Optimizing molecules using efficient queries from property evaluations

Samuel C. Hoffman, Vijil Chenthamarakshan, Kahini Wadhawan, Pin-Yu Chen and Payel Das

Machine learning-based methods have shown potential for optimizing existing molecules with more desirable properties, a critical step towards accelerating new chemical discovery. Here we propose QMO, a generic query-based molecule optimization framework that exploits latent embeddings from a molecule autoencoder. QMO improves the desired properties of an input molecule based on efficient queries, guided by a set of molecular property predictions and evaluation metrics. We show that QMO outperforms existing methods in the benchmark tasks of optimizing small organic molecules for drug-likeness and solubility under similarity constraints. We also demonstrate substantial property improvement using QMO on two new and challenging tasks that are also important in real-world discovery problems: (1) optimizing existing potential SARS-CoV-2 main protease inhibitors towards higher binding affinity and (2) improving known antimicrobial peptides towards lower toxicity. Results from QMO show high consistency with external validations, suggesting an effective means to facilitate material optimization problems with design constraints.
to better bind to the SARS-CoV-2 target protein sequence while preserving the desired drug properties, a sufficient number of such paired molecule sequences is unavailable. We also note that these two categories are not exclusive. Guided search can be jointly used with translation.

In this Article, we propose a novel query-based molecule optimization (QMO) framework, as illustrated in Fig. 1. In this context, a query to a designed loss function for QMO gives the corresponding numerical value obtained through the associated property evaluations. Efficiency refers to the performance of the optimization results given a query budget. QMO uses an encoder–decoder and external guidance, but it differs from existing works in the following aspects. (1) QMO is a generic end-to-end optimization framework that reduces the problem complexity by decoupling representation learning and guided search. It applies to any plug-in (pre-trained) encoder–decoder with continuous latent representations. It is also a unified and principled approach that incorporates multiple predictions and evaluations made directly at the molecule sequence level into guided search without further model fitting. (2) To achieve efficient end-to-end optimization with discrete molecule sequences and their continuous latent representations, QMO adopts a novel query-based guided search method based on zeroth-order optimization\(^{39,40}\), which performs efficient mathematical optimization using only function evaluations (more details are provided in Supplementary Section 1). Its query-based guided search enables direct optimization over the property evaluations provided by chemical informatics/simulation software packages or prediction application programming interfaces (APIs), and it supports guided search with exact property evaluations that only operate at the molecular sequence level instead of latent representations or surrogate models. To the best of our knowledge, this work is the first study that facilitates molecule optimization by disentangling molecule representation learning and guided search, and by exploiting zeroth-order optimization for efficient search in the molecular property landscape. The success of QMO can be attributed to its data efficiency, achieved by exploiting the latent representations learned from abundant unlabelled data and the guidance for property prediction trained on relatively limited labelled data.

We first demonstrate the effectiveness of QMO by means of two sets of standard benchmarks. On two existing (and simpler) MO benchmark tasks—optimizing QED (quantitative estimate of drug-likeness)\(^41\) and penalized log\(P\) (reflecting the octanol–water partition coefficient)\(^42\) with similarity constraints—QMO attains superior performance over existing baselines, showing at least 15% higher success on QED optimization and an absolute improvement of 1.7 on penalized log\(P\). The performance of QMO on these molecular physical property benchmarks shows its potential for optimizing material design before synthesis, which is critical in many applications and fields, such as the food industry, agrochemicals, pesticides, drugs, catalysts and waste chemicals.

Next, as a motivating discovery use case that also, at least to some extent, reflects the complexity of real discovery problems\(^42\), we demonstrate how QMO can be used to improve the binding affinity of a number of existing inhibitor molecules for the SARS-CoV-2 main protease (M\(^{pp}\)), one of the most extensively studied drug targets for SARS-CoV-2. As an illustration, Fig. 2 shows the top docking poses of dipyridamole and its QMO-optimized variant with SARS-CoV-2 M\(^{pp}\). We formulate this task as an optimization over predicted binding affinity (obtained using a pre-trained machine learning model) starting from an existing molecule of interest (a lead mol-
A molecule). Because experimental half-maximal inhibitory concentration (IC₅₀) values are widely available, we use them as a measure for protein–ligand binding affinity. The pIC₅₀ of the optimized molecule (where pIC₅₀ = −log₁₀(IC₅₀)) is constrained to be above 7.5, a sign of good affinity, while the Tanimoto similarity between the optimized molecule and the original molecule is maximized. Retaining high similarity while optimizing the initial lead molecule means important chemical characteristics can be maximally preserved. Moreover, a high similarity to existing leads is important for a rapid response to a novel pathogen such as SARS-CoV-2, as it is more likely to leverage existing knowledge and manufacturing pipelines for the synthesis and wet-lab evaluation of the optimized variants. Moreover, the chance of optimized variants inducing adverse effects is potentially low. Our results show that QMO can find molecules with high similarity and improved affinity, while preserving other properties of interest such as drug-likeness.

We also consider the task of optimizing existing antimicrobial peptides towards lower selective toxicity, which is critical for accelerating safe antimicrobial discovery. In this task, QMO shows a high success rate (~72%) in improving the toxicity of antimicrobial peptides, and the properties of optimized molecules are consistent with external toxicity and antimicrobial activity classifiers. Finally, we perform property landscape visualization and trajectory analysis of QMO to illustrate its efficiency and diversity in finding improved molecules with desired properties.

We emphasize that QMO is a generic-purpose optimization algorithm that enables optimization over discrete spaces (for example, sequences and graphs), which involves searching over a latent space of the system by using guidance from (expensive) black-box function evaluations. Beyond the organic and biological molecule optimization applications considered in this study, QMO can be applied to the optimization of other classes of material, such as inorganic solid-state materials like metal alloys or metal oxides.

**Results**

**Representations of molecules.** In our QMO framework, we model a molecule as a discrete string of chemical or amino-acid characters (a sequence). Depending on the downstream MO tasks, the sequence representation can either be a string of natural amino acids or a string designed for encoding small organic chemicals. In particular, the simplified molecular input line entry (SMILES) representation describes the structure of small organic molecules using short ASCII strings. Without loss of generality, we define X := X × X × ... × X as the product space containing every possible molecule sequence of length m, where X denotes the set of all chemical characters. To elucidate the problem complexity, considering the 20 protein-building amino acids as characters in a peptide sequence, the number of possible candidates in the space of sequences with length m = 60 is already reaching the number of atoms in the known universe (~10⁸⁰). Similarly, the space of small molecules with therapeutic potential is estimated to be on the order of 10⁶⁰ (refs. ⁴¹,⁴²). Therefore, the problem of MO in the ambient space X can be computationally inefficient as the search space grows combinatorially with sequence length m.

**Encoder–decoder for learning latent molecule representations.** To address the issue of the large search space for molecule sequences, QMO adopts an encoder–decoder framework. The encoder Enc : Xᵐ ↦ ℝᵈ encodes a sequence x ∈ Xᵐ to a low-dimensional continuous real-valued representation of dimension d, denoted by an embedding vector z = Enc(x). The decoder Dec : ℝᵈ ↦ Xᵐ decodes the latent representation z of x back to the sequence representation, denoted by ẑ = Dec(z). We note that, depending on the encoder–decoder implementation, the input sequence x and the decoded sequence ẑ may have different lengths. On the other hand, the latent dimension d is universal (fixed) to all sequences. Winter et al.⁴⁶ proposed a novel molecular descriptor and used it for an autoencoder to learn latent representations featuring high similarity between the original and reconstructed sequences. QMO applies to any plug-in (pre-trained) encoder–decoder with continuous latent representations and thus decouples representation learning and guided search, reducing the problem complexity of MO.
MO formulation via guided search. In addition to leveraging learned latent representations from a molecule encoder–decoder, our QMO framework incorporates molecular property prediction models and similarity metrics at the sequence level as external guidance. Specifically, for any given sequence \( x \in \mathbb{X}^m \) we use a set of \( I \) separate prediction models \( \{ f_i(x) \}_{i=1}^m \) to evaluate the properties of interest for MO. In principle, for a candidate sequence \( x \), a set of thresholds \( \{ \tau_i \}_{i=1}^m \) on its property predictions \( \{ f_i(x) \}_{i=1}^m \) is used for validating the condition \( f_i(x) \geq \tau_i \) for all \( i \in [I] \), where \( [I] \) denotes the integer set \( \{1, 2, \ldots, I\} \). Moreover, we can simultaneously impose a set of \( I \) separate prediction models \( \{ g_i(x; S) \} \) in the optimization process, such as molecular similarity, relative to a set of reference molecule sequences denoted by \( S \).

Our QMO framework covers two practical cases in MO: (1) optimizing molecular similarity while satisfying the desired chemical properties and (2) optimizing chemical properties with similarity constraints. It can be easily extended to other MO settings that can be formulated via \( \{ f_i(x) \}_{i=1}^m \) and \( \{ g_i(x; S) \} \). In what follows, we formally define our designed loss function for QMO as case (1). Given a starting molecule sequence \( x_0 \) (a lead molecule) and a pre-trained encoder–decoder, let \( x = \text{Dec}(z) \) denote a candidate sequence decoded from a latent representation \( z \in \mathbb{R}^d \). Our QMO framework aims to find an optimized sequence by solving the following continuous optimization problem:

\[
\text{Minimize}_{z \in \mathbb{R}^d} \quad \sum_{i=1}^I \max \{ \tau_i - f_i(\text{Dec}(z)), 0 \}
\]

Property validation loss (to be minimized)

\[
- \sum_{j=1}^J d_j \times g_j(\text{Dec}(z)|S)
\]

Molecular score (to be maximized)

The first term, \( \sum_{i=1}^I \max \{ \tau_i - f_i(\text{Dec}(z)), 0 \} \), quantifies the loss of property constraints and is presented as the sum of hinge loss over all property predictions, which approximates the binary property validation relative to the required thresholds \( \{ \tau_i \}_{i=1}^m \). It achieves the optimal value (that is, 0) only when the candidate sequence \( x = \text{Dec}(z) \) satisfies all the desired properties, which is equivalent to the condition \( f_i(\text{Dec}(z)) \geq \tau_i \) for all \( i \in [I] \). The second term, \( \sum_{j=1}^J d_j \times g_j(\text{Dec}(z)|S) \), corresponds to a set of molecular similarity scores to be maximized (therefore a minus sign in the minimization formulation). The reference sequence set \( S \) can be the starting sequence such that \( S = \{ x_0 \} \), or a set of molecules. The positive coefficients \( \{ d_j \}_{j=1}^J \) are associated with the set of molecular similarity scores \( \{ g_j(\text{Dec}(z)|S) \} \), respectively. It is worth mentioning that the use of the latent representation \( z \) as the optimization variable in a low-dimensional continuous space greatly facilitates the original MO problem in a high-dimensional discrete space. The optimization variable \( z \) can be initialized as the latent representation of \( x_0 \), denoted by \( z_0 = \text{Enc}(x_0) \).

Similarly, for case (2), the optimization problem is formulated as

\[
\text{Minimize}_{z \in \mathbb{R}^d} \quad \sum_{i=1}^J \max \{ \eta_i - g_i(\text{Dec}(z)|S), 0 \}
\]

Molecular constraint loss (to be minimized)

\[
- \sum_{i=1}^I \gamma_i \times f_i(\text{Dec}(z))
\]

Property score (to be maximized)

where \( \{ \eta_i \}_{i=1}^I \) are the similarity score constraints and \( \{ \gamma_i \}_{i=1}^I \) are positive coefficients of the property scores \( \{ f_i(\text{Dec}(z)) \} \).

QMO procedure. Although we formulate MO as an unconstrained continuous minimization problem, we note that solving it for a feasible candidate sequence \( x = \text{Dec}(z) \) is not straightforward, for two reasons. (1) The output of the decoder \( x = \text{Dec}(z) \) is a discrete sequence, which imposes challenges on any gradient-based (and high-order) optimization method because acquiring the gradient of \( z \) becomes non-trivial. Even by resorting to the Gumbel-softmax sampling trick for discrete outputs\(^\text{10}\), the large output space of the decoder may render it ineffective. (2) In practice, many molecular property prediction models and molecular metrics are computed in an access-limited environment, such as prediction APIs and chemical software, which only allow inference on a queried sequence and prohibit other functionalities such as gradient computation. To address these two issues, we use zeroth-order optimization in our QMO framework (details of the procedure are provided in the Methods) to provide a generic and model-agnostic approach for solving the problem formulation in equations (1) and (2), using only the inference results of \( \{ f_i \}_{i=1}^I \) and \( \{ g_i \}_{i=1}^J \) on queried sequences.

Let \( \text{Loss}(z) \) denote the objective function to be minimized, as defined in either equation (1) or equation (2). Our QMO framework uses zeroth-order gradient descent to find a solution, which mimics the descent steps on the loss landscape in gradient-based solvers but only uses the function values \( \text{Loss}(z) \) of queried sequences. Specifically, at the \( t \)th iteration of the zeroth-order optimization process, the iterate (candidate embedding vector) \( z^{t+1} \) is updated by

\[
z^{t+1} = z^t - \alpha_t \times \nabla \text{Loss}(z^t),
\]

where \( \alpha_t \geq 0 \) is the step size at the \( t \)th iteration, and the true gradient \( \nabla \text{Loss}(z^t) \) (which is challenging or infeasible to compute) is approximated by the pseudo gradient \( \nabla \text{Loss}(z^t) \). The pseudo gradient \( \nabla \text{Loss}(z^t) \) is estimated by \( Q \) independent random directions

\[
\nabla \text{Loss}(z^t) = \frac{d}{\beta} \cdot \sum_{q=1}^Q \left[ \text{Loss}(z^t + \beta u^q) - \text{Loss}(z^t) \right] \times u^q,
\]

where \( d \) is the dimension of the latent space of the encoder–decoder used in QMO, and \( \beta > 0 \) is a smoothing parameter used to perturb the embedding vector \( z^t \) for neighbourhood sampling with \( Q \) random directions \( \{ u^q \}_{q=1}^Q \) that are independently and identically sampled on a \( d \)-dimensional unit sphere. Figure 1 presents an illustration of random neighbourhood sampling. In our implementation, we sample \( \{ u^q \}_{q=1}^Q \) using a zero-mean \( d \)-dimensional isotropic Gaussian random vector divided by its Euclidean norm, such that the resulting samples are drawn uniformly from the unit sphere. Intuitively, the gradient estimator in equation (4) can be viewed as an average of \( Q \) random directional derivatives along the sampled directions \( \{ u^q \}_{q=1}^Q \). The constant \( \frac{d}{\beta} \) in equation (4) ensures the norm of the estimated gradient is of the same order as that of the true gradient\(^\text{10a}\).

A schematic example of the QMO procedure is illustrated in Fig. 1 using binding affinity and Tanimoto similarity as property evaluation criteria. Note that, based on the iterative optimization step in equation (3), QMO only uses function values queried at the original and perturbed sequences for optimization. The query count made on the Loss function for computing \( \nabla \text{Loss}(z^t) \) is \( Q + 1 \) per iteration. Larger \( Q \) further reduces the gradient estimation error at the price of increased query complexity. When solving equation (1), an
iterate $z^{(t)}$ is considered as a valid solution if its decoded sequence $\text{Dec}(z^{(t)})$ satisfies the property conditions $f_i(\text{Dec}(z^{(t)})) \geq \tau_i$ for all $i \in [I]$. Similarly, when solving equation (2), a valid solution $z^{(t)}$ means $g_i(\text{Dec}(z^{(t)})) \geq \eta_i$ for all $j \in [J]$. Finally, QMO returns a set of found solutions (returning null if in vain). Detailed descriptions for the QMO procedure are provided in the Methods.

**Three sets of MO tasks with multiple property evaluation criteria.** In what follows, we demonstrate the performance of our proposed QMO framework on three sets of tasks that aim to optimize molecular properties with constraints, including standard MO benchmarks and challenging tasks relating to real-world discovery problems. The pre-trained encoder–decoder and the hyperparameters of QMO for each task are specified in the Methods and in Supplementary Section 3.

**Benchmarks on QED and penalized logP optimization.** We start by testing QMO on two single property targets: penalized logP and QED\(^{41}\). LogP is the logarithm of the partition ratio of the solute between octanol and water. Penalized logP is defined as the logP minus the synthetic accessibility (SA) score\(^{42}\). Given a similarity constraint, finding an optimized molecule that improves the drug-likeness of compounds using the QED score (from a range of $[0.7, 0.8]$ to $[0.9, 1.0]$\(^{43}\)) or improves the penalized logP score\(^{44}\) are two widely used benchmarks. For a pair of original and optimized sequences $(x_0, x)$, we use the QMO formulation in equation (2) with the Tanimoto similarity (ranging from 0 to 1) over Morgan fingerprints\(^{45}\) as $g_{\text{Tanimoto}}(x(x_0))$ and the interested property score (QED or penalized logP) as $f_{\text{score}}(x)$. Following the same setting as existing works, the threshold $\delta$ for $g_{\text{Tanimoto}}(x(x_0))$ is set as either 0.4 or 0.6. We use RDKIt (open-source cheminformatics; http://www.rdkit.org) to compute QED and logP, and use MOSES\(^{45}\) to compute SA, where $f_{\text{penalized logP}}(x) = \log P(x) - \text{SA}(x)$.

In our experiments, we use the same set of 800 molecules with low penalized logP scores and 800 molecules with QED in $[0.7, 0.8]$ chosen from the ZINC test set\(^{46}\) as in Jin et al.\(^{21}\) as our starting sequences. We compare QMO with various guided-search and translation-based methods in Extended Data Figs. 1 and 2. Baseline results are obtained from the literature\(^{15,18}\) that use machine learning for solving the same task.

For the QED optimization task, the success rate, defined as the percentage of improved molecules having similarity greater than $\delta=0.4$, is shown in Extended Data Fig. 1. QMO outperforms all baselines by at least 15%. For penalized logP task, the molecules optimized by QMO outperform the baseline results by a notable margin, as shown in Extended Data Fig. 2. The increased standard deviation in QMO is an artefact of having some molecules with much improved penalized logP scores (Supplementary Section 4).

Although the above-mentioned molecular property optimization tasks provide well-defined benchmarks for testing our QMO algorithm, it is well recognized that such tasks are easy to solve and do not capture the complexity associated with real-world discovery\(^{41}\). For example, it is trivial to achieve state-of-the-art results for logP optimization by generating long saturated hydrocarbon chains\(^{35}\). Coeley et al.\(^{41}\) have proposed that MO goals that better reflect the complexity of real discovery tasks might include binding or selectivity attributes. Therefore, in the remainder of this Article we consider two such tasks: (1) optimizing the binding affinity of existing SARS-CoV-2 M\(^{\text{TM}}\) inhibitor molecules and (2) lowering the toxicity of known antimicrobial peptides.

**Optimizing existing SARS-CoV-2 main protease inhibitor molecules toward better IC\(_{50}\).** To provide a timely solution and accelerate the drug discovery against a new virus such as SARS-CoV-2, it is a sensible practice to optimize known leads to facilitate design and production as well as to minimize the emergence of adverse effects. Here we focus on the task of optimizing the parent-molecule structure of a set of existing SARS-CoV-2 M\(^{\text{TM}}\) inhibitors. Specifically, we use the QMO formulation in equation (1), a pre-trained binding affinity predictor\(^{49}\) $f_{\text{affinity}}$ (the output is the pIC\(_{50}\) value) and the Tanimoto similarity $g_{\text{Tanimoto}}$ between the original and optimized molecules. Given a specific inhibitor molecule $x_0$, we aim to find an optimized molecule $x$ such that $f_{\text{affinity}}(x) \geq \tau_{\text{affinity}}$ while $g_{\text{Tanimoto}}(x(x_0))$ is maximized.

For this task, we start by assembling 23 existing molecules shown to have weak to moderate affinity with SARS-CoV-2 M\(^{\text{TM}}\)\(^{34,35}\). These are generally in the micromolar range of IC\(_{50}\), a measure of inhibitory potency (Supplementary Section 3 describes the experimental IC\(_{50}\) values). We choose the target affinity threshold $\tau_{\text{affinity}}$ as pIC\(_{50} \geq 7.5$, which implies strong affinity. Table 1 shows the final optimized molecules compared to their initial state (that is, the original lead molecule). We highlight common substructures and show a similarity map to emphasize the changes. The results of all 23 inhibitors are summarized in Supplementary Table 2.

Because all of these 23 inhibitors are reported to bind to the substrate-binding pocket of M\(^{\text{TM}}\), we investigate possible binding mode alterations of the QMO-optimized molecules. It should be noted that a direct comparison of IC\(_{50}\) with binding free energy (BFE) is not always possible, as the relationship of binding affinity and IC\(_{50}\) for a given compound varies depending on the assay conditions and the compound’s mechanism of inhibition\(^{49}\). Furthermore, high-fidelity BFE estimation requires accounting for factors such as conformational entropy and the explicit presence of the solvent. Nevertheless, we report the BFE and mode for the QMO-optimized variants. For simplicity, we limit the analysis to achiral molecules. First, we run blind docking simulations using AutoDock Vina\(^{46}\) over the entire structure of M\(^{\text{TM}}\) with the exhaustiveness parameter set to 8. We further refine the top three docking poses for each of the original and QMO-optimized molecules using the molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) method and AMBER forcefield\(^{47}\), which is known to be more rigorous and accurate than the scoring function used in docking. Next we inspect if any of the top three docking poses of the original as well as of QMO-optimized variants involve the substrate-binding pocket of M\(^{\text{TM}}\), as favourable interaction with that pocket is crucial for M\(^{\text{TM}}\) function inhibition.

As an illustration, Fig. 2 shows the top docking pose of dipyridamol and its QMO-optimized variant to the M\(^{\text{TM}}\) substrate-binding pocket. Consistent with the more favourable MM/PBSA BFE, the QMO-optimized variant forms 14% more contacts (with a 5-Å distance cutoff between heavy atoms) with the M\(^{\text{TM}}\) substrate-binding pocket compared to dipyridamol. Some of the M\(^{\text{TM}}\) residues that explicitly form contacts with the dipyridamol variant are Leu167, Asp187, Arg188 and Gln192. Similar observations, for example, a higher number of contacts with the M\(^{\text{TM}}\) substrate-binding pocket that involve Tyr54, were found for other exemplars of QMO variants, such as lavipiravir and umifenovir. Supplementary Section 4 provides an extended analysis of blind docking.

**Optimization of existing antimicrobial peptides towards improved toxicity.** As an additional motivating use case, discovering new antibiotics at rapid speed is critical for tackling the looming crisis of a global increase in antimicrobial resistance\(^{11}\). Antimicrobial peptides (AMPs) are considered promising candidates for next-generation antibiotics. Optimal AMP design requires balancing between multiple, tightly interacting attribute objectives\(^{9,12}\), such as high potency and low toxicity. In an attempt to address this challenge, we show how QMO can be used to find improved variants of known AMPs with reported/predicted toxicity, such that the variants have lower predicted toxicity and high sequence similarity compared to original AMPs.

For the AMP optimization task, a peptide molecule is represented as a sequence of 20 natural amino-acid characters. Using
the QMO formulation in equation (1), subject to the constraints of the toxicity prediction value \( f_{\text{tox}} \) and the AMP prediction value \( f_{\text{AMP}} \), we aim to find most similar molecules for a set of toxic AMPs. The sequence similarity score \( g_{\text{sim}} \) to be maximized is computed using Biopython (http://www.biopython.org), which uses global alignment between two sequences (normalized by the length of the starting sequence) to evaluate the best concordance of their characters. Detailed descriptions are provided in the Methods.

The objective of QMO is to search for improved AMP sequences by maximizing similarity while satisfying the AMP activity and toxicity predictions (that is, classified as being an AMP and non-toxic based on predictions from pre-trained deep learning models30).

In our experiments, we use QMO to optimize 150 experimentally verified toxic AMPs collected from public databases61,62 by Das et al.30 as starting sequences. Note that the toxic annotation here does not depend on a specific type of toxicity, such as haemolytic toxicity. Extended Data Fig. 3 shows their cumulative success rate (turning toxic AMPs into non-toxic AMPs) using QMO up to the \( t \)th iteration. Within the first few iterations, more than 60% molecules are successfully optimized. Eventually, \( \sim 72.67 \% \) (109/150) molecules can be successfully optimized. Analysis over all 109 original–improved pairs reveals notable physicochemical changes, for example, lowering of hydrophobicity and hydrophobic moment in the QMO-optimized AMP sequences (Fig. 3a,b and Supplementary Table 8). This trend is consistent with the reported positive correlation of hydrophobicity and hydrophobic moment with cytotoxicity and haemolytic activity63,64. Figure 3c shows examples of known AMPs and their QMO-optimized variant sequences. Sequence alignment and similarity ratio relative to the original sequence are also shown, indicating that sequences resulting from QMO differ widely from the initial ones. Supplementary Fig. 3 depicts the optimization process of some AMP sequences. QMO can further improve similarity while maintaining low predicted toxicity and high AMP values for the specified thresholds after the first success.

We perform additional validation of our optimization results by comparing QMO-optimized sequences using a number of state-of-the-art AMP and toxicity predictors that are external classifiers not used in the QMO framework. Extended Data Fig. 4 shows the external classifiers’ prediction results on 109 original and improved sequence pairs that are successfully optimized by QMO. We note that these external classifiers vary in terms of

| Table 1 | The final QMO-optimized molecules compared with their initial state for SARS-CoV-2 main protease inhibitor molecules |
|-----------------|----------------|----------------|----------------|----------------|
| Dipyridamole    | Favipiravir    | Umifenovir     | Kaempferol     |
| Original        | Original       | Original       | Original       |
| ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) | ![Image](image4.png) |
| −11.49          | −0.77          | −16.08         | −11.86         |
| Similarity      | Similarity     | Similarity     | Similarity     |
| 0.58            | 0.46           | 0.73           | 0.67           |
| Optimized       | Optimized      | Optimized      | Optimized      |
| ![Image](image5.png) | ![Image](image6.png) | ![Image](image7.png) | ![Image](image8.png) |
| −25.65          | −10.93         | −20.87         | −13.48         |

Rows from top to bottom (images): the original molecule with the common substructure highlighted, a Tanimoto similarity map of the improved molecule with respect to the original molecule (green indicates similar regions; purple indicates dissimilar regions), and the QMO-optimized molecule with the common substructure highlighted. Rows from top to bottom (numbers): original molecule MM/PBSA BFE estimate, overall Tanimoto similarity and QMO-optimized molecule BFE estimate. Both BFE calculations are performed with the molecules docked to the substrate-binding pocket of Mpro.
training data size and type, as well as on model architecture, and report a range of accuracy. Data and models for the toxicity prediction task are more rare than those for the AMP classification problem. Furthermore, external toxicity classifiers such as HAPPENN and HLPPred-fuse explicitly target predicting the haemolytic nature. For these reasons, the predictions of the external classifiers on the original lead sequences may vary when compared to ground-truth labels (third column in Supplementary Table 7). Nonetheless, predictions on the QMO-optimized sequences using external classifiers show high consistency in terms of toxicity improvement when compared with the predictors used in QMO. Specifically, the predictions from iAMP-2L and HAPPENN (haemolytic toxicity prediction) show that 56.88% (62/109) of QMO-optimized molecules are predicted as non-toxic AMPs. In Supplementary Section 7.1, we also show that the use of a better encoder–decoder helps the optimization performance of QMO.

**Property landscape visualization and trajectory analysis.** To gain a better understanding of how QMO optimizes a lead molecule with respect to the property constraints and objectives, we provide visual illustration of the property landscapes and search trajectories via QMO using a two-dimensional (2D) local interpolation on the molecule embedding space. Specifically, given the original embedding \( z_0 \) and the embedding of the best candidate \( z^* \) returned by QMO, we perform local grid sampling following two selected directions \( v_i \) and \( v_j \), and then evaluate the properties of the decoded sequences from the sampled embeddings for property landscape analysis. For the purpose of visualizing the property landscape in low dimensions, we project the high-dimensional search trajectories \( \{ z_t \}_{t=1} \) to the two directions \( v_i \) and \( v_j \). Figure 4a shows the landscape of Tanimoto similarity versus binding affinity when using remdesivir as the lead molecule, with the optimization objective of maximizing Tanimoto similarity while ensuring the predicted binding affinity is above a defined threshold of 7.5. The two directions are the principal vector \( z^* \) and a random vector orthogonal to the principal vector (more details are provided in the Methods). The trajectory shows how QMO leverages the evaluations of similarity and binding affinity for optimizing the lead molecule. Figure 4b displays the common substructure of candidate molecules in comparison to the remdesivir molecule in terms of subgraph similarity and their predicted properties over sampled iterations in QMO.

In addition to demonstrating the efficiency in optimizing lead molecules, we also study the diversity of the optimized molecules by varying the random seed used in QMO for query-based guided search. Figure 5 shows three different sets of trajectory on the landscape of predicted binding affinity when using remdesivir as the lead molecule (more details are provided in Methods). The optimization objective is the same as that of Fig. 4. The visualization suggests that the trajectories are distinct and the best candidate molecules in each trajectory are distant from each other in the embedding space, suggesting that QMO can find a diverse set of improved molecules with desired properties. In Supplementary Section 6.1 we also provide a quantitative study on the diversity and novelty of the QMO-optimized sequences when varying the similarity threshold. Setting a lower similarity threshold in QMO results in more novel and diverse sequences.

**Discussion and conclusion**

In this Article we have proposed QMO, a generic MO framework that readily applies to any pre-trained molecule encoder–decoder with continuous latent molecule embeddings and any set of property predictions and evaluation metrics. It features efficient guided search with molecular property evaluations and constraints obtained using predictive models and cheminformatics software. More broadly, QMO is a machine learning tool that can be incorporated into different scientific discovery pipelines with deep generative models, such as generative adversarial networks, for efficient guided optimization with constraints. As a demonstration, Supplementary Sections 6.2 and 6.3 show the QMO results on the SARS-CoV-2 main protease inhibitor optimization task, with alternative objectives and randomly generated lead sequences, respectively. QMO is able to perform successful optimization with respect to different objectives, constraints and starting sequences. The proposed QMO framework can be applied, in principle, to other classes of material, such as metal oxides, alloys and genes.

On the simpler benchmark tasks for optimizing drug-likeness and penalized logP scores with similarity constraints, QMO demonstrates superior performance over baseline results. We also apply QMO to improve the binding affinity of existing inhibitors of the SARS-CoV-2 main protease and to improve the toxicity of AMPs. The QMO-optimized variants of existing drug molecules show a favourable BFE with SARS-CoV-2 main protease upon blind
docking and MM/PBSA re-scoring, whereas the QMO-optimized peptides are consistently predicted to be antimicrobial and non-toxic by external peptide property predictors. The property landscape analysis and low-dimensional visualization of the optimization trajectories provide insights on how QMO efficiently navigates in the property space to find a diverse set of improved molecules with the desired properties. Our results show strong evidence that QMO can serve as a novel and practical tool for molecule optimization and other process/product design problems as well to aid in accelerating chemical discovery with constraints. In Supplementary Section 7 we provide an ablation study of QMO for additional performance analysis, including the effect of the encoder–decoder, the difference between sequence-level and latent-space classifiers, and a comparison between different gradient-free optimizers. The results show
that QMO is a query-efficient end-to-end MO framework, and a better encoder–decoder can further improve its performance. Future work will include integrating multi-fidelity expert feedback into the QMO framework for human–AI collaborative material optimization, and using QMO for accelerating the discovery of novel, high-performance and low-cost materials.

Methods

Procedure descriptions for the QMO framework.

- Procedure inputs: pre-trained encoder–decoder; molecular property predictors \( \{ f_i \}_{i=1} \) and thresholds \( \{ \tau_i \}_{i=1} \); molecular similarity metrics \( \{ q_i \}_{i=1} \) and thresholds \( \{ \eta_i \}_{i=1} \); total search iteration \( T \); step size \( \{ \alpha_i \}_{i=1} \); starting lead molecule sequence \( x_0 \); reference sequence set \( S \); Loss function from equation (1) or (2).
- Procedure initialization: \( z_0 = \text{Enc}(x_0) \); \( \mathbf{Z}_{\text{solution}} \leftarrow \emptyset \).
- Repeat the following steps \( T \) times, starting from \( T = 0 \):
  - Gradient estimation: generate \( Q \) random unit-norm perturbations \( \{ \varepsilon^q_i \} \) and compute \( \nabla Q \text{Loss}(z(t)) = \frac{1}{Q} \sum_{q=1}^{Q} [\text{Loss}(z(t) + \varepsilon^q_i) - \text{Loss}(z(t))] \times \varepsilon^q_i \).
  - Pseudo gradient descent: \( z(t+1) = z(t) - \alpha \times \nabla Q \text{Loss}(z(t)) \).
  - Molecular property and constraint verification: if solving for formulation (1), check \( f_j(\text{Dec}(z(t))) \geq \tau_j \) for all \( j \). If solving for formulation (2), check \( g(\text{Dec}(z(t))) \leq \eta_j \) for all \( j \).
  - Update valid molecule sequence: \( \mathbf{Z}_{\text{solution}} \leftarrow \mathbf{Z}_{\text{solution}} \cup \{ z(t) \} \).

Procedure convergence guarantee and implementation details for QMO. Inherited from zeroth-order optimization, QMO has algorithmic convergence guarantees. Under mild conditions for the true gradient (Lipschitz continuous and bounded gradient), the zeroth-order gradient descent following equation (3) ensures QMO takes at most \( \mathcal{O}(\epsilon) \) iterations to be sufficiently close to a local optimum in the loss landscape for a non-convex objective function\(^{14} \), where \( T \) is the number of iterations. In addition to the standard zeroth-order gradient descent method, our QMO algorithm can naturally adopt different zeroth-order solvers, such as zeroth-order stochastic and accelerated gradient descent. Our implementation of gradient estimation gives \( Q \) + 1 loss function queries per iteration. If the decoder outputs a SMILES string, we pass the string to RDKit for validity verification and disregard invalid strings.

In our QMO implementation, we use the zeroth-order version of the popular Adam optimizer\(^{18} \) that automatically adjusts the step sizes \( \{ \alpha_i \}_{i=1} \) with an initial learning rate \( \alpha \) (further details are provided in Supplementary Section 2). Empirically, we find that Adam performs better than stochastic gradient descent in our tasks. The convergence of the zeroth-order Adam-type optimizer is described in ref.\(^{18} \). We will specify experimental settings, data descriptions and QMO hyperparameters for each task. In all settings, QMO hyperparameters were tuned to a narrow range and then all the reported combinations were tried for each starting sequence. Among all feasible solutions returned by QMO, we report the one having the best molecular score given the required constraints. The stability analysis of QMO is studied in Supplementary Section 5.

Machine learning experimental settings. In our experiments, we run the QMO procedure based on the reported hyperparameter values and report the results of the best molecule found in the search process. The procedure will return null (that is, an unsuccessful search) if it fails to find a valid molecular sequence.

**References**

1. Bartók, A. P. et al. Machine learning unifies the modeling of materials and molecules. Sci. Adv. 3, e1701816 (2017).
2. Tkatchenko, A. Machine learning for chemical discovery. Nat. Commun. 11, 4125 (2020).
3. Button, A., Merk, D., Hiss, J. A. & Schneider, G. Automated de novo molecular design by hybrid machine intelligence and rule-driven chemical synthesis. Nat. Mach. Intell. 1, 307–315 (2019).
4. Kotsas, P.-C. et al. Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. Nat. Mach. Intell. 2, 254–260 (2020).
5. Polshchuk, P. G., Madzhidov, T. I. & Varnek, A. Estimation of the size of drug-like chemical space based on GDB-17 data. J. Comput. Aided Mol. Des. 27, 675–679 (2013).
6. Zhan andrew, A. Artificial intelligence for drug discovery, biomarker development and generation of novel chemistry. Mol. Pharm. 15, 4311–4313 (2018).
7. Sun, W. et al. Machine learning-assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. Sci. Adv. 5, eaay4275 (2019).
8. Ekina, S. et al. Exploiting machine learning for end-to-end drug discovery and development. Nat. Mater. 18, 435–441 (2019).
9. Wu, C. et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharm. Sin. B 10, 766–780 (2020).
10. Yang, J. et al. Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. Nat. Commun. 11, 4541 (2020).
11. Coates, A. R., Halls, G. & Hu, Y. Novel classes of antibiotics or more of the same? Br. J. Pharmac. 163, 184–194 (2011).
12. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J. Chem. Inf. Comput. Sci. 28, 31–36 (1988).
13. Reutlinger, M., Rodrigues, T., Schneider, P. & Schneider, G. Multi-objective molecular de novo design by adaptive fragment prioritization. Angew. Chem. Int. Ed. 53, 4244–4248 (2014).
14. Yuan, Y., Pei, J. & Lai, L. LigBuilder 2: a practical de novo drug design approach. J. Chem. Inf. Model. 51, 1083–1091 (2011).
15. Nigam, A., Friederich, P., Krenn, M. & Aspuru-Guzik, A. Augmenting genetic algorithms with deep neural networks for exploring the chemical space. In Proc. International Conference on Learning Representations (2020).
16. Korovina, K. et al. ChemBio: Bayesian optimization of small organic molecules with synthesizable recommendations. In Proc. International Conference on Artificial Intelligence and Statistics 3393–3403 (PMLR, 2020).
17. Gómez-Bombarelli, R. et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent. Sci. 4, 268–276 (2018).
18. Skalic, M., Jiménez, J., Sabbadin, D. & De Fabritiis, G. Shape-based generative modeling for de novo drug design. J. Chem. Inf. Model. 59, 1205–1214 (2019).
19. Griffiths, R.-R. & Hernández-Lobato, J. M. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. Chem. Sci. 11, 577–586 (2020).
20. Jiménez-Luna, J. et al. DeltaDelta neural networks for lead optimization of small molecule potency. Chem. Sci. 10, 10911–10918 (2019).
21. Boitreet, J., Mallet, V., Oliver, C. & Waldispühl, J. Optimal: optimization of binding affinities in chemical space for drug discovery. J. Chem. Inf. Model. 59, 6556–6566 (2019).
22. Jin, W., Barzilay, R. & Jaakkola, T. Junction tree variational autoencoder for molecular graph generation. In Proc. International Conference on Machine Learning 2323–2332 (PMLR, 2018).
23. Fu, T., Xiao, C. & Sun, J. CORE: automatic molecule optimization using copy & refine strategy. In Proc. AAAI Conference on Artificial Intelligence 638–645 (AAAI, 2020).
24. Winter, R. et al. Efficient multi-objective molecular optimization in a continuous latent space. Chem. Sci. 10, 8016–8024 (2019).
25. Oliveira-Mor, N., Blaschke, T., Engkvist, O. & Chen, H. Molecular de-novo design through deep reinforcement learning. J. Cheminform. 9, 48 (2017).
26. Guimaraes, G. L., Sanchez-Lengeling, B., Outeiral, C., Farias, P. L. C. & Aspuru-Guzik, A. Objective-reinforced generative adversarial networks (organ) for sequence generation models. Preprint at https://arxiv.org/abs/1705.10843 (2017).
27. Sanchez-Lengeling, B., Outeiral, C., Guimaraes, G. L. & Aspuru-Guzik, A. Optimizing distributions over molecular space: An objective-reinforced generative adversarial network for inverse-design chemistry (organic). Preprint at https://doi.org/10.26434/c3x9396688.v3 (2018).
28. You, J., Liu, B., Yang, Z., Pande, V. & Leskovec, J. Graph convolutional policy network for goal-directed molecular graph generation. In Advances in Neural Information Processing Systems 6410–6421 (NIPS, 2018).
29. Zhou, Z., Kearnes, S., Li, L., Zare, R. N. & Riley, P. Optimization of molecules via deep reinforcement learning. Sci. Rep. 9, 10752 (2019).
30. Das, P. et al. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. Nat. Biomed. Eng. 5, 613–623 (2021).
31. Griffen, E., Leach, A. G., Robb, G. R. & Warner, D. J. Matched molecular pairs as a medicinal chemistry tool: miniperspective. J. Med. Chem. 54, 7779–7780 (2011).
32. Dossetter, A. G., Griffen, E. J. & Leach, A. G. Matched molecular pair analysis in drug discovery. Drug Discov. Today 18, 724–731 (2013).
33. Dalke, A., Hert, J., Kramer, C. & mmpdb: an open-source matched molecular pair platform for large multiproperty data sets. J. Chem. Inf. Model. 58, 992–910 (2018).
34. Bahdanau, D., Cho, K. & Bengio, Y. Neural machine translation by jointly learning to align and translate. In Proc. International Conference on Learning Representations (2015).
35. Jin, W., Yang, K., Barzilay, R. & Jaakkola, T. Learning multimodal graph-to-graph translation for molecule optimization. In Proc. International Conference on Learning Representations (2019).
36. Yang, K., Jin, W., Swanson, K., Barzilay, R. & Jaakkola, T. Improving molecular design by stochastic iterative target augmentation. In Proc. International Conference on Machine Learning 10716–10726 (PMLR, 2020).
37. Jin, W., Barzilay, R. & Jaakkola, T. Hierarchical graph-to-graph translation for molecules. Preprint at https://arxiv.org/abs/1907.11223 (2019).
38. Maragakis, P., Nisonoff, H., Cole, B. & Shaw, D. E. A deep-learning view of chemical space designed to facilitate drug discovery. J. Chem. Inf. Model. 60, 4487–4496 (2020).
39. Ghadimi, S. & Lan, G. Stochastic first-and zeroth-order methods for convex stochastic optimization. SIAM J. Opt. 23, 2341–2368 (2013).
40. Liu, S. et al. A primer on zeroth-order optimization in signal processing and machine learning. In IEEE Signal Processing Magazine 43–54 (IEEE, 2020).
41. Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S. & Hopkins, A. L. Quantifying the chemical beauty of drugs. Nat. Chem. 4, 90–98 (2012).
42. Coley, C. W., Eyke, N. S. & Jensen, K. F. Autonomous discovery in the chemical sciences. Part II: outlook. Angew. Chem. Int. Ed. 59, 23414–23436 (2019).
43. Qin, Z. et al. Artificial intelligence method to design and fold alpha-helical structural proteins from the primary amino acid sequence. Extreme Mech. Lett. 36, 100652 (2020).
44. Bohacek, R. S., McMartin, C. & Guida, W. C. The art and practice of structure-based drug design: a molecular modeling perspective. Med. Res. Rev. 16, 53–50 (1996).
45. Reymond, J.-L., Rüdiger, K., Blum, L. & van Deursen, R. The enumeration of chemical space. Wiley Inter. Rev. Comput. Mol. Sci. 2, 717–733 (2012).
Competing interests
The authors have submitted a patent application on the method and system for query-based molecule optimization (US patent application number 17/016640).

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s42256-021-00422-y.
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42256-021-00422-y.

Correspondence and requests for materials should be addressed to Pin-Yu Chen or Payel Das.

Peer review information Nature Machine Intelligence thanks Amir Barati Farimani and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021
Extended Data Fig. 1 | Cumulative success rate of antimicrobial peptide (AMP) sequence optimization v.s. iterations using QMO.
Extended Data Fig. 2 | AMP and toxicity analysis. Reported accuracy, prediction rate, and property improvement for 109 pairs of starting and QMO-optimized sequences based on different AMP (a) and toxicity (b) classifiers. The 109 starting sequences are experimentally verified toxic AMPs and are correctly predicted by the AMP and toxicity classifiers used in QMO. The external classifiers have varying prediction accuracy as they may yield incorrect predictions on some of starting sequences. The prediction rate on QMO-optimized sequences is defined as the fraction of AMP and/or toxin predictions. About 56.88% of QMO-optimized sequences are predicted as non-toxic AMPs by iAMP-2L + HAPPENN, showing high agreement with the classifiers used in QMO. The complete results are reported in Supplementary Table 7.
| Method          | Success (%) |
|-----------------|-------------|
| MMPA$^{33}$     | 32.9        |
| JT-VAE$^{22}$   | 8.8         |
| GCPN$^{28}$     | 9.4         |
| VSeq2Seq$^{34}$ | 58.5        |
| VJTNNGAN$^{35}$ | 60.6        |
| AtomG2G$^{37}$  | 73.6        |
| HierG2G$^{37}$  | 76.9        |
| DESMILES$^{38}$ | 77.8        |
| **QMO**         | **92.8**    |

Extended Data Fig. 3 | Performance of drug likeness (QED) task with Tanimoto similarity constraint $\delta = 0.4$. 
Extended Data Fig. 4 | Performance of penalized logP task at various Tanimoto similarity constraint value $\delta$.

| Method       | Improvement $\delta = 0.6$ | Improvement $\delta = 0.4$ |
|--------------|-----------------------------|-----------------------------|
| JT-VAE$^{22}$| 0.28 ± 0.79                 | 1.03 ± 1.39                 |
| GC$^{28}$    | 0.79 ± 0.63                 | 2.49 ± 1.30                 |
| MolDQN$^{29}$| 1.86 ± 1.21                 | 3.37 ± 1.62                 |
| VSeq2Seq$^{34}$| 2.33 ± 1.17                 | 3.37 ± 1.75                 |
| VJTN$^{35}$  | 2.33 ± 1.24                 | 3.55 ± 1.67                 |
| GA$^{15}$    | 3.44 ± 1.09                 | 5.93 ± 1.41                 |
| QMO          | 3.73 ± 2.85                 | 7.71 ± 5.65                 |