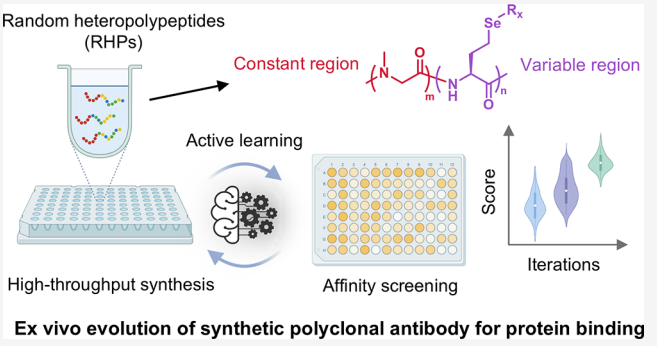
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Data-Driven Design of Random Heteropolypeptides as Synthetic Polyclonal Antibodies

Haisen Zhou, Guangqi Wu,* Zuo Zhang, Zhen Zhu, Tanfeng Zhao, Qiankang Zhou, Haolin Chen, Zhenzhong Zhao, Yuanhao Dai, Xiaodong Jing, Wei Ji, Xuehai Yan, Lijiang Yang, Yiqin Gao, Huan Wang,* and Hua Lu*

Cite This: https://doi.org/10.1021/jacs.5c06240

ABSTRACT: Antibodies are indispensable in biomedicine. However, conventional antibody development faces challenges, including high costs and lengthy production timelines spanning months. Here, we present a data-driven workflow for engineering random heteropolypeptides (RHPs) as synthetic polyclonal antibodies (SpAbs) with programmable binding properties. By combining high-throughput synthesis of selenopolypeptide derivatives with algorithm-assisted optimization, we rapidly identified SpAbs targeting human interferon-α (IFN) and tumor necrosis factor-α (TNF-α) within 2 weeks. The SpAbs exhibited binding affinities comparable to natural antibodies, with the top candidate achieving a dissociation constant of 7.9 nM for TNF-α and 418-fold selectivity over human serum albumin, effectively neutralizing TNF-α-induced cytotoxicity. Liquid-phase electron microscopy revealed flexible, intrinsically disordered protein-like conformations and folding-upon-binding dynamics. This study establishes a robust framework for SpAb discovery, demonstrating that sequence-independent RHPs can serve as functional antibody mimics with tunable binding properties, rapid optimization, and broad potential in diagnostics and therapeutics.



Ex vivo evolution of synthetic polyclonal antibody for protein binding

INTRODUCTION

Molecular recognition is essential for many vital biological processes such as signal transduction, immune responses, and enzyme catalysis.[1−3] Antibodies, the most widely used biomolecules for molecular recognition, play critical roles and possess an immense global market in diagnostics, therapeutics, and biomedical research.[4] Peptides and aptamers, produced by solid phase synthesis, are synthetic counterparts of antibodies for molecular recognition.[5−8] However, all of them face high costs, long development cycles that typically span several months, and limited stability under harsh conditions.[9−11] These challenges have spurred the pursuit of artificial alternatives to antibodies that offer reduced costs, a faster production time, and enhanced stability. Pioneering studies on molecularly imprinted polymers (MIPs), which employed a template-directed strategy to form structurally complementary binding sites within a cross-linked polymer network, provided a cost-effective solution with modest affinity for the template.[12−16] However, the limitations of MIPs—including the absence of affinity maturation mechanisms, difficulties in optimizing selectivity, and challenges in imprinting and removing large protein templates—have spurred the exploration of alternative approaches.[13]

Conventional strategies rely on macromolecular binders with precise sequences and more ordered structures to achieve strong, selective molecular recognition. In contrast, we propose that synthetic antibodies can be engineered using the ensemble properties of polymers—particularly random heteropolypeptides (RHPs)—bypassing the need for exact sequence control or rigid structural design. One rare but inspiring example is glatiramer acetate, a synthetic polypeptide composed of four amino acids with random sequences made by the copolymerization of four different amino acids, which has been developed as a blockbuster drug for treating multiple sclerosis.[17] Designed to emulate the composition of myelin antigens, glatiramer acetate is found to compete with myelin for binding to major histocompatibility complex class II and stimulate the production of Th2 type suppressor cells.[18,19] Recent studies have also demonstrated that random heteropolymers can perform functions akin to those of proteins,[20]

Received: April 13, 2025
Revised: May 24, 2025
Accepted: May 27, 2025

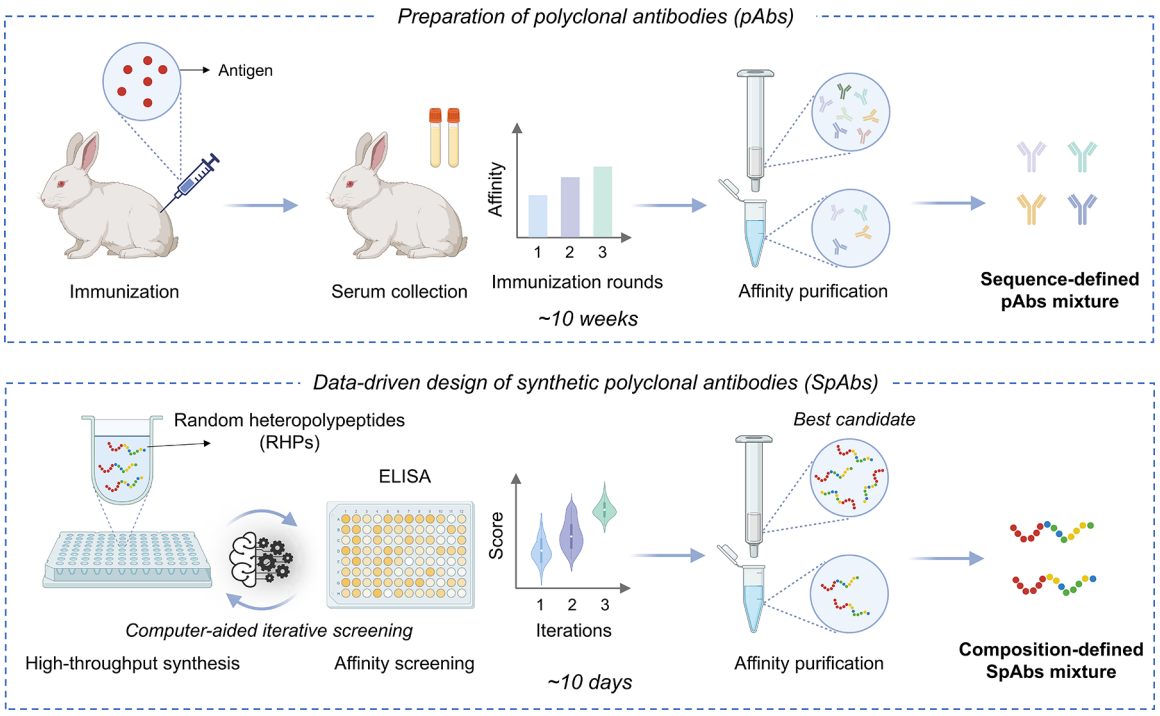


Figure 1. Analogy of polyclonal antibody (pAb) production and the data-driven design of random heteropolypeptides (RHPs) as synthetic polyclonal antibodies (SpAbs). The top panel illustrates the pAb production process, including animal immunization, serum collection, and affinity purification, to obtain antibodies targeting diverse epitopes. The bottom panel outlines the SpAb design workflow, incorporating high-throughput synthesis, quantitative binding evaluation, and optimization algorithms to identify RHPs with enhanced and selective binding toward the target protein. The highly specific SpAb can be purified via affinity chromatography from the best-performing candidate to further enhance binding affinity, demonstrating their potential as cost-effective and stable alternatives to pAbs. Created in BioRender.

such as catalysis,[21] proton transport,[22] and mimicking cytosol components.[23]

RHPs share inherent structural heterogeneity with polyclonal antibodies—mixtures of diverse antibodies that collectively target multiple epitopes. While pAbs are produced through animal immunization, somatic hypermutation, and affinity-based purification (Figure 1, top panel),[24,25] RHPs could theoretically achieve similar antigen selectivity without the expensive and tedious biological processes. However, the vast design space of RHPs, such as degree of polymerization (DP), dispersity ($Đ$), and monomer composition, poses challenges to engineer RHPs with tailored properties.[26] Recent advances combining high-throughput experimentation with machine learning have enabled accelerated functional materials discovery.[27−44] The closed-loop workflows (integrating synthesis, screening, and active learning) have successfully identified functional polymers for applications ranging from antimicrobial agents,[45] delivering materials,[46] protein enzyme stabilizers,[47] fMRI reagents,[48] and drug excipients.[49]

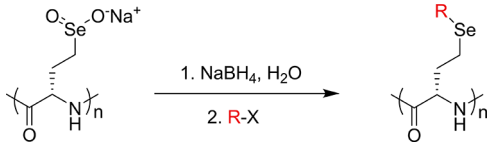
In this study, we demonstrate that synthetic polyclonal antibodies (SpAbs) identified from the RHP library achieve antigen-binding affinities and selectivities rivaling those of natural pAbs. To do this, we develop a novel workflow combines high-throughput synthesis (HTS), quantitative binding assessments using enzyme-linked immunosorbent assay (ELISA), and algorithm-driven affinity evolution using Bayesian optimization (BO) and genetic algorithm (GA) (Figure 1, bottom panel), mirroring in vivo development of pAbs. The effectiveness of the workflow was validated using two model antigens, namely, human interferon-α (IFN) and tumor necrosis factor-α (TNF-α). Remarkably, one selected SpAb_T1 possessed a dissociation constant (K_D) of 7.9 nM on TNF-α, ~418-fold selectivity over the control protein human serum albumin (HSA), and considerable neutralization ability in cellular assays. Liquid-phase electron microscopy (LP-EM) captures the flexible conformation of SpAb and the single-molecule binding dynamics of the SpAb-TNF-α pair. Notably, the time frame for the SpAb discovery takes only 14 days, bypassing the need of animal immunization. Given the critical role of pAbs in immunoassays and clinical treatments,[50] the rapid, scalable, and animal-free production of SpAbs could offer a more efficient and ethical approach to antibody development, expanding their potential of SpAbs in biomedical research and clinical use.

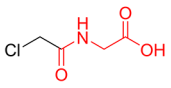
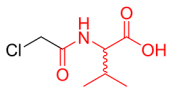
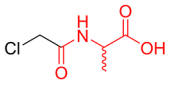
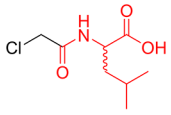
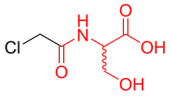
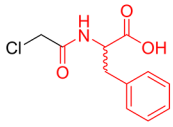
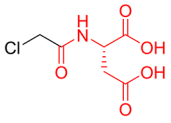
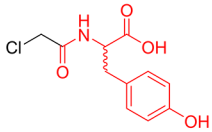
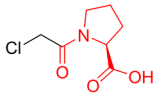
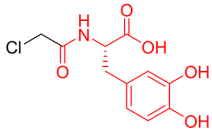
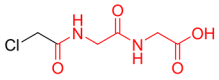
RESULTS

Synthesis of RHP Library and Regulation of Secondary Structures. To build the RHP library, we tried the modification of a selenopolypeptide precursor $PSeO_2Na$ with over 10 different chloroacetyl amino acids (Table 1). All of the modifiers gave anionic water-soluble polypeptides with high grafting efficiency under optimized conditions (Supplementary Figures S1−S11), including those with a DOPA (3,4-dihydroxyphenyl $_L$-alanine) moiety featuring two phenolic hydroxyl groups (entry 11, Table 1) or dipeptides such as glycyl-glycine (entry 6, Table 1). The introduction of different modification groups led to distinct secondary structures in the polypeptides (Table 1, Supplementary Figure S12). Specifically, hydrophilic amino acid side chains tended to offer random coil structures, while more hydrophobic side chains progressively shifted the secondary structure toward α-helices with increasing helicity. The results indicated that, by

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Table 1. Post-Polymerization Modification of $PSeO_2Na$ with Various *N*-Chloroacetyl Amino Acids



	Modifier R-X	Product (GE*)	Helicity**		Modifier R-X	Product (GE*)	Helicity**
1		PSe-Gly (>95%)	4%	7		PSe-Val (>95%)	65%
2		PSe-Ala (>95%)	6%	8		PSe-Leu (>95%)	80%
3		PSe-Ser (>95%)	4%	9		PSe-Phe (>95%)	108%
4		PSe-Asp (>95%)	14%	10		PSe-Tyr (>95%)	91%
5		PSe-Pro (>95%)	64%	11		PSe-Dopa (>90%)	ND***
6		PSe-GlyGly (>95%)	17%				

*GE = grafting efficiency, quantified by ^{1}H NMR spectroscopy. **Helicity was calculated based on the molar ellipticity at 222 nm ($[\theta_{222}]$) from the CD spectrum. Helicity % = $((-[\theta_{222}] + 3000)/39000) \times 100\%$. Test condition: 0.1−0.25 mg/ml polypeptides (DP = 70) in H_2O, pH 7.0. ***ND = not determined.

incorporating a diverse array of side groups, tailored secondary structures that closely mimic protein-like characteristics could be achieved.

The ionizable carboxyl groups on the side chains render the polypeptide conformation sensitive to pH changes. For example, the circular dichroism (CD) spectroscopy depicted a conformational change of PSe-Leu (entry 8, Table 1) from a typical α-helical secondary structure at neutral pH to a β-sheet structure precipitable as the pH further decreased to 3 or lower (Supplementary Figure S13). PSe-Tyr, with a phenolic hydroxyl group, adopted a α-helical structure at neutral pH (entry 10, Table 1) and shifted to a random coil and a β-sheet conformation by increasing and decreasing the pH, respectively (Supplementary Figure S14).

The redox sensitivity of selenoether introduced another dimension for secondary structure control. For example, the conformation of PSe-Leu (entry 8, Table 1) switched from an α-helix to a random coil when the selenide was oxidized to selenoxide. Reduction of the selenoxide restored the structure to α-helix (Supplementary Figure S15). Similar transformations were also observed for PSe-Phe (entry 9, Table 1, and Supplementary Figure S16).

Workflow for Screening RHPs as SpAbs. We designed the precursor copolymer taking inspiration from antibodies that usually comprise a constant region for stabilization and solubility and a variable region responsible for antigen binding. The block copolymer, PSar-*b*-$PSeO_2Na$, contains of a polysarcosine (PSar) block for enhancing the overall solubility and minimizing nonspecific interactions (constant region), and a variable $PSeO_2Na$ block for introducing diversity (Figure 2a). The block copolymer precursor was characterized with size-exclusion chromatography (SEC) and proton nuclear magnetic resonance (^{1}H NMR) spectroscopy (Supplementary Figures S17 and S18). The terminal of PSar-*b*-$PSeO_2Na$ was capped with a biotin for the affinity assay with ELISA.

Five modifiers from Table 1: Gly (R_1), Ser (R_2), Phe (R_3), Leu (R_4), and Val (R_7) were used to construct the RHP library. These modifiers were chosen to represent a broad spectrum of hydrophilic and hydrophobic side chains, enabling various interactions such as hydrogen bonding, hydrophobic interactions, and $\pi-\pi$ stacking. To introduce charge diversity, we supplemented the selected amino acids with three additional modifiers, which include a neutral chloroacetamide (R_5), a positively charged *N*-(2-aminoethyl)-2-chloroacetamide (R_6), and an ionizable 4-(chloromethyl)-1*H*-imidazole (R_8) with a p*K*a near 7 (Figure 2b). All modifiers are water-soluble and exhibit similarly high reactivity, enabling quantitative reaction with selenolate at 1.2 equivalent of total modifiers.

A closed-loop workflow that integrates build, test, and design was implemented to optimize the ratios of various modifiers and accelerate SpAbs discovery (Figure 2c). RHPs were synthesized by reacting the reduced precursor polypeptide with a premixed aqueous solution of the modifiers, prepared using a liquid handling workstation for precise control and consistency. The binding affinity of the RHPs to the target antigen was evaluated by using indirect ELISA (Supplementary Scheme S1). HSA, the most abundant protein in serum, were used as the negative control. The goal of the screening assay was to identify RHPs with high specificity for the target antigen while reducing off-target to HSA. Two algorithms,

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

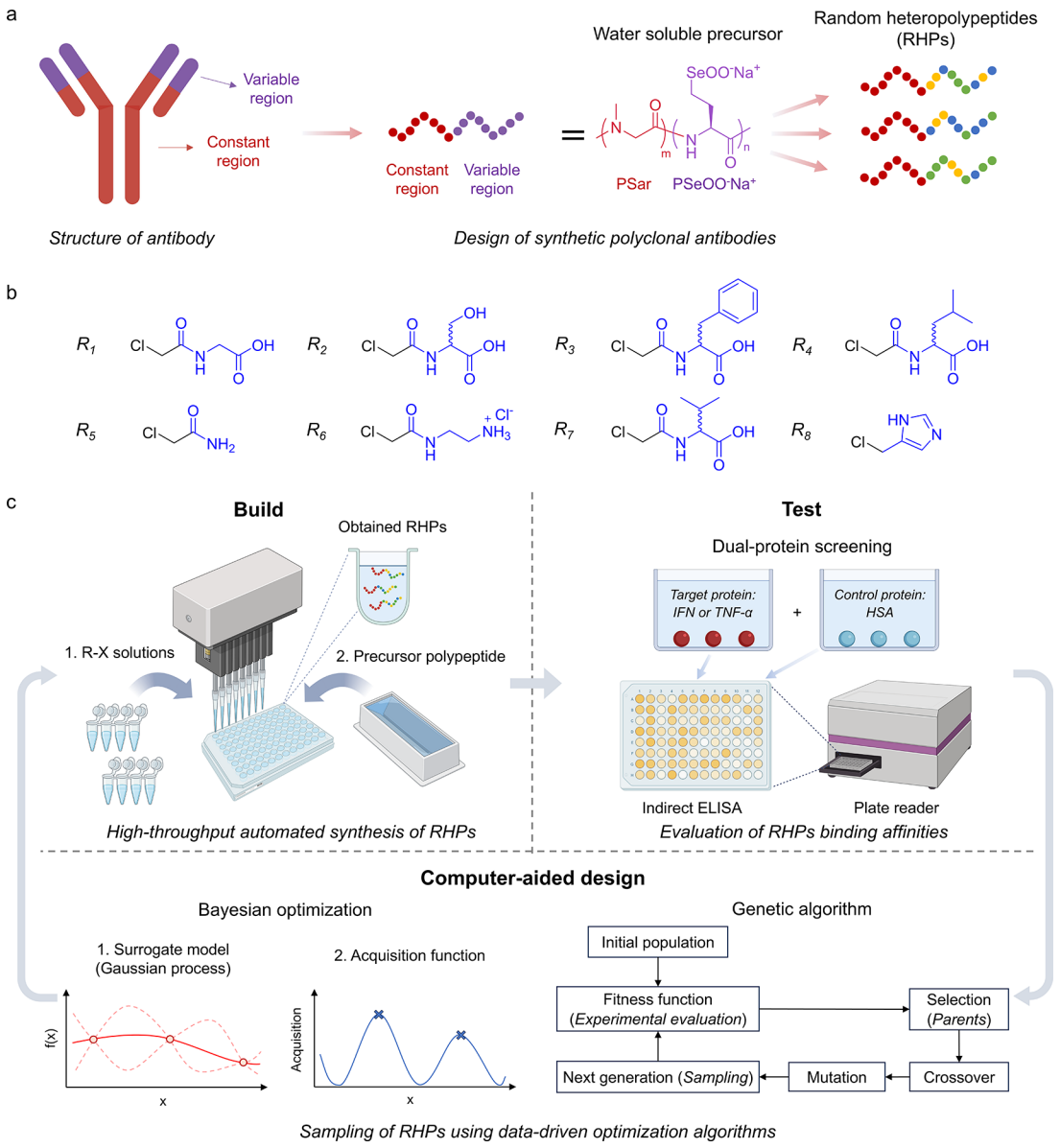


Figure 2. (a) Design of the precursor polypeptide mimicking structures and functions of the antibody. The block copolymer contains a modifiable selenium-containing polypeptide to mimic variable-region mimicry and polysarcosine to mimic constant region. (b) Selected modifier structures for RHP synthesis. (c) The close-looped workflow including build, test, and design for accelerated discovery of RHPs with selective protein binding abilities. Icons created in BioRender.

namely, the model-based strategy BO and the computationally efficient model-free GA strategy that simulates natural selection, were introduced to refine the sampling of candidates during the data-driven optimization process. The ELISA results were transformed into a single target, *Score*, that reflects selective binding ability of SpAb (see Methods).

Workflow Validation Using IFN as the Model Antigen. IFN, an immune regulatory cytokine, was employed as a model target antigen to validate the workflow. The screening process started with 384 randomly sampled candidates and was optimized for six iterations. For simplicity, the *Score* in this validation campaign was defined as the difference value of *Target* and *Control*. Each optimization iteration contains 30 random points, 30 points generated by BO, 30 points generated by GA, and 6 replicates of the highest *Score* points from the last iteration. Both GA and BO demonstrated superior performance in identifying RHPs with a stronger selective binding affinity for IFN compared to random sampling (Figure 3a). As iterations progressed, sampled points increasingly clustered within region IV, which is associated with stronger binding to the target (Figure 3b) compared to the HSA. From all data set, the top six scoring candidates were selected for large-scale batch synthesis and further validated using ELISA (Supplementary Table S1), with their performances compared to that of control polypeptides (Homo_1−Homo_8) that are modified by just single modifiers. All six selected SpAbs exhibited a higher *Score* than the control polypeptides (Figure 3c). To further validate the binding affinity and selectivity of the hits, the binding to IFN and HSA for one of the selected candidates, SpAb_I2, the SpAb with highest *Target* value, was tested using biolayer interferometry (BLI). SpAb_I2 exhibited a K_D of 103 nM for IFN and negligible binding to HSA (Figure 3d). These results indicated that the workflow successfully optimized the RHP

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

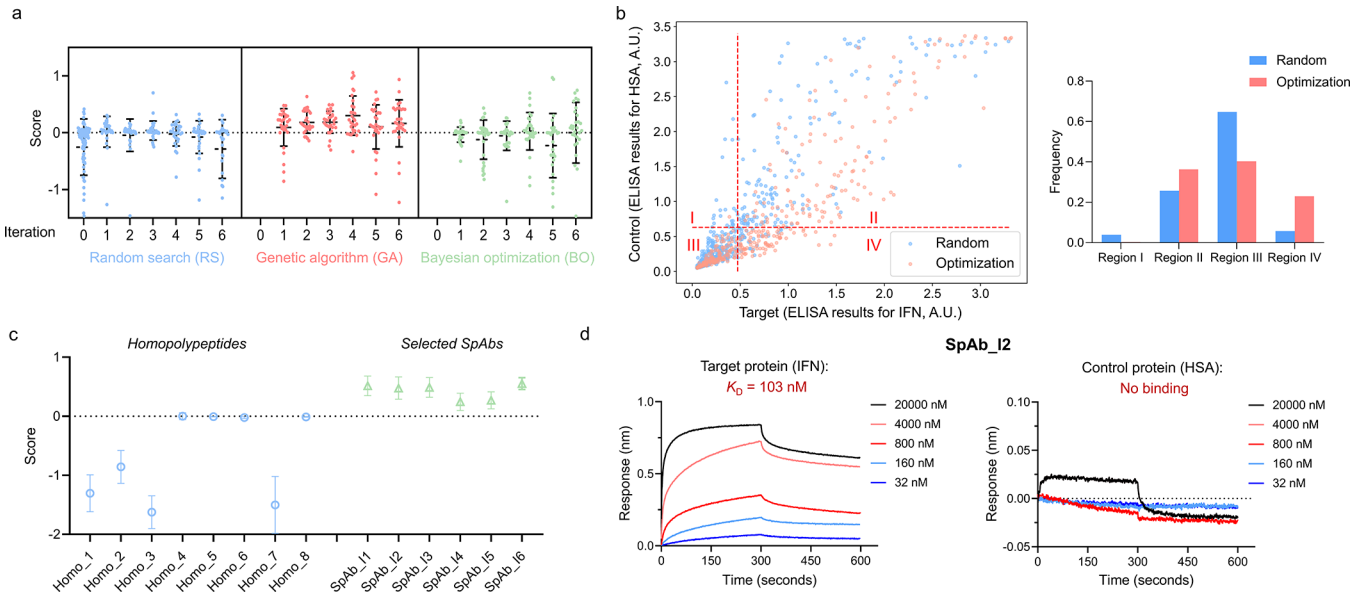


Figure 3. Workflow validation using IFN as a model target protein. (a) Performance of RHPs in each iteration via random search (RS, blue), genetic algorithm (GA, red), and Bayesian optimization (BO, green). (b) ELISA value distribution of all randomly sampled points (random) and points generated by GA and BO (optimization). The intersection points of the four quadrants are set as the average value of all random points for *Target* (0.471) and the average value of all random points for *Control* (0.630). (c) Validation of the top six SpAbs (SpAb_I1−SpAb_I6) compared to homopolypeptides (Homo_1−Homo_8) by ELISA. (d) Binding affinity validation of SpAb_I2 for IFN and HSA using BLI.

compositions to enhance selective binding to the target protein.

Data-Driven Optimization of RHP for Targeting Tumor Necrosis Factor α (TNF-α). We then focused on screening SpAbs for TNF-α, a key mediator and an important therapeutic target in many inflammatory diseases. Several anti-TNF-α monoclonal antibodies, such as adalimumab[51] and infliximab,[52] were developed to combat rheumatoid arthritis (RA) and other autoimmune disorders. We wish to identify SpAb compositions that can selectively bind and blocking its interaction with TNF-α receptor (TNFR), and also mitigating inflammation through leveraging the anti-inflammatory and antioxidant properties of selenium.[53]

Using the same workflow as mentioned earlier, we conducted six rounds of screening with slightly modified *Score* for data standardization (see Methods), starting with an expanded set of 524 randomly sampled points to refine the baseline mean and enhance robustness. Significant performance improvements were observed with GA and BO compared to random sampling from rounds 3−6 (Figure 4a). The *Scores* of the top six candidates increased steadily with each iteration (Supplementary Figure S19). Analyzing the origins of the top 20 candidates revealed that early iterations (1−2) were dominated by random search contributions, while later iterations (3−6) witnessed GA, followed by BO, emerge as the primary contributors (Figure 4b). Component distribution analysis (Figure 4c) showed that from iterations 3 to 6, GA exhibited noticeable convergence, favoring regions dominated by R_1, R_3, and R_6. In contrast, BO remained more exploratory but still outperformed the random search.

We retrospectively analyzed the contributions of different modifiers. A linear regression was first applied to predict the outcomes, but only achieve R^2 values of 0.70, 0.75, and 0.25 on test set for *Target* (TNF-α), *Control* (HSA), and *Score*, respectively (Supplementary Figure S20). Then, a Gaussian regression model was applied to predict the outcomes. The optimized model achieved R^2 values of 0.96, 0.97, and 0.69 on the training set for *Target*, *Control*, and *Score*, respectively (Supplementary Figure S21). The same R^2 values on the test set were 0.89, 0.90, and 0.40, respectively (Figure 4d and Supplementary Figure S22). While the model performed well for predicting ELISA results (*Target* and *Control*), the *Score* prediction was less accurate, likely due to cumulative noise. SHAP (SHapley Additive exPlanations) analysis identified R_6 and R_8 as the most influential components on *Score* (Figure 4e,f). Increasing the R_6 content, featuring a positively charged side chain (lysine mimic), was associated with improved *Score*, while higher R_8 content, containing an imidazole group (histidine mimic), negatively impacted the *Score*.

Validation and Optimization of RHPs as Functional SpAbs to Neutralize TNF-α. The best-performing RHP, SpAb_T1, was synthesized in batch and validated using BLI. Several RHPs with lower scores were also synthesized as controls. SpAb_T1 showed a strong affinity for TNF-α, with a K_D of 7.9 nM, and much weaker binding to HSA (K_D = 3.3 μM, Figure 5a), which resulted in an estimated selectivity for TNF-α over HSA of around 418-fold. In contrast, a lower-performing candidate, RHP_T2, showed binding constants of 413 nM for TNF-α and 2.1 μM for HSA (Figure 5b).

SpAbs can undergo affinity purification to enrich populations with improved affinity and selectivity within an RHP of fixed composition. Briefly, SpAb_T1 was first passed through a column with HSA-immobilized resin, and the flow-through was then incubated with TNF-α-immobilized resin to enrich RHPs with high affinity to TNF-α (Figure 5c). The purified SpAb_T1, with a yield of 18%, showed significantly reduced nonspecific binding and enhanced TNF-α binding affinity, achieving a K_D of less than 1.6 nM compared to the original 7.9 nM before purification. As a comparison, we evaluated the affinity of a commercial rabbit polyclonal anti-TNF-α antibody. The K_D for TNF-α binding was measured as 18.8 nM, and

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

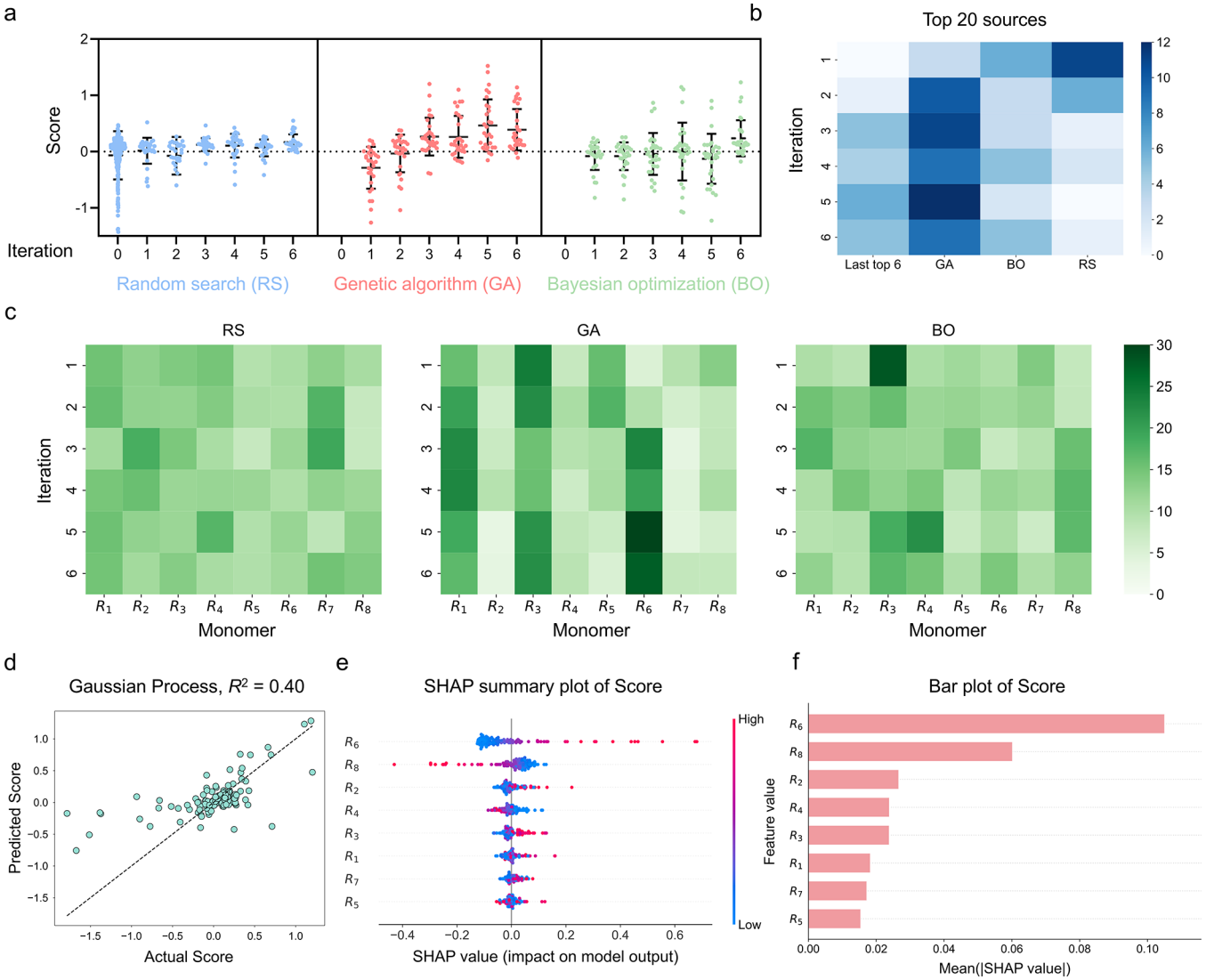


Figure 4. Data-driven optimization workflow for identifying SpAbs targeting TNF-α. (a) Performance of RHPs in each iteration via RS (blue), GA (red), and BO (green). (b) Heatmap of origins of the top 20 candidates in each iteration. (c) Heatmap of component distribution across iterations with different optimization methods. (d) Performance of the Gaussian regression model for predicting *Score* on the test set, with a test set size of 0.1. (e) SHAP (SHapley Additive exPlanations) value analysis for each modifier. Each dot represents the influence of a modifier on the prediction of an individual RHP's *Score*, with colors indicating the level of feature value from high (pink) to low (blue). (f) Bar plot visualizing the impact of individual modifiers on the *Score*, with higher values corresponding to greater influence on the result.

there was no detectable binding to HSA under the same condition (Supplementary Figure S23).

Next, we examined whether the anti-TNF-α SpAbs possessed a neutralizing ability besides binding to the antigen. Specifically, TNF-α induces cytotoxicity in mouse fibroblast L929 cells in the presence of the metabolic inhibitor actinomycin D. SpAb_T1 was found to reduce TNF-induced toxicity in L929 cells in a dose-dependent manner, with greater toxicity reductions observed for higher concentrations of SpAb_T1 (Figure 5d). The toxicity-reduction ability was also associated with the affinity of SpAbs to TNF-α. Affinity purification of SpAb_T1 significantly enhanced its neutralizing capability, reducing the IC_{50} from 70 to 272 pg/mL, whereas the randomly sampled RHP_T2 exhibited negligible toxicity reduction (IC_{50} = 0.37 pg/mL, comparable to TNF-α alone; Figure 5e).

Observation of the Binding Process of SpAb_T1 and TNF-α. CD analysis revealed that upon binding of TNF-α to SpAb_T1, the resulting helicity was higher than the linear combination of the individual components (Figure 6a), implying that the overall helicity increased upon binding. The results suggested that the interaction between SpAb_T1 and TNF-α may resemble the folding-upon-binding characteristic observed in IDPs.[54,55] Dynamic light scattering (DLS) measurements showed an increase in hydrodynamic size following the binding of SpAb_T1 to TNF-α (Supplementary Figure S24), indicating substantial associations occurred between the molecules.

To investigate the binding dynamics in more detail, we employed liquid-phase electron microscopy (LP-EM) to directly observe the conformational dynamics and binding process at the single-molecule level in situ.[56,57] We first characterized the single-molecule conformation dynamics of SpAb_T1 and TNF-α separately. Both molecules were individually observed in graphene liquid cell using transmission electron microscopy (TEM), where SpAb_T1 displayed a

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

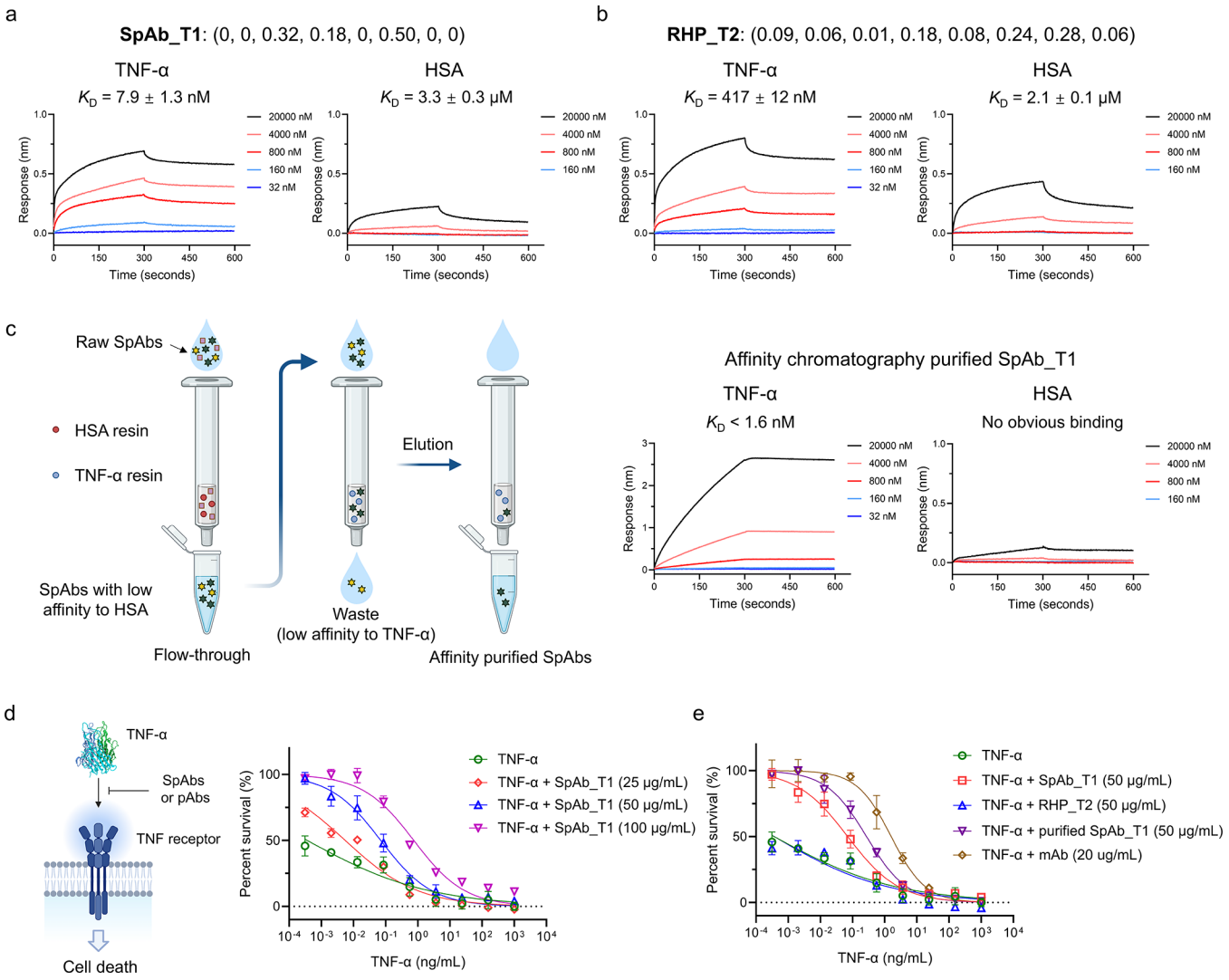


Figure 5. Validation and optimization of RHPs as functional pAb mimics. BLI validation of the highest-scoring SpAb_T1 (a) and randomly sampled RHP_T2 (b) binding affinities for TNF-α and HSA, respectively. (c) Schematic of affinity purification of RHPs using TNF-α or HSA-immobilized resins, resulting in purified SpAb_T1 with enhanced binding affinity and selectivity for TNF-α. (d) Neutralization of TNF-α-induced cytotoxicity in L929 cells by SpAb_T1 with varying concentrations. (e) Neutralization of TNF-α-induced cytotoxicity in L929 cells by different RHPs with varying binding affinities. The IC_{50} for TNF-α, TNF-α + SpAb_T1, TNF-α + RHP_T2, TNF-α + purified SpAb_T1, and TNF-α + anti-TNF-α monoclonal antibody (mAb) are 0.36, 70, 0.37, 272, and 1591 pg/mL, respectively. Icons created in BioRender.

more flexible conformational dynamic than TNF-α (Supplementary Movie S1). We calculated R_g* as the approximate radii of gyration (R_g) of the molecules using projected 2D LP-EM images as a quantitative characterization of conformational dynamics (method see the Supporting Information (SI)). Proteins with fixed conformations exhibited narrower R_g* distributions during the observation time window, while SpAb_T1, resembling IDPs, displayed slightly broader distributions (Figure 6b,c). The coefficient of variation of R_g* in SpAb_T1 was significantly higher than that of TNF-α, further supporting the notion that SpAb_T1 adopts a conformation that is more flexible than that of TNF-α.

Next, we examined the binding dynamics between SpAb_T1 and TNF-α. A representative binding event showed the approach, contact, and binding stages (Figure 6d and Supplementary Movie S2). The two molecules approached each other at an initial distance of about 12 nm, contacted each other at 14 s, and finally formed a TNF-α/SpAb_T1 complex. The R_g* of both molecules remained stable before contact, and the R_g* of the THF-α/SpAb_T1 complex decreased first after binding and then reached a stable value (Figure 6d), reflecting the conformational fitting of SpAb_T1 upon binding to TNF-α. Although it is challenging to distinguish between SpAb_T1 and TNF-α in mixed samples due to their similar size, data shows that in mixed samples, molecules exhibited significantly more binding events compared with separate samples (Supplementary Figure S25), which confirms that the observed interactions were much more likely to be heterologous SpAb_T1/TNF-α binding, rather than self-interactions. Throughout all the LP-EM experiments of TNF-α/SpAb_T1 interactions, we observed only one-to-one binding events, with no evidence of multivalent interactions. These results demonstrate the dynamic interaction occurred between SpAb_T1 and TNF-α, indicating that RHPs exhibit flexible conformation similar to IDPs, facilitating biomacromolecular interactions.

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

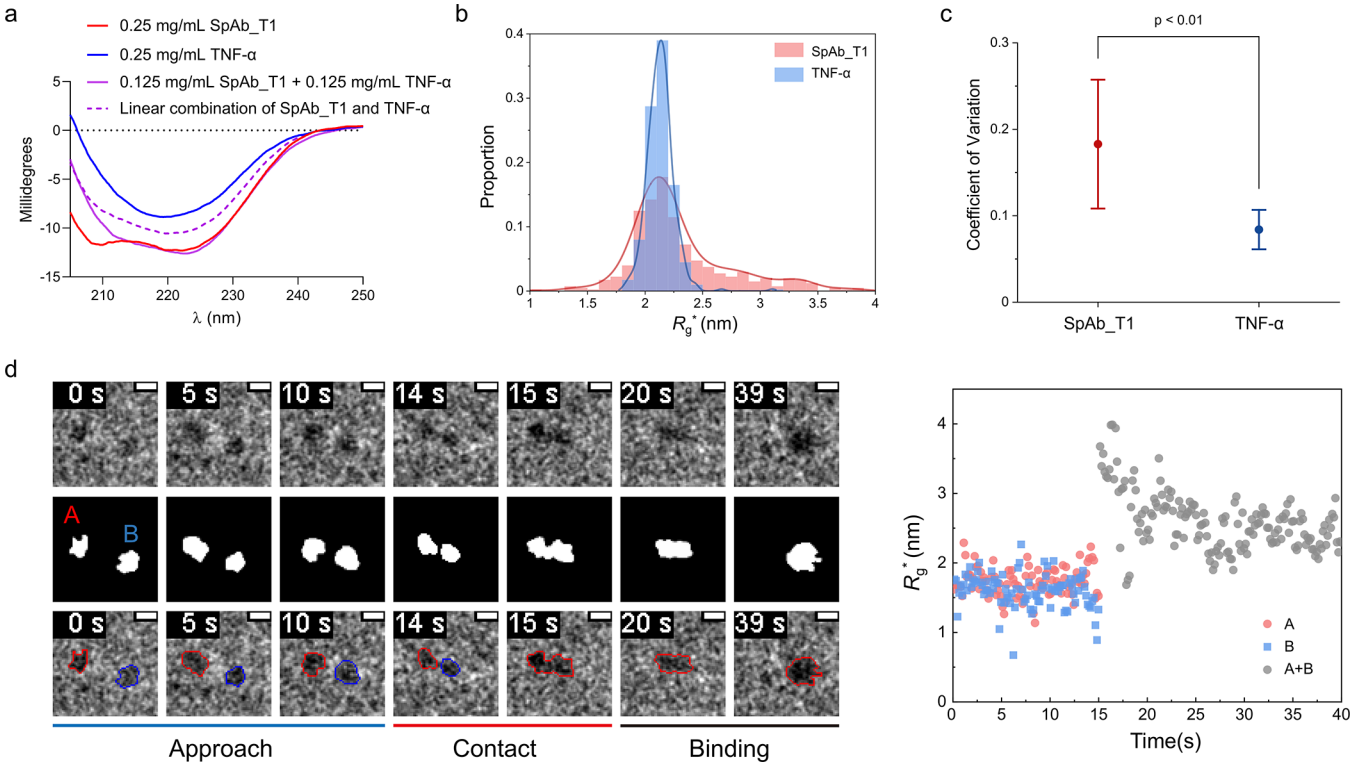


Figure 6. Observation of the binding processes of SpAb_T1 and TNF-α. (a) CD spectra of 0.25 mg/mL SpAb_T1, 0.25 mg/mL TNF-α, and their equal volume mixed solution in PBS buffer. Dashed curve showed linear combination of spectra of 0.25 mg/mL RHP_T1 and 0.25 mg/mL TNF-α. (b) SpAb_T1 showed a wider R_g* distribution than TNF-α. R_g* is calculated from one representative SpAb_T1 and one TNF-α molecule in two separate samples. The coefficient of variation (c) serves as a measure of the range of conformational distribution. Five of SpAb_T1 molecules with total 2000 frames and seven of TNF-α molecules with total 3150 frames were calculated. (d) Representative binding event captured by LP-EM in SpAb_T1/TNF-α mixed sample. The capture interval between each LP-EM image is 0.1594 s, and representative images are selected to show the binding process. Scale bar = 5 nm. The R_g* variation with time of the two molecules before and after binding.

DISCUSSION

Compared to sequence-defined peptides and antibodies, RHPs are generally considered more challenging to achieve high affinity and selectivity for a specific target due to their structural heterogeneity. This study, however, demonstrates the potential of RHPs as synthetic pAbs (SpAbs), achieved through a data-driven workflow combining high-throughput synthesis, algorithm-assisted design, and systematic experimental validation. The successful identification of compositions that selectively bind valuable target proteins—such as TNF-α and IFN—convincingly proved this concept. Compared to polyclonal antibodies, which are typically produced at milligram scales relying on costly animal immunization, RHPs can be easily produced chemically at more than decagram scales[58] without the use of experimental animals. Furthermore, our SpAb maintained its binding affinity after being stored at 4 °C for over a year (Supplementary Figure S26), demonstrating exceptional stability and robustness. Additionally, SpAbs offer unique advantages in specific contexts. For instance, the incorporation of selenium into SpAbs imparts oxidation-responsiveness (Supplementary Figures S27−S29 and Supplementary Table S2), allowing for modulation of properties at another dimension.[59−61]

Unlike most other synthetic polymers, synthetic polypeptides can adopt diverse secondary structures,[62−66] making them ideal candidates for mimicking protein-like conformations. A particularly notable feature of our system is its ability to modulate secondary structures by adjusting the side-chain amino acid compositions (Supplementary Figure S12). Interestingly, we observed a weak correlation between secondary structure of RHPs and their protein binding ability. RHPs with considerable helicity (e.g., SpAb_T1, RHP_T2, RHP_T3, and RHP_T4) exhibited measurable binding to TNF-α, while those with negligible helicity (e.g., RHP_T5, RHP_T6, and RHP_T7) showed no binding to both proteins (Supplementary Figures S30 and S31 and Supplementary Table S2). It is worth mentioning that the relatively high ELISA signals observed for RHP_T5, RHP_T6, and RHP_T7 may be attributed to self-aggregation, leading to false positive results. These findings suggest that secondary structures of RHPs play a critical role in protein interaction. Nevertheless, the deep understanding of the correlation between the dynamics of RHP secondary structure and the binding affinity needs further exploration.

Beyond the selenium-containing block, the poly(sarcosine) block in our RHPs is crucial for enhancing solubility and stability and minimizing nonspecific interactions. Without the poly(sarcosine) block, RHPs, such as SpAb_T1, precipitated in aqueous solution due to self-interactions. The DP of the selenium block positively correlated with binding affinity (Supplementary Figure S32), although it showed little impact on binding selectivity. Overall, the diblock architecture and DP of both blocks are essential for RHPs participating in biomacromolecular interaction.

Results from LP-EM and CD experiments indicate that the binding dynamics of SpAbs aligns with the folding-upon-binding process commonly seen in IDPs. The structural

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

flexibility of SpAbs enables adaptive conformational adjustments upon target surface engagement, thereby facilitating binding. To further validate this mechanism, we performed molecular dynamics (MD) simulations, including umbrella sampling to probe the binding and dissociation processes of RHPs with TNF-α. Root-mean-square fluctuation (RMSF) values revealed that RHP exhibits minimal conformational changes in the bound state, whereas a marked increase in RMSF values was observed after dissociation, indicating enhanced flexibility in the unbound state (Supplementary Figures S33 and S34). We propose that initial long-range electrostatic or salt bridge interactions anchor the SpAb to the protein surface, followed by conformational changes that enable tighter surface complementarity and displacement of the hydration layers. SpAbs combine the binding properties of pAbs with the dynamic conformational behavior of IDPs, offering new possibilities for modulating biomacromolecular interactions. Further experimental and simulative studies are still ongoing to offer new details and insight regarding the binding process.

While this study establishes a robust framework for RHP discovery, there are still several limitations. Some RHPs exhibited strong ELISA signals but failed to show binding in BLI. Complementary binding assays and improved screening workflows are needed to address these issues. Additionally, the predictive accuracy of our optimization models for the composite "*Score*" metric was lower than for individual ELISA results, likely due to noise accumulation. Incorporating structural information, such as polymer descriptors and secondary structure data, could enhance the model performance.

Automation remains a critical area for high-throughput and high-quality data, which is indispensable for machine learning and artificial intelligence.[33] Our current workflow is a significant step toward fully automated discovery platforms. We are actively developing a self-driving experimental setup for automated RHP synthesis and ELISA detection, to improve data quality, consistency, and throughput (Supplementary Movie S3). Such platforms will accelerate the discovery of functional polymers for a wide range of applications beyond molecular recognition including drug delivery, antimicrobial materials, and polymer adjuvants.

CONCLUSIONS

In summary, our study demonstrates the potential of sequence-independent RHP to serve as synthetic pAbs, offering a promising alternative to antibodies for molecular recognition. Through a data-driven workflow combining high-throughput synthesis, systematic validation, and algorithm-led design, we successfully identified SpAbs capable of selectively binding protein targets, such as TNF-α and IFN, with affinities comparable to those of pAbs. The ability to modulate secondary structures of RHPs via side-chain composition, coupled with redox-responsive selenium chemistry, highlights the versatility of RHPs in emulating protein-like conformations and enabling dynamic responsiveness to oxidative environments. By leveraging flexible binding dynamics akin to IDP, as revealed by LP-EM and MD simulations, along with selective binding properties tailored through compositional tuning, SpAbs open new opportunities in modulating protein−protein interactions and addressing complex biological targets. While challenges in specificity and workflow optimization remain, advancements in automation and machine learning are expected to further enhance the efficiency of RHP evolution. Overall, this work establishes a robust framework for the development of functional RHPs for participating in protein interaction, providing a foundation for expanding their applications in therapeutic, diagnostic, etc.

METHODS

Data Analysis of ELISA Assay.

1. In a workflow validation campaign for model target antigen IFN, we simply defined the following *Score* (*Fitness*) function to evaluate the selective binding ability of RHP:

$$Score(Fitness) = Target - Control$$

Here, *Target* and *Control* represent the ELISA results for the target protein and the control protein of RHPs, respectively. RHPs with higher *Score* indicate a stronger selective binding ability to the *Target*.

2. In the campaign of optimizing RHPs for targeting TNF-α, we slightly modified the *Score* (*Fitness*) function for data standardization:

$$Score(Fitness) = \frac{Target}{Target_{mean}} - \frac{Control}{Control_{mean}}$$

Here, $Target_{mean}$ and $Control_{mean}$ refer to the average ELISA results calculated across all randomly sampled points. This evaluation method ensures a consistent and normalized comparison of the RHPs' selective affinity for the target protein, accounting for variability across different experimental batches.

Sampling Methods in Data-Driven Workflow. All methods were implemented in Python (see https://github.com/HaisenZhou/Optimization_SpAb for the source code), and were performed on a Lenovo Legion Y9000P laptop with i9-14900HX CPU and NVDIA GeForce RTX 4090 Laptop GPU (16GB).

Random Search. RHPs for random search were generated as a $n \times 8$ array (n: number of RHPs) of uniformly sampled values between 0 and 1. Each row (relative abundances) was normalized by its sum to give the composition of RHP in terms of mole fractions. These mole fractions were multiplied by a factor of 100 μL to yield the volume of the organohalide solutions that was required for each well.

Genetic Algorithm. A genetic algorithm was implemented to optimize the candidate solutions within the design space. The top-performing candidates in each iteration, determined by their *Scores* (*Fitness*), were selected as parents for generating the next generation. The algorithm employed both crossover and mutation operations to explore the solution space. Crossover was performed by randomly selecting genes from the two parents. Two types of mutations were applied: composition mutation, where Gaussian noise was added to nonzero genes, and gene switch mutation, where genes were randomly swapped between positions. To prevent convergence to suboptimal solutions, random mutations were also introduced to generate new candidates when necessary. The generated offspring were normalized to ensure that the sum of each row equaled 1. This process was repeated iteratively, with each generation producing a new population of candidate solutions and progressively refining the quality of the solutions over time.

Bayesian Optimization. Bayesian optimization was carried out using the BoTorch library in Python, specifically focusing on optimizing a given objective function within the design space. The Gaussian process (SingleTaksGP) with a ConstantMean() and squared-exponential (RBF) kernel was chosen as the surrogate model owing to its suitability for low-data learning and its inherent ability to estimate uncertainty. The observation-noise variance of the training data was set to 0.12 according to the repeat experiments. After each iteration, all data was randomly split into 90/10 training/testing for surrogate model training. The model with the lowest loss function (negative marginal log likelihood) on the test set was chosen.

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

To propose new candidate solutions, the optimization process uses the qNoisyExpectedImprovement (qNEI) acquisition function, which is optimized to generate new candidates. The optimization is constrained by the design space, ensuring that the sum of all standardized variables equals 1. The acquisition function is optimized using a multistart approach, where 20 different starting points are explored, and 512 samples are used for initialization. After optimization, 30 new candidate solutions are proposed, with the final results rounded and returned for subsequent iterations.

Both optimization methods were conducted using *Score* (or *Fitness*) as the single objective.

Please see the Supporting Information for other methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.5c06240.

Materials and instruments, experimental procedures, computational details, additional ^{1}H NMR data, CD spectra, BLI results, SEC traces, DLS measurements, and data analysis (PDF)

Conformational dynamic of TNF-α and SpAb_T1 (AVI)

Binding dynamics between SpAb_T1 and TNF-α (AVI)

Automated experimental setup (AVI)

■ AUTHOR INFORMATION

Corresponding Authors

Guangqi Wu − *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, China;* orcid.org/0000-0002-9763-6982; Email: guangqiwu@pku.edu.cn

Huan Wang − *Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Key Laboratory of Polymer Chemistry & Physics, National Biomedical Imaging Center, Peking University, Beijing 100871, China;* orcid.org/0000-0002-2542-936X; Email: wanghuan_ccme@pku.edu.cn

Hua Lu − *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, China;* orcid.org/0000-0003-2180-3091; Email: chemhualu@pku.edu.cn

Authors

Haisen Zhou − *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, China;* orcid.org/0000-0002-3767-5525

Zuo Zhang − *Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Key Laboratory of Polymer Chemistry & Physics, National Biomedical Imaging Center, Peking University, Beijing 100871, China;* orcid.org/0009-0000-2763-0123

Zhen Zhu − *Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China; Changping Laboratory, Beijing 102200, China*

Tanfeng Zhao − *State Key Laboratory of Biopharmaceutical Preparation and Delivery, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China*

Qiankang Zhou − *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, China*

Haolin Chen − *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, China*

Zhenzhong Zhao − *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, China*

Yuanhao Dai − *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, China*

Xiaodong Jing − *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, China*

Wei Ji − *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, China;* orcid.org/0000-0002-3332-4795

Xuehai Yan − *State Key Laboratory of Biopharmaceutical Preparation and Delivery, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China;* orcid.org/0000-0002-0890-0340

Lijiang Yang − *Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China*

Yiqin Gao − *Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; Changping Laboratory, Beijing 102200, China;* orcid.org/0000-0002-4309-9376

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.5c06240

Author Contributions

These authors contributed equally: Haisen Zhou, Guangqi Wu, and Zuo Zhang.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (22331002 and 22125101 to H.L.), Beijing Municipal Science and Technology Commission (Z220023 to H.L.), the Natural Science Foundation of Beijing Municipality (Z240010 to H.W.), and Open Funding Project

of the State Key Laboratory of Biopharmaceutical Preparation and Delivery (No. 2023KF-01 to H.L.). The automation platform was financially supported by Beijing AI for Science Institute. Figures 1, 2c, 5c, and 5d were created with BioRender.com. We appreciate valuable discussion and suggestions from Prof. Peng R. Chen and Prof. Chengqi Yi.

REFERENCES

(1) Groves, J. T.; Kuriyan, J. Molecular mechanisms in signal transduction at the membrane. *Nat. Struct. Mol. Biol.* **2010**, *17*, 659−665.
(2) Motwani, M.; Pesiridis, S.; Fitzgerald, K. A. DNA sensing by the cgas−sting pathway in health and disease. *Nat. Rev. Genet.* **2019**, *20*, 657−674.
(3) Brubaker, S. W.; Bonham, K. S.; Zanoni, I.; Kagan, J. C. Innate immune pattern recognition: A cell biological perspective. *Annu. Rev. Immunol.* **2015**, *33*, 257−290.
(4) Elgundi, Z.; Reslan, M.; Cruz, E.; Sifniotis, V.; Kayser, V. The state-of-play and future of antibody therapeutics. *Adv. Drug Del. Rev.* **2017**, *122*, 2−19.
(5) Zhou, J.; Rossi, J. Aptamers as targeted therapeutics: Current potential and challenges. *Nat. Rev. Drug Discov.* **2017**, *16*, 181−202.
(6) Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic peptides: Current applications and future directions. *Signal Transduction Targeted Ther.* **2022**, *7*, 48.
(7) Zhang, P.; Ye, X.; Wang, J. C. K.; Baddock, H. T.; Jensvold, Z.; Foe, I. T.; Loas, A.; Eaton, D. L.; Hao, Q.; Nile, A. H.; Pentelute, B. L. Reversibly reactive affinity selection−mass spectrometry enables identification of covalent peptide binders. *J. Am. Chem. Soc.* **2024**, *146*, 15627−15639.
(8) Ellington, A. D.; Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818−822.
(9) Zhuo, Z.; Yu, Y.; Wang, M.; Li, J.; Zhang, Z.; Liu, J.; Wu, X.; Lu, A.; Zhang, G.; Zhang, B. Recent advances in selex technology and aptamer applications in biomedicine. *Int. J. Mol. Sci.* **2017**, *18*, 2142.
(10) Le, S. P.; Krishna, J.; Gupta, P.; Dutta, R.; Li, S.; Chen, J.; Thayumanavan, S. Polymers for disrupting protein−protein interactions: Where are we and where should we be? *Biomacromolecules* **2024**, *25*, 6229−6249.
(11) Frokjaer, S.; Otzen, D. E. Protein drug stability: A formulation challenge. *Nat. Rev. Drug Discov.* **2005**, *4*, 298−306.
(12) Vlatakis, G.; Andersson, L. I.; Müller, R.; Mosbach, K. Drug assay using antibody mimics made by molecular imprinting. *Nature* **1993**, *361*, 645−647.
(13) Haupt, K.; Medina Rangel, P. X.; Bui, B. T. S. Molecularly imprinted polymers: Antibody mimics for bioimaging and therapy. *Chem. Rev.* **2020**, *120*, 9554−9582.
(14) Hoshino, Y.; Kodama, T.; Okahata, Y.; Shea, K. J. Peptide imprinted polymer nanoparticles: A plastic antibody. *J. Am. Chem. Soc.* **2008**, *130*, 15242−15243.
(15) Mahon, C. S.; Fulton, D. A. Mimicking nature with synthetic macromolecules capable of recognition. *Nat. Chem.* **2014**, *6*, 665−672.
(16) Ghosh, A.; Sharma, M.; Zhao, Y. Cell-penetrating protein-recognizing polymeric nanoparticles through dynamic covalent chemistry and double imprinting. *Nat. Commun.* **2024**, *15*, 3731.
(17) Teitelbaum, D.; Meshorer, A.; Hirshfeld, T.; Arnon, R.; Sela, M. Suppression of experimental allergic encephalomyelitis by a synthetic polypeptide. *Eur. J. Immunol.* **1971**, *1*, 242−248.
(18) Teitelbaum, D.; Arnon, R.; Sela, M. Copolymer 1: From basic research to clinical application. *Cell. Mol. Life Sci.* **1997**, *53*, 24−28.
(19) Fridkis-Hareli, M.; Teitelbaum, D.; Gurevich, E.; Pecht, I.; Brautbar, C.; Kwon, O. J.; Brenner, T.; Arnon, R.; Sela, M. Direct binding of myelin basic protein and synthetic copolymer 1 to class ii major histocompatibility complex molecules on living antigen-presenting cells−specificity and promiscuity. *Proc. Natl. Acad. Sci.* **1994**, *91*, 4872−4876.
(20) Jin, T.; Coley, C. W.; Alexander-Katz, A. Sequence-sensitivity in functional synthetic polymer properties. *Angew. Chem., Int. Ed.* **2025**, *64*, No. e202415047.
(21) Wu, G.; Zhou, H.; Zhang, J.; Tian, Z.-Y.; Liu, X.; Wang, S.; Coley, C. W.; Lu, H. A high-throughput platform for efficient exploration of functional polypeptide chemical space. *Nat. Synth.* **2023**, *2*, 515−526.
(22) 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.; Xu, T. Single-chain heteropolymers transport protons selectively and rapidly. *Nature* **2020**, *577*, 216−220.
(23) Ruan, Z.; Li, S.; Grigoropoulos, A.; Amiri, H.; Hilburg, S. L.; Chen, H.; Jayapurna, I.; Jiang, T.; Gu, Z.; Alexander-Katz, A.; Bustamante, C.; Huang, H.; Xu, T. Population-based heteropolymer design to mimic protein mixtures. *Nature* **2023**, *615*, 251−258.
(24) Rajewsky, K. Clonal selection and learning in the antibody system. *Nature* **1996**, *381*, 751−758.
(25) Odegard, V. H.; Schatz, D. G. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* **2006**, *6*, 573−583.
(26) Upadhya, R.; Kosuri, S.; Tamasi, M.; Meyer, T. A.; Atta, S.; Webb, M. A.; Gormley, A. J. Automation and data-driven design of polymer therapeutics. *Adv. Drug Del. Rev.* **2021**, *171*, 1−28.
(27) Gormley, A. J.; Webb, M. A. Machine learning in combinatorial polymer chemistry. *Nat. Rev. Mat.* **2021**, *6*, 642−644.
(28) Soheilmoghaddam, F.; Rumble, M.; Cooper-White, J. High-throughput routes to biomaterials discovery. *Chem. Rev.* **2021**, *121*, 10792−10864.
(29) Yang, L.; Pijuan-Galito, S.; Rho, H. S.; Vasilevich, A. S.; Eren, A. D.; Ge, L.; Habibović, P.; Alexander, M. R.; de Boer, J.; Carlier, A.; van Rijn, P.; Zhou, Q. High-throughput methods in the discovery and study of biomaterials and materiobiology. *Chem. Rev.* **2021**, *121*, 4561−4677.
(30) Hong, X.; Yang, Q.; Liao, K.; Pei, J.; Chen, M.; Mo, F.; Lu, H.; Zhang, W.-B.; Zhou, H.; Chen, J.; Su, L.; Zhang, S.-Q.; Liu, S.; Huang, X.; Sun, Y.-Z.; Wang, Y.; Zhang, Z.; Yu, Z.; Luo, S.; Fu, X.-F.; You, S.-L. Ai for organic and polymer synthesis. *Sci. China Chem.* **2024**, *67*, 2461−2496.
(31) Coley, C. W.; Eyke, N. S.; Jensen, K. F. Autonomous discovery in the chemical sciences part I: Progress. *Angew. Chem., Int. Ed.* **2020**, *59*, 22858−22893.
(32) Griffin, D. J.; Coley, C. W.; Frank, S. A.; Hawkins, J. M.; Jensen, K. F. Opportunities for machine learning and artificial intelligence to advance synthetic drug substance process development. *Org. Process Res. Dev.* **2023**, *27*, 1868−1879.
(33) Seifrid, M.; Pollice, R.; Aguilar-Granda, A.; Morgan Chan, Z.; Hotta, K.; Ser, C. T.; Vestfrid, J.; Wu, T. C.; Aspuru-Guzik, A. Autonomous chemical experiments: Challenges and perspectives on establishing a self-driving lab. *Acc. Chem. Res.* **2022**, *55*, 2454−2466.
(34) Bannigan, P.; Aldeghi, M.; Bao, Z.; Häse, F.; Aspuru-Guzik, A.; Allen, C. Machine learning directed drug formulation development. *Adv. Drug Del. Rev.* **2021**, *175*, No. 113806.
(35) Pollice, R.; dos Passos Gomes, G.; Aldeghi, M.; Hickman, R. J.; Krenn, M.; Lavigne, C.; Lindner-D'Addario, M.; Nigam, A.; Ser, C. T.; Yao, Z.; Aspuru-Guzik, A. Data-driven strategies for accelerated materials design. *Acc. Chem. Res.* **2021**, *54*, 849−860.
(36) Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; Zhao, S. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **2019**, *18*, 463−477.
(37) Zhang, Z.; Cai, Z.; Zhang, W.; Lu, H.; Chen, M. Machine learning-assisted investigations toward polymer synthesis. *Chin. Sci. Bull.* **2025**, *70*, 471−480.
(38) Fransen, K. A.; Av-Ron, S. H. M.; Buchanan, T. R.; Walsh, D. J.; Rota, D. T.; Van Note, L.; Olsen, B. D. High-throughput experimentation for discovery of biodegradable polyesters. *Proc. Natl. Acad. Sci.* **2023**, *120*, No. e2220021120.
(39) Dai, T.; Vijayakrishnan, S.; Szczypiński, F. T.; Ayme, J.-F.; Simaei, E.; Fellowes, T.; Clowes, R.; Kotopanov, L.; Shields, C. E.;

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Zhou, Z.; Ward, J. W.; Cooper, A. I. Autonomous mobile robots for exploratory synthetic chemistry. *Nature* **2024**, *635*, 890−897.
(40) Chen, T.; Pang, Z.; He, S.; Li, Y.; Shrestha, S.; Little, J. M.; Yang, H.; Chung, T.-C.; Sun, J.; Whitley, H. C.; Lee, I. C.; Woehl, T. J.; Li, T.; Hu, L.; Chen, P.-Y. Machine intelligence-accelerated discovery of all-natural plastic substitutes. *Nat. Nanotechnol.* **2024**, *19*, 782−791.
(41) Li, B.; Raji, I. O.; Gordon, A. G. R.; Sun, L.; Raimondo, T. M.; Oladimeji, F. A.; Jiang, A. Y.; Varley, A.; Langer, R. S.; Anderson, D. G. Accelerating ionizable lipid discovery for mRNA delivery using machine learning and combinatorial chemistry. *Nat. Mater.* **2024**, *23*, 1002−1008.
(42) Wu, G.; Jin, T.; Alexander-Katz, A.; Coley, C. W. Autonomous discovery of functional random heteropolymer blends through evolutionary formulation optimization. *ChemRxiv* **2024**. DOI:.
(43) Gormley, A. J.; Yeow, J.; Ng, G.; Conway, Ó.; Boyer, C.; Chapman, R. An oxygen-tolerant PET-RAFT polymerization for screening structure−activity relationships. *Angew. Chem., Int. Ed.* **2018**, *57*, 1557−1562.
(44) Van de Reydt, E.; Marom, N.; Saunderson, J.; Boley, M.; Junkers, T. A predictive machine-learning model for propagation rate coefficients in radical polymerization. *Polym. Chem.* **2023**, *14*, 1622−1629.
(45) Judzewitsch, P. R.; Zhao, L.; Wong, E. H. H.; Boyer, C. High-throughput synthesis of antimicrobial copolymers and rapid evaluation of their bioactivity. *Macromolecules* **2019**, *52*, 3975−3986.
(46) Kumar, R.; Le, N.; Oviedo, F.; Brown, M. E.; Reineke, T. M. Combinatorial polycation synthesis and causal machine learning reveal divergent polymer design rules for effective pDNA and ribonucleoprotein delivery. *JACS Au* **2022**, *2*, 428−442.
(47) Tamasi, M. J.; Patel, R. A.; Borca, C. H.; Kosuri, S.; Mugnier, H.; Upadhya, R.; Murthy, N. S.; Webb, M. A.; Gormley, A. J. Machine learning on a robotic platform for the design of polymer−protein hybrids. *Adv. Mater.* **2022**, *34*, No. 2201809.
(48) Reis, M.; Gusev, F.; Taylor, N. G.; Chung, S. H.; Verber, M. D.; Lee, Y. Z.; Isayev, O.; Leibfarth, F. A. Machine-learning-guided discovery of 19F MRI agents enabled by automated copolymer synthesis. *J. Am. Chem. Soc.* **2021**, *143*, 17677−17689.
(49) Di Mare, E. J.; Punia, A.; Lamm, M. S.; Rhodes, T. A.; Gormley, A. J. Data-driven design of novel polymer excipients for pharmaceutical amorphous solid dispersions. *Bioconjugate Chem.* **2024**, *35*, 1363−1372.
(50) Dixit, R.; Herz, J.; Dalton, R.; Booy, R. Benefits of using heterologous polyclonal antibodies and potential applications to new and undertreated infectious pathogens. *Vaccine* **2016**, *34*, 1152−1161.
(51) Bain, B.; Brazil, M. Adalimumab. *Nat. Rev. Drug Discov.* **2003**, *2*, 693−694.
(52) Lipsky, P. E.; van der Heijde, D. M. F. M.; Clair, E. W. S.; Furst, D. E.; Breedveld, F. C.; Kalden, J. R.; Smolen, J. S.; Weisman, M.; Emery, P.; Feldmann, M.; Harriman, G. R.; Maini, R. N. Infliximab and methotrexate in the treatment of rheumatoid arthritis. *N. Engl. J. Med.* **2000**, *343*, 1594−1602.
(53) Rayman, M. P. Selenium and human health. *The Lancet* **2012**, *379*, 1256−1268.
(54) Kiefhaber, T.; Bachmann, A.; Jensen, K. S. Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr. Opin. Struct. Biol.* **2012**, *22*, 21−29.
(55) Arai, M.; Suetaka, S.; Ooka, K. Dynamics and interactions of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2024**, *84*, No. 102734.
(56) Zhang, Z.; Zhou, H.; Xu, Z.; Wu, G.; Lu, H.; Wang, H. Chemical reactions on a single linear chain show cooperativity. *JACS Au* **2025**, *5*, 261−270.
(57) Zhang, D.-Y.; Xu, Z.; Li, J.-Y.; Mao, S.; Wang, H. Graphene-assisted electron-based imaging of individual organic and biological macromolecules: Structure and transient dynamics. *ACS Nano* **2025**, *19*, 120−151.
(58) Campos-García, V. R.; Herrera-Fernández, D.; Espinosa-de la Garza, C. E.; González, G.; Vallejo-Castillo, L.; Avila, S.; Muñoz-García, L.; Medina-Rivero, E.; Pérez, N. O.; Gracia-Mora, I.; Pérez-Tapia, S. M.; Salazar-Ceballos, R.; Pavón, L.; Flores-Ortiz, L. F. Process signatures in glatiramer acetate synthesis: Structural and functional relationships. *Sci. Rep.* **2017**, *7*, 12125.
(59) Wu, G.; Ge, C.; Liu, X.; Wang, S.; Wang, L.; Yin, L.; Lu, H. Synthesis of water soluble and multi-responsive selenopolypeptides via ring-opening polymerization of N-carboxyanhydrides. *Chem. Commun.* **2019**, *55*, 7860−7863.
(60) Ge, C.; Zhu, J.; Wu, G.; Ye, H.; Lu, H.; Yin, L. Ros-responsive selenopolypeptide micelles: Preparation, characterization, and controlled drug release. *Biomacromolecules* **2022**, *23*, 2647−2654.
(61) Li, T.; Xu, H. Selenium-containing nanomaterials for cancer treatment. *Cell Rep. Phys. Sci.* **2020**, *1*, No. 100111.
(62) Baumgartner, R.; Fu, H.; Song, Z.; Lin, Y.; Cheng, J. Cooperative polymerization of α-helices induced by macromolecular architecture. *Nat. Chem.* **2017**, *9*, 614−622.
(63) Scott, W. A.; Gharakhanian, E. G.; Bell, A. G.; Evans, D.; Barun, E.; Houk, K. N.; Deming, T. J. Active controlled and tunable coacervation using side-chain functional α-helical homopolypeptides. *J. Am. Chem. Soc.* **2021**, *143*, 18196−18203.
(64) Song, Z.; Tan, Z.; Cheng, J. Recent advances and future perspectives of synthetic polypeptides from N-carboxyanhydrides. *Macromolecules* **2019**, *52*, 8521−8539.
(65) Grazon, C.; Salas-Ambrosio, P.; Antoine, S.; Ibarboure, E.; Sandre, O.; Clulow, A. J.; Boyd, B. J.; Grinstaff, M. W.; Lecommandoux, S.; Bonduelle, C. Aqueous ropisa of α-amino acid N-carboxyanhydrides: Polypeptide block secondary structure controls nanoparticle shape anisotropy. *Polym. Chem.* **2021**, *12*, 6242−6251.
(66) Badreldin, M.; Salas-Ambrosio, P.; Bourasseau, S.; Lecommandoux, S.; Harrisson, S.; Bonduelle, C. Toward synthetic intrinsically disordered polypeptides (idps): Controlled incorporation of glycine in the ring-opening polymerization of N-carboxyanhydrides. *Biomacromolecules* **2024**, *25*, 3033−3043.

https://doi.org/10.1021/jacs.5c06240
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX