

Fast Structure-Based *De Novo* Drug Design with Molecular Encoders and Computational Intelligence

Silvia Multari^{*}, Matteo Grazioso[†], Leone Bacciu[†], Daniele M. Papetti^{‡§}, Daniela Besozzi^{‡§}, Marco S. Nobile^{†§}

^{*}Department of Molecular Sciences and Nanosystems, Ca' Foscari University of Venice, Venice, Italy
Email: silvia.multari@unive.it

[†]Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, Venice, Italy
Email: matteo.grazioso@unive.it, leone.bacciu@unive.it, marco.nobile@unive.it

[‡]Department of Informatics, Systems and Communication, University of Milano-Bicocca, Milan, Italy
Email: daniele.papetti@unimib.it, daniela.besozzi@unimib.it

[§]Bicocca Bioinformatics, Biostatistics and Bioimaging Research Center, Monza, Italy

Abstract—Drug design is the first step of the long and expensive process, named drug development, aimed at launching a new drug on the market. Due to the high cost and time required by experimental techniques, *in silico* methods represent a valuable tool to help speed up and make the development more efficient. In this work, we propose EVOMOL-FLOW, an approach that combines evolutionary computation and deep learning models, specifically tailored to this problem. A multi-objective optimization algorithm – with novel evolutionary operators – is used to design novel drug candidates, maximizing their binding affinity to the target while minimizing the difficulty of their synthesis in the laboratory. To reduce the bottlenecks of the fitness function evaluation, we take advantage of DrugCLIP, a molecular encoder that maps the chemical structures of the ligand and the pocket into a latent space, and determines the binding affinity by measuring the cosine distance between the embeddings. This approach reduces the overall running time by orders of magnitude, paving the way to large-scale and more accurate evolution of candidate drugs. As a case study, we applied EVOMOL-FLOW to the design of putative molecules targeting Vascular Endothelial Growth Factor Receptor (VEGFR), a membrane protein responsible for vascularisation in solid tumors. To show the effectiveness of the approach, we further validate the produced compounds by means of molecular dynamics and docking experiments.

Index Terms—Molecular Transformers, Neural Networks, Evolutionary Computation, Multi-objective Optimization, Drug Design

I. INTRODUCTION

The design of new drugs for specific diseases is the first step of drug development, a process that typically requires billions of dollars and more than a decade to put a drug on the market [1]. One of the principal causes of this expensiveness lies in the dimensionality of the search space of small molecules that can be considered as drug candidates: a number estimated to be in the order of 10^{60} . This space is usually narrowed by taking advantage of knowledge of biological targets and rational drug design [2]. However, designing even a single drug candidate still remains an expensive task, despite recent advancements in computational methods for *in silico* screening.

To enhance computer-aided drug design, several Deep Learning approaches have been proposed [3], [4]; among them, Chemical Language Models (CLMs) have gained popularity for

their ability to manipulate 2D molecules using different string representations. CLMs are widely leveraged for a variety of tasks, ranging from the *de novo* generation of novel molecular structures [5], [6], to the prediction of metabolic processes related to drug pharmacokinetics and pharmacodynamics [7]. Although CLMs proved to be capable of creating novel molecules by mimicking structural patterns learned during training, they showed several drawbacks that limit their benefits: (i) they struggle at generating small molecules for under-represented targets in the training set, resulting in poor generalization to previously unseen cases; (ii) they tend to generate small molecules that are unfeasible from a chemical point of view; (iii) they often fail at creating small molecules that satisfy chemical characteristics required for drugs (*e.g.*, the possibility to synthesize the molecule in laboratory).

Optimization methods – in particular, multi-objective evolutionary approaches – have been combined with CLMs to refine their output, and enhance generational capabilities by means of operators that introduce useful novelty in the chemical structures [8]–[10]. Taking inspiration from Darwinian theories, evolutionary approaches evolve a set of candidate solutions (*i.e.*, small molecules) that try to optimize either a single objective or a set of conflicting objectives. The drawback of such approaches is that the operators used to evolve the solutions depend on the representation of the solutions themselves. Moreover, the objectives to optimize need to be defined according to the problem at hand, with improper selection of the objectives resulting in unfeasible solutions.

When dealing with structure-based *de novo* drug design [11], the main objective is that the evolved small molecule strongly binds to the target receptor; several ancillary objectives can also be considered, such as the possibility of synthesizing the small molecule. The reliable prediction of binding affinity is therefore an essential step, as we rely on it to determine whether the binding takes place, and to rank candidate molecules from the most to the least promising based on the free energy of the bound complex (the higher the energy, the worse the binding) [12], [13]. Traditional methods for free energy estimation are:

- Molecular docking, which predicts plausible binding poses (*i.e.*, 3D configurations of the ligand with the respect to the target) and their affinity scores [14], relying on limited conformational sampling and often neglecting protein flexibility and explicit solvent effects [15].
- Molecular Dynamics (MD), which explicitly models atomic motions, solvent, and conformational fluctuations over time [16]. Whilst it offers a more accurate and physically grounded description of ligand-protein binding, it carries a significantly higher computational burden.

To overcome the issues related to the high computational resources and time required by both methods, a Deep Learning (DL) model named DrugCLIP has recently been introduced to perform binding predictions by computing a distance between embeddings [17]. DrugCLIP computes the embeddings of both the biological target (also referred to as receptor or protein) and small molecules (or ligands), allowing a comparison in terms of distance in a geometric space where the closer the embeddings, the stronger the binding.

Considering the obstacles hindering the possibility to couple molecular docking and dynamics with evolutionary approaches, and the promise held by DL-based techniques, we developed EVOMOL-FLOW (EVolutionary MOlecular Fragment-enhanced Ligand Optimization Workflow), a multi-objective evolutionary approach that evolves a set of novel candidate small molecules using a structure-based *de novo* drug design approach [11]. EVOMOL-FLOW simultaneously optimizes two objectives: (1) the affinity between the small molecule and the target, computed by DrugCLIP, and (2) the synthetic accessibility to ensure that evolved candidates remain viable for laboratory production. To navigate this chemical space effectively, the framework employs domain-specific operators tailored for molecular representation, resulting in drug candidates that are both computationally optimized and practically attainable. Since the evolution process could generate a large set of equally optimal solutions, a subset of the most chemically feasible solutions is determined by using thresholds related to each objective. The selected molecules are then validated by molecular docking and molecular dynamics experiments, and their 3D structures analyzed to understand the chemical features underlying their favorable performance.

To assess the effectiveness of EVOMOL-FLOW, we performed our experiments on the Vascular Endothelial Growth Factor Receptor (VEGFR), a well known tyrosine kinase that plays a major role in pathological angiogenesis, such as the one promoting solid tumor growth [18]. Blocking the access of Vascular Endothelial Growth Factor (VEGF) to its receptor would thus be crucial to prevent tumor cells proliferation. The 3D molecular structure of the receptor was retrieved by the RCSB Protein Data Bank (RCSB PDB) [19].

In this structure, the receptor is bound to a known VEGFR inhibitor, Sunitinib. This reference ligand allows us to locate the ATP-binding site, which is relevant in this approach to drive the evolution of drug candidates: the pocket is, in fact, the portion of the receptor encoded by DrugCLIP used for the evaluation of the binding.

The paper is structured as follows. Section II introduces the components of EVOMOL-FLOW. In Section III, we present the experimental setup and the obtained results. Finally, in Section IV we draw the conclusions of the work and present future developments.

II. METHODS

A. Multi-Objective Optimization

To evolve the most suitable small molecules, we formalized the structure-based *de novo* drug design as a multi-objective optimization problem. Formally, a multi-objective optimization problem can be defined, without loss of generality, as:

$$\underset{x \in S}{\text{minimize}} F(x) = (f_1(x), \dots, f_K(x))^T,$$

where S is the search space and $x \in S$ is a candidate solution. $F(x)$ is a set of K -objective functions (also referred to as fitness functions), each defined as $f_i : S \rightarrow \mathbb{R}$, $i \in \{1, \dots, K\}$. Since these objectives are often conflicting, and a single solution that simultaneously optimizes all of them does not exist [20], multi-objective approaches aim to find multiple equally optimal solutions [21]. This circumstance is defined by the strong and weak domination rules:

- a solution $x_1 \in S$ strongly dominates another solution $x_2 \in S$ if and only if $f_i(x_1) \leq f_i(x_2)$, $\forall i \in \{1, \dots, K\}$;
- a solution $x_1 \in S$ weakly dominates another solution $x_2 \in S$ if $f_j(x_1) < f_j(x_2)$, for at least one index $j \in \{1, \dots, K\}$.

The goal of a multi-objective optimization algorithm is to identify the Pareto-optimal front, that is, a set of candidate solutions that weakly dominate each other, while strongly dominating all other candidate solutions. The quality of the Pareto Front (PF) can be assessed using the hypervolume of the hyper-polytope identified by the Pareto-optimal solutions and a reference point (typically, the worst possible solution). Since the domination rules induce a partial order on the solutions, the concept of PF can be extended to dominated solutions by defining a ranking of PFs. For instance, the second-ranked PF include the dominated solutions that dominate all other solutions with the exception of those belonging to the Pareto-optimal front. Among the algorithms introduced to tackle these problems, Evolutionary Computation (EC) obtained a large success in the literature, thanks to its approach inspired by Darwinian theories. EC algorithms implement an iterative approach that, starting from a population of individuals (*i.e.*, candidate solutions), aims to improve their quality according to some given objectives. During each iteration, also named generation, the best individuals are selected to act as parents and produce an offspring of novel individuals by means of evolutionary operators. Commonly, such operators are crossover and mutation; the first takes two individuals and generates a new individual that combines (often randomly) the characteristics of both parents, while the second randomly changes a small part of one individual. Then, the quality of the offspring is evaluated and a new generation is performed, until a termination criterion is met (*e.g.*, a certain number of generations is run).

Since the representation and initialization of candidate solutions, as well as the definition of evolutionary operators, have to be tailored for the problem at hand, the main difference among the EC algorithms proposed in the literature is mainly related to the selection strategy. EVOMOL-FLOW uses NSGA-II, a multi-objective evolutionary approach that performs selection according to two criteria: (i) the ranking of the PFs detected during each generation, and (ii) the crowding distance [22]. In particular, this type of selection works as an iterative process that, starting from the first PF, adds all solutions of that front to the set of surviving solutions. If adding all the solutions of the k -th PF exceeds the number of surviving solutions, then only the most isolated solutions, according to the crowding distance, will survive.

B. Population initialization

EVOMOL-FLOW leverages the SMILES (Simplified Molecular Input Line Entry System) encoding, a widely used chemical notation that represents molecular structures as short ASCII strings [23]. SMILES encode the molecule's connectivity and characteristics in a linear format, making it easy for computers to store, process, and interpret the chemical information.

EVOMOL-FLOW can leverage any set of valid SMILES strings – produced by generative models or retrieved by databases – to start the evolution. In the experiments reported in this work, the initial population used by NSGA-II was based on the molecules available on ZINC15 [24], a free repository of commercially-available compounds for virtual screening. ZINC15 contains more than 230 million purchasable compounds in ready-to-dock format, also represented as SMILES notation. In particular, we selected the tranches tagged as “lead-like” – namely, compounds that are considered suitable as starting points for drug development – and with availability “in-stock”, for a total of 2 076 017 unique SMILES. The rationale behind this selection is to begin the evolutionary process from a readily available population of small molecules that have already shown safety and therapeutical activity. EVOMOL-FLOW automatically fetches the molecules from the database during the initialization stage. Then, at run-time, it randomly samples 500 molecules, which are exploited as initial population during the evolutionary phase.

C. DrugCLIP

DrugCLIP is an AI-platform created to accelerate *in silico* screening in drug discovery [25]. Based on a molecular encoder, and trained using a contrastive learning approach, DrugCLIP generates an embedding of the candidate small molecules (also named ligands) and the binding pocket of the protein target. One key advantage of DrugCLIP is its execution speed: according to developers, it can calculate the embedding of millions of molecules in about one minute. By quickly mapping protein pockets and small molecules in a shared latent space, the alignment between two molecules can be assessed by comparing their embeddings. Thus, DrugCLIP represents an extremely efficient alternative to molecular docking, paving the way to massive testing of drug candidates.

The goal of DrugCLIP was to accelerate virtual screening: given a target molecule, retrieve all compatible molecules across all public databases as efficiently as possible. In this work we propose a different strategy: explore the space of possible drugs with an evolutionary approach, and assess the pocket-ligand affinity using cosine similarity as the first objective function of EVOMOL-FLOW.

D. Objectives

EVOMOL-FLOW simultaneously optimizes two objectives: (i) the binding affinity between the candidate small molecule and the ligand under investigation; (ii) the synthesizability of the candidate small molecule.

To compute the binding affinity, we calculate the dot product between the DrugCLIP embeddings of the candidate small molecule and the ligand's binding pocket: the larger the dot product the greater the affinity, that is, the binding produces small values of free energy. Because the embeddings are normalized, the value of the product ranges between -1 and 1.

To maximize the synthesizability of the candidate molecules, EVOMOL-FLOW aims to minimize the Synthetic Accessibility (SA) score [26] – a value that ranges from 1 (easy to synthesize) to 10 (hard to synthesize) – computed by RDKit [27]. The SA score is based on the analysis of one million representative molecules from PubChem [28] and takes into account several characteristics of the molecule, ranging from the presence of non-standard structural features to the overall molecule's size.

E. Evolutionary Operators

When applying an evolutionary approach on SMILES strings, the operators must be carefully tailored, as not every arbitrary combination of SMILES characters corresponds to a valid 2D molecule [3]. SMILES strings follow a strict syntax: atoms must respect their correct valence, *e.g.*, four bonds for carbon (C), two bonds for oxygen (O). Rings and branches, that are respectively encoded with numbers and parentheses, must remain properly matched. Any violation of this syntax during offspring generation can produce invalid molecules, potentially reducing the population size and compromising the evolutionary process. In EVOMOL-FLOW, the population is evolved using a combination of specific evolutionary operators, as follows.

The crossover operator is adopted from MolFinder [9], which performs syntax-aware cuts on the parents: one parent is truncated from the left side, and the other from the right side, both avoiding ring substructures and re-balancing parenthesis when needed, thus reaching a validity higher than 80%. In case of a failed crossover, the parents are placed back in the population without any modification. The atom-based mutation is taken from MolFinder, as well. Two lists of possible atoms for substitution are set (one for generic atoms, and one for aromatic atoms), ensuring the core features of the molecule are preserved, and the valence of the atoms is checked, guaranteeing that every modification is chemically valid. To enhance the exploration of the chemical space, we also implemented a brand new fragment-based mutation. The idea is to preserve the core scaffold of the molecule and mutate a randomly selected side-chain

Algorithm 1 EVOMOL-FLOW Evolutionary Optimization**Input:** N (Population size), G (Generations), M_{target} (Target)**Output:** S^* (Validated candidate scaffolds)

```

1:  $\mathcal{P}_0 \leftarrow \text{SampleLeadLike}(\text{ZINC15}, N)$ 
2:  $\mathbf{e}_{\text{target}} \leftarrow \text{DrugCLIP}(\text{Encode}(M_{\text{target}}))$ 
3: for  $t = 0$  to  $G - 1$  do
4:   for each  $m \in \mathcal{P}_t$  do
5:      $f_1(m) \leftarrow \text{Affinity}(\mathbf{e}_m, \mathbf{e}_{\text{target}})$ 
6:      $f_2(m) \leftarrow \text{SAScore}(m)$ 
7:   end for
8:    $\mathcal{P}_{\text{parents}} \leftarrow \text{NSGA-II}(\mathcal{P}_t, f_1, f_2)$ 
9:    $\mathcal{O}_t \leftarrow \emptyset$ 
10:  while  $|\mathcal{O}_t| < N$  do
11:     $m_i, m_j \in \mathcal{P}_{\text{parents}}$ 
12:    if with probability  $P_{\text{cross}}$  then
13:       $m_{\text{off}} \leftarrow \text{SyntaxAwareCrossover}(m_i, m_j)$ 
14:    else
15:      if with probability  $P_{\text{atom}}$  then
16:         $m_{\text{off}} \leftarrow \text{MutateAtom}(m_i)$ 
17:      else
18:         $m_{\text{off}} \leftarrow \text{MutateFragment}(m_i)$ 
19:      end if
20:    end if
21:    if  $\text{IsValid}(m_{\text{off}})$  then
22:       $\mathcal{O}_t \leftarrow \mathcal{O}_t \cup \{m_{\text{off}}\}$ 
23:    end if
24:  end while
25:   $\mathcal{P}_{t+1} \leftarrow \text{SelectTop}(\mathcal{P}_t \cup \mathcal{O}_t, N)$ 
26: end for
27:  $\mathcal{F} \leftarrow \{m \in \mathcal{P}_G \mid f_1(m) > T_{\text{aff}} \wedge f_2(m) < T_{\text{SA}}\}$ 
28:  $\mathcal{C} \leftarrow \text{Cluster}(\mathcal{F})$ 
29:  $S^* \leftarrow \text{Validate}(\mathcal{C})$ 
30: return  $S^*$ 

```

Variables: N, G : Search hyperparameters; M_{target} : Target protein structure; $\mathcal{P}_t, \mathcal{O}_t$: Current and offspring populations; m_{off} : Offspring candidate molecule; f_1, f_2 : Affinity and synthetic accessibility; $P_{\text{cross}}, P_{\text{atom}}$: Crossover and atom mutation probabilities; $\mathbf{e}$: Latent embeddings; $T_{\text{aff}}, T_{\text{SA}}$: Affinity and SA score thresholds; $\mathcal{F}$: Filtered Pareto front; $\mathcal{C}$: Structural clusters; S^* : Final validated leads.

with a compatible one. Also in this case, every substitution is chemistry-informed: the replacement of a side-chain is chosen from a curated database of side-chains commonly used in medicinal chemistry, together with their preferred substitutions [29].

At each step, the operators are applied according to a user-specified probability. After a preliminary investigation of the parameters (data not shown due to space constraints), one of the most promising combinations was: crossover probability: 0.6; mutation probability: 0.4. When mutation is selected, the specific strategy is chosen according to a second conditional probability distribution: atom-based mutation probability: 0.3; fragment-based mutation probability: 0.7. The complete EVOMOL-FLOW framework is formalized in Algorithm 1.

F. Metrics for Docking

After the evolution, we further assess the quality of the most promising candidates by leveraging molecular docking and molecular dynamics simulations for their ability to approximate the experimental binding modes of the ligands. While the output of the docking is the binding conformation of the ligand together with its affinity score, MD provides a trajectory, that is, a time-ordered series of the ligand-receptor complex conformations generated during the simulation. This

assessment allows us to directly rationalize the chemical features underlying the predicted scores, and validate our results. Docking experiments were performed with AutoDock Vina [30], using an exhaustiveness of 8 and a grid box centered based on the reference ligand with dimensions 20 x 20 x 20. Molecular dynamics simulations were performed with OpenMM [31] for a total simulation time of 100 ns, using a time step of 2 fs and maintaining the system at 300 K. Pressure was controlled at 1 bar. To estimate the binding affinity of the complex, we used the Vina scoring function for both docking and MD experiments. In this latter case, the function was applied on a specific pose of the ligand-receptor complex. This pose corresponds to the most frequently sampled conformation during the MD simulation, representing the most stable complex. In practical terms, we find this structure by clustering all the frames of the MD trajectory [32]. The clustering is performed using DBSCAN [33], a density based algorithm that groups structures based on their structural similarity; in this specific case, each frame represent a pose of the ligand-receptor complex. After computing a pair-wise Root Mean Squared Deviation (RMSD) between all pairs of frames, DBSCAN uses the resulting matrix – where each element represents the RMSD between two conformations – as a distance metric to identify dense regions of similar structures, without requiring a predefined number of clusters. Conformations within a certain similarity cutoff will fall in the same cluster. The centroid of the most populated cluster corresponds to the most frequently sampled conformational states, and is selected as the representative structure for binding energy evaluation and for analyzing key chemical interactions. This ensures that the evaluation accounts for conformational stability and dynamical effects, and is therefore more physically grounded than docking-based scoring alone. The selected molecules are then clustered based on structural similarity to understand if the model converged towards a shared set of pharmacophoric features, such as hydrogen donors or acceptors, aromatic rings, and hydrophobic groups. This would be a suggestion of the characteristics to consider as crucial for the binding. The structural similarity is computed using 2D pharmacophore fingerprints (Pharm2D) [34], which are able to capture the spatial relation of the pharmacophoric features, combined with the Tanimoto coefficient [35], a score ranging from 0 to 1 that, in this context, reflects the degree of overlap between pharmacophoric patterns. Then, clustering is performed again with DBSCAN, this time using the similarity matrix resulted from the calculation of Tanimoto coefficients between pairs of molecules. The resulting clusters of molecules are then visually inspected with PyMOL [36] to relate their shared molecular structures and pharmacophoric features to binding properties. For example, key contacts (*i.e.*, those known to be important for the binding, as established by the reference inhibitor) maintained by the evolved molecules, and possibly new pocket hotspots found by EVOMOL-FLOW, have been reported.

III. RESULTS

We retrieved the 3D model of the VEGFR-Sunitinib complex (PDB code: 4AGD), and used the receptor’s binding pocket as a target for the EC experiments. We performed 30 independent runs of NSGA-II, using 500 individuals and 200 generations per run. Autodock-Vina-GPU (v2.1) and DrugCLIP were run on a machine with a GPU Nvidia A40 with 46 GB of GDDR6 VRAM, CUDA version 12.8. EVOMOL-FLOW also uses RDKit (v2025.9.3) [27], DEAP (v1.4.3) [37], and PyTorch (v2.9.1) [38].

To investigate the convergence of EVOMOL-FLOW, we computed the hypervolume of the PF hyper-polytope – as defined in Section II-A – of the initial and the final populations. In principle, the hypervolume is supposed to increase during the evolutionary process, so that the higher the hypervolume, the better the solutions. Figure 1 shows the distribution of the hypervolumes at the beginning (left boxplot) and at the end of the optimization process (right boxplot), showing that EVOMOL-FLOW successfully evolves the random population towards Pareto-optimal solutions.

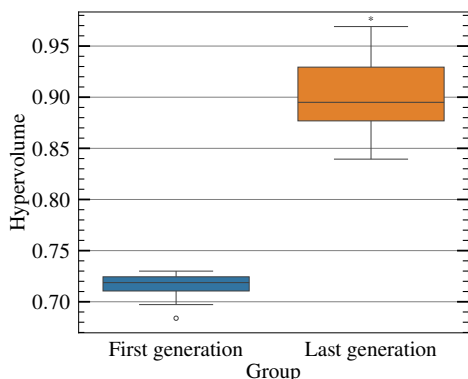


Fig. 1. Distribution of the hypervolumes at the beginning (left) and at the end (right) of the optimization with EVOMOL-FLOW across 30 repetitions.

As a criterion for the selection of the most promising molecules among the ones in the optimal PF, we filtered each final PF by using a threshold for each objective. Specifically, we filtered out solutions having the SA score ≥ 4 (as suggested in [39]) and the affinity score ≤ 0.7 . The latter threshold was defined considering the maximum value DrugCLIP’s affinity score can reach (*i.e.*, 1) to effectively discriminate high-affinity candidates within the upper range of the score distribution while preserving chemical diversity; consistently, molecules above this threshold always display docking scores below -6 kcal/mol, further strengthening the choice of the cutoff. In Figure 2 we plot an example of the PF solutions during run #4 of EVOMOL-FLOW. The figure shows two relevant outcomes. First, the evolved molecules (orange dots) are shifted towards a more promising tradeoff between the two objectives, when compared to the starting molecules (blue dots). This pattern holds across all runs (data not shown due to space constraints). Second, all solutions in the PF dominate the known drug for this protein, Sunitinib (represented by the black star). Sunitinib

is characterized by a worse SA score, and it is dominated by a subset of evolved molecules with respect to the affinity score.

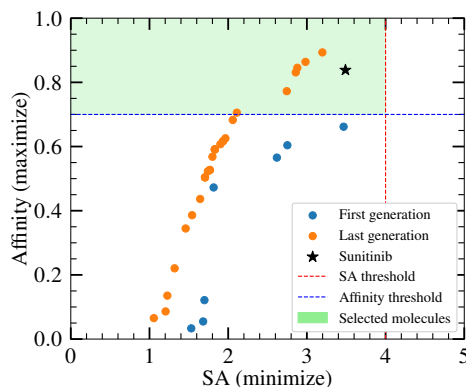


Fig. 2. One example of PF evolved by EVOMOL-FLOW. The dominant molecules in the first and last generation, respectively, are denoted by blue and orange dots. From the Pareto-optimal solutions, we extract only the molecules having a SA score lower than 4 (red dashed line) and an affinity score higher than 0.7 (blue dashed line), that is, the solutions contained in the green area. The known drug for this protein, Sunitinib, is denoted by a black star.

By filtering the final PFs across all EVOMOL-FLOW runs, we obtained a subset of 110 candidate molecules to be further investigated. Each molecule underwent docking and MD simulations of the corresponding 3D structures with respect to the 3D structure of the target receptor. Figure 3 shows that the distributions of Vina and MD scores for the filtered molecules are always lower than -6 kcal/mol, as expected for good binders, reaching even values ≤ -10 kcal/mol, typical of an excellent binding mode. As a comparison, we performed a docking evaluation on a set of drug candidates considered unfavorable based on the DrugCLIP score. These molecules exhibited significantly higher docking scores, highlighting a clear separation between the two sets. Due to computational constraints, we did not perform an MD simulation for the unfavorable set. However, Vina and MD scores consistently agreed in determining the quality of the good ligands. Therefore, docking scores were considered sufficient to confirm the discriminative power of the deep learning-based fitness function. To better understand the chemical features of our drug candidates and rationalize their promising scores, we clustered them as explained in Section II-F, and analyzed how the binding-related metrics change with different scaffolds. From the subset of 110 promising molecules, we identified 4 structural clusters and 29 outliers. This means that, across 30 independent runs of EVOMOL-FLOW, the model consistently converged on a number of core scaffolds, each representing a key pharmacophore driving binding affinity, while also finding alternative solutions that explore chemical space beyond the main pharmacophoric cores. Figure 4 shows the best structure (*i.e.*, the one with the highest DrugCLIP score) for each cluster (A to D) and the outliers (Ω). The top left panel in Figure 4 represents the reference inhibitor Sunitinib, which is known to interfere with the ATP-binding site of VEGFR, with the halogen substituted oxindole ring penetrating the cavity. Specifically,

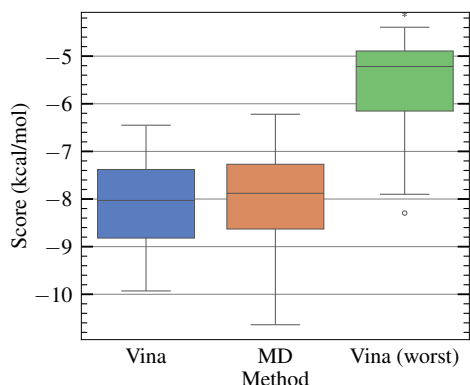


Fig. 3. Distribution of the Vina (blue boxplot) and MD scores (orange boxplot) of the filtered molecules (across all runs), against the Vina score of the worst molecules (*i.e.*, those with the lowest DrugCLIP score, green boxplot). The lower the scores, the better the solutions.

Glu917 and Cys919 residues of VEGFR have been reported to be important for the inhibition [40], [41].

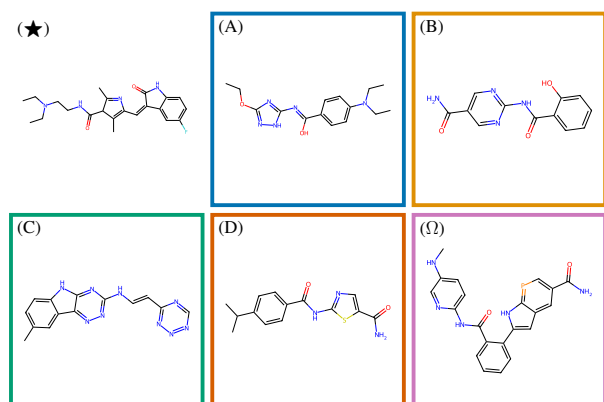


Fig. 4. (★) Sunitinib VEGFR inhibitor molecule. (A-D) Representative molecules for all clusters obtained from the set of filtered dominant molecules. (Ω) Representative molecule of the outliers.

To characterize the binding-behavior of each core scaffold, we observed the distribution of the affinity score across the clusters, shown in Figure 5. Despite some outliers, Cluster C is the one showcasing the most promising results: the molecule with the highest score (0.89) is represented in Figure 4C. This molecule mimics the oxindole substructure of Sunitinib, but introduces a third fused ring in place of the oxygen atom, thereby increasing rigidity, which may reduce conformational flexibility and stabilize binding within the ATP pocket.

Importantly, this molecule maintains the ability to deeply penetrate the pocket with the indolic moiety, as shown in Figure 6, and preserves the fundamental bonds with the receptor, as per Figure 7. The net of interactions here displayed is compatible with the binding mode of known inhibitors [42]: the key hydrogen bonds with the glutamine and cysteine within the hinge region are maintained, along with the hydrophobic contacts established with two valines in the back of the pocket and a leucine functioning as an anchoring point towards the

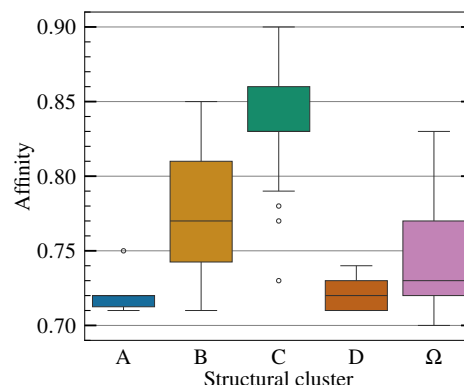


Fig. 5. Cluster-based distribution of the affinity scores computed by DrugCLIP. The higher the score, the better the solution. (A-D) Representative molecules for all clusters obtained from the set of filtered dominant molecules. (Ω) Representative molecule of the outliers.

surface. Interestingly, some of the hydrogen bonds are mediated by the third ring with two nitrogens. Despite not being part of Sunitinib, this substructure is present in Pazopanib, another known VEGFR-inhibitor, and has been investigated to produce novel potential anti-angiogenic compounds within a quinazoline moiety [43]. The diethylaminoethyl tail is replaced by a benzene moiety (Figure 4C), keeping a fair balance between polar and hydrophobic nature due to nitrogen substitutions.

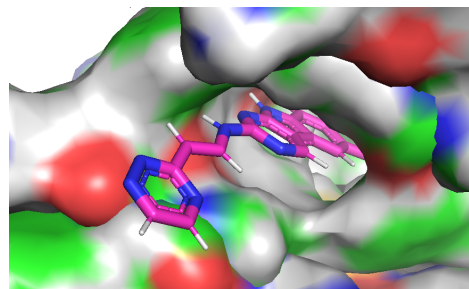


Fig. 6. Binding pose of the top-ranked molecule from Cluster C in the VEGFR binding pocket. The ligand is shown in pink as sticks, and the receptor is displayed as a surface, highlighting the shape and accessibility of the binding cavity. Visualization was performed using PyMOL.

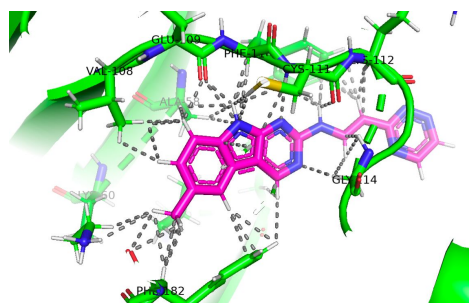


Fig. 7. Interaction network of the top-ranked molecule from Cluster C in the VEGFR binding pocket. The ligand is shown in pink (sticks), the receptor in green (ribbons), with pocket residues represented as sticks. Visualization was performed using PyMOL.

The other clusters also present some interesting features. A two-rings pattern is recurrent, and help the novel compounds to partially superimpose with the pharmacophoric region of Sunitinib. In particular, Cluster D presents a thiazole moiety, a substructure that was identified as promising in mimicking Sunitinib mode of action [44]. Overall, these features showcase EVOMOL-FLOW’s ability of evolving molecules with key substructures, mimicking the behavior of known inhibitors within the hinge region, and preserving key contacts.

Concerning the outliers, the conformation of the best molecule in Cluster Ω represents a possible binder but fails in penetrating the pocket, and therefore taking contact with the key amino acids, because of its bulky scaffold (Figure 4 Ω). Although this could be interpreted as a negative result, it may also provide a valuable indication. The fact that EVOMOL-FLOW classifies this molecule as an outlier suggests that this conformation is only marginally favored during evolution. This means it recognizes it is not the best conformation overall, but it may suggest another type of inhibition mechanism, that should not be discarded *a priori*.

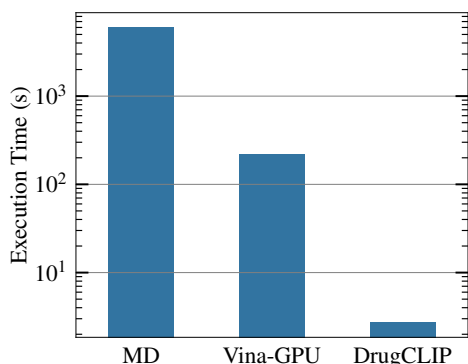


Fig. 8. Execution time of the objective function over a single generation of EVOMOL-FLOW when using MD, Vina-GPU and DrugCLIP for the fitness evaluation. The y axis is in logarithmic scale.

Although the evolutionary approach is computationally demanding in terms of number of individuals and generations performed, EVOMOL-FLOW remains feasible thanks to DrugCLIP, which reduces the running time by three orders of magnitude in comparison with a conventional approach based on MD, and two orders of magnitude with respect to simple docking with Vina accelerated on GPU, as shown in Figure 8. Specifically, a whole run of EVOMOL-FLOW with DrugCLIP requires 550 seconds, while the running time using Vina for the fitness evaluations is 44 394 seconds. This work shows that DrugCLIP’s AI-based technology is extremely efficient not only for virtual screening but also for any alternative application where binding prediction is central, like the exploration of the molecular search space performed by EVOMOL-FLOW.

IV. CONCLUSION

In this paper we presented a novel methodology for structure-based *de novo* drug design based on evolutionary computation and neural networks used, respectively, to perform multi-objective optimization and molecular encoding. The former

allows to identify the set of Pareto-dominant solutions, which are optimal from the point of view of binding affinity and synthetic accessibility. The latter represents a more computationally efficient alternative to MD or docking, and allows to strongly reduce the running time of the overall methodology.

We tested EVOMOL-FLOW on the VEGFR transmembrane protein, showing that it can identify novel promising drugs – dominating a known inhibitor of this molecule, Sunitinib – in a matter of a few minutes. Performing docking and MD analysis on non-dominated solutions not only ensures the reliability and chemical relevance of these predictions, but also assesses the consistency and effectiveness of the deep learning-based fitness function employed during evolution. Moreover, the observed convergence towards a limited number of core structures highlights EVOMOL-FLOW’s ability to identify robust scaffold architectures while still allowing exploration of structurally diverse candidates. Currently, the main bottleneck of EVOMOL-FLOW is the conversion, performed by RDKit, of the small molecule from its SMILES format to the 3D structure that is then processed by DrugCLIP. We accelerate this step by using multi-threading to convert every molecule, but only linearly with the number of individuals in NSGA-II. DrugCLIP, on the other hand, can process in parallel all individuals of a generation keeping the GPU still far from resources saturation. As a future work, we will investigate how to reduce RDKit-related computation time, in order to improve efficiency and scalability to larger populations. Further extensions will include the introduction of adaptive operator probabilities to enhance early exploration and progressively refine candidates. To better characterize the entire drug design pipeline, we plan to meta-optimize parameters, including assessing the impact of the initial population on outcomes.

ACKNOWLEDGMENT

This work is supported by the high-performance computing infrastructure developed under the project “CONVECS”, funded by the PR Veneto FESR 2021-2027 program, Priority 1 – Specific Objective 1.1 – Action 1.1.2. This work is also partially supported by the MUR under the grant “Dipartimenti di Eccellenza 2023-2027” of the Department of Informatics, Systems and Communication of the University of Milano-Bicocca, Milan, Italy, and by the National Plan for NRRP Complementary Investments (established with the decree-law May 6, 2021, n. 59, converted by law n. 101 of 2021) in the call for the funding of research initiatives for technologies and innovative trajectories in the health care sectors (Directorial Decree n. 931 of 06-06-2022) – project n. PNC0000003 – Advanced Technologies for Human-centred Medicine (project acronym: ANTHEM). This work reflects only the authors’ views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them.

REFERENCES

- [1] O. J. Wouters, M. McKee, and J. Luyten, “Estimated research and development investment needed to bring a new medicine to market, 2009-2018,” *JAMA*, vol. 323, no. 9, pp. 844–853, 2020.

- [2] S. Mandal, M. Moudgil, and S. K. Mandal, "Rational drug design," *European Journal of Pharmacology*, vol. 625, no. 1-3, pp. 90-100, 2009.
- [3] L. Bacciu, M. Grazioso, S. Multari, F. Grisoni, A. Mazzolari, and M. S. Nobile, "Improving the efficiency and the validity of molecular transformers," in *2025 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*, pp. 1-9, IEEE, 2025.
- [4] M. Moret, I. Pachon Angona, L. Cotos, S. Yan, K. Atz, C. Brunner, M. Baumgartner, F. Grisoni, and G. Schneider, "Leveraging molecular structure and bioactivity with chemical language models for *de novo* drug design," *Nature Communications*, vol. 14, no. 1, p. 114, 2023.
- [5] F. Grisoni, "Chemical language models for *de novo* drug design: Challenges and opportunities," *Current Opinion in Structural Biology*, vol. 79, p. 102527, 2023.
- [6] R. Özçelik, S. de Ruiter, and F. Grisoni, "Chemical language modeling with structured state space sequence models," *Nature Communications*, vol. 15, no. 1, p. 6176, 2024.
- [7] S. Multari, R. Özçelik, A. Mazzolari, M. S. Nobile, and F. Grisoni, "Predicting metabolic reactions with a molecular transformer for drug design optimization," in *2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1-8, IEEE, 2024.
- [8] J. C. Fromer and C. W. Coley, "Computer-aided multi-objective optimization in small molecule discovery," *Patterns*, vol. 4, no. 2, p. 100678, 2023.
- [9] Y. Kwon and J. Lee, "MolFinder: an evolutionary algorithm for the global optimization of molecular properties and the extensive exploration of chemical space using SMILES," *Journal of Cheminformatics*, vol. 13, no. 1, p. 24, 2021.
- [10] E. van der Horst, P. Marqués-Gallego, T. Mulder-Krieger, J. van Veldhoven, J. Kruisselbrink, A. Aleman, M. T. Emmerich, J. Brussee, A. Bender, and A. P. IJzerman, "Multi-objective evolutionary design of adenosine receptor ligands," *Journal of Chemical Information and Modeling*, vol. 52, no. 7, pp. 1713-1721, 2012.
- [11] Y. Tang, R. Moretti, and J. Meiler, "Recent advances in automated structure-based *de novo* drug design," *Journal of Chemical Information and Modeling*, vol. 64, no. 6, pp. 1794-1805, 2024.
- [12] X.-Y. Meng, H.-X. Zhang, M. Mezei, and M. Cui, "Molecular docking: a powerful approach for structure-based drug discovery," *Current Computer-Aided Drug Design*, vol. 7, no. 2, pp. 146-157, 2011.
- [13] S. Grewal, G. Deswal, A. S. Grewal, and K. Guarve, "Molecular dynamics simulations: Insights into protein and protein ligand interactions," in *Advances in Pharmacology*, vol. 103, pp. 139-162, Elsevier, 2025.
- [14] Z. Ma, A. Ajibade, and X. Zou, "Docking strategies for predicting protein-ligand interactions and their application to structure-based drug design," *Communications in Information and Systems*, vol. 24, no. 3, p. 199, 2024.
- [15] M. K. Sahu, A. K. Nayak, B. Hailemeskel, and O. E. Eyupoglu, "Exploring recent updates on molecular docking: Types, method, application, limitation & future prospects," *International Journal of Pharmaceutical Research and Allied Sciences*, vol. 13, no. 2-2024, pp. 24-40, 2024.
- [16] P. Śledź and A. Caflisch, "Protein structure-based drug design: from docking to molecular dynamics," *Current Opinion in Structural Biology*, vol. 48, pp. 93-102, 2018.
- [17] B. Gao, B. Qiang, H. Tan, Y. Jia, M. Ren, M. Lu, J. Liu, W.-Y. Ma, and Y. Lan, "DrugCLIP: contrastive protein-molecule representation learning for virtual screening," in *NeurIPS 2023*, 2023.
- [18] M. Shibuya, "Vascular endothelial growth factor (VEGF) and its receptor (VEGFR) signaling in angiogenesis: a crucial target for anti-and pro-angiogenic therapies," *Genes & Cancer*, vol. 2, no. 12, pp. 1097-1105, 2011.
- [19] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The protein data bank," *Nucleic Acids Research*, vol. 28, no. 1, pp. 235-242, 2000.
- [20] K. Deb, "Multi-objective optimization," in *Search methodologies: Introductory tutorials in optimization and decision support techniques*, pp. 403-449, Springer, 2014.
- [21] A. Eiben and J. Smith, *Introduction to evolutionary computing*. Springer, 2015.
- [22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.
- [23] D. Weininger, "SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules," *Journal of Chemical Information and Computer Sciences*, vol. 28, no. 1, pp. 31-36, 1988.
- [24] T. Sterling and J. J. Irwin, "Zinc 15—ligand discovery for everyone," *Journal of Chemical Information and Modeling*, vol. 55, no. 11, pp. 2324-2337, 2015.
- [25] Y. Jia, B. Gao, J. Tan, J. Zheng, X. Hong, W. Zhu, H. Tan, Y. Xiao, L. Tan, H. Cai, *et al.*, "Deep contrastive learning enables genome-wide virtual screening," *Science*, vol. 391, no. 6781, p. eads9530, 2026.
- [26] P. Ertl and A. Schuffenhauer, "Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions," *Journal of Cheminformatics*, vol. 1, pp. 1-11, 2009.
- [27] RDKit Community, "RDKit: Open-source cheminformatics," 2023.
- [28] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, *et al.*, "Pubchem 2025 update," *Nucleic Acids Research*, vol. 53, no. D1, pp. D1516-D1525, 2025.
- [29] K. Takeuchi, R. Kunitomo, and J. Bajorath, "Searchable database of frequent R-groups in medicinal chemistry and their preferred replacements," *Data in Brief*, vol. 39, p. 107456, 2021.
- [30] O. Trott and 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*, vol. 31, no. 2, pp. 455-461, 2010.
- [31] P. Eastman, R. Galvelis, R. P. Peláez, C. R. Abreu, S. E. Farr, E. Gallicchio, A. Gorenko, M. M. Henry, F. Hu, J. Huang, *et al.*, "OpenMM 8: molecular dynamics simulation with machine learning potentials," *The Journal of Physical Chemistry B*, vol. 128, no. 1, pp. 109-116, 2023.
- [32] J.-h. Peng, W. Wang, Y.-q. Yu, H.-l. Gu, and X. Huang, "Clustering algorithms to analyze molecular dynamics simulation trajectories for complex chemical and biological systems," *Chinese Journal of Chemical Physics*, vol. 31, no. 4, pp. 404-420, 2018.
- [33] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, p. 226-231, AAAI Press, 1996.
- [34] A. Gobbi and D. Poppinger, "Genetic optimization of combinatorial libraries," *Biotechnology and Bioengineering*, vol. 61, no. 1, pp. 47-54, 1998.
- [35] D. J. Rogers and T. T. Tanimoto, "A computer program for classifying plants: The computer is programmed to simulate the taxonomic process of comparing each case with every other case," *Science*, vol. 132, no. 3434, pp. 1115-1118, 1960.
- [36] W. L. DeLano *et al.*, "PyMOL: An open-source molecular graphics tool," *CCP4 Newsl. protein crystallogr.*, vol. 40, no. 1, pp. 82-92, 2002.
- [37] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: Evolutionary algorithms made easy," *Journal of Machine Learning Research*, vol. 13, pp. 2171-2175, jul 2012.
- [38] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
- [39] A. Kerstjens, *Computational design of synthesizable molecules by imitating reference chemistry*. PhD thesis, University of Antwerp, 2024.
- [40] T.-H. Yang, C.-I. Lee, W.-H. Huang, and A.-R. Lee, "Synthesis and evaluation of novel 2-pyrrolidone-fused (2-oxoindolin-3-ylidene) methylpyrrole derivatives as potential multi-target tyrosine kinase receptor inhibitors," *Molecules*, vol. 22, no. 6, p. 913, 2017.
- [41] P. C. Tang, Y. D. Su, J. Feng, J. H. Fu, J. L. Yang, L. Xiao, J. H. Peng, Y. L. Li, L. Zhang, B. Hu, *et al.*, "Novel potent orally active multitargeted receptor tyrosine kinase inhibitors: synthesis, structure-activity relationships, and antitumor activities of 2-indolinone derivatives," *Journal of Medicinal Chemistry*, vol. 53, no. 22, pp. 8140-8149, 2010.
- [42] S. J. Modi and V. M. Kulkarni, "Exploration of structural requirements for the inhibition of vegfr-2 tyrosine kinase: Binding site analysis of type ii, 'dfg-out' inhibitors," *Journal of Biomolecular Structure and Dynamics*, vol. 40, no. 12, pp. 5712-5727, 2022.
- [43] M. Moradi, A. Mousavi, Z. Emamgholipour, J. Giovannini, S. Moghimi, F. Peytam, A. Honarmand, S. Bach, and A. Foroumadi, "Quinazoline-based VEGFR-2 inhibitors as potential anti-angiogenic agents: A contemporary perspective of SAR and molecular docking studies," *European Journal of Medicinal Chemistry*, vol. 259, p. 115626, 2023.
- [44] A. A. Abd Elhameed, A. R. Ali, H. A. Ghabbour, S. M. Bayomi, and N. S. El-Gohary, "Probing structural requirements for thiazole-based mimetics of sunitinib as potent vegfr-2 inhibitors," *RSC Medicinal Chemistry*, vol. 16, no. 3, pp. 1329-1349, 2025.