

SUPPORTING INFORMATION: Data-driven discovery of innate immunomodulators via machine learning-guided high throughput screening

Yifeng Tang,^{†,¶} Jeremiah Y. Kim,^{†,¶} Carman KM IP,[‡] Azadeh Bahmani,[‡] Qing
Chen,[†] Matthew G. Rosenberger,[†] Aaron P. Esser-Kahn,^{*,†} and Andrew L.
Ferguson^{*,†}

[†]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637 USA*

[‡]*Cellular Screening Center, University of Chicago, Chicago, IL 60637 USA*

[¶]*These authors contributed equally to this paper.*

E-mail: aesserkahn@uchicago.edu; andrewferguson@uchicago.edu

Contents

S1 Supplementary Computational Methods	3
S1.1 Definition of ~1M compound augmented ZINC library	3
S1.2 Definition of ~140k compound immunomodulator candidate library	4
S1.3 Deep representational learning using variational autoencoders (VAEs)	5
S1.3.1 Model architecture	5
S1.3.2 Model training over augmented ZINC library	6
S1.4 Gaussian process regression (GPR) surrogate models	8
S1.5 Multi-objective Bayesian optimization (BO)	11
S1.6 VAE-GPR surrogate model ablation test	13
S1.7 Inference of chemical design rules using LASSO regression	14
S2 Supplementary Experimental Methods and Results	23
S2.1 High throughput screening experiments	23
S2.1.1 Immunomodulators delivered alone do not affect immune responses of NF- κ B and IRF pathways	23
S2.2 Low throughput immunomodulation validation experiments	23
S2.2.1 Assay used for cytokine release profile testing	23
S2.2.2 Immunomodulators delivered alone do not affect cytokine production	24
S2.2.3 Comparing PME-4007 and MSA-2	24
S2.3 Safety statement	26
S3 Supplementary Figures	28
S4 Supplementary Tables	34
References	35

S1 Supplementary Computational Methods

S1.1 Definition of \sim 1M compound augmented ZINC library

We defined an augmentation of the ZINC small molecule library¹⁻⁴ to train our VAE network for deep representational embeddings of our immunomodulator candidates. We combined the 2,674 molecules employed in our prior screening work,⁵ with a subset of 924,870 molecules from the ZINC libraries¹⁻⁴ comprising those molecules that containing a biphenyl scaffold, inspired by the structure of previous small molecule immunomodulator discovery of Honokiol.⁶ We also incorporated some other generic molecular libraries from various vendors. As a result, our augmented ZINC library consists of 1,262,866 small molecule compounds in total. A list of the molecular libraries compiled to define the initial augmented ZINC library is provided in Table S1.

We then filtered the library under a number of criteria. First, we represented each compound as simplified molecular-input line-entry system (SMILES) strings.⁷⁻¹⁰ We then canonicalized these SMILES representations and removed duplicates. Second, we translated each SMILES string into self-referencing embedded strings (SELFIES) as a more robust representation of molecules that are guaranteed to define valid chemical structures.¹¹ We then eliminated compounds that produced an inconsistent SMILES string upon back-translation from the SELFIES representation to ensure a strict one-to-one mapping between SMILES representations and SELFIES representations of the compounds, thus enforcing a strict one-to-one mapping between SELFIES representations and the structure of the compounds. With the first two steps of filtering, we removed 8734 entries and were left with 1,254,132 compounds in the library. Third, we capped the maximum SELFIES string length to 137 characters and eliminated compounds containing tokens that appear in fewer than 500 compounds. Since the SELFIES strings will ultimately be represented to the VAE input layer as fixed-length vectors, this step constrained the dimension of the representation space so as to limit the size of the network and its associated training costs. These constraints resulted in

the elimination of only 5468 compounds, corresponding to fewer than 0.5% of candidates in the immunomodulator library (Section S1.2). After performing these filtering operations, we were left with 1,248,664 compounds in the augmented ZINC library. We note that a number of candidate compounds within the library exist as salts or with other components that can serve as, for example, stabilizers. The SELFIES representation is sufficiently flexible to support representation of such compounds by joining the two molecules within a single string using a special “linking token”. We elected to take this approach as it more faithfully represented the substance actually delivered in the experimental assay and also eliminated the need for manual curation of the data to enhance the scalability of the approach.

Finally, we digitized the SELFIES representations into one-hot matrices of dimension 137-by-55, where 137 represents the maximum possible string length and 55 are the number of possible SELFIES tokens, including the null token. SELFIES strings less than the maximum length are padded with null tokens. We then flattened the one-hot matrices into 7535-element vectors to provide a fixed length representation of each molecule to be passed to the VAE.

S1.2 Definition of \sim 140k compound immunomodulator candidate library

We selected seven generic commercial small molecule compound libraries to define an initial pool of 139,998 candidate immunomodulators to draw candidates from and bring to high throughput screening experimentation. These libraries were a subset of our previously assembled 1,248,664 compound augmented ZINC library. These screening compound libraries were designed with the intention of enabling cell-based and target-based high throughput screening initiatives by making a diverse range of small molecules readily available, which make it easier for us to access the molecules for screening experiments. This immunomodulator candidate library went through the same filtering steps as introduced in Section S1.1, with the only exception that we were not using a new set of SELFIES tokens, but we used the token dictionary we used for the augmented ZINC library. This is to ensure the map-

Table S1: **Compound libraries used to assemble the augmented ZINC library.**

Company/Institution	Compound Library	# Compounds	URL
UChicago	Our Prior Screen Work ⁵	2674	N/A
Analyticon	Synthetic Screening Compounds	5000	URL
ChemBridge	DS550-3	29882	URL
ChemBridge	ES550-1	49966	URL
ChemBridge	ES550-2	49958	URL
ChemRoutes	ChemArray	21135	URL
Life Chemicals	Pre-plated Diversity Sets PS4	50200	URL
Life Chemicals	Pre-plated Diversity Sets PS5	50168	URL
Life Chemicals	Biologically Active Compound Library	7991	URL
MedChemExpress	STAT Inhibitors	58	URL
Micro Source	The Spectrum Collection	1975	URL
Prestwick Chemical	Prestwick Chemical Library	1276	URL
Selleckchem	FDA-approved Drug Library	2667	URL
Selleckchem	STAT Inhibitors	18	URL
UCSF SMDC	Various Compound Libraries	65790	URL
ZINC¹⁻⁴	Biphenyl Scaffold	924870	URL
	Sum	1262866	
	Sum (After Filtering)	1248664	

ping between the network parameters of VAE and the one-hot encodings of compounds. We transformed the SELFIES representations into one-hot matrices, with a dimension of 137-by-55 then flattened the one-hot matrices into 7535-element vectors to obtain a standardized representation of each molecule. A list of the molecular libraries compiled to define the initial immunomodulator candidate library is provided in Table S2. A CSV file providing SMILES strings of all compounds of the 139,998 compound library is provided in the Supporting Information.

S1.3 Deep representational learning using variational autoencoders (VAEs)

S1.3.1 Model architecture

We employ a VAE architecture inspired by prior work by Aspuru-Guzik and co-workers^{11,12} and implemented in PyTorch.¹³ The encoder consists of three fully-connected (FC) layers

Table S2: **Compound libraries used to assemble the immunomodulator candidate library.**

Company	Compound Library	Number of Molecules	URL
ChemBridge	DS550-3	29547	URL
ChemBridge	ES550-3	49584	URL
Life Chemicals	Biologically Active Compound Library	7938	URL
Life Chemicals	Pre-plated Diversity Sets PS4	50233	URL
Micro Source	The Spectrum Collection	1821	URL
Prestwick Chemical	Prestwick Chemical Library	1233	URL
Selleckchem	FDA-approved Drug Library	1539	URL
	Sum (After Filtering)	139998	

that passes into a fourth fully-connected bottleneck layer corresponding to the latent space embedding. For the decoder, we employed a stack of gated recurrent units (GRU).¹⁴ The architecture of the VAE is illustrated in Figure S1. The hyperparameters of the model architecture were optimized over the ranges reported in Table S3.

Table S3: **Architecture and hyperparameter optimization ranges for the VAE.**

Hyper-parameter description	Value	Optimization range
The number of neurons in the first FC layer	500	[250, 1000]
The number of neurons in the second FC layer	200	[200, 500]
The number of neurons in the third FC layer	100	[50, 250]
Latent space dimension	100	[3, 250]
The number of recurrent layers of GRU	3	[1, 3]
The number of features in the hidden state of GRU	200	[20, 400]

S1.3.2 Model training over augmented ZINC library

The VAE loss function \mathcal{L}_{VAE} comprises two components: (1) the reconstruction loss \mathcal{L}_{Rec} measured by the cross-entropy between the input and output one-hot SELFIES vectors to enforce reconstruction fidelity and (2) the Kullback-Leibler divergence (KLD)¹⁵ \mathcal{L}_{KLD} of the latent vectors relative to the standard normal distribution to regularize the latent space,^{12,16}

$$\mathcal{L}_{\text{VAE}} = \mathcal{L}_{\text{Rec}} + \mathcal{L}_{\text{KLD}}, \tag{S1}$$

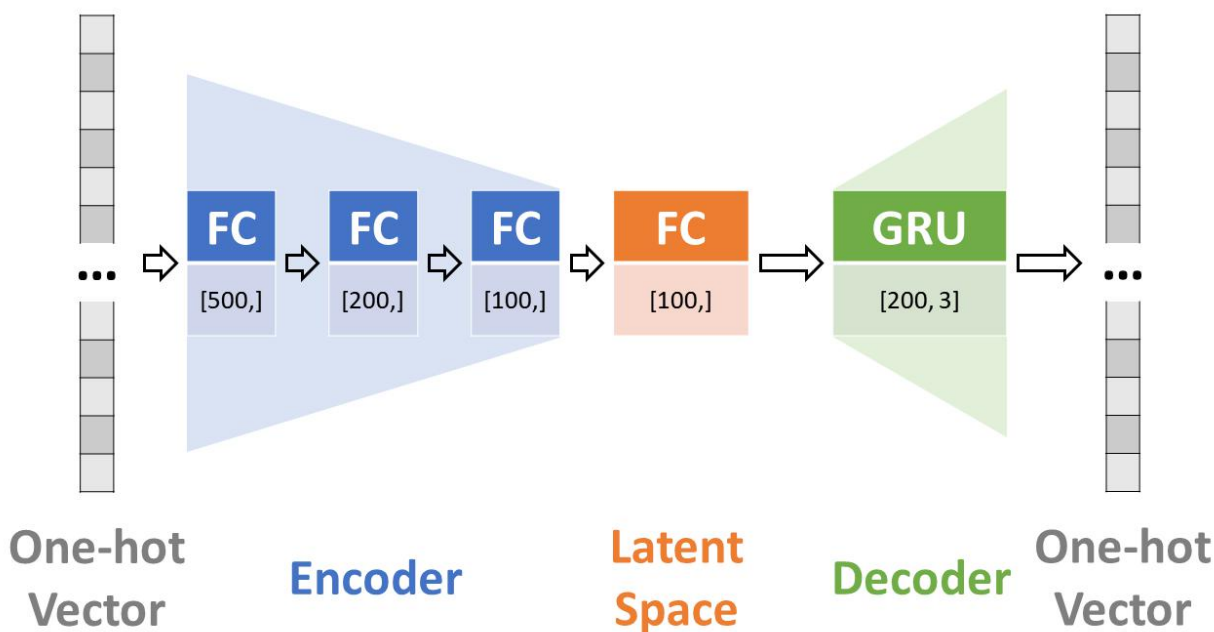


Figure S1: **Architecture of VAE for deep representational learning of molecular latent space.** The input and output of the VAE model are 7535-element one-hot vector that are one-to-one with the SELFIES string representation of each molecule. The encoder stack comprises three fully-connected feedforward layers with a 500-200-100 architecture. A bottleneck layer comprising 100 neurons defines the 100D latent space embedding. The decoder comprises three stacked GRUs each containing 200 neurons.

where,

$$\mathcal{L}_{\text{Rec}} = - \sum_{i=1}^N \sum_{j=1}^D x_{ij} \log(\hat{x}_{ij}) + (1 - x_{ij}) \log(1 - \hat{x}_{ij}), \quad (\text{S2})$$

where N is the number of samples, D is the dimension of the data, x_{ij} is the j^{th} component of the i^{th} input vector, and \hat{x}_{ij} is the corresponding reconstructed output of the autoencoder, and,

$$\mathcal{L}_{\text{KLD}} = -\frac{1}{2} \sum_{j=1}^J (1 + \log(\sigma_j^2) - \mu_j^2 - \sigma_j^2), \quad (\text{S3})$$

where μ_j and σ_j^2 are the mean and variance of the j^{th} element of the latent vector z , respectively, and J is the dimensionality of z .

Network training is conducted by minimizing Equation S1 using the Adam optimizer.¹⁷ We first train the VAE over the augmented ZINC library for 10,000 epochs. To guard

against posterior collapse,¹⁸ we employ an initial KLD loss coefficient of $\alpha = 10^{-2}$ that we schedule to reduce to $\alpha = 10^{-3}$ at epoch 2000, and to $\alpha = 10^{-4}$ at epoch 5500. The training hyperparameters comprising batch size and learning rates were optimized over the ranges reported in Table S4. We assess model performance by computing the exact reconstruction accuracy defined as the fraction of molecules in the training set whose SELFIES strings can be reconstructed with 100% fidelity. Since the intended application of the trained VAE is to define a latent space embedding of our candidate molecules to support an active learning search, we use the decoder in service of learning this latent embedding but, in the present work, do not make use of its generative capacity to extend the model beyond the training data and produce novel synthetic molecules. As such, we assess model performance on the training data rather than the usual practice of employing a hold-out test set. Our model achieves an exact reconstruction accuracy of 97.3% at epoch 9032 after more than 1344 GPU-hours of training. This high accuracy indicates that the network has discovered a 100D latent space embedding that preserves the salient information necessary for accurate molecular reconstruction, and we terminate training at this point. The training curves for the model are illustrated in Figure S2.

Table S4: **VAE training hyperparameters and optimization ranges.**

Training parameter	Value	Optimization range
Batch size	2500	[2000, 3000]
Initial learning rate of the encoder	0.0001	[0.0001, 0.01]
Initial learning rate of the decoder	0.0001	[0.0001, 0.01]

S1.4 Gaussian process regression (GPR) surrogate models

In our present application, we seek to identify immunomodulators capable of enhancing or suppressing the immune activity of a certain pathway when the pathway is activated by a certain agonist. This leads to our optimization in maximizing or minimizing the immunomodulation values as fold changes. To guide our search in the chemical design space

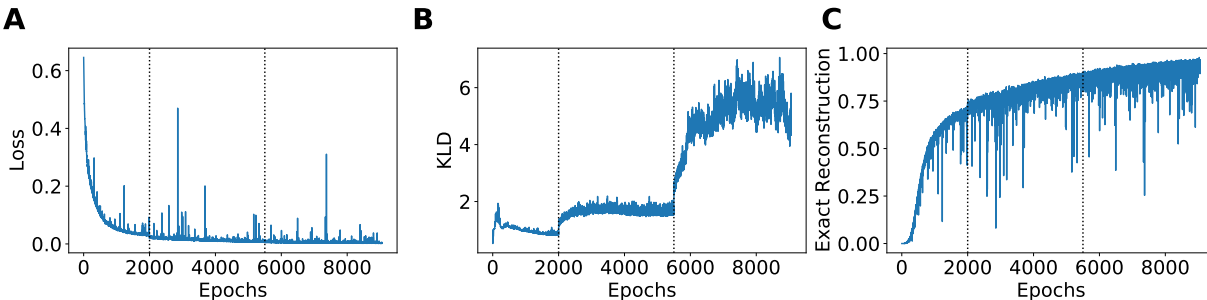


Figure S2: **VAE training curves over the augmented ZINC training library.** (a) VAE loss function (Equation S1). (b) KLD component of VAE loss (Equation S3). The dashed vertical lines in each panel denote the epochs at which the KLD loss coefficient α was tuned from an initial value of 10^{-2} to 10^{-3} then 10^{-4} . (c) Exact reconstruction accuracy reporting the fraction of molecules in the training data whose SELFIES representations are exactly reconstructed by the trained network.

towards profitable regions, we train a surrogate model to predict the fitness of immunomodulator candidates that have not been experimentally tested. We include an uncertainty measure that reflects our confidence in the predictions, enabling us to select points in the space that balance both high predicted fitness (exploitation) and high uncertainty (exploration). For this supervised regression task, we use Gaussian Process Regression (GPR) as our surrogate model, as it is a non-parametric Bayesian regression technique that provides built-in uncertainty estimates.¹⁹ We have chosen to use the radial basis function (RBF) kernel (also known as the squared exponential kernel) as the covariance function within the GPR implementation in scikit-learn,²⁰

$$k(x_i, x_j) = \sigma^2 \exp\left(-\frac{|x_i - x_j|^2}{2\ell^2}\right), \quad (\text{S4})$$

where x_i and x_j are input data points, σ^2 is the signal variance, ℓ is the length scale of the kernel, and $|\cdot|$ represents the Euclidean distance between two points. To account for uncertainties associated with each experimental measurement σ^n , we adopt Tikhonov regularization²¹ such that an uncertainty vector $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_n]^T$ is added to the diagonal of the kernel matrix K during fitting. Although in each experiment, the standard deviation

for each sample being tested was different, we took the average of the experimental errors of all samples $\bar{\sigma} = \frac{1}{n} \sum_{i=1}^n \sigma_i$ and fed into the GPR models as the uncertainty vector $\sigma = [\bar{\sigma}, \bar{\sigma}, \dots, \bar{\sigma}]^T$.

As we were seeking immunomodulators with various capabilities – different kinds of immunomodulation (enhancement, suppression), different pathways (NF- κ B, IRF), and different agonists (LPS, MPLA, CpG, cGAMP) – we pursue a multi-objective optimization. As such, we constructed multiple GPR models, with each GPR corresponding to enhancement/suppression with a specific agonist on a specific pathway. In addition to specialists – immunomodulators that induce large changes in enhancement or suppression when co-delivered with one particular agonist – we also sought to identify generalists – immunomodulators that lead to large effects when co-delivered with any one of a group of agonists. We constructed GPR models for generalists using the arithmetic mean of the immunomodulation fold change values of corresponding specialist objectives. For example, as one of the objectives for enhancers of the NF- κ B response, the immunomodulation of generalist over LPS, MPLA and CpG agonists can be expressed as

$$M_{\text{generalist(LPS, MPLA, CpG)}}^{\text{NF-}\kappa\text{B Enhancer}} = \frac{1}{3} \left(M_{\text{LPS}}^{\text{NF-}\kappa\text{B Enhancer}} + M_{\text{MPLA}}^{\text{NF-}\kappa\text{B Enhancer}} + M_{\text{CpG}}^{\text{NF-}\kappa\text{B Enhancer}} \right), \quad (\text{S5})$$

where $M_{\text{LPS}}^{\text{NF-}\kappa\text{B Enhancer}}$, $M_{\text{MPLA}}^{\text{NF-}\kappa\text{B Enhancer}}$ and $M_{\text{CpG}}^{\text{NF-}\kappa\text{B Enhancer}}$ are immunomodulation fold change values for modulator specialist with agonist LPS, MPLA and CpG, respectively. We initially have 14 objectives, each representing a GPR model: (1) Enhancers of the NF- κ B response (4 objectives): specialist for LPS agonist, specialist for MPLA agonist, specialist for CpG agonist, and generalist over LPS, MPLA, and CpG agonists; (2) Inhibitors of the NF- κ B response (4 objectives): specialist for LPS agonist, specialist for MPLA agonist, specialist for CpG agonist, and generalist over LPS, MPLA, and CpG agonists; (3) Enhancers of the IRF response (6 objectives): specialist for LPS agonist, specialist for MPLA agonist, specialist for CpG agonist, specialist for cGAMP agonist, generalist over LPS, MPLA, and

CpG agonists, and generalist over LPS, MPLA, CpG and cGAMP agonists. We did not seek to identify inhibitors of the IRF response because of limited clinical significance of such immunomodulation. We did not optimize for enhancers or inhibitors of the NF- κ B response for specialist for cGAMP agonist because we knew from prior work that cGAMP does not strongly stimulate the NF- κ B pathway.²² Following our screening experiment, we found that the specific CpG agonist we were using (CpG ODN 1826) demonstrated minimal IRF stimulation, as can be seen in Figure S7. As a result, we only report specialist for CpG agonist as enhancers and inhibitors of the NF- κ B response, and the 14 initial objectives were reduced to the 12 objectives as shown in Figure 1E.

S1.5 Multi-objective Bayesian optimization (BO)

After building GPR surrogate models to predict various forms of immunomodulation for every compound, our objective is to explore the design space by iteratively querying new compounds to experimentally test. This involves striking a balance between selecting the most promising candidates with high expected performance (exploitation) and potentially valuable candidates with high uncertainty associated with their prediction (exploration). To achieve this balance, we utilize Bayesian optimization (BO), which is a powerful black-box optimization technique that utilizes the posterior mean and uncertainties derived from GPR predictions to identify the most promising candidates for testing.²³ BO can optimize costly black-box functions by optimizing the surrogate objectives that are defined by GPR. An acquisition function is needed to determine the next best candidate for evaluation. To meet our goal of discovering more effective immunomodulators, we have chosen the Expected Improvement (EI) acquisition function.²⁴ This function guides the selection of candidates that are predicted to result in maximum expected gains in fitness. The

EI acquisition function can be expressed as,

$$\text{EI}(\mathbf{x}) = \begin{cases} (\mu(\mathbf{x}) - f(\mathbf{x}^+) - \xi)\Phi(Z) + \sigma(\mathbf{x})\phi(Z), & \text{if } \sigma(\mathbf{x}) > 0, \\ 0, & \text{otherwise} \end{cases} \quad (\text{S6})$$

where \mathbf{x} is the input point, $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ are the mean and standard deviation of the surrogate model prediction at \mathbf{x} , $f(\mathbf{x}^+)$ is the best function value observed so far, ξ is a trade-off parameter that balances exploration and exploitation, Φ and ϕ are the standard normal cumulative distribution function and probability density function, and $Z = \frac{\mu(\mathbf{x}) - f(\mathbf{x}^+) - \xi}{\sigma(\mathbf{x})}$ is the standard normal random variable. We select the next best candidate point \mathbf{x}^* with the highest value of the acquisition function from the set of available candidates \mathbf{x}_k that have not yet been sampled, i.e., $\mathbf{x}^* = \text{argmax}_k \text{EI}(\mathbf{x}_k; \mathbf{x}_k \notin \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\})$. In this work, we employ $\xi = 0.01$ as a good general purpose choice commonly employed in the literature.^{24,25}

Traditional multi-objective Bayesian optimization (MOBO) strategies typically aim to balance multiple objectives to efficiently map out Pareto optimal solutions.^{26,27} In this work, this would be valuable in identifying potent generalists, but may sacrifice the discovery of potent specialists that reside at the ‘‘corners’’ of the Pareto frontier. Furthermore, some of our objectives are mutually incompatible, such as seeking to find enhancers and inhibitors for the same pathway with the same agonist. Accordingly, we adopted an alternative MOBO strategy in which we polled each of our 12 GPR models independently to collect the best candidates recommended by each model under the EI acquisition function and eliminated duplicates. To make efficient use of the experimental assay, we followed a batched sampling Kriging believer approach^{28,29} wherein we asked each model for its next top-ranked candidates until we had collected a batch of 720 molecules for experimental testing. Under this protocol, we temporarily augment the training data with the molecules selected by the GPRs annotated with the GPR activity predictions, the GPR models are retrained on this augmented data, and the models polled again for their next top-ranked prediction under the

EI acquisition function. This process is repeated until a batch of 720 molecules has been collated for experimental testing. This batched sampling approach sacrifices efficiency from an information theoretic perspective – each model is asked to choose a series of candidates without first receiving back experimental information on the previously selected candidates – but increases temporal efficiency by testing multiple molecules in an experimental batch. Importantly, all selected molecules, regardless of their source GPR models, were subjected to comprehensive experimental testing to evaluate their immunomodulatory profiles across all agonist types and activation of both NF- κ B and IRF pathways. As such, once the experiments are complete, the activity entries in the augmented training data that were completed with the GPR activity predictions in service of the Kriging believer batching are corrected with the measured activities, and these data used to retrain all GPR model. In this manner, molecules selected by one GPR model are not only used to retrain that particular GPR model, but also used to retrain all other models.

S1.6 VAE-GPR surrogate model ablation test

Our choice of constructing a low-dimensional embedding using VAE and training a GPR surrogate model was motivated, in part, by our previous experience in the success of the VAE+GPR+BO paradigm in a number of recent applications in our lab^{30–35} and others.^{36–39} The GPR paradigm is also a natural fit for BO due to its intrinsic uncertainty estimates. There are, however, a number of other surrogate modeling approaches that do not require learning low-dimensional embeddings, but rather can operate directly on molecular featurizations,^{40–42} and that may employ simpler regression models such as linear regression or support vector regression. To test the value of the learned low-dimensional embedding and nonlinear and nonparametric nature of the GPR in elevating the predictive performance, we conducted a simple ablation test in which we evaluated model performance upon replacing the learned VAE featurization with a simple and popular 2048-bit Morgan topological fingerprints (FP) ECFP4⁴³ computed using RDKit,⁴⁴ and replacing the GPR regression model

with a simple linear regression (LR). Considering these two ablations either independently or in combination results in four possible models: (1) VAE-GPR; (2) VAE-LR; (3) FP-GPR; and (4) FP-LR. We use the MAE calculated from a 5-fold cross-validation using each QSAR model over six immunological functional goals to compare model performance.

The performance comparison is presented in Figure S3. It is immediately apparent that the performance of the FP-LR model is significantly poorer than the other three. Between the remaining three, the VAE-LR tends to perform marginally less well than the VAE-GPR and FP-GPR. These results indicate that the simple featurization combined with the simple regression strategy (FP-LR) is outperformed by both the learned featurization employing the simple regression strategy (VAE-LR) and the simple featurization employing the sophisticated regression strategy (FP-GPR). Interestingly, the performance of the FP-GPR is on par with the VAE-GPR, suggesting that the sophisticated GPR regression model performs equally well under a simple (FP) or learned (VAE) featurization. The VAE featurization, however, possesses two significant advantages over the Morgan fingerprint featurizations. First, the 100D VAE embedding has a relatively modest dimensionality compared to the 2048D Morgan fingerprints. The relatively poor scaling of GPR models with dimensionality, means that this results in a $\sim 3\times$ slowdown in training and deployment of the FP-GPR model relative to the VAE-GPR. Second, the VAE embeddings are invertible in the sense that molecules can be generated from VAE vectors, whereas it is not generally considered possible to straightforwardly invert Morgan fingerprints into molecular structures.^{43,45} Although we do not use the generative capacity of the VAE in this work, in future work we wish to use the model to generate new synthetic molecular candidates with potentially superior performance than those contained in the current screening libraries.

S1.7 Inference of chemical design rules using LASSO regression

A comprehensive list of descriptors that RDKit is able to compute is available at <https://www.rdkit.org/docs/GettingStartedInPython.html#list-of-available-descriptors>,

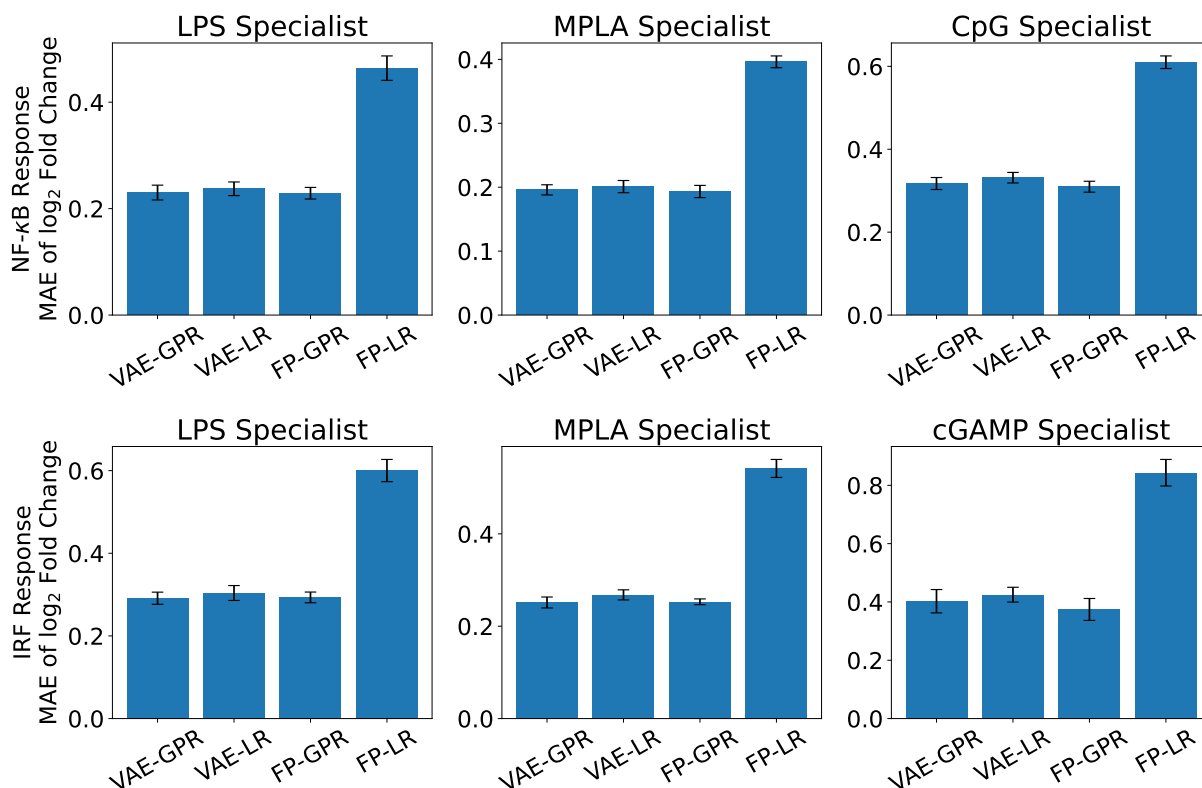


Figure S3: **Analysis of the influence of featurization and regression model upon predictive performance of the surrogate model.** We compare all combinations of featurizations – VAE and Morgan fingerprints (FP) – and regression model – GPR and linear regression (LR): (1) VAE-GPR; (2) VAE-LR; (3) FP-GPR; and (4) FP-LR. Performance is compared based on the MAE in the log₂-fold change in six immunological functional goals. The mean MAE is shown as the bars and the standard deviation of the MAE computed from 5-fold cross validation is shown as the error bars. The simplest model, FP-LR, has a poor MAE compared to the other three. The three remaining model combinations exhibit similar performance.

which includes a total of 208 2D descriptors. We elected not to include descriptors since 3D molecular structures were not readily available for most of the compounds in our study. Although RDKit is able to generate 3D conformers,⁴⁴ we judged for the purposes of an interpretable LASSO model that the 208 2D descriptors were sufficient. We selected those 85 descriptors denoting the occurrence of chemical groups for better interpretability. During feature selection, eight irrelevant descriptors were removed because they are invariant across all 3560 viable compounds in our study. To eliminate redundant descriptors, we looked at

the 13 descriptors with Pearson correlation coefficients higher than 0.95 with any other descriptors. We then used clustering analysis to group highly correlated descriptors together into larger families or categories (Figure S4). In each group of highly correlated descriptors, we selected one descriptor as representative of the group. In this way, seven additional descriptors were marked redundant and removed. In sum, we retained 70 ($= 85 - 8 - 7$) effective descriptors that were kept for further analysis. A CSV file listing all 85 descriptors is provided in the the Supporting Information, within which we designate them as “Irrelevant” or “Redundant” if they were eliminated in the corresponding feature selection step, and “Effective” if they were retained in further investigations.

Having defined the featurization for each immunomodulator, we construct a straightforward and interpretable model to predict the log2-fold change in immunomodulatory activity $\log_2 M$ as a linear function of the $k = 70$ standardized features F_k ,

$$\log_2 M_{\text{predicted}} = \frac{1}{N} \sum_{N=1}^{3560} \log_2 M_{\text{exp}} + \sum_{k=1}^{70} \theta_k F_k, \quad (\text{S7})$$

where $\frac{1}{N} \sum_{N=1}^{3560} \log_2 M_{\text{exp}}$ is the arithmetic mean of the log fold change in immunomodulatory activity over all $N = 3560$ experimental measurements and $\theta \in \mathbb{R}^{70}$ is a vector regression coefficients assigning weights to the different features. The coefficients θ are learned by minimizing the LASSO regression loss between the predicted and experimentally measured log2-fold change in immunomodulatory activities,

$$\mathcal{L}_{\text{LASSO}}(\theta; \alpha_L) = \frac{1}{2} \|\log_2 M_{\text{predicted}} - \log_2 M_{\text{exp}}\|_2^2 + \alpha_L \|\theta\|_1. \quad (\text{S8})$$

The first term on the right side of the equation is the mean squared error between the predicted and experimental immunomodulation values, and the second term is the L1 regularization penalty term with hyperparameter α_L . The $\|\cdot\|_2$ and $\|\cdot\|_1$ represent the L2 and L1 norms, respectively. The L1 regularization penalty term encourages the model to have fewer non-zero coefficients or parameters, thereby promoting a sparse model that can

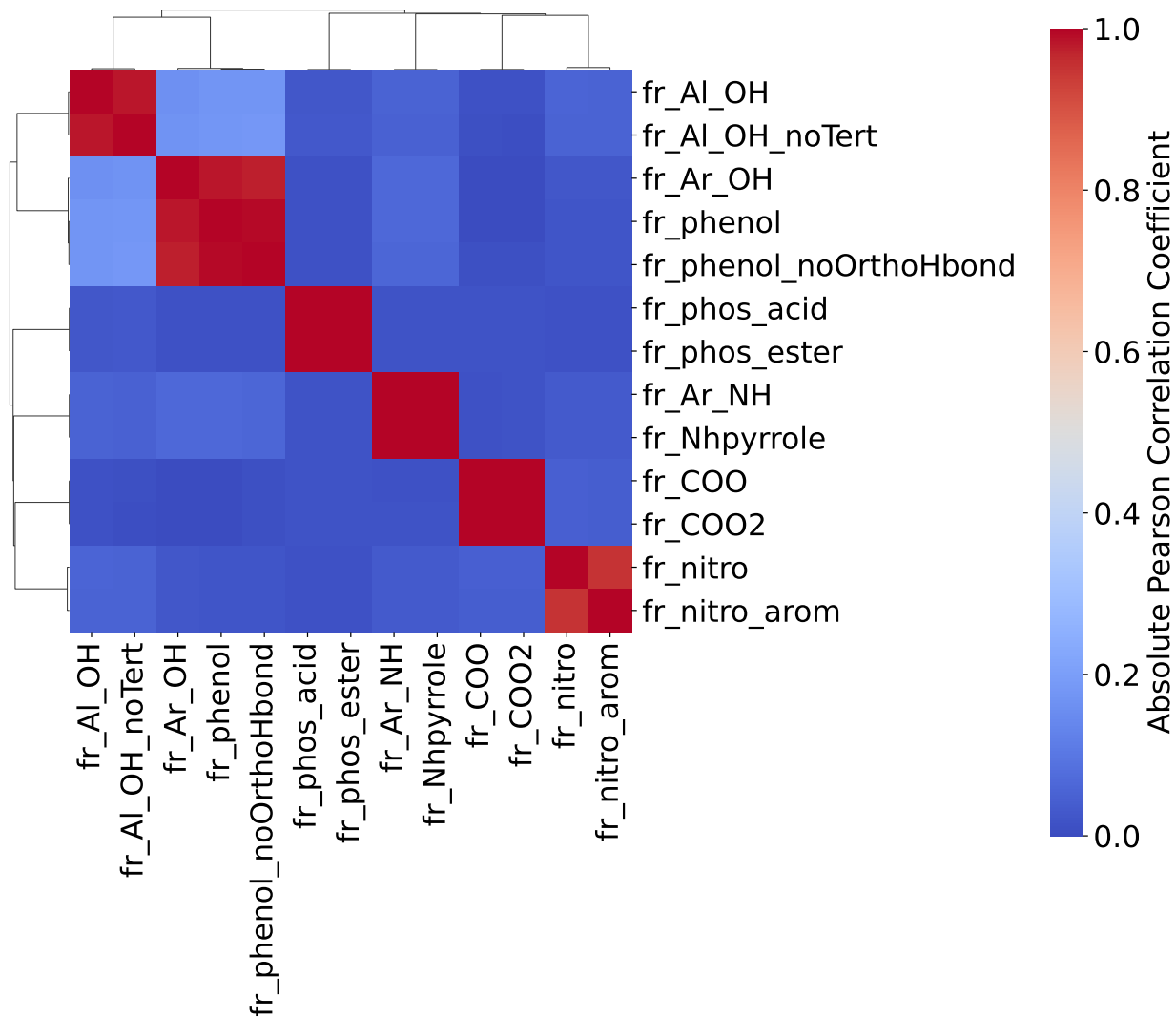


Figure S4: **Correlation heatmap and cluster dendrogram of highly correlated descriptors.** We selected 13 descriptors that had a Pearson correlation coefficient greater than 0.95 with at least one other descriptor. Clustering analyses were conducted to organize highly correlated descriptors into larger groups/families. In this way, we identified six descriptor groups with high correlation (red blocks shown in the clustered heat map). For each of the six descriptor groups, we select one representative descriptor. Thus, we retained six descriptors, namely “fr_Al_OH”, “fr_Ar_OH”, “fr_phos_acid”, “fr_Ar_NH”, “fr_COO”, “fr_nitro”.

offer a more concise and interpretable representation of the data. The optimal value for the hyperparameter α_L is chosen using cross-validation, and the resulting non-zero coefficients in θ can be interpreted as the most critical features for immunomodulation, as shown in Figure S5.

The linear nature of Equation S7 makes it easy to interpret the sign of the learned coefficients: large negative weights $\theta_k < 0$ indicate features that are negatively correlated with immunomodulation, while large positive weights $\theta_k > 0$ indicate features that are positively correlated with immunomodulation. As shown in Figure S6, by examining the rank-ordering of the coefficients with the largest absolute values, we can identify structural fragments that are most informative for predicting immunomodulation. There were, in total, 33 chemical fragment descriptors that were retained by at least one of the LASSO models for different immunological objectives. In addition, our trained LASSO model can be used to predict immunomodulation for larger molecules and/or molecules not contained within the training data. A CSV file showing the code names and the corresponding chemical interpretations of these 33 chemical fragment descriptors is provided in the Supporting Information.

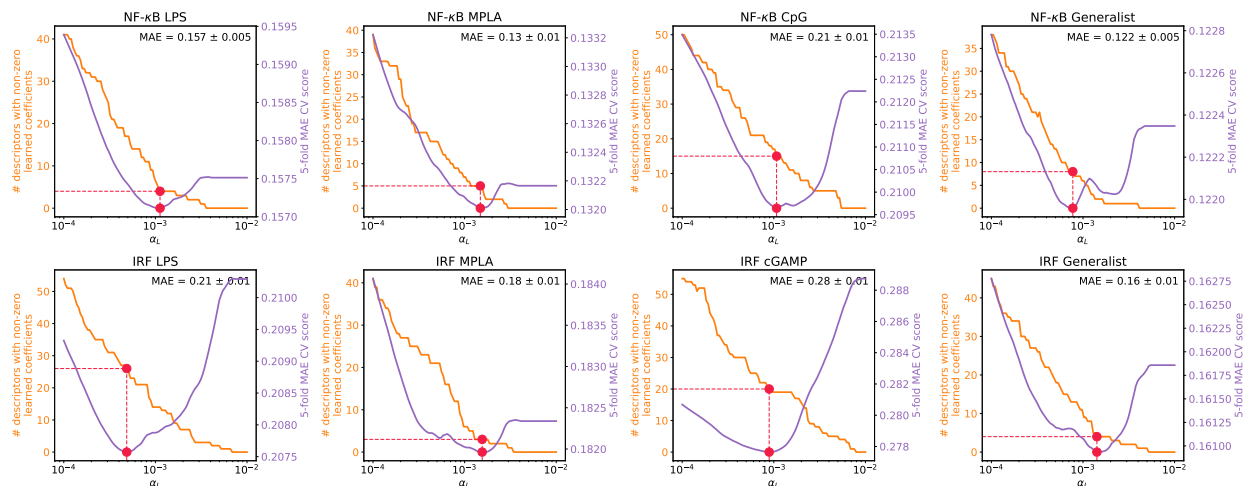
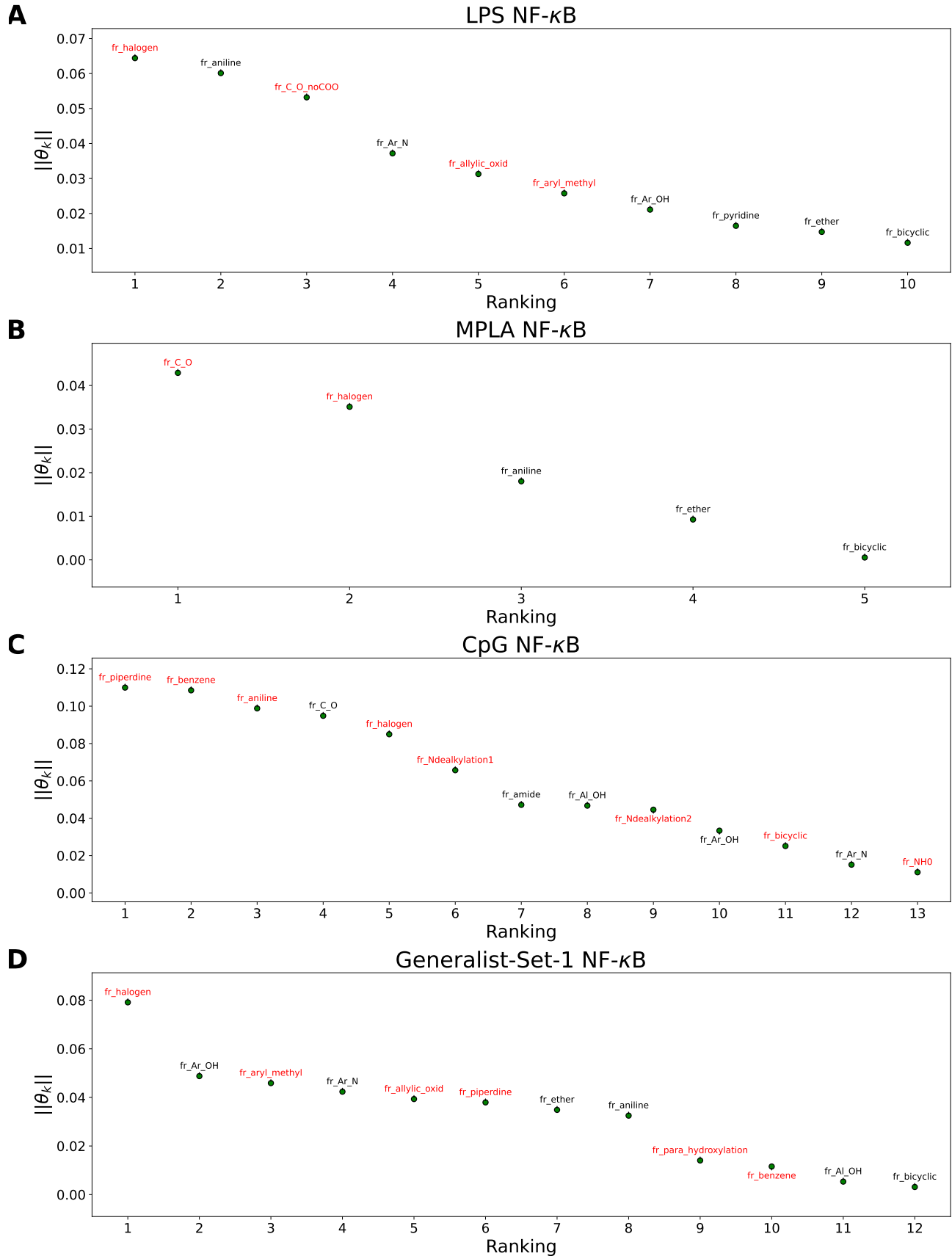


Figure S5: **The performance of LASSO regression models evaluated with respect to sparsity regularization parameter α_L .** The plot presents the number of molecular descriptors that have non-zero learned coefficient values, which were identified by training the model using a range of α_L values (displayed on the left y-axis in orange). Additionally, the plot shows the 5-fold cross-validation mean absolute error (MAE) score of the LASSO regression model in the log2-fold change at each corresponding α_L value (displayed on the right y-axis in purple). The optimal α_L is shown in red points where the corresponding MAE is the lowest. The title of each plot denotes the corresponding immune signaling pathway and agonist used for stimulation. The MAE of the model with the optimal α_L is reported in the upper-right corner of each panel with error bars estimated by 5-fold cross validation. In all cases, the MAE of the LASSO regression models are poorer than those for the corresponding GPR (cf. Figure 2), but the performance differential is at most only 32.3%.



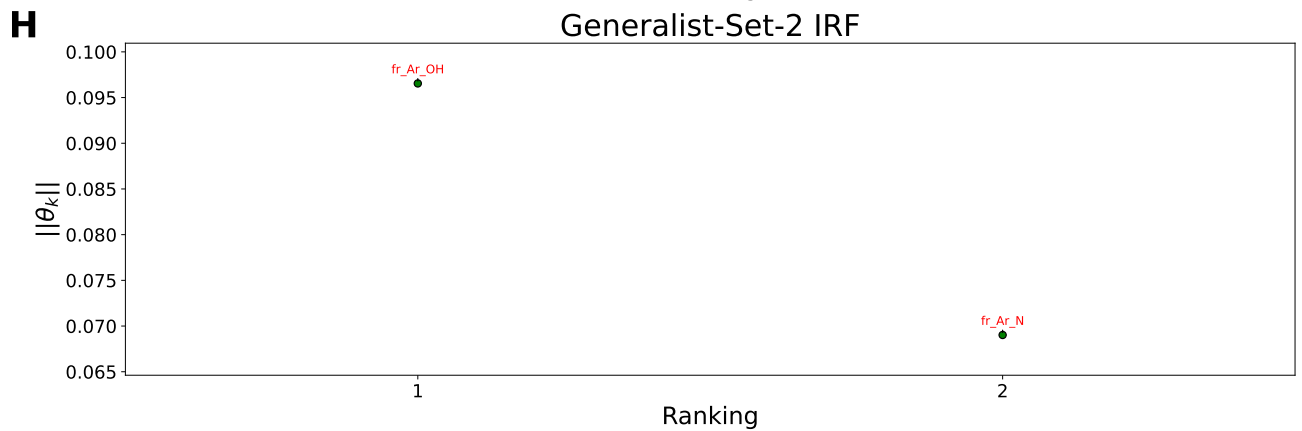
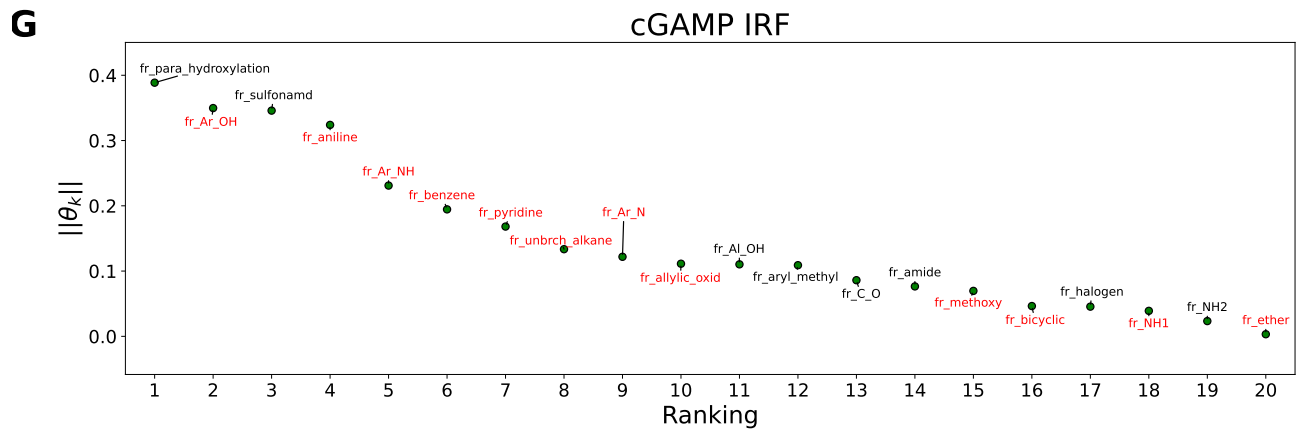
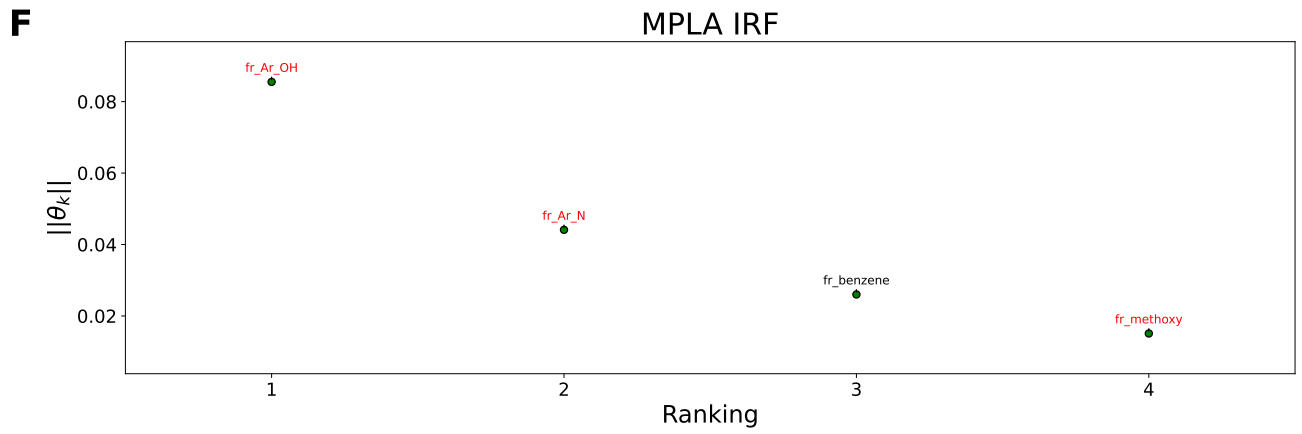
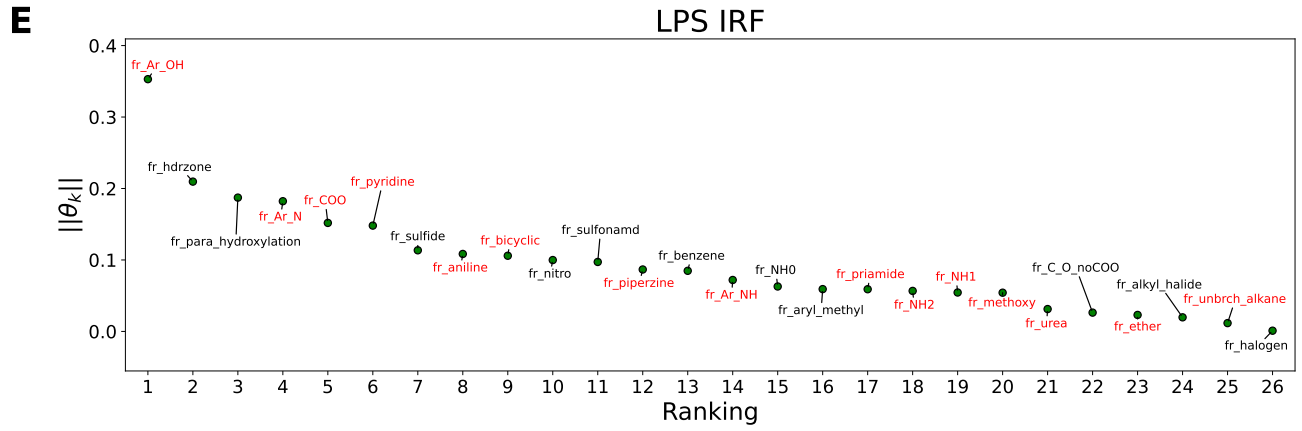


Figure S6: **Full accounting of non-zero learned coefficient weights θ_k associated with substructure features ranked by their magnitudes of weights in units of log2-fold change.** For each molecule n , we generate a standardized feature vector over all 70 features k , and collect these within a feature matrix $F_{n,k}$. Using this feature matrix, we apply LASSO regression to predict the calculated immunomodulation acquired from high throughput screening experiments and identifying the learned non-zero coefficient weights θ_k with the highest magnitude from the reduced feature set retained by the LASSO model. The features with positive weights are displayed in black text, while the ones with negative weights are displayed in red text. Positive weights imply that there is a positive correlation between the feature values and enhanced immunomodulation, while negative weights indicate a positive correlation between the feature values and inhibitory immunomodulation.

S2 Supplementary Experimental Methods and Results

S2.1 High throughput screening experiments

S2.1.1 Immunomodulators delivered alone do not affect immune responses of NF- κ B and IRF pathways

The raw data captured by the plate reader in our high throughput screening (HTS) experiments, which reflects the magnitude of the immune response induced by each agonist or the combination of an agonist and an immunomodulator, is presented in Figure S7. LPS, MPLA and CpG are generally good NF- κ B agonists while cGAMP struggles in activating NF- κ B pathway. LPS, MPLA and cGAMP are typically effective in activating the IRF pathway, while CpG faces difficulties in initiating the IRF pathway. Most candidate immunomodulators do not stimulate an immune response on their own, without the presence of agonists. There were only 27/2880 (=0.9%) immunomodulator candidates with which the samples show absorbance > 0.5 .

S2.2 Low throughput immunomodulation validation experiments

S2.2.1 Assay used for cytokine release profile testing

Monocytes were harvested from 6-week-old C57BL/6 mice and were differentiated into dendritic cells (BMDCs) using supplemented culture medium: RPMI 1640 (Life Technologies), 10% HIFBS (Sigma-Aldrich), Recombinant Mouse GM-CSF (carrier-free)(20 ng/ml; BioLegend), 2 mM l-glutamine (Life Technologies), 1% antibiotic- antimycotic (Life Technologies), and 50 μ M β -mercaptoethanol (Sigma-Aldrich). After 6 days of culture, BMDCs were plated at 100,000 cells per well and incubated with modulator (10 μ M). After 1 hour, agonist was added. Cells were incubated for 24 hours at 37°C and 5% CO₂. Supernatant cytokines were measured using Legendplex Mouse Inflammation Cytokine Kit (Biolegends) or a VeriKine IFN- β ELISA (PBL Assay Science).

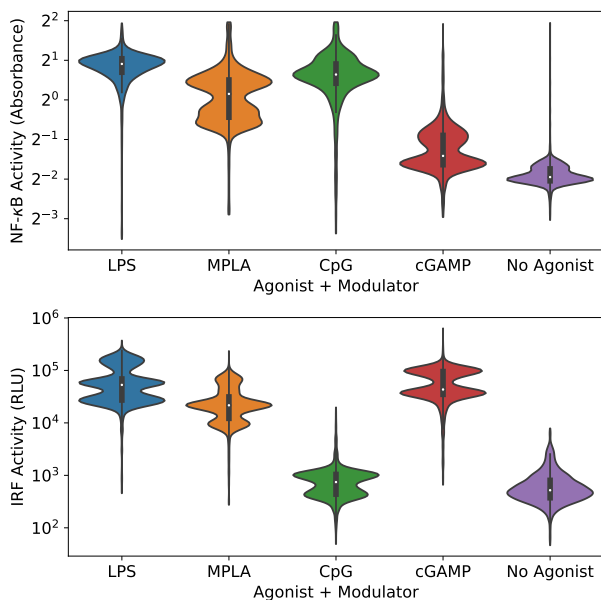


Figure S7: **Results of active learning screen as raw data.** NF- κ B activity was measured as absorbance readings at 620 nm, and IRF activity was measured as luminescence readings in the units of relative light unit (RLU). The higher the absorbance or luminescence, the stronger the corresponding immune activity. The CpG agonist shows minimal stimulation of the IRF pathway relative to no agonist present and was dropped from subsequent analyses.

S2.2.2 Immunomodulators delivered alone do not affect cytokine production

For transcription factor activity, modulators alone do not influence the immune responses (Figure S7). This also applies to cytokine secretion profile in that modulator do not stimulate cytokine release in the absence of agonist stimulation (Figure S8). It is the combination of agonist and immunomodulators that allow for significant enhancement and inhibition of both transcription factor activity and cytokine secretion.

S2.2.3 Comparing PME-4007 and MSA-2

MSA-2 is a recently discovered potent STING agonist that can stimulate STING pathway and was identified as a chemical inducer of interferon- β secretion⁴⁶ via a high throughput screening process which involved ~ 2.4 M compounds. We subjected our leading IFN- β inducing compound, PME-4007, to additional comparisons of its cytokine profile in the presence of cGAMP to that of MSA-2. We largely followed the protocol mentioned in Section 2.3

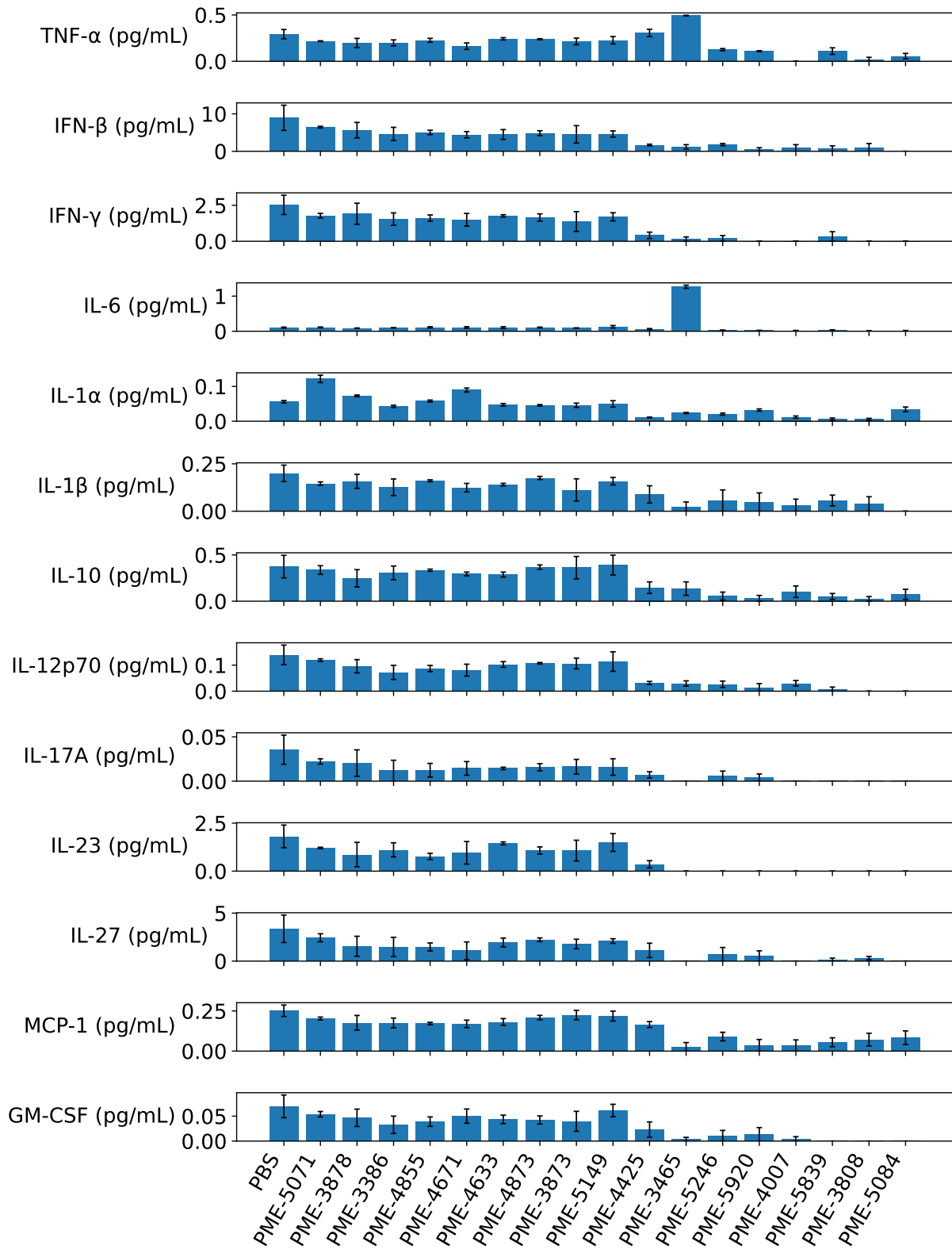


Figure S8: **Modulators (10 μM) alone minimally affects cytokine production 24 hours after addition.** Each panel shows the secreted concentration for a specific cytokine stimulated by PBS (as negative control) and 17 selected top-performing immunomodulator candidates (without the addition of agonists), addressed for 13 cytokines. The cytokine production result from the addition of modulators in the absence of agonists is within the same order of magnitude as that of the PBS negative control, showing that the addition of immunomodulators alone minimally affects the production of cytokines. Although for IL-6 production, there is one molecule PME-3465 that has much higher IL-6 secretion than PBS control, the magnitude of the amount of IL-6 being released is still minimal immunologically. The cytokine production is measured by LegendPlex.

in the main text. To obtain more stringent results for this low throughput comparison, we used more replicates ($N=6$). Considering the moderate cytotoxicity exhibited by PME-4007, we used a lower concentration at 2 μM instead of the generic concentration of 10 μM for all immunomodulators in all preceding experiments. To measure the cell viability of this experiment, we utilized the MTT assay. After BMDCs were incubated with or without the immunomodulator PME-4007 and with the agonist cGAMP or MSA-2 overnight, the cell supernatant was removed and 70 μL of PBS was added. Cells were incubated with 10 μL of 5 mg/mL (4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) for 3 hours at 37°C. The medium was then removed, and 70 μL of DMSO was added to solubilize formazan crystals. Plates were then read at 540 nm using a BioTek plate reader. Cell viability was calculated as a percentage of proliferation versus PBS-treated negative control cells. The results of MTT assay are shown in Figure S9A. Figure S9B and Figure S9C, as extensions to Figure 6B in the main text, show that the modulation profile of PME-4007 over cytokine secretion is not limited to IFN- β and TNF- α .

S2.3 Safety statement

During all wet lab experiments involved in this work, there were no unexpected or unusually high safety hazards encountered.

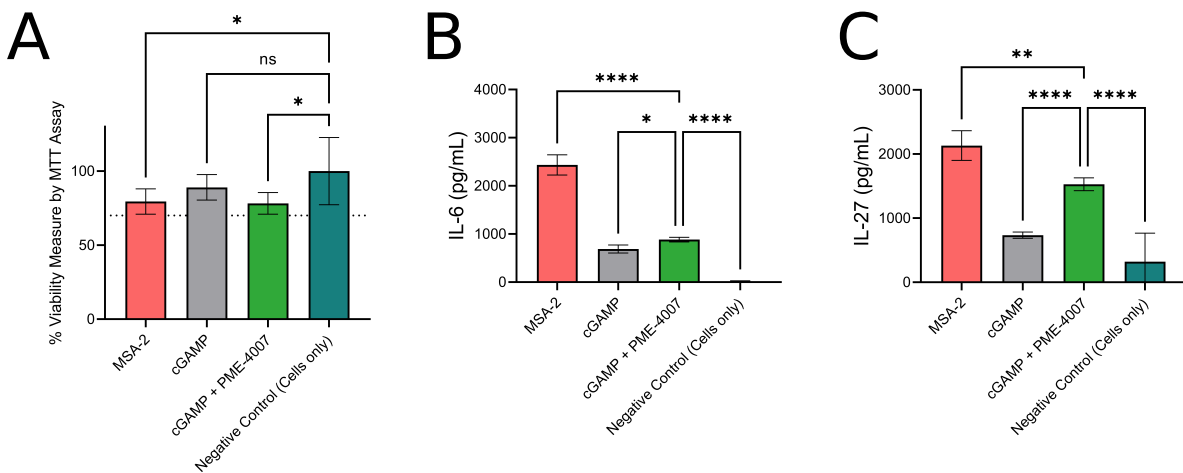


Figure S9: **PME-4007 induces minimal cytotoxicity while enhancing the secretion of IL-6 and IL-27 stimulated by cGAMP.** (A) MTT Assay indicates that MSA-2, cGAMP, and the combination of cGAMP and PME-4007 induce minimal cytotoxicity to the cells as the cell viability is all higher than the cutoff 70%, and close to the level of negative control resting cells. (B) PME-4007 slightly enhances the secretion of IL-6. MSA-2 strongly stimulates the secretion of IL-6 and is much stronger than the stimulation induced by cGAMP. (C) PME-4007 significantly enhances the secretion of IL-27 with cGAMP, and it is close to the level of stimulation induced by MSA-2.

S3 Supplementary Figures

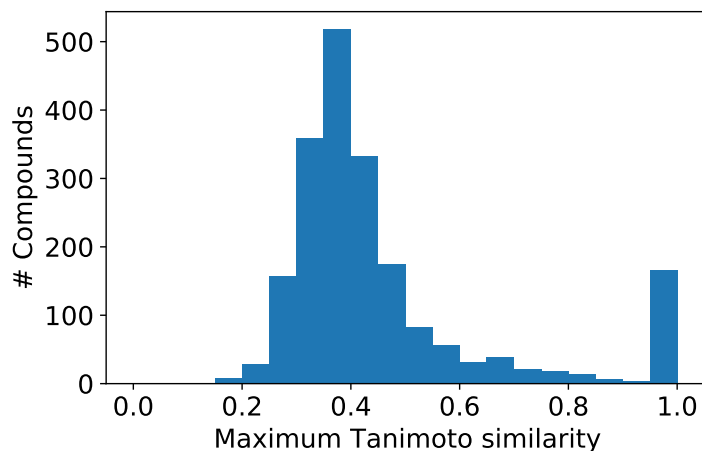


Figure S10: **Maximum pairwise Tanimoto similarities between the 2674 previously screened molecules used to train our initial GPR models and the top-performing candidates capable of >2-fold enhancement/inhibition identified by our active learning screen.** The maximum Tanimoto similarities for each of the molecules identified in the active learning screen were determined by $\max_j \text{Tanimoto}(\mathbf{x}_i, \mathbf{x}_j)$, where \mathbf{x}_i denoted the i^{th} molecule in the active learning screen and \mathbf{x}_j denoted the j^{th} molecule in the previous screen. The Tanimoto similarity quantifies the proportion of chemical substructures shared by a pair of molecules and it is a continuous number between 0 and 1.0, where 1.0 denotes complete topological identity. The peak near 1.0 indicates that the model samples molecules with similar structures to known good immunomodulators, but the large support for the histogram in the 0-0.6 region indicates that the active learning screen is also exploring molecules that are substantially diverse from the initial training data.

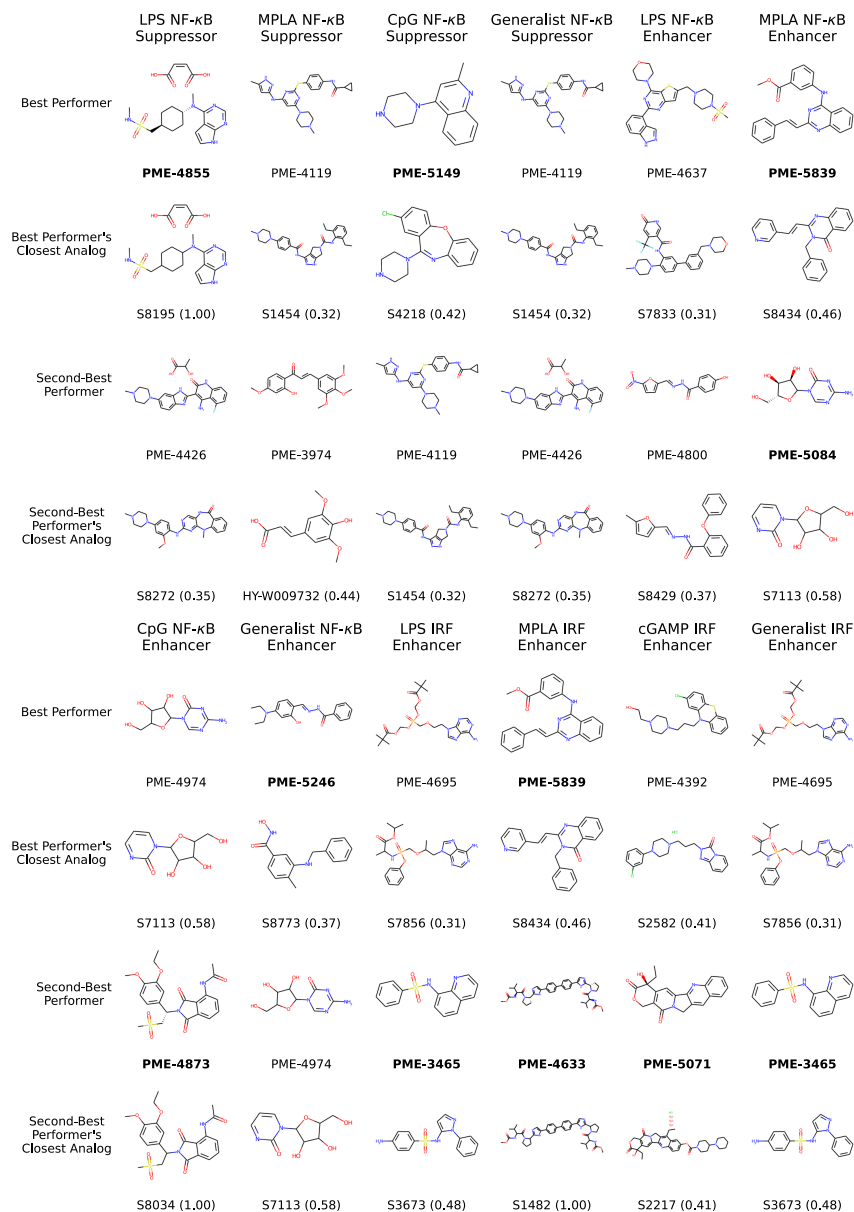


Figure S11: **Top performers identified in the active learning-assisted screen and their closest analog in the 2674 molecules from our prior screen.** The two top-performing immunomodulator candidates in each of the 12 functional objectives. We present for each molecule its chemical structure along with their code names. Below each best performer or second-best performer, we present the chemical structure, code names, and the Tanimoto similarity score of the closest analog to it. The Tanimoto similarity quantifies the proportion of chemical substructures shared by a pair of molecules and it is a continuous number between 0 and 1.0, where 1.0 denotes complete topological identity. Of the twelve closest analogs to the top performing molecules, 6/12 (50%) possess a Tanimoto similarity score to the top performer of less than 0.4 and 10/12 (83%) possess a score of less than 0.5, indicative of a significant level of dissimilarity in the chemical structure.

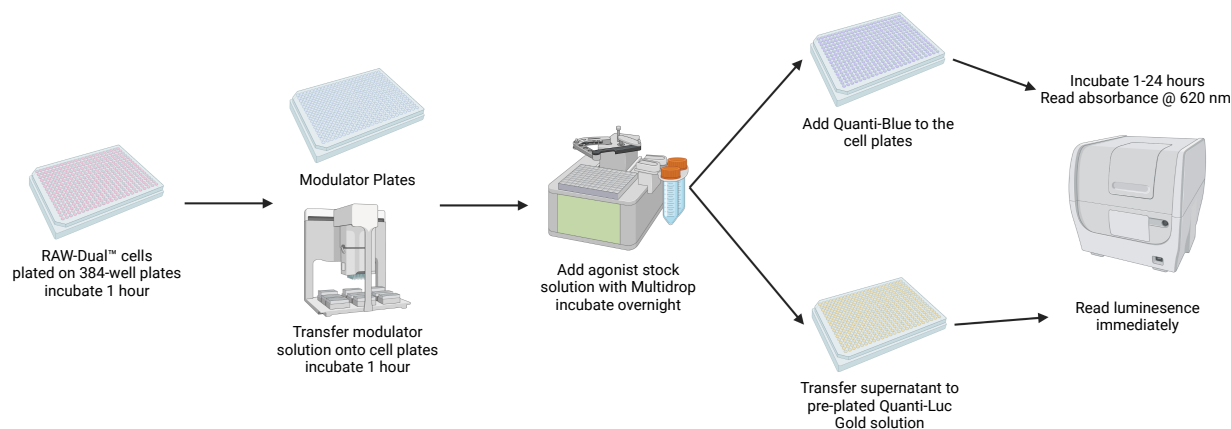
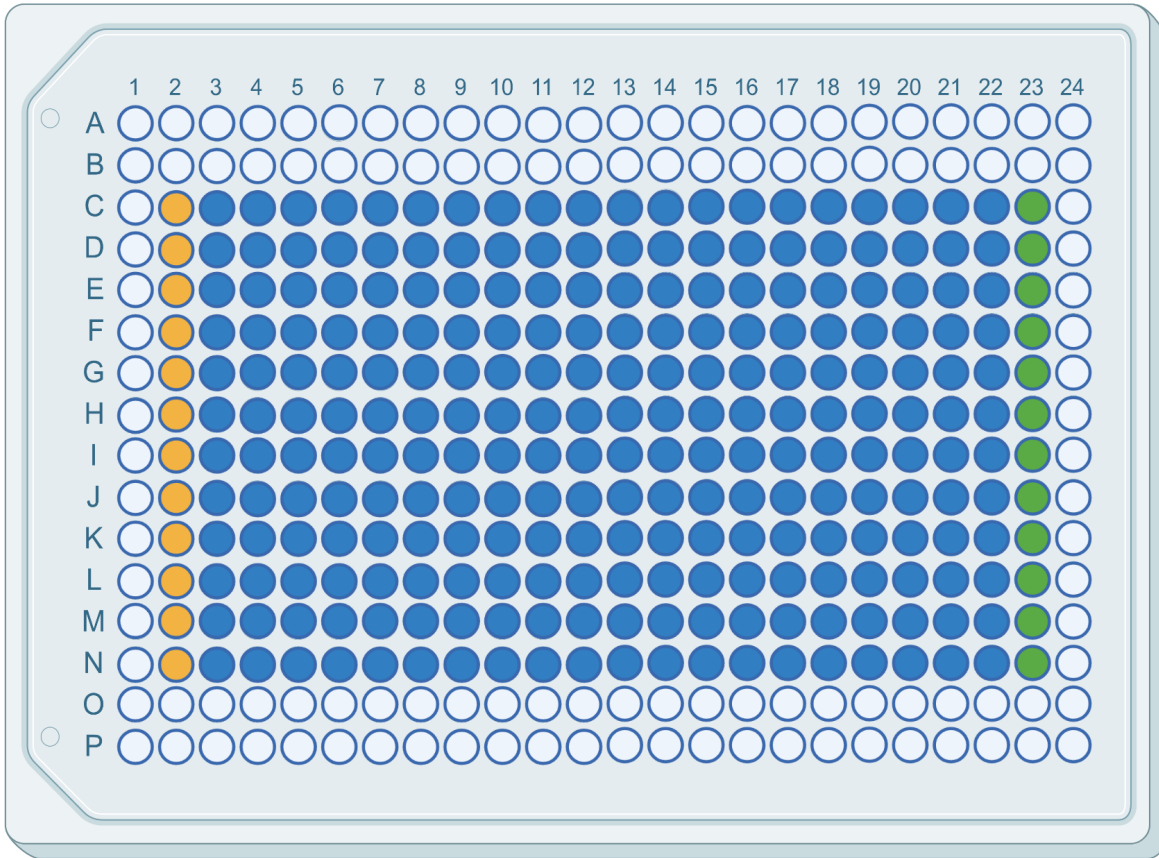


Figure S12: **High-throughput experimental screening protocol for NF- κ B transcription factor levels and IRF activity.** We seeded 50,000 RAW-Dual™ macrophages in 384-well plates in 45 μ L of complete media using a MultiDrop™ Combi liquid handler, then incubated for 1 hour at 37°C. We transferred immunomodulator compounds from source plates (10mM dissolved in dimethyl sulfoxide, DMSO) to a final concentration of 10 μ M using a Janus® G3 via pintool, then incubated for another 1 hour at 37°C. One of the four PRR agonists was added in 5 μ L of media to achieve the desired concentration into corresponding wells using a MultiDrop™ Combi liquid handler. For control plates where we test the effect of modulators in the absence of agonists, we added same volume of media instead. We incubated the cells overnight. We then add QUANTI-Blue™ and QUANTI-Luc™ reagents as appropriate, and read the absorbance at 620 nm and the luminescence using a BioTek Synergy™ Neo2 Hybrid multimode microplate reader, which correspond to the NF- κ B transcription factor levels and the IRF activity levels, after necessary calibration and standardization. Portions of this figure were created with BioRender.com.







-  Cell culture media
-  DMSO negative control
-  Agonist positive control
-  Testing modulators

Figure S13: **384-well plate layout.** One column and two rows on edges are not used to test and filled with cell culture media to avoid error induced by evaporation of water. Portions of this figure were created with BioRender.com.

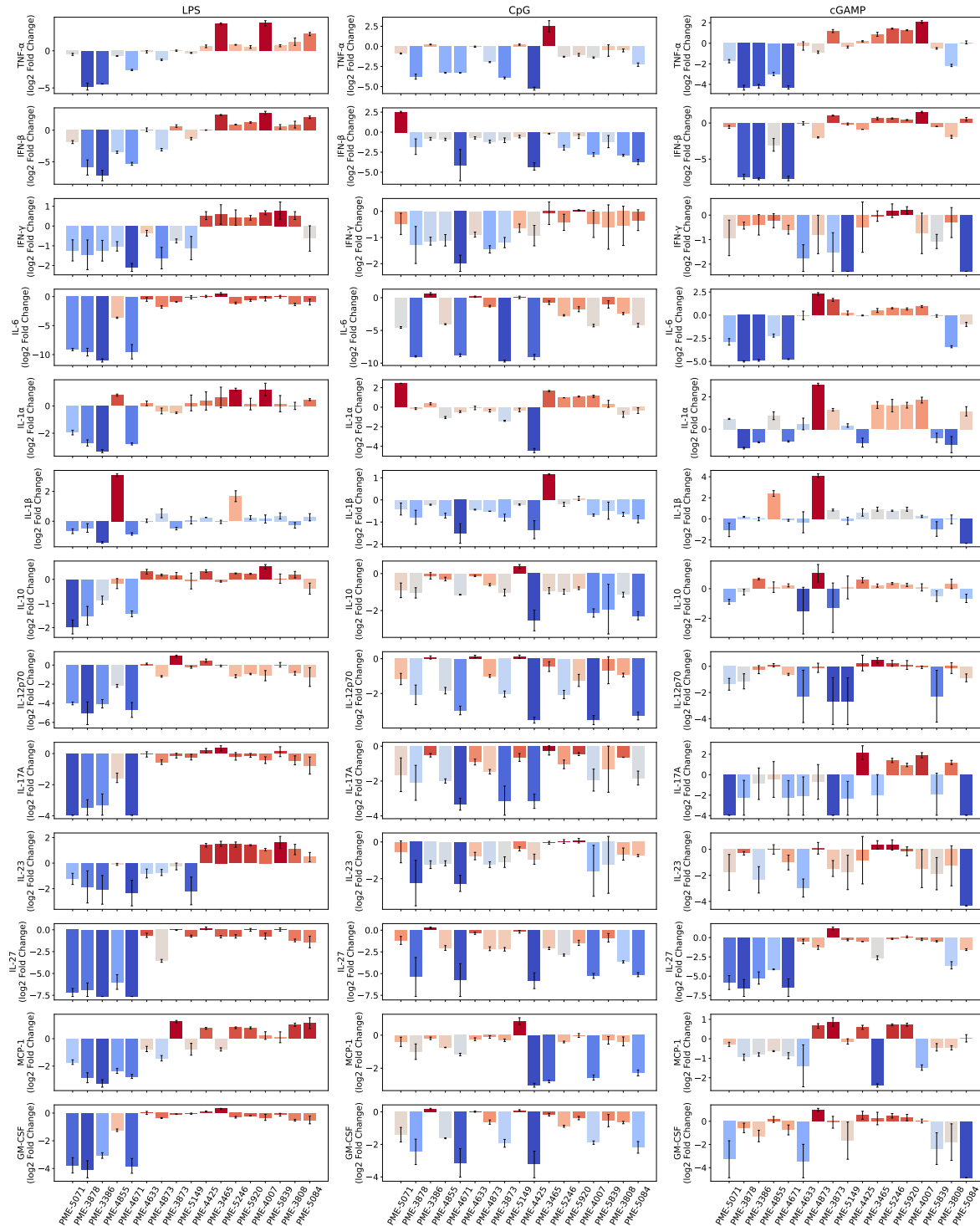


Figure S14: Immunomodulation of 17 selected top-performing candidates over cytokine release profiles activated by LPS, CpG and cGAMP. The extent of immunomodulation is visualized as bar plots with error bars indicating standard errors associated with each experiment. The bars are also color-coded by the log₂-fold change values for better clarity in showing those immunomodulation with large extent, representing potent enhancers or suppressors.

NF- κ B Suppressor (LPS Specialist)	53	60	23	25	0	6	0	1
NF- κ B Suppressor (MPLA Specialist)	24	37	7	18	1	0	0	0
NF- κ B Suppressor (CpG Specialist)	142	190	76	91	14	32	0	8
NF- κ B Suppressor (Generalist)	36	52	13	15	1	3	0	0
NF- κ B Enhancer (LPS Specialist)	172	22	60	4	7	0	0	0
NF- κ B Enhancer (MPLA Specialist)	65	114	22	42	0	1	0	0
NF- κ B Enhancer (CpG Specialist)	48	54	17	13	0	0	0	0
NF- κ B Enhancer (Generalist)	76	51	26	9	0	0	0	0
IRF Enhancer (LPS Specialist)	29	128	10	11	2	2	0	0
IRF Enhancer (MPLA Specialist)	35	210	7	27	2	1	0	0
IRF Enhancer (cGAMP Specialist)	11	16	0	5	0	0	0	0
IRF Enhancer (Generalist)	10	28	6	7	0	0	0	0
All Functional Goals	382	554	159	167	23	36	0	9
	1.5x		2x		5x		10x	
	Fold Change Cutoff							

Figure S15: **Comparison of the number of immunomodulators with 1.5 \times , 2 \times , 5 \times , and 10 \times enhancement or suppression identified in the present work compared to our prior screen.**⁵ For each immunological functional goal (rows), we present the number of compounds meeting the fold-change threshold (columns). In each cell, the left number represents the number of candidates meeting this criterion identified from our prior work,⁵ and the right number represents the number identified in the present work. In many combinations of functional goals and cutoffs, the active learning-guided screen yielded a significant increase in the population of immunomodulator candidates, and in five instances identified candidates with activity levels that were not achieved in our prior screen.

S4 Supplementary Tables

Table S5: **Source library of screened molecules.** The table shows (1) the number of compounds in each chemical screening library, (2) the number of compounds experimented with HTS in each library, (3) the number of good modulators (with >2-fold modulation) found in each library and (4) the ratio of the number of good modulators identified within our screen of that library to the total number of compounds within the library (i.e., the 3rd column divided by the 1st column). The last three libraries, namely Microsource Spectrum Collection, Prestwick Chemical Library and Selleckchem FDA-approved Drug Library, possess relatively higher ratios of good modulators identified. The statistics excluded non-viable compounds.

Chemical Screening Library	# compounds	# experimented	# good modulators	% good modulators
ChemBridge DS550-3	29547	571	52	0.18 %
ChemBridge ES550-2	49584	657	54	0.11 %
Life Chemicals Biologically Active Compound Library	7938	176	27	0.34 %
Life Chemicals Pre-plated Diversity Sets PS4	50233	546	22	0.04 %
Micro Source The Spectrum Collection	1821	350	10	0.55 %
Prestwick Chemical Library	1219	276	9	0.74 %
Selleckchem FDA-approved Drug Library	1499	708	26	1.73 %

Table S6: **List of agonists studied and their working concentration.**

Agonist	Target	Concentration ($\mu\text{g}/\text{mL}$)	Reference
LPS	TLR4	0.1	47
MPLA	TLR4	0.25	48
CpG (ODN 1826)	TLR9	0.5	49
cGAMP	STING	10	22

References

- (1) Irwin, J. J.; Shoichet, B. K. *Journal of Chemical Information and Modeling* **2005**, *45*, 177–182.
- (2) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. *Journal of Chemical Information and Modeling* **2012**, *52*, 1757–1768.
- (3) Sterling, T.; Irwin, J. J. *Journal of Chemical Information and Modeling* **2015**, *55*, 2324–2337.
- (4) Irwin, J. J.; Tang, K. G.; Young, J.; Dandarchuluun, C.; Wong, B. R.; Khurelbaatar, M.; Moroz, Y. S.; Mayfield, J.; Sayle, R. A. *Journal of Chemical Information and Modeling* **2020**, *60*, 6065–6073.
- (5) Kim, J. Y.; Rosenberger, M. G.; Chen, S.; IP, C. K.; Bahmani, A.; Chen, Q.; Shen, J.; Tang, Y.; Wang, A.; Kenna, E.; Son, M.; Tay, S.; Ferguson, A. L.; Esser-Kahn, A. P. *ACS Central Science* **2023**, *9*, 427–439.
- (6) Moser, B. A.; Escalante-Buendia, Y.; Steinhardt, R. C.; Rosenberger, M. G.; Cas-saidy, B. J.; Naorem, N.; Chon, A. C.; Nguyen, M. H.; Tran, N. T.; Esser-Kahn, A. P. *Frontiers in Immunology* **2020**, *11*, 511513.
- (7) Weininger, D. *Journal of Chemical Information and Computer Sciences* **1988**, *28*, 31–36.
- (8) Weininger, D.; Weininger, A.; Weininger, J. L. *Journal of Chemical Information and Computer Sciences* **1989**, *29*, 97–101.
- (9) Weininger, D. *Journal of Chemical Information and Computer Sciences* **1990**, *30*, 237–243.
- (10) O’Boyle, N. M. *Journal of Cheminformatics* **2012**, *4*, 22.

- (11) Krenn, M.; Häse, F.; Nigam, A.; Friederich, P.; Aspuru-Guzik, A. *Machine Learning: Science and Technology* **2020**, *1*, 045024.
- (12) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. *ACS Central Science* **2018**, *4*, 268–276.
- (13) Paszke, A. et al. *PyTorch: An imperative style, high-performance deep learning library*; Curran Associates, Inc., 2019; pp 8024–8035.
- (14) Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. 2014.
- (15) Kullback, S.; Leibler, R. A. *The Annals of Mathematical Statistics* **1951**, *22*, 79–86.
- (16) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. 2022.
- (17) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. 2017.
- (18) Lucas, J.; Tucker, G.; Grosse, R.; Norouzi, M. **2019**,
- (19) Rasmussen, C. E.; Williams, C. K. I. In *Gaussian Processes for Machine Learning*; Bach, F., Ed.; MIT Press: Cambridge, MA, USA, 2005.
- (20) Pedregosa, F. et al. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
- (21) Mohammadi, H.; Riche, R. L.; Durrande, N.; Touboul, E.; Bay, X. An analytic comparison of regularization methods for Gaussian Processes. 2017.
- (22) Zhang, X.; Shi, H.; Wu, J.; Zhang, X.; Sun, L.; Chen, C.; Chen, Z. *J. Molecular Cell* **2013**, *51*, 226–235.
- (23) Močkus, J. *On Bayesian Methods for Seeking the Extremum*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1975; pp 400–404.

- (24) Brochu, E.; Cora, V. M.; de Freitas, N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. 2010.
- (25) Snoek, J.; Larochelle, H.; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. 2012.
- (26) Van Veldhuizen, D. A.; Lamont, G. B. Evolutionary computation and convergence to a pareto front. Proceedings of the Genetic Programming Conference. 1998; pp 221–228.
- (27) Irshad, F.; Karsch, S.; Döpp, A. Expected hypervolume improvement for simultaneous multi-objective and multi-fidelity optimization. 2021.
- (28) Wang, B.; Singh, H. K.; Ray, T. Bridging kriging believer and expected improvement using bump hunting for expensive black-box optimization. Lille France, 2021; pp 211–212.
- (29) Ginsbourger, D.; Le Riche, R.; Carraro, L. *A multi-points criterion for deterministic parallel global optimization based on gaussian processes*; 2008.
- (30) Mohr, B.; Shmilovich, K.; Kleinwächter, I. S.; Schneider, D.; Ferguson, A. L.; Bereau, T. *Chemical Science* **2022**, *13*, 4498–4511.
- (31) Shmilovich, K.; Mansbach, R. A.; Sidky, H.; Dunne, O. E.; Panda, S. S.; Tovar, J. D.; Ferguson, A. L. *The Journal of Physical Chemistry B* **2020**, *124*, 3873–3891.
- (32) Ferguson, A. L.; Ranganathan, R. *ACS Macro Letters* **2021**, *10*, 327–340.
- (33) Shmilovich, K.; Yao, Y.; Tovar, J. D.; Katz, H. E.; Schleife, A.; Ferguson, A. L. *Molecular Systems Design & Engineering* **2022**, *7*, 447–459.
- (34) Shmilovich, K.; Panda, S. S.; Stouffer, A.; Tovar, J. D.; Ferguson, A. L. *Digital Discovery* **2022**, *1*, 448–462.

- (35) Panda, S. S.; Shmilovich, K.; Herringer, N. S. M.; Marin, N.; Ferguson, A. L.; To-
var, J. D. *Langmuir* **2021**, *37*, 8594–8606.
- (36) Koscher, B. et al. **2023**,
- (37) Dave, A.; Mitchell, J.; Kandasamy, K.; Wang, H.; Burke, S.; Paria, B.; Póczos, B.;
Whitacre, J.; Viswanathan, V. *Cell Reports Physical Science* **2020**, *1*, 100264.
- (38) Agarwal, G.; Doan, H. A.; Robertson, L. A.; Zhang, L.; Assary, R. S. *Chemistry of
Materials* **2021**, *33*, 8133–8144.
- (39) Deshwal, A.; Simon, C. M.; Doppa, J. R. *Molecular Systems Design & Engineering*
2021, *6*, 1066–1086.
- (40) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G.
Chemical Reviews **2021**, *121*, 10073–10141.
- (41) Raddi, R. M.; Voelz, V. A. *Journal of Computer-Aided Molecular Design* **2021**, *35*,
953–961.
- (42) Krynski, M.; Rossi, M. *npj Computational Materials* **2021**, *7*.
- (43) Rogers, D.; Hahn, M. *Journal of Chemical Information and Modeling* **2010**, *50*, 742–
754.
- (44) Landrum, G. et al. rdkit/rdkit: 2021_09_2 (Q3 2021) Release. 2021; [https://doi.
org/10.5281/zenodo.5589557](https://doi.org/10.5281/zenodo.5589557).
- (45) Le, T.; Winter, R.; Noe, F.; Clevert, D.-A. **2020**,
- (46) Pan, B.-S. et al. *Science* **2020**, *369*, eaba6098.
- (47) Vargas-Caraveo, A.; Sayd, A.; Robledo-Montaña, J.; Caso, J. R.; Madrigal, J. L. M.;
García-Bueno, B.; Leza, J. C. *Journal of Neuroinflammation* **2020**, *17*, 6.

- (48) Mata-Haro, V.; Cekic, C.; Martin, M.; Chilton, P. M.; Casella, C. R.; Mitchell, T. C. *Science* **2007**, *316*, 1628–1632.
- (49) Ballas, Z. K.; Krieg, A. M.; Warren, T.; Rasmussen, W.; Davis, H. L.; Waldschmidt, M.; Weiner, G. J. *Journal of Immunology (Baltimore, Md.: 1950)* **2001**, *167*, 4878–4886.