2013). https://doi.org:10.1021/ma4014669

56 Engler, A. C., Lee, H. I. and Hammond, P. T. Highly Efficient "Grafting onto" a Polypeptide Backbone Using Click Chemistry. *Angew. Chem. Int. Ed.* **48**, 9334-9338 (2009). https://doi.org:10.1002/anie.200904070

57 Krannig, K. S. and Schlaad, H. pH-Responsive Bioactive Glycopolypeptides with Enhanced Helicity and Solubility in Aqueous Solution. *J. Am. Chem. Soc.* **134**, 18542-18545 (2012). https://doi.org:10.1021/ja308772d

58 Cao, J. B., Hu, P., Lu, L., Chan, B. A., Luo, B. H. and Zhang, D. H. Non-ionic water-soluble "clickable" alpha-helical polypeptides: synthesis, characterization and side chain modification. *Polym. Chem.* **6**, 1226-1229 (2015). https://doi.org:10.1039/c4py01560f

59 Liu, J., Chen, Q. Q. and Rozovsky, S. Utilizing Selenocysteine for Expressed Protein Ligation and Bioconjugations. *J. Am. Chem. Soc.* **139**, 3430-3437 (2017). https://doi.org:10.1021/jacs.6b10991

60 Zhao, Z. G., Shimon, D. and Metanis, N. Chemoselective Copper-Mediated Modification of Selenocysteines in Peptides and Proteins. *J. Am. Chem. Soc.* **143**, 12817-12824 (2021). https://doi.org:10.1021/jacs.1c06101

61 Quaderer, R., Sewing, A. and Hilvert, D. Selenocysteine-mediated native chemical ligation. *Helv. Chim. Acta* **84**, 1197-1206 (2001). https://doi.org:Doi 10.1002/1522-2675(20010516)84:5<1197::Aid-Hlca1197>3.0.Co;2-#

62 Li, X., Nelson, C. G., Nair, R. R., Hazlehurst, L., Moroni, T., Martinez-Acedo, P., Nanna, A. R., Hymel, D., Burke, T. R., Jr. and Rader, C. Stable and Potent Selenomab-Drug Conjugates. *Cell Chem Biol* **24**, 433-442 e436 (2017). https://doi.org:10.1016/j.chembiol.2017.02.012

63 Li, X. L., Patterson, J. T., Sarkar, M., Pedzisa, L., Kodadek, T., Roush, W. R. and Rader, C. Site-Specific Dual Antibody Conjugation via Engineered Cysteine and Selenocysteine Residues. *Bioconjugate Chem.* **26**, 2243-2248 (2015). https://doi.org:10.1021/acs.bioconjchem.5b00244

64 Sayers, J., Karpati, P. M. T., Mitchell, N. J., Goldys, A. M., Kwong, S. M., Firth, N., Chan, B. and Payne, R. J. Construction of Challenging Proline-Proline Junctions via Diselenide-Selenoester Ligation Chemistry. *J. Am. Chem. Soc.* **140**, 13327-13334 (2018). https://doi.org:10.1021/jacs.8b07877

65 Zhao, Z., Shimon, D. and Metanis, N. Chemoselective Copper-Mediated Modification of Selenocysteines in Peptides and Proteins. *J. Am. Chem. Soc.* **143**, 12817-12824 (2021). https://doi.org:10.1021/jacs.1c06101

66 Flemer, S., Jr. Selenol protecting groups in organic chemistry: special emphasis on selenocysteine Se-protection in solid phase peptide synthesis. *Molecules* **16**, 3232-3251 (2011). https://doi.org:10.3390/molecules16043232

67 Sharpless, K. B., Lauer, R. F. and Teranishi, A. Y. Electrophilic and nucleophilic organoselenium reagents. New routes to α,β-unsaturated carbonyl compounds. *J. Am. Chem. Soc.* **95**, 6137-6139 (1973). https://doi.org:10.1021/ja00799a062

68 Wu, J. A., Ding, C. L., Xing, D., Zhang, Z. B., Huang, X. F., Zhu, X. L., Pan, X. Q. and Zhu, J. The functionalization of poly(epsilon-caprolactone) as a versatile platform using epsilon-(alpha-phenylseleno) caprolactone as a monomer. *Polym. Chem.* **10**, 3851-3858 (2019). https://doi.org:10.1039/c9py00467j

69 Yu, L., Zhang, M., Du, F. S. and Li, Z. C. ROS-responsive poly(epsilon-caprolactone) with pendent thioether and selenide motifs. *Polym. Chem.* **9**, 3762-3773 (2018). https://doi.org:10.1039/c8py00620b

70 Wang, L., Zhu, K. X., Cao, W., Sun, C. X., Lu, C. J. and Xu, H. P. ROS-triggered degradation of selenide-containing polymers based on selenoxide elimination. *Polym. Chem.* **10**, 2039-2046 (2019). https://doi.org:10.1039/c9py00171a

71 Reich, H. J., Wollowitz, S., Trend, J. E., Chow, F. and Wendelborn, D. F. Syn Elimination of Alkyl Selenoxides - Side Reactions Involving Selenenic Acids - Structural and Solvent Effects on Rates. *J. Org. Chem.* **43**, 1697-1705 (1978). https://doi.org:DOI 10.1021/jo00403a016

72 Yang, Y., Song, H. P., He, D., Zhang, S., Dai, S. Z., Xie, X., Lin, S. X., Hao, Z. Y., Zheng, H. T. and Chen, P. R. Genetically encoded releasable photo-cross-linking strategies for studying protein-protein interactions in living cells. *Nat. Protoc.* **12**, 2147-2168 (2017).

73 Yang, Y., Song, H. P., He, D., Zhang, S., Dai, S. Z., Lin, S. X., Meng, R., Wang, C. and Chen, P. R. Genetically encoded protein photocrosslinker with a transferable mass spectrometry-identifiable label. *Nat. Commun.* **7**, 12299 (2016). https://doi.org:10.1038/ncomms12299

74 Lin, S., He, D., Long, T., Zhang, S., Meng, R. and Chen, P. R. Genetically encoded cleavable protein photo-cross-linker. *J. Am. Chem. Soc.* **136**, 11860-11863 (2014). https://doi.org:10.1021/ja504371w

75 Tian, Z. Y., Zhang, Z. C., Wang, S. and Lu, H. A moisture-tolerant route to unprotected alpha/beta-amino acid N-carboxyanhydrides and facile synthesis of hyperbranched polypeptides. *Nat. Commun.* **12** (2021). https://doi.org:ARTN 5810 10.1038/s41467-021-25689-y

76 Wu, G., Ge, C., Liu, X., Wang, S., Wang, L., Yin, L. and Lu, H. Synthesis of water soluble and multi-responsive selenopolypeptides via ring-opening polymerization of N-carboxyanhydrides. *Chem. Commun. (Camb.)* **55**, 7860-7863 (2019). https://doi.org:10.1039/c9cc03767e

77 Lin, Y. Y. A., Boutureira, O., Lercher, L., Bhushan, B., Paton, R. S. and Davis, B. G. Rapid Cross-Metathesis for Reversible Protein Modifications via Chemical Access to Se-Allyl-selenocysteine in Proteins. *J. Am. Chem. Soc.* **135**, 12156-12159 (2013).

78 Reddy, K. M. and Mugesh, G. Application of dehydroalanine as a building block for the synthesis of selenocysteine-containing peptides. *RSC Adv.* **9**, 34-43 (2019). https://doi.org:10.1039/c8ra09880h

79 Banik, S. M., Pedram, K., Wisnovsky, S., Ahn, G., Riley, N. M. and Bertozzi, C. R. Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature* **584**, 291-297 (2020). https://doi.org:10.1038/s41586-020-2545-9

80 Ahn, G., Banik, S. M., Miller, C. L., Riley, N. M., Cochran, J. R. and Bertozzi, C. R. LYTACs that engage the asialoglycoprotein receptor for targeted protein degradation. *Nat. Chem. Biol.* **17**, 937-946 (2021). https://doi.org:10.1038/s41589-021-00770-1

81 Delaveris, C. S., Chiu, S. H., Riley, N. M. and Bertozzi, C. R. Modulation of immune cell reactivity with cis-binding Siglec agonists. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021). https://doi.org:ARTN e2012408118 10.1073/pnas.2012408118

82 Du, J. J., Xin, L. M., Lei, Z., Zou, S. Y., Xu, W. B., Wang, C. W., Zhang, L., Gao, X. F. and Guo, J. Glycopeptide ligation via direct aminolysis of selenoester. *Chin. Chem. Lett.* **29**, 1127-1130 (2018). https://doi.org:10.1016/j.cclet.2018.04.016

83 Temperini, A., Piazzolla, F., Minuti, L., Curini, M. and Siciliano, C. General, Mild, and Metal-Free Synthesis of Phenyl Selenoesters from Anhydrides and Their Use in Peptide Synthesis. *J. Org. Chem.* **82**, 4588-4603 (2017). https://doi.org:10.1021/acs.joc.7b00173

84 Chen, L. H., Pilania, G., Batra, R., Huan, T. D., Kim, C., Kuenneth, C. and Ramprasad, R. Polymer informatics: Current status and critical next steps. *Materials Science & Engineering R-Reports* **144** (2021). https://doi.org:ARTN 100595 10.1016/j.mser.2020.100595

85 Barbosa, N. V., Nogueira, C. W., Nogara, P. A., de Bem, A. F., Aschner, M. and Rocha, J. B. T. Organoselenium compounds as mimics of selenoproteins and thiol modifier agents. *Metallomics* **9**, 1703-1734 (2017). https://doi.org:10.1039/c7mt00083a

86 Huang, X., Liu, X. M., Luo, Q. A., Liu, J. Q. and Shen, J. C. Artificial selenoenzymes: Designed and redesigned. *Chem. Soc. Rev.* **40**, 1171-1184 (2011).

87 Seibt, T. M., Proneth, B. and Conrad, M. Role of GPX4 in ferroptosis and its pharmacological implication. *Free Radic. Biol. Med.* **133**, 144-152 (2019). https://doi.org:10.1016/j.freeradbiomed.2018.09.014

88 Yant, L. J., Ran, Q. T., Rao, L., Van Remmen, H., Shibatani, T., Belter, J. G., Motta, L., Richardson, A. and Prolla, T. A. The selenoprotein GPX4 is essential for mouse development and protects from radiation and oxidative damage insults. *Free Radic. Biol. Med.* **34**, 496-502 (2003). https://doi.org:10.1016/S0891-5849(02)01360-6

89 Xu, C. X., Sun, S. G., Johnson, T., Qi, R., Zhang, S. Y., Zhang, J. and Yang, K. The glutathione peroxidase Gpx4 prevents lipid peroxidation and ferroptosis to sustain Treg cell activation and suppression of antitumor immunity. *Cell Rep.* **35** (2021). https://doi.org:ARTN 109235 10.1016/j.celrep.2021.109235

90 Parnham, M. and Sies, H. Ebselen: prospective therapy for cerebral ischaemia. *Expert Opinion on Investigational Drugs* **9**, 607-619 (2000). https://doi.org:Doi 10.1517/13543784.9.3.607

91 Landgraf, A. D., Alsegiani, A. S., Alaqel, S., Thanna, S., Shah, Z. A. and Sucheck, S. J. Neuroprotective and Anti-neuroinflammatory Properties of Ebselen Derivatives and Their Potential to Inhibit Neurodegeneration. *ACS Chem. Neurosci.* **11**, 3008-3016 (2020). https://doi.org:10.1021/acschemneuro.0c00328

92 Yamagata, K., Ichinose, S., Miyashita, A. and Tagami, M. Protective effects of ebselen, a seleno-organic antioxidant on neurodegeneration induced by hypoxia and reperfusion in stroke-prone spontaneously hypertensive rat. *Neuroscience* **153**, 428-435 (2008). https://doi.org:10.1016/j.neuroscience.2008.02.028

93 Paglia, D. E. and Valentine, W. N. Studies on the quantitative and qualitative characterization of erythrocyte glutathione peroxidase. *J. Lab. Clin. Med.* **70**, 158-169 (1967).

94 Shahriari, B., Swersky, K., Wang, Z. Y., Adams, R. P. and de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the Ieee* **104**, 148-175 (2016). https://doi.org:10.1109/Jproc.2015.2494218

95 Shields, B. J., Stevens, J., Li, J., Parasram, M., Damani, F., Alvarado, J. I. M., Janey, J. M., Adams, R. P. and Doyle, A. G. Bayesian reaction optimization as a tool for

chemical synthesis. *Nature* **590**, 89-96 (2021). https://doi.org:10.1038/s41586-021-03213-y

96 Nambiar, A. M. K., Breen, C. P., Hart, T., Kulesza, T., Jamison, T. F. and Jensen, K. F. Bayesian Optimization of Computer-Proposed Multistep Synthetic Routes on an Automated Robotic Flow Platform. *ACS Cent. Sci.* **8**, 825-836 (2022). https://doi.org:10.1021/acscentsci.2c00207

97 Hase, F., Roch, L. M., Kreisbeck, C. and Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **4**, 1134-1145 (2018). https://doi.org:10.1021/acscentsci.8b00307

98 Balandat, M., Karrer, B., Jiang, D., Daulton, S., Letham, B., Wilson, A. G. and Bakshy, E. BoTorch: a framework for efficient Monte-Carlo Bayesian optimization. *Adv. Neural Inf. Process. Syst.* **33**, 21524-21538 (2020).