

先导优化自动化工作流程设计

艾彦志 1900011718

2019级2班

E-mail:1900011718@pku.edu.cn

提交日期: 2023年1月13日

摘要 先导物优化时药物发现中的关键环节, 如果可以实现先导优化的流程在计算机上的模拟, 将大大降低先导优化的实验成本。本工作结合现有的分子性质预测工具和相似分子生成手段, 设计了一种用于先导优化自动化的工作流程。在实验部分中, 测试了上述算法流程应用到上市药物先导优化实例上的效果, 并研究了部分工作流程组件(图翻译模型、合成可及性合意函数所占权重系数、从次优分子中的采样)对算法效果(化学空间探索的高效性、算法输出化学结构的合理性等)的影响。

关键词 先导优化, 性质预测, 亲和力预测, 分子生成

Workflow Design of Automated Lead Optimization

AI Yanzhi 1900011718

Class 2, Grade 2019

Abstract Lead optimization is a key link in drug discovery. If the process of lead optimization can be simulated on a computer, the experimental cost of lead optimization will be greatly reduced. In this work, a workflow for automated lead optimization is designed by utilizing existing molecular property prediction tools and models which generate molecule similar to the input. In the experiment, the performance of the algorithm on lead optimization is examined with examples of listed drugs, and the effect of some workflow components (graph translation model, weight coefficient of the desirability function of SA score, sampling from suboptimal molecules) on the performance of the algorithm (the efficiency of chemical space exploration, the rationality of the output chemical structure, etc.) is studied.

Keywords Lead optimization, property prediction, affinity prediction, molecule generation

1 引言

先导物优化(lead optimization)是药物发现过程中的重要环节。先导化合物(lead compound)

是指，从在虚拟筛选或高通量筛选中获得的一些苗头化合物（hit）出发，经过初步遴选和优化后，得到的具有一定药理学或生物活性的化合物。一般而言，由于缺乏足够的活性或者部分药理学性质（如毒性、水溶性、细胞膜通透性等）不合要求，需要对先导化合物进行进一步修饰和改造，即先导物的优化，才可以获得具有成药潜力的分子。在现代药物的研发中，先导物优化相当耗时耗力，因为待考虑的生理药理性质有很多，通常需要对化合物结构进行多次迭代优化，才能获得各方面都满足要求的药物分子。^[1]

多年以来，研究人员一直致力于借助物理学、统计学知识以及计算机算法以降低药物研发的实验成本。其中，与先导物优化相关的工作主要包括预测药物分子与靶标蛋白的亲和力^[2]、预测分子的药理学性质^[3]、对化学空间的探索^[4]以及有机合成路线的设计^[5]。

本工作主要把目标放在化学空间的探索上，即在先导化合物和靶标蛋白的结构已知的条件下，试图通过算法给出一些具有较好活性和药理学性质的分子，以指导先导物优化的方向。对于化学空间的探索，目前已有一些关于深度学习生成模型的研究，尽管生成模型目前还没有在实际的药物研发中得到太多应用，但它们未来可期。^[4]而本工作主要采用多目标优化的思路，结合已有的亲和力预测、药理学性质预测工具，借助匹配分子对转换、图生成模型来达到探索化学空间的目的，并以上市药物奥美沙坦和马拉维若为例，对算法效果进行了初步验证。

2 实验方法

2.1 背景和动机

传统的先导物优化主要通过研究构效关系（SAR）来进行，此外针对某些特定药理学性质的改造，有特定的修饰策略。由于需要考虑的方面比较多，一般而言，在每一步优化中药物化学家会把重点放在其中一个分子性质上。这样做一方面便于建立构效关系，但另一方面先导优化的困难与曲折也随之而来：例如，优化活性使得 hERG 毒性增强，优化 hERG 毒性又使得吸收率下降，先导优化陷入反复的迭代中。另外，传统的先导优化思路类似于使用贪心（greedy）策略来解决优化问题，具体而言，就是指在每一步优化中，总是选择表现最好的

化合物作为下一代先导物。贪心策略未必能获得最优解，可能遇到诸如陷入局部最优之类的问题。

考虑到已开发出一些预测配体-靶标亲和力和分子药理学性质的工具，以及根据输入分子生成相似分子的方法，一个自然的想法是，能否将上述的先导优化过程放到计算机中进行？由 Schrodinger 公司开发的 AutoDesigner^[6]是在这方面上的一个尝试，AutoDesigner 的算法主要由两类步骤组成：在生成步骤中，从少量分子出发，产生一系列与之前分子结构相似的分子；在筛选步骤中，根据某些限制条件，去除掉不合条件的分子。算法的主要流程是从先导化合物的结构出发，生成步骤与筛选步骤交替进行，最后产生出一系列具有成药潜力的分子。本质上，AutoDesigner 是一个广度优先搜索算法，属于高度计算密集型任务，在高性能计算集群上完成一次算法流程需要 24-48 h，中途使用最多 150000 个 cpu 核心。为了更好地在实际项目中实施先导优化自动化算法，需要根据算法输出的结果，结合实验验证，不断调整算法中的参数和模型，并再次运行算法。有鉴于此，有必要研究出更加高效的算法。

如果先导物优化的传统思路可以借鉴，那么搜索范围就可以得到限制：从生成步骤产生的分子中选出最优的分子，排除所有次优解，然后依此反复迭代。但这样做显然存在一些问题：首先，现有的亲和力预测和药理性质预测工具准确性有限，很难保证程序选出的最优分子确实客观上也最优；其次是前面提到的关于传统先导优化思路的一些问题，即每一步优化中只关注单一性质导致优化步骤冗长，以及贪心策略的局限性。

为了解决这些问题，本工作在算法上的主要思想可以概括为三个方面：（1）多目标优化：具体做法是对分子各个性质的合意函数进行加权几何平均，将多目标优化转换为单目标优化问题；（2）集束搜索：每轮迭代后保留的是最优的一系列分子，而不仅仅是一个分子；（3）效仿模拟退火：在每轮迭代开始时，部分次优解会被放入搜索队列中。

2.2 算法设计

2.2.1 多目标优化

后续实验过程中，算法考虑了以下优化目标：药物分子与靶标的亲和力、定量评估类药性（QED）^[7]、合成可及性打分（SA score）^[8]、hERG 毒性（hERG 钾通道 IC₅₀）、Caco-2 细胞

通透性。给定一个分子，可以使用一系列工具预测出它的上述性质。对于每种性质，设定一个值域在 0 到 1 之间的合意函数。例如，对于 hERG 毒性，合意函数的形式如下：

$$f(x) = \begin{cases} \varepsilon, & x < -6 + \varepsilon \\ x + 6, & -6 + \varepsilon \leq x < -5 \\ 1, & x \geq -5 \end{cases} \quad (eq. 1)$$

其中 $\varepsilon = 10^{-8}$ （作用是避免取对数时报错），自变量 x 是 hERG 钾通道 IC_{50} 的对数值。选取这个合意函数的依据是：一般认为 hERG $IC_{50} < 1 \mu M$ 的分子对 hERG 钾通道表现出强抑制，hERG $IC_{50} > 10 \mu M$ 的分子为弱抑制或无抑制^[9]。类似地，可以依据经验对其他分子性质定义出各自的合意函数。

知道了每个性质对应的合意函数值后，对它们求加权几何平均值，就得到一个分子的综合评估打分：

$$\text{evaluation} = \exp\left(\frac{\sum_i w_i \ln(d_i)}{\sum_i w_i}\right) \quad (eq. 2)$$

式中 d_i 是性质 i 的合意函数值， w_i 是性质 i 的权重系数。之所以采用几何平均值，是因为这样的话如果有一个合意函数值接近 0（换言之，有一个条件不满足），那么总体的评估打分也会接近于 0。^[10]

关于每一种分子性质的详细信息列在表 1 中。值得注意的是，之所以选择这几个分子性质，只是因为它们比较有代表性。本工作在程序设计上保证了使用者可以任意选择他们所关注的分子性质，只要给出预测程序、合意函数和权重系数即可。

Table 1. Details of the Considered Molecular Properties

Property	Predicting Tool	Form of Desirability function	Weight
Binding Affinity ^a	Vina ^[11, 12] /GNINA ^{[13]a}	sigmoid function	1.0
QED	RDKit	linear	0.2
SA Score	RDKit	sigmoid function	0.3
log hERG IC_{50}	QikProp ^[14]	piecewise linear	0.3
Caco-2 Cell Permeability	QikProp	piecewise linear	0.3

^a GNINA is the default tool to predict binding affinity

关于靶标结合亲和力的预测，此处做一些细节说明。标准流程为：首先根据分子的 SMILES 表达式，使用距离几何获得分子的初始三维坐标，然后使用 MMFF94 力场^[15]对分子构型进行

优化，随后由对接程序获得分子与靶标蛋白的对接姿势，最后由对接姿势得出亲和力打分。程序支持 Autodock Vina^[11, 12]和 GNINA^[13]两种对接工具¹。由于 Autodock Vina 的亲和力打分主要用来评估姿势而不是预测亲和力，而 GNINA 自带亲和力预测，所以默认使用后者。为优化算法性能，曾尝试过使用 Vina-GPU^[16]，不过由于对接结果可重复性较差，最终没有采用。

2.2.2 生成步骤

本工作主要使用了两种生成相似分子的方法：匹配分子对转换^[17]和图生成模型^[18]。

为了执行匹配分子对转换，需要先准备匹配分子对数据库 (mmpdb)。在本实验中，mmpdb 由 500000 个从 ChEMBL 数据库中随机挑选出的分子产生，它们大约占整个 ChEMBL 数据库的 $1/4^2$ 。在执行匹配分子对转换时，程序会首先对输入的分子进行切割，然后将分子的一部分标记为可变，之后从 mmpdb 中搜寻合适的分子片段来代替可变部分，最后输出转换后的分子³。每次匹配分子对转换会生成数千个片段替换后的分子，由于算力有限，只能随机挑选出其中一部分来进行搜索。

图生成模型的细节详见文献[18]。在这篇文献中，作者将这个生成模型的编码器和解码器用于图到图的翻译任务，后者适用于先导优化的情形。本实验中，使用文献作者给出的数据集训练了图到图的翻译模型。在本实验中，该模型将一个输入分子翻译成 4 个具有更高 QED、且结构与输入分子相似的分子。

关于上面模型的数据集，有一点补充说明：对于本实验而言，使用文献[18]作者给出的数据集来训练模型并不完全合适。作者给出的数据集包含 87226 个分子对(X, Y)，其中 X 的 QED 在 0.7 到 0.8 之间，Y 的 QED 在 0.9 到 1.0 之间，X 与 Y 分子指纹的 Tanimoto 相似度大于 0.4。然而，本工作的实验中发现，进行化学空间搜索时，遇到的多数分子 QED 在 0.3-0.5 之间，与训练集的分布有差异，这导致由模型翻译出的分子与输入分子之间的相似度较小，一般只有 0.3 左右。⁴

¹ GNINA 是 SMINA 的一个分支，而 SMINA 是 Vina 的一个分支。

² 之所以没有用整个 ChEMBL 数据库来构建 mmpdb，是由于受到个人电脑硬盘容量的限制。

³ 可变部分的大小可由一些参数指定。例如，可以设置最多有 N 个重原子被包含在可变部分中（本实验中 N = 4）。

⁴ 这个错误在撰写此报告时才被发现。若时间充裕，可以通过自制数据集来避免。

2.2.3 效仿模拟退火

在每轮搜索迭代开始时，此前被抛弃的次优分子将有机会加入到搜索队列中。基本想法参考模拟退火中的 Metropolis 准则，简而言之就是次优分子中越“好”的分子，被选中的概率越大。具体做法如下：随机在次优分子中选择其中一个分子 X ，然后生成一个 0 到 1 之间的随机数 r ，考虑下列不等式是否成立：

$$r < \exp\left(\frac{X_{\text{evaluation}} - \min(Y_{\text{evaluation}})}{T}\right) \quad (eq.3)$$

式中 $X_{\text{evaluation}}$ 是分子 X 的综合评估打分（见 2.2.1 节）， $\min(Y_{\text{evaluation}})$ 是上一轮迭代结束时，搜索前沿所有分子中最低的评估打分， T 代表接纳评估打分较差分子的意愿， T 值越大，越有可能接纳一些打分较差的分子（实验中 $T=1/40$ ）。若上面的不等式成立，就将分子 X 放入搜索前沿中，反之则不放入。然后重复前面的步骤，直到放入了一定比例的次优分子为止。

2.2.4 算法流程

表 2 展示出整个算法流程的伪代码。首先，将搜索前沿（fringe，评估最优的若干个分子）和标记分子（marked，次优的分子将被标记，可以避免重复搜索；这些分子有机会重新加入搜索前沿）初始化为空列表，然后将已知的先导化合物加入搜索前沿。接下来进入迭代步骤，首先从被标记的次优分子中选出一部分放入搜索前沿（见 2.2.3 节），随后对于搜索前沿中的每一个分子，找出一定数量与之结构相似的分子（2.2.2 节），并对这些分子进行性质预测和综合评估打分（2.2.1 节），然后将这些分子与搜索前沿合并，选出评估打分最高的几个分子（这一步选择分子的数量即集束宽度 beam width），并将标记其余分子。

此外，程序还包含一些额外内容，用于优化算法的性能。实际上，程序的绝大多数运行时间都用于分子对接。同时开启多个子进程用于对接，可以更有效地利用计算机的全部 cpu 核心⁵。最有效的优化手段是如下所述的剪枝（clipping）步骤：对于一个待处理的分子，首先预测它除了结合亲和力以外的性质，然后假设它的亲和力非常好（亲和力合意函数值为 1），如

⁵ 在较低的 exhaustiveness 下，Vina 不会使用全部的 cpu 核心。

果即使这样，这个分子的评估打分仍小于上一轮迭代结束时，搜索前沿所有分子中最低的评估打分，那么就不进行对接，直接标记这个分子。对上述做法合理性的解释是，由于亲和力在综合评估中的权重较大（表 1），因此满足剪枝条件的分子一般都具有过于不合理的 QED、hERG 毒性或合成难度等。剪枝步骤能够使大约 25%-50% 的分子不参与对接。

Table 2. Pseudocode of the Workflow

```
def algorithm(lead_compound, iter_num, beam_width, ...):
    fringe = []
    marked = []
    fringe.append(lead_compound)
    clip_bound = 0.0 # for algorithm optimization

    for i in range(iter_num):

        fringe.extend(sample_from(marked)) # see 2.2.3

        predictMedicinalProperty = []
        for molecule in fringe:
            mols = randomly_choose_from(molecule.mmpdb_transform(), for molecules not in marked)
            predictMedicinalProperty.extend(mols)
            predictMedicinalProperty.extend(molecule.graph2graph_translate()) # see 2.2.2

        for molecule in predictMedicinalProperty:
            molecule.predict_medicinal_property()
            if molecule.evaluate() > clip_bound:
                molecule.predict_affinity()
                molecule.evaluate() # see 2.2.1
            else:
                marked.append(molecule)

        fringe.extend(predictMedicinalProperty)
        fringe = sorted(fringe, key=molecule.evaluation, reverse=True)
        fringe = predictMedicinalProperty[:beam_width]
        marked.extend(predictMedicinalProperty[beam_width:])
        clip_bound = min_evaluation_for_molecules_in(fringe)
```

3 结果与讨论

3.1 药物发现实例：奥美沙坦和马拉维若

为了验证算法的效果，考察了两个药物发现的实例：奥美沙坦（Olmesartan）能与血管紧张素 II 受体（angiotensin II receptor）结合，阻碍血管紧张素 II 发挥作用，起到降低血压的作用；马拉维若（Maraviroc）是趋化因子受体 5（CCR5）的别构调节物，它与 CCR5 的结合会阻碍 HIV 病毒上的蛋白 gp120 与 CCR5 的结合，从而阻止病毒进入巨噬细胞和 T 细胞。它们与它们先导化合物的结构被展示在图 1 中。^[19]

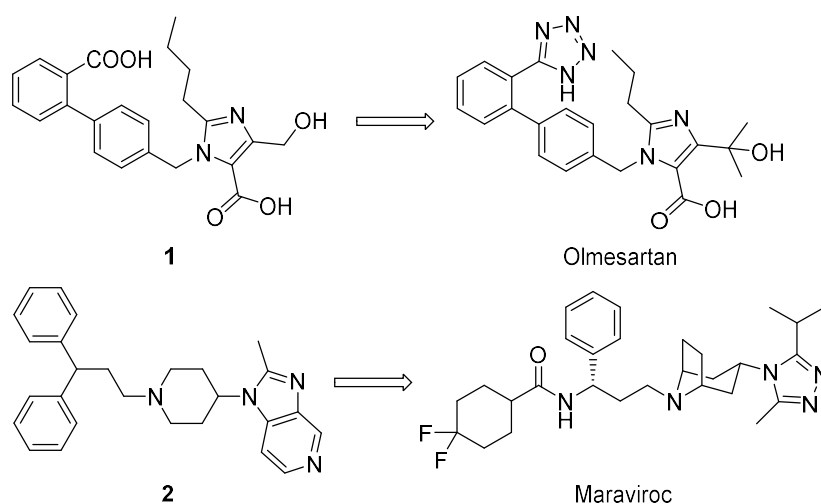


Figure 1. Olmesartan, Maraviroc and their lead compound

将图 1 中的化合物 1 和化合物 2 分别设置为算法优化的起点，对应的受体蛋白结构由分别由文献[20]和[21]给出（PDB id: 4ZUD、4MBS）。将集束宽度设为 5，每次迭代开始时放入 5 个次优分子，每次迭代由匹配分子对转换生成 60 个分子，由图翻译模型生成 40 个分子，迭代次数设为 16。图 2a 和图 2b 列出了两次算法运行结束后搜索前沿中分子。

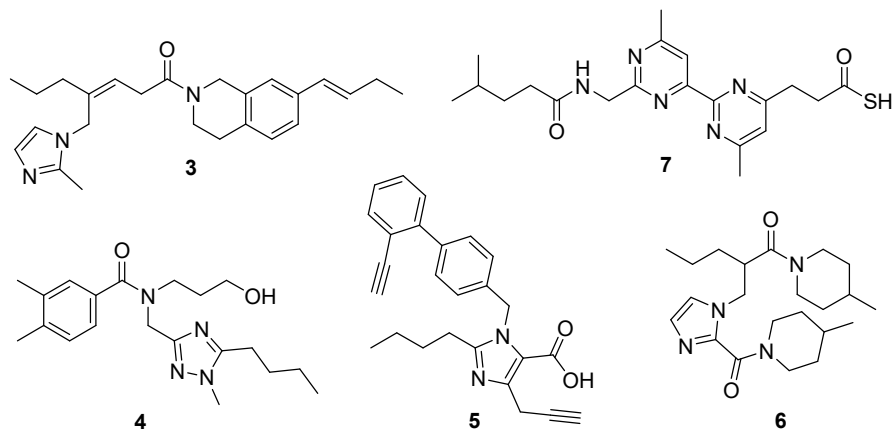


Figure 2a. Result of Optimization from Lead Compound 1

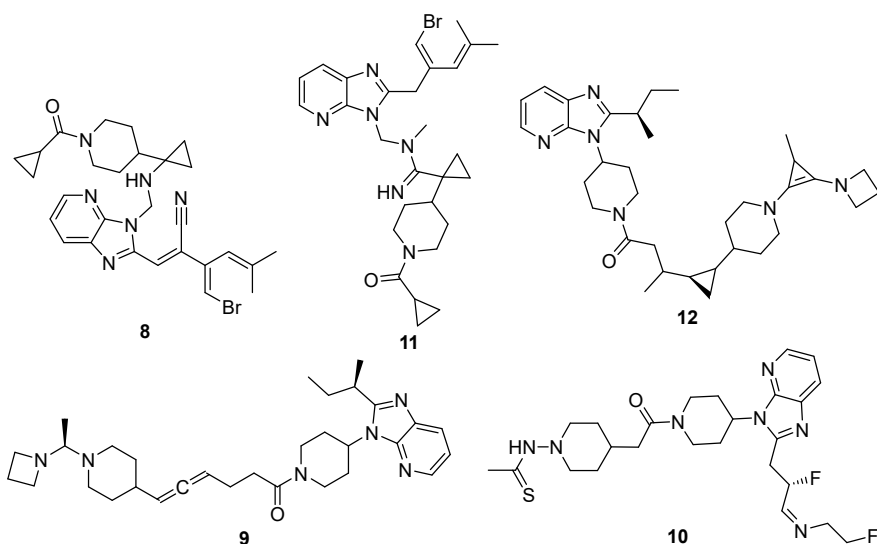


Figure 2b. Result of Optimization from Lead Compound 2

可见生成的分子具有比较丰富的结构，同时也各自保留了一部分先导化合物中的结构特征。例如，化合物 **5**, **7**, **9**, **10**, **12** 保留了先导化合物中的部分骨架，除了化合物 **7** 以外，所有化合物都保留了先导物中含氮杂环的结构。然而，从化合物化学稳定性和可合成性的角度来看，对化合物 **2** 的优化结果并不令人满意，因为生成的化合物 **8-12** 中包含了大量不合理的结构，例如 **8**, **11** 中拥挤的三元环，**9** 中的联烯，**10** 中的亚胺和 **12** 中三元环内的双键。

表 3 列出了上述这些化合物的详细数据。下面是对这些数据的一些分析：

从综合评估打分来看，在 16 个迭代之后，搜索前沿中的分子打分已经高于对应上市药物

的打分，这一方面说明算法有效，但另一方面也暗示 2.2.1 节中设计的评估打分尚有考虑不周之处，毕竟真实的药物设计需要面面俱到，简单地通过几个性质来进行评估不能满足实际药物设计的需求。例如，Maraviroc 实际开发中的先导优化步骤还考虑了药物的抗病毒活性（antiviral activity IC₉₀）和人肝微粒体清除半衰期（HLM t_{1/2}）等^[19]。

从亲和力分数来看，搜索前沿中的分子具有与上市药物相当或更高的亲和力打分，暗示着这些分子可能也具有能与受体蛋白结合的活性。

对于药理学性质，算法优化了化合物 **1** 较差的 Caco-2 细胞通透性⁶以及化合物 **2** 对 hERG 的强抑制，后者是 Maraviroc 实际开发中的一个真实问题^[19]。

Table 3. Detailed Data for each Compound

Compound	Binding Affinity ^a	QED	SA Score	log hERG IC ₅₀	Caco-2 Cell Permeability ^b	Evaluation
1	6.2437	0.4960	2.3327	-2.426	5.008	0.4835
Olmesartan	7.2670	0.3765	2.8070	-3.772	23.473	0.7506
random100 ^c	6.1052	0.5533	3.0283	-4.963	1024.311	0.3492
3	7.6689	0.5716	3.1000	-4.639	2303.299	0.8292
4	7.0237	0.7481	2.4872	-4.604	1585.441	0.8284
5	7.3719	0.5568	2.7999	-4.463	529.978	0.8268
6	7.3675	0.7501	3.1961	-3.159	3169.147	0.8266
7	7.3086	0.6270	2.8938	-4.229	586.628	0.8259
2	7.7144	0.4149	2.3494	-7.623	821.588	0.0448
Maraviroc	8.3403	0.4738	4.6906	-5.244	259.110	0.5922
random100	6.1030	0.5533	3.0283	-4.963	1024.311	0.1284
8	8.1872	0.3563	3.7214	-4.627	101.786	0.6404
9	8.3245	0.5100	4.7086	-5.00	1772.009	0.6348
10	8.3403	0.3910	4.6212	-4.774	719.042	0.6281
11	8.0794	0.2656	3.6692	-4.724	632.549	0.6276
12	8.5645	0.4868	5.1514	-4.887	2434.520	0.6247

^aThe binding affinity here is the CNNAffinity given by GNINA through an ensemble of CNN models. Larger value implies better binding, which is different from Vina

^bin nm·s⁻¹

^cAverage values of 100 randomly picked molecules from ChEMBL

3.2 一些组件对算法效果的影响

⁶ Caco-2 细胞通透性在 25 nm·s⁻¹ 左右被视为中等水平，500 nm·s⁻¹ 以上为良好。

3.2.1 图翻译模型

考察了算法在不使用图翻译模型时的效果。结果列在图 3 和表 4 中。可见，此时产生的分子都由化合物 1 的骨架衍生出来，它们结构单一且具有不稳定的结构；另外，从综合评估打分的角度来看，也明显不如使用图翻译模型时的结果。这反映出生成模型确实是探索化学空间的有力工具，它比起单纯使用片段替换或修饰而言，更具有“想象力”，换言之，它能够使算法不局限于分子局部的改变，从而扩大有效的搜索范围。

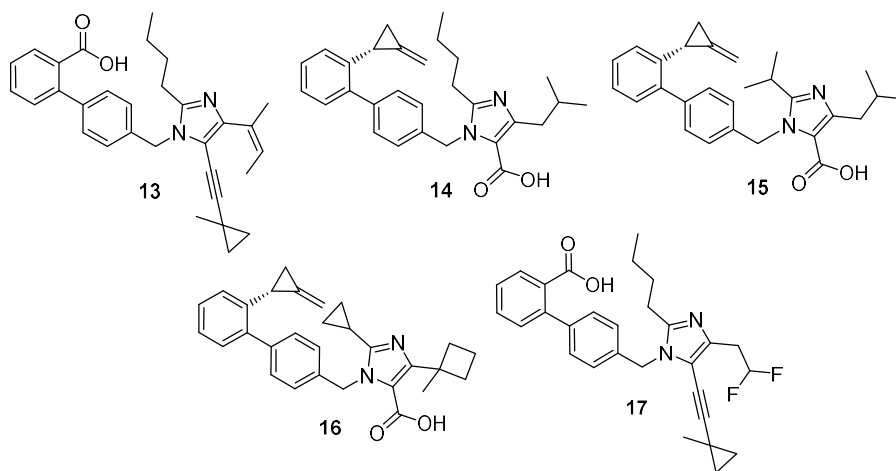


Figure 3. Result of optimization from lead compound 1, without the usage of graph-to-graph translation

Table 4. Detailed Data for Compound 13-17

Compound	Binding Affinity ^a	QED	SA Score	log hERG IC ₅₀	Caco-2 Cell Permeability	Evaluation
13	7.6723	0.3568	3.0606	-4.8030	358.1280	0.7847
14	7.5980	0.3482	3.1262	-4.6070	496.7590	0.7839
15	7.4169	0.4081	3.1784	-3.1590	573.0990	0.7816
16	7.5789	0.4197	3.3819	-4.0010	629.8960	0.7809
17	7.6747	0.4197	3.0630	-4.7680	300.7660	0.7804

3.2.2 合成可及性 (SA score)

实验发现，调整合成可及性 (SA score) 在综合评估打分中所占的权重系数，会显著地影响优化的结果。图 4 展示出将 SA score 权重系数设为 0 时，先导物 1 的优化结果。可见，相比于之前的结果 (化合物 3-7)，此时的结果中多出了许多不稳定或难以合成的结构，例如三

元或四元小环，与小环相邻的双键，与双键和三元环直接相连的氟原子等。此外，产生的分子结构也比较单一。一个可能的解释是，如果没有 SA score 的限制，算法就倾向于从 mmpdb 中搜寻罕见的片段（通常不稳定或难合成），来追求更高的亲和力分数（表 5）；与此同时，算法将较少地依赖图翻译模型来获取新颖的结构，导致最终产生的分子结构单一。

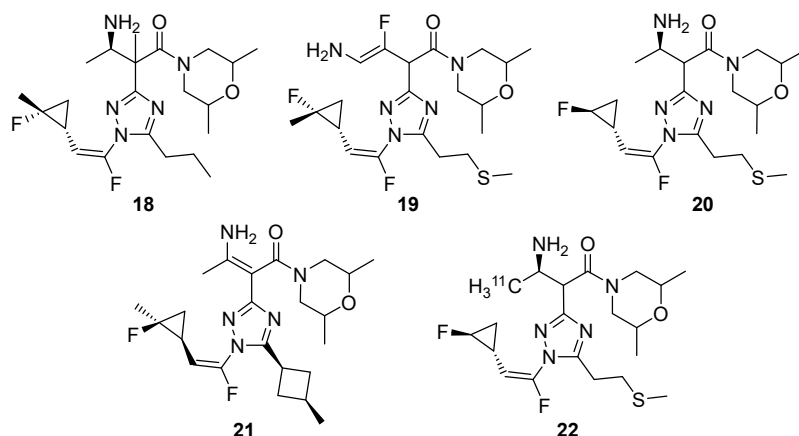


Figure 4. Result of optimization from lead compound **1**, when the weight of SA score is set to 0

Table 5. Detailed Data for Compound 18-22

Compound	Binding Affinity ^a	QED	SA Score	log hERG IC ₅₀	Caco-2 Cell Permeability	Evaluation
18	7.7876	0.7059	5.0839	-3.9100	642.0750	0.9340
19	8.0269	0.6242	5.2884	-3.7430	852.7880	0.9316
20	7.9670	0.6640	5.0913	-4.0490	438.4530	0.9307
21	7.7352	0.6940	5.0140	-4.8750	1673.9530	0.9293
22	7.9319	0.6653	5.0913	-4.5390	432.5390	0.9291

根据这个思路，提高 SA score 的权重系数应该有助于提高结果的化学稳定性和可合成性，获得更加合理的结构。然而，将这个想法应用于化合物 **2** 的优化，并未取得满意的结果，实验中将 SA score 的权重系数提高至 0.6，生成的分子确实避免了拥挤三元环、联烯等不稳定结构，但又出现了裸露巯基、与双键直接相连的卤原子等不合理的结构，此外，这样做使得生成分子的亲和力打分只有 8 左右，显著低于表 3 中的结果。也许扩大搜索规模（即增加集束宽度、匹配分子对转换得到的分子数量、图翻译模型生成的分子数量以及迭代次数）后，能够取得更满意的结果⁷。

⁷ 此实验尚未进行。

值得指出，SA score 本身并不是用于评估化学结构合理性的工具。从表 3 中可以看出，Maraviroc 的 SA score 实际上相当高。虽然化合物 **8-12** 的 SA score 基本与之相当，但是从直观上看来后者化学结构的合理性远远不如前者。这带来的启示是，或许有必要开发一些预测化学结构合理性的工具，用来协助分子生成或化学空间探索的工作。

3.2.3 对模拟退火的效仿：从次优分子中采样

下面的实验考察了，如果算法不在每轮迭代开始时，从被标记的次优分子中选择一些放入搜索前沿，结果会如何。为了公平起见，此时将集束宽度设为 10，以保证进行预测和打分的分子数目相同。结果列在图 5 和表 6 中。可以看出，尽管从评估打分的角度来看，此次实验的结果比表 3 中的结果更好，但是生成的结构比较单一。原因类似于数值优化中的局部最优：总是在目前已知的最优解附近搜索更好的解，从来不去远处寻找更优的解。本工作中加入从次优分子采样的机制，就是为了避免算法局限于化学空间中的特定区域，从而进行更加广泛、更加有效的探索。

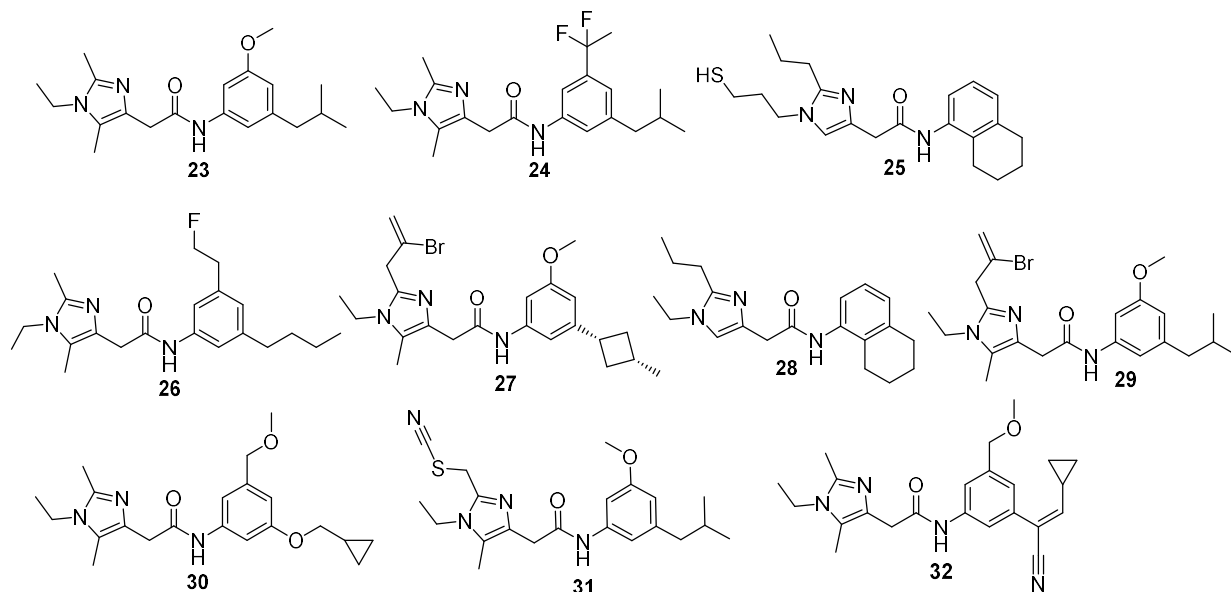


Figure 5. Result of optimization from lead compound **1**, without sampling from suboptimal molecules at the beginning of each iteration

Table 6. Detailed Data for Compound 23-32

Compound	Binding Affinity ^a	QED	SA Score	log hERG IC ₅₀	Caco-2 Cell Permeability	Evaluation
23	7.1382	0.8291	2.4499	-4.8740	3072.6230	0.8468
24	7.2812	0.7371	2.7186	-4.6120	3365.2650	0.8409

25	7.4600	0.6828	2.6803	-5.0920	3562.2270	0.8405
26	7.5201	0.7135	2.6687	-5.1850	3071.1660	0.8358
27	7.5560	0.5768	2.9986	-4.9500	3102.8670	0.8276
28	6.9540	0.8749	2.4335	-4.7140	3309.3870	0.8243
29	7.4068	0.5842	2.8496	-5.0250	3086.0970	0.8237
30	7.3489	0.7309	2.5932	-5.2270	3098.8940	0.8221
31	7.2810	0.6344	2.8848	-4.8470	1592.8800	0.8201
32	7.4731	0.6841	2.9781	-5.1450	830.3120	0.8183

4 结论

本工作提出了一种用于先导物优化的工作流程，其内容包括多目标优化、生成步骤、集束搜索，并参考模拟退火的思想对化学空间进行探索。实验中，探索了本算法在先导优化实例上的效果，其中对 **Olmesartan** 先导物优化效果尚可，对 **Maraviroc** 先导物的优化效果不佳，产生了许多不合理的分子结构。另外，探索了工作流程中一些组件对算法效果的影响，其中图翻译模型和从次优分子中采样的方法有助于更广泛有效地探索化学空间，生成比较丰富的化学结构，合成可及性的权重系数会影响算法产生分子的化学结构合理程度。

为了改进算法的效果，还有许多工作有待完成。例如使用自制的数据集来训练图翻译模型、扩大算法的搜索规模等。考虑到图生成模型生成的分子一般亲和力打分较低，可以添加一些机制，使得从次优分子中采样时更容易采到这些分子，从而进一步降低算法局限于特定化学空间的可能性。

除了算法本身以外，这个工作流程实际上十分依赖于其各个组件的性能，例如各种预测工具和模型的准确性，以及表现更好的生成模型。遗憾的是目前很多这方面的工作离实际应用尚且有一些差距，不过相信这些差距将会在不远的未来中被一一克服。

REFERENCES

1. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. *Br J Pharmacol*. **2011**, 162(6), 1239-1249.
2. Derek Jones, Hyojin Kim, Xiaohua Zhang, Adam Zemla, Garrett Stevenson, W. F. Drew Bennett, Daniel Kirshner, Sergio E. Wong, Felice C. Lightstone, Jonathan E. Allen. Improved Protein-Ligand Binding Affinity Prediction with Structure-Based Deep Fusion Inference. *Journal of Chemical Information and Modeling*. **2021**, 61(4), 1583-1592.

3. Douglas E. V. Pires, Tom L. Blundell, David B. Ascher. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *Journal of Medicinal Chemistry*. **2015**, 58(9), 4066-4072.
4. Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R., Jensen, K. F. Generative models for molecular discovery: Recent advances and challenges. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, 12(5), e1608.
5. Segler, M., Preuss, M., Waller, M. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*. **2018**, 555, 604–610.
6. Pieter H. Bos, Evelyne M. Houang, Fabio Ranalli, Abba E. Leffler, Nicholas A. Boyles, Volker A. Eyrich, Yuval Luria, Dana Katz, Haifeng Tang, Robert Abel, and Sathesh Bhat. AutoDesigner, a De Novo Design Algorithm for Rapidly Exploring Large Chemical Space for Lead Optimization: Application to the Design and Synthesis of d-Amino Acid Oxidase Inhibitors. *Journal of Chemical Information and Modeling* **2022**, 62(8), 1905-1915.
7. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. *Nat Chem*. **2012**, 4(2), 90-98.
8. Ertl, P., Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Cheminform*. **2009**, 1, 8.
9. Yu, Hb., Zou, By., Wang, Xl. et al. Investigation of miscellaneous hERG inhibition in large diverse compound collection using automated patch-clamp assay. *Acta Pharmacol Sin*. **2016**, 37, 111–123.
10. Lambrinidis, G., Tsantili-Kakoulidou, A. Multi-objective optimization methods in novel drug design. *Expert Opinion on Drug Discovery*. **2021**, 16(6), 647–658.
11. O. Trott, A. J. Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. **2010**, 31(2), 455-461.
12. J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*. **2021**, 61(8), 3891–3898.
13. McNutt, A.T., Francoeur, P., Aggarwal, R. et al. GNINA 1.0: molecular docking with deep learning. *J Cheminform*. **2021**, 13, 43.
14. Schrödinger Release 2022-2: QikProp, Schrödinger, LLC, New York, NY, 2021.
15. Thomas A. Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*. **1996**, 17(5-6), 490-519.
16. Tang, S.; Chen, R.; Lin, M.; Lin, Q.; Zhu, Y.; Ding, J.; Hu, H.; Ling, M.; Wu, J. Accelerating AutoDock Vina with GPUs. *Molecules* **2022**, 27, 3041.
17. Dalke, A.; Hert, J.; Kramer, C. mmpdb: An Open-Source Matched Molecular Pair Platform for Large Multiproperty Data Sets. *J. Chem. Inf. Model*. **2018**, 58(5), 902–910.
18. Jin, W.; Barzilay, R.; Jaakkola, T. Hierarchical Generation of Molecular Graphs using Structural Motifs. *arXiv preprint*, **2020**, arXiv:2002.03230v2
19. Huang, X.; Aslanian, R. G. *Case Studies in Modern Drug Discovery and Development*; Wiley: **2012**.
20. H. Zhang, et al. Structural basis for ligand recognition and functional selectivity at angiotensin receptor. *J. Biol. Chem*. **2015**, 290(49), 29127-29139.
21. Tan, Q., Zhu, Y., Li, J., Chen, Z., Han, G.W., Kufareva, I., Li, T., Ma, L., Fenalti, G., Li, J., Zhang, W., Xie, X., Yang, H., Jiang, H., Cherezov, V., Liu, H., Stevens, R.C., Zhao, Q., Wu, B. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*. **2013**, 341, 1387-1390.