Optimal Molecular Design: Generative Active Learning Combining REINVENT with Absolute Binding Free Energy Simulations

Hannes H Loeffler[1]*, Shunzhou Wan[2], Marco Klähn[1], Agastya P Bhati[2]*, Peter V Coveney[2,3,4]

Abstract

Active learning (AL) is a specific instance of sequential experimental design and uses machine learning to intelligently choose the next data point or batch of molecular structures to be evaluated. In this sense it closely mimics the iterative design-make-test-analysis cycle of laboratory experiments to find optimized compounds for a given design task. Here we describe an AL protocol which combines generative molecular AI, using REINVENT, and physics-based absolute binding free energy molecular dynamics simulation, using ESMACS, to discover new ligands for two different target proteins, 3CLpro and TNKS2. We have deployed our generative active learning (GAL) protocol on Frontier, the world's only exa-scale machine. We show that the protocol can find better binders compared to baseline, a surrogate ML docking model for 3CLpro and compounds with experimentally determined binding affinities for TNKS2. The ligands found are also chemically diverse and occupy a different chemical space than the baseline. We vary the batch sizes that are put forward for free energy assessment in each GAL cycle to assess the impact on their efficiency on the GAL protocol and recommend their optimal values in different scenarios. Overall, we demonstrate a powerful capability of the combination of physics-based and AI methods which yields effective chemical space sampling at an unprecedented scale, of immediate and direct relevance for modern, data-driven drug discovery.

1. Introduction

Designing optimised molecules for specific purposes is fundamental to chemistry and is central to the discovery of new medicines and new materials. In practice, experimentally this is an iterative, slow, and expensive process (Schlander et al., 2021; Wouters et al., 2020) commonly achieved through a DMTA (design-make-test-analysis) cycle where newly created compounds need to be tested for their properties and inform the design used in the next iteration (Wesolowski & Brown, 2016; Schneider et al., 2020). Thus, there is a natural desire to turn to computation to find better compounds more quickly and cheaply.

Active learning (AL) is a commonly applied machine learning (ML) method that queries an information source, which could be a human expert (human-in-the-loop) or accurate computational predictor, to label large amounts of data interactively. An AL algorithm seeks to find an optimal way to label only a small subset of the data out of a vast amount of unlabeled data and thus accelerate

Corresponding authors: hannes.loffler@astrazeneca.com, agastya.bhati.14@ucl.ac.uk
[1] Molecular AI, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden
[2] Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, U. K.
[3] Advanced Research Computing Centre, University College London, London WC1H 0AJ, U.K.
[4] Institute for Informatics, Faculty of Science, University of Amsterdam, 1098XH Amsterdam, The Netherlands.

the learning process. The subset is selected *actively*, *i.e*. intelligently, to minimize the number of iterations and only forwards samples to the expert which the algorithm predicts will increase knowledge or is uncertain about. A passive algorithm to construct a prediction model would choose the subset at random and often not in an iterative manner. AL is thus classed as a *semi-supervised* learning strategy as the expert can only explicitly label a small amount of the dataset. These labels inform a predictive model (the *surrogate*) which can eventually be used to label the whole dataset as needed.

The relationship of AL to Bayesian optimization (BO) (Garnett, 2023) has been noted, in particular to pool-based AL (Di Fiore et al., 2023). AL algorithms define an efficient way to label data and so improve the predictive model while BO seeks to find the global optimum of an unknown ("black box") function. In pool-based AL the distribution of the data is unknown and consequently only a discrete subset of examples can be queried or *vice versa* rather than the whole dataset being available upfront. In contrast, in population-based AL the distribution is known, and the goal is to find the optimal density of the distribution (Sugiyama & Nakajima, 2009). Here, we will discuss an algorithm that seeks to optimize for ligand-protein binding affinity while still maintaining a high level of exploration (that is, of diversity to promote discovery of a wide range of new molecules).

AL for molecule design often make use of large libraries or vendor catalogs as the pool as the pool (Gentile et al., 2020; Graff et al., 2021; Marin et al., 2023), when the size of the data set is known *a priori*. The size of these libraries can easily range into the millions and, indeed, even billions of compounds (Bellmann et al., 2022). In AL a, typically, computationally expensive method like docking or MD simulation is used to assign new labels e.g., a free energy of binding (also known as the binding affinity) or a docking score as proxy for the affinity. AL approaches that efficiently optimize compounds for binding affinity with RBFE (relative binding free energy) methods, (also referred to as FEP=free energy perturbation) have only appeared recently in the literature (Crivelli-Decker et al., 2023; de Oliveira et al., 2023; Gorantla et al., 2024; Gusev et al., 2023; Khalak et al., 2022; Knight et al., 2021; Konze et al., 2019; Mohr et al., 2022; Thompson et al., 2022) but have also been used to optimize an RBFE protocol itself (de Oliveira et al., 2023). Combining Generative AI with active learning (GAL) has only started very recently (Filella-Merce et al., 2023) including a proof-of-concept study with peptides (Hernandez-Garcia et al., 2023) and an application with ABFE (absolute binding free energy) (Eckmann et al., 2024). The latter focuses on multi-fidelity surrogate modelling: docking, experimental results from BindingDB (Liu et al., 2007), and a double-decoupling ABFE method (Heinzelmann & Gilson, 2021).

AL relies on an oracle (a term which refers to either a human expert or a more expensive computational method) for labelling which is the ground-truth in this scheme. A surrogate model is created with the aim to reproduce the predictions of the oracle but at a much lower computational cost. Typical algorithms used here are classical ML methods used in QSAR modeling such as random forest or state vector machines, but they can also be more a sophisticated artificial neural networks or Gaussian processes, the latter especially being often the preferred choice with Bayesian optimization. The surrogate model can then be used to compute a much larger subset of the library or even its entirety to create the labels and so eventually replaces the expensive oracle. An acquisition function (often called infill sampling criteria in BO) would then be applied to select a new, small subset for evaluation with the oracle. Informativeness, representativeness and diversity have been proposed previously as three criteria for the acquisition function [Wu2018]. The quality of the surrogate model and the acquisition strategy are crucial to the design of the AL algorithm as the final goal, as in this study, is to find optimized molecules while using the minimal resources and/or time possible. Alternatively, the final goal may be the construction of an optimal surrogate model to

replace the oracle in making further predictions—for instance, the surrogate model could be used as the starting model for a related target.

In this work we replace the fixed-size library with a generative model to create molecules on-the-fly drawing from a distributional description (Kearns et al., 1994) of chemical space (Loeffler et al., 2024). A generative model can produce a substantial subset of a chemical space (Arús-Pous et al., 2019) whereas vendor libraries are naturally limited to the molecules contained in the library which is defined and restrained in terms of their synthesis protocols (Bellmann et al., 2022). To this end, we apply REINVENT (Loeffler et al., 2024) which uses reinforcement learning (RL) to generate optimal molecules subject to external "information" i.e. scoring functions which evaluate each compound for its fitness. The RL algorithm drives a "prior" model of general chemical knowledge towards a specialized model representing the chemical space of the task (objective) at hand. The scoring function can be an agglomeration (weighted arithmetical or geometrical mean) of scoring components. The surrogate model is here the most crucial scoring component as it informs the RL algorithm about the label which in this case is the binding free energy of the compound to the target as computed using ESMACS (ensemble simulation-based MM-PBSA approach). Other, secondary, scoring components are also used; these are described in the Methods section.

Our work demonstrates the power of combining AL with physics-based methods to effectively sample the vast chemical space. The major advantages of our approach over previous, similar approaches are (a) the novelty of generating high-quality small molecules with generative AI, (b) the innovation of combining this with a physics-based model in an AL workflow and thus the direct importance to *in silico* drug discovery, (c) the reliability of our physics-based oracle which is based on ensemble simulations that minimizes false positives/negatives, (d) the scale of operation using a batch size as large as 1000 to provide a more comprehensive picture, (e) deployment on very large scale HPC platforms such as Frontier, and (f) short wall clock time requirements.

2. Methods

Here we describe the detailed protocol for RL with REINVENT, the simulation protocol for ESMACS, the GAL workflow and provide some details on the protein targets used in this work.

2.1 Sampling Chemical Space with REINVENT

REINVENT's main run mode is Reinforcement Learning (RL) (Loeffler et al., 2024). With this approach a model (the *agent*) is iteratively biased towards a target profile which is the aggregation of scoring components describing the desirability of a molecule. Here we use the weighted geometric mean to aggregate the components into one total score for each generated molecule. The main scoring component is a QSAR-like response model (scoring weight of 0.6) created using ChemProp 1.5.2 (Heid et al., 2024; Heid & Green, 2021; Stokes et al., 2020; Yang et al., 2019). ChemProp uses a directed message-passing neural network (D-MPNN) to produce a property prediction model. This model serves here as a surrogate model, and it is fixed in its algorithm and hyper-parameters. We thus assume that the model's predictive power will remain mostly constant. It is updated with a new batch of molecules with predicted binding affinities from ESMACS simulations in each GAL iteration. The initial model for $3CL^{Pro}$ is based on about 10,000 structures found with a surrogate docking model (Clyde et al., 2022, 2023). This ChemProp model was created after hyper-parameter optimization using cross-validation (Stone, 1974) with 5 folds and 5 ensembles (5 models with different initial weight setup) each, the maximum number of iterations being set to 30. The initial data set was split in the ratio 0.8: 0.1: 0.1 for training, test, and validation respectively. RDKit 2D

normalization without feature scaling (Heid et al., 2024) was used for this step as well as for all model updates. In each GAL step we update the model as described in 2.3. See (Loeffler et al., 2024) for details of the REINVENT protocol.

To bootstrap a model for TNKS2 the 27 known compounds with experimental affinity measurements from the benchmark set (Schindler et al., 2020) were handled with Qptuna (Mervin et al., 2024) which allows automatic model selection from a series of classical machine learning algorithms. The best Qptuna model was a random forest model which was subsequently used to generated 10,000 structures with REINVENT for evaluation with ESMACS (described in section 2.2). The resulting binding free energies were used to train another ChemProp model. The initial model was created in the same way as the one for $3CL^{Pro}$ and updated as described in 2.3.

The three other scoring components were QED (Bickerton et al., 2012) with weight 0.2, stereochemistry with weight 0.2 and structural alerts which acts as a filter meaning that compounds which match an undesired structural pattern receive a score of zero, one otherwise. QED assigns a drug-likeness score to each molecule. This is needed to produce reasonable drug-like molecules and stay within the applicability domain of the ChemProp model as generative models very quickly optimize towards the weaknesses of a response model. For example, we found that if molecule generation were otherwise unrestrained, REINVENT would start creating molecules with very long alkyl chains because the ChemProp model scores these highly. This is, however, clearly undesirable. As the REINVENT prior does not support stereochemistry, molecules with stereo-centres are scored with zero to encourage the agent to generate compounds without stereo-centres. Structural alerts are a small set of SMARTS pattern to filter and suppress unwanted substructures (functional groups, ring sizes); see input configuration in the Supplementary Information (SI) for details.

RL was run in two stages (Loeffler et al., 2024). The first stage uses QED, stereo-centr and structural alert scoring components to create a sensible drug-like agent to reduce the need to query the surrogate model with less useful compounds. This only needs to be done once as preparation for the GAL protocol. In the second stage these three scoring components together with the ChemProp model were applied. Only this second stage was run in each GAL step. For molecule generation we use the classical Reinvent prior (Blaschke, Arús-Pous, et al., 2020; Olivecrona et al., 2017) which is a *de novo* model that creates SMILES sequences starting from scratch. The number of new compounds generated in each GAL step is termed ”batch” and its size was set to 100. As learning strategy, we use DAP (Fialková et al., 2021) with a sigma of 128 and a learning rate of 0.0001. A diversity scaffold filter was used to encourage exploration of a wide range of unique scaffolds. The scaffolds were generated using the Bemis-Murcko algorithm (Bemis & Murcko, 1996) as implemented in RDKit (*RDKit: Open-Source Cheminformatics*, n.d.). The diversity filter used a memory size of 10 scaffolds and enforced a minimum score of 0.7 meaning that molecules having that particular scaffold receive a zero score in the following steps, provided the total score exceeds 0.7. Identical SMILES were downscored to zero starting with the second occurrence in the RL run. An inception memory was used as replay memory (Blaschke, Engkvist, et al., 2020) with a memory size of 50 and a sample size of 10. In case of $3CL^{pro}$ we seeded the inception memory with the best binders of the current GAL iteration and in the case of TNKS2 the memory was seeded with the 27 experimental verified structures. The input configurations are shown in the SI.

2.2 Binding Affinity Prediction with ESMACS

We use the ESMACS (enhanced sampling of molecular dynamics with approximation of continuum solvent) protocol (Wan et al., 2015, 2020) for the binding free energy calculations. ESMACS is based on MMPBSA (molecular mechanics Poisson-Boltzmann surface area) (Homeyer & Gohlke, 2012) calculations but uses ensembles of replicas to obtain reproducible binding affinity estimates with robust uncertainty estimates. It is an endpoint free energy calculation, in which the binding free energy is calculated as the free energy changes of bound (compound-protein complex) and unbound (separated protein and compounds) states. Here conformations of the complex, protein and compound are all extracted from simulation of the complex, a commonly used protocol for the end-point free energy methods. Such a protocol is well suited for a rational drug screening project, in which the correct ranking of binding affinities is more important than the calculation of accurate binding affinities for the selection of compounds for further investigation.

Model preparation. The compounds generated from REINVENT were first processed using FixpKa (version 2.1.3.0, OpenEye Scientific Software, Inc.) (*Molecular Modeling Software | OpenEye Scientific*, 2024) to obtain the correct protonation state at pH 7.4. Subsequently, up to 200 conformers per compound were generated using OMEGA (version 4.1.2.0, OpenEye Scientific Software, Inc.) (*Molecular Modeling Software | OpenEye Scientific*, 2024). The prepared structural library of the compounds was then docked to the protein (PDB IDs: 6W63 for $3CL^{pro}$ and 4UI5 for TNKS2) using FRED (version 4.1.1.0, OpenEye Scientific Software, Inc.) (*Molecular Modeling Software | OpenEye Scientific*, 2024). The docking poses with the best Chemgauss4 docking scores were used for the ensuing ESMACS simulations.

Preparation and setup of the ESMACS simulations were implemented using BAC (binding affinity calculator) (Sadiq et al., 2008). The AmberTools23 package (Case et al., 2023) was used for the set-up of the systems and the calculations of binding free energies. Parameterizations of the compounds were produced using the general Amber force field 2 (GAFF2). The Amber ff14SB force field was used for the protein, and TIP3P for water molecules. Partial charges of the compounds were generated using the AM1-BCC method. The protonation states of the protein residues were assigned using the "reduce" module of AmberTools (Case et al., 2023). All water molecules in the pdb files were retained. All ligand-protein complexes were solvated in orthorhombic water boxes with a minimum distance from the protein of 10 Å. Counterions were added to electrostatically neutralize the systems.

Molecular simulation. The standard ESMACS protocol (Wan et al., 2015) uses an ensemble of 25 replicas for each compound-protein complex to get precise predictions. For a drug screening project, a huge number of compounds need to be evaluated. A coarse-grained protocol may be used on grounds of speed, node hours required and throughput of many ligands (Wan et al., 2023). Here we used an ensemble of 10 replicas, with a smaller box size and a buffer distance of 10 Å instead of 14 Å compared to that in the standard ESMACS protocol. Such a coarse-grained protocol inevitably decreases the precision of the predictions but is well-suited for high-throughput virtual screening (Wan et al., 2023). NAMD3 (Phillips et al., 2020) was used as the MD engine for all of the equilibration and production runs. For each individual simulation, energy minimizations were first performed with heavy protein atoms restrained at their initial positions. The initial velocities were then generated independently from a Maxwell−Boltzmann distribution at 50 K, and the systems were heated up to and kept at 300K within 60 ps. Finally, 4 ns production simulations were run for each replica for the ESMACS simulations. For this study, we ran ~750 number of ESMACS calculations overall aggregating to a total simulation time of 37.5 μs for each iteration. However, we would like to point out here that all these calculations were performed on the world's first exascale machine, Frontier, allowing us to run them all concurrently. Given that the performance of MD simulations for our systems was 150

ns/day using a single GPU (AMD Instinct MI250X), we were able to get results for the entire batch of compounds in 50 minutes of wall clock time for each iteration. Further, we note that the present study forms part of a much bigger workflow involving multiple additional steps (to be published elsewhere).

2.3 Generative Active Learning

The main components of the GAL loop (see Figure 1) are ESMACS as the oracle, ChemProp as the surrogate model while REINVENT generates molecules on-the-fly and so replaces the more common fixed-size library (Konze et al., 2019; Thompson et al., 2022; Warmuth et al., 2003). After evaluation of the current batch with ESMACS the surrogate model is updated. In principle, it is possible to freeze a selected set of layers in a ChemProp model to avoid "catastrophic forgetting" (also catastrophic interference) (McCloskey & Cohen, 1989). However, we found that training a new model from scratch from the accumulating set of molecules from each GAL cycle yielded somewhat better results than selectively freezing neural network layers. This is of course more expensive to train but it may also counter covariate shift to some degree. We used 5-fold cross validation and 5 ensembles in each fold to create the surrogate model. Model update was started from the model of the previous step such that no more than 6 epochs (model optimization iterations) were needed. But for performance reasons we only used the best model out of these 25 models as determined by the smallest test RMSE. RL was then run with the ChemProp model as free energy predictor together with the other scoring components (see above) in 300-500 iterations. REINVENT forms an inner optimization loop as RL is used for molecule optimization while the outer loop is the AL algorithm itself (see Figure 1).

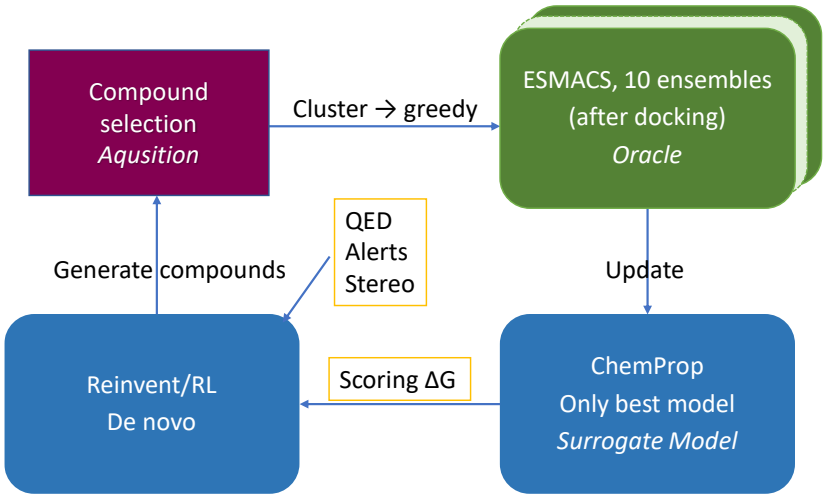

Figure 1. Schematic of overall workflow. Starting from the top right, ESMACS assesses the binding affinity of a set of compounds which are then used together with their ΔG_{ESMACS} to update a predictive ChemProp model (bottom right). REINVENT using the classical Reinvent prior recruits this model together with other scoring components (QED, Alerts, Stereo) to generate tens of thousands of new, potentially good binders (bottom left). The acquisition step (top left) chooses a subset of the generated compounds by clustering and selecting the best predicted binder by ΔG_{ESMACS} (greedy selection) to feed into the next GAL iteration.

The main parameters for GAL were the choice of the batch size and the selection criteria for compound progress. For $3CL^{pro}$ we used batch sizes of 250 and 500 molecules, respectively. For TNKS2 we used 5 batch sizes: 100, 300, 500, 700, and 1000. The acquisition of compounds from each RL run was done after clustering and choosing the top binders from each cluster (cluster-greedy approach) as done by Gusev et al., 2023. In this way we promote a level of exploration which is also supported by using the *de novo* REINVENT prior. The clusters were generated after reducing the predictor space to a 2D representation with UMAP (McInnes et al., 2018). As descriptors we employed RDKit fingerprints and measured their similarity using the Tanimoto distance (Bajusz et al., 2015).

The clustering approach we adopt here places emphasis on diversity and to a smaller degree on representativeness (by choosing the lowest ΔG value per cluster) but it neglects informativeness (Wu, 2019) i.e., we do not include e.g., uncertainty quantification or information gain in our GAL protocol.

2.4 Description of Test Targets

2.4.1 $3CL^{pro}$

C30 endopeptidase, usually referred to as $3CL^{pro}$ or M^{pro}, is the main protease in coronaviruses and plays a vital role in the lifecycle of SARS-CoV-2 (severe acute respiratory syndrome-coronavirus-2) which is the cause of Covid-19 (coronavirus disease 2019). Covid-19 led the WHO to declare a worldwide epidemic from 30 January 2020 to 5 May 2023 (WHO, 2023) after the first case was confirmed in November 2019. Hundreds of millions of cases have been confirmed since then and several million deaths have been directly linked to the virus (WHO, 2023). The virus and its various mutations are still a global threat to health. Several medications such as remdesivir (Veklury), nirmatrelvir / ritonavir (Paxlovid), baricitinib (Olumiant), tocilizumab (Actemra) and molnupiravir (Lagevrio) are available to treat various stages and forms of Covid-19. However, there is still a need to find new active compounds to deal with the continuously mutating virus and future related diseases. Preparation of $3CL^{pro}$ for simulations is described in section 2.2.

2.4.2 Tankyrase-2

Tankyrase-2 (gene TNKS2) was taken as a further test case from a large-scale binding free energy calculation benchmark data set (Schindler et al., 2020). TNKS2 is oncogenic and regulates various cellular processes, such as telomere maintenance, mitosis and glucose metabolism (Kim, 2018). It was chosen to study performance of GAL for a target with a more closed and confined binding pocket, compared to $3CL^{pro}$. It was also selected as a suitable test-case for a target for which 27 experimentally confirmed ligands from a congeneric series were available for use as seed structures to initialize GAL in order to see if binding affinities could be improved further and how far these generated structures would diverge structurally from the original ligands. Preparation of TNKS2 for simulations is described above in section 2.2.

3. Results and Discussion

In this section we show how GAL performance has been evaluated when applied to two different test targets, $3CL^{pro}$ and TNKS2, in terms of chemical space exploration, structural diversity of generated compounds, and the effect of compound training batch size on obtained results.

3.1 $3CL^{pro}$ Target

3.1.1 Evaluation of Surrogate Model

Initial training of the surrogate model based on ChemProp to predict binding affinities was achieved with a first batch of 10,000 structures whose binding free energies were calculated with ESMACS. These initial structures were selected according to their docking scores from a large library of molecules using a surrogate docking model (Clyde et al., 2022, 2023). In subsequent iteration steps, those compounds that were acquired from ESMACS calculations were added to the training set and the surrogate model was trained from scratch at each step.

Values of the binding free energy, ΔG_b, predicted by the surrogate model are compared with values derived with ESMACS, for all compounds that were acquired for oracle evaluation during GAL iterations, are shown in Fig. 2. Results are shown for each iteration step for small (250 molecules) and large (500 molecules) training batch sizes, respectively. The usually narrower range of ΔG_b values from the surrogate model is a result of the oracle acquisition procedure, whereby only compounds with the lowest surrogate ΔG_b values were sent to the ESMACS oracle for evaluation.

According to the results, even though the surrogate model started with a reasonable ability to rank compounds according to ΔG_b after initial training, the ability to rank newly generated compounds from iteration one onwards was lost. For batch size 250, this ranking ability recovered mostly towards the end of the GAL procedure, whereas a similar recovery was not observed for batch size 500.

Nevertheless, already in the first iteration step, the surrogate model did identify numerous good binders, according to surrogate scores, some of which were true positives. The number of these true positives, i.e. compounds with good binding affinity according to surrogate *and* oracle scores, steadily increased from thereon. The ratio of true positives over true and false positives, i.e. the precision of the surrogate model, also steadily improved, as shown in Fig. 2 and indicated by decreasing ΔG_b values predicted with ESMACS (light blue circle) that gradually approached initially overoptimistic ΔG_b values estimated by the surrogate model. Precision values that started from 0.11 and 0.07 for smaller and larger training batch sizes, respectively, decreased after the first iteration and eventually increased to values of 0.62 and 0.39, respectively, with qualitatively similar trends for both training batch sizes. A deterioration of surrogate model quality was not unsurprising given the fixed model capacity and expected covariate shift in chemical space (*cf*. diversity analysis below) which is of principal concern in AL. Subsequent recovery of prediction quality was facilitated by structural convergence of generated structures towards the end of the AL process.

In the following sections we will see that even though the ability of the surrogate model to rank compounds was very limited, the precision achieved for finding good binders was sufficient to generate new compounds with increasingly favorable binding free energies after each iteration.

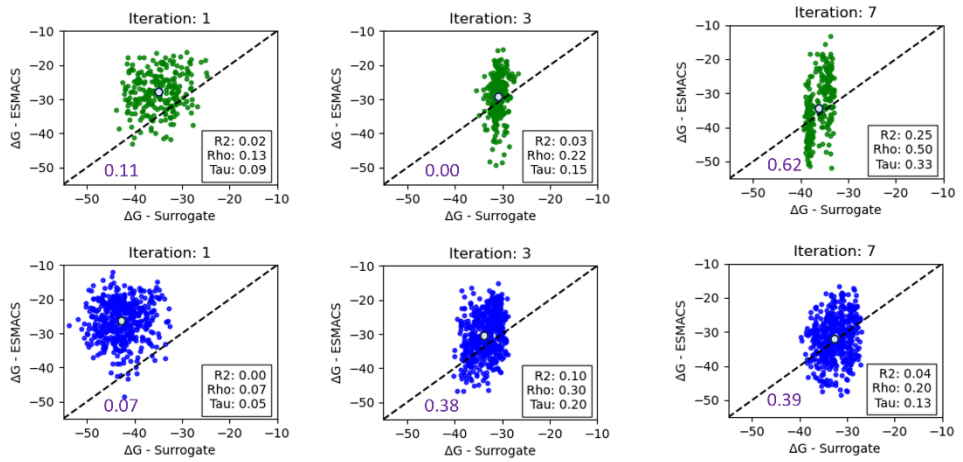

Figure 2. Comparison of surrogate model predictions of ΔG_b with calculated ESMACS values for training batch sizes of n = 250 (in blue) and n = 500 (in green) at selected GAL iteration steps for $3CL^{pro}$. R^2-coefficient as well as Spearman and Kendall rank correlation coefficients rho and tau are given in the insets of each plot. The average ΔG_b of all surrogate model predictions and ESMACS calculations within an iteration is shown as a light blue circle. All energies are given in units of kcal/mol. Results for all iteration steps are given in Figs. S2 and S3. The precision of the surrogate model is given in purple for each batch size and iteration step, where a true positive compound was defined here as a compound with ΔG_b < -35 kcal/mol according to the surrogate model and ESMACS prediction.

3.1.2 Distribution of Binding Free Energy

The binding free energy distributions derived with ESMACS, ΔG_{ESMACS}, are shown in Fig. 3. Additionally, in Fig. S5 in the SI the ΔG_{ESMACS} distribution of the 100 best binders are shown for all compounds combined after each GAL cycle.

Overall, a steady shift of energy distributions towards lower values is observed with convergent behavior towards the end of GAL. A strong enrichment of the low energy tail with more better binders, compared to the original seed structures, is evident. For small training batch size additionally an elongation of the lower energy tail towards lower values can be seen. The ΔG_{ESMACS} distributions for all acquired compounds combined in Fig. S5 also show a steady shift of the 100 best binders towards lower binding energies until the end of the GAL run.

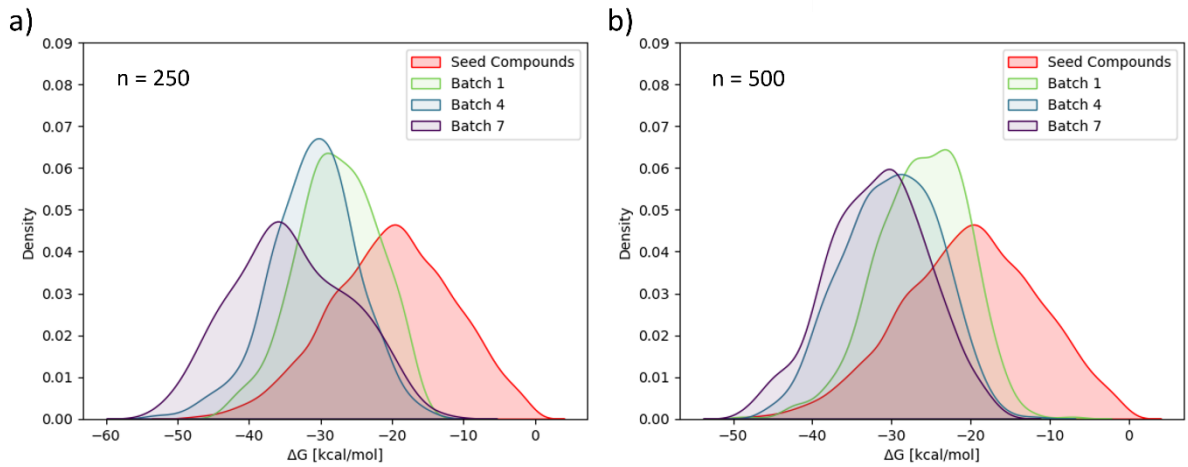

Figure 3. Distribution of calculated ΔG_{ESMACS} after a selected number of GAL iteration steps for (a) batch size 250 and for (b) batch size 500 for $3CL^{pro}$. The ΔG_{ESMACS} distribution of seed compounds used to train the initial surrogate model is shown in red. ΔG_{ESMACS} distributions for all iteration steps are shown in Fig. S4.

In Fig. 4a, the average ΔG_{ESMACS} after each GAL cycle is also shown. An almost linear decrease of ΔG_{ESMACS} was obtained even though the more relevant low energy tail of the distribution, according to Fig. 3, did not move further towards lower ΔG_{ESMACS} during the last cycles. This means that we see and enrichment of good binders rather than finding compounds with stronger binding affinities. A possible explanation for this behavior is that GAL converged to a local minimum in chemical space, where structural variation cannot lower ΔG_{ESMACS} any further. However, we also note that predicted binding free energies are below -50 kcal/mol which, in our experience, is the lower limit of ESMACS free energies. Achieving structural diversity with comparably low ΔG_{ESMACS} is very valuable for drug design as it increases the number of possible leads for further optimization.

Overall, the results clearly demonstrate that each GAL cycle produced more structures with strong binding affinities with ΔG_{ESMACS} values similarly low or even up to 5 kcal/mol lower than the best binders from the original 10,000 seed compounds. It also appears that, after about 6-7 GAL cycles, compound generation has more or less converged, i.e. only a few new compounds with relatively lower ΔG_{ESMACS} are found after each cycle

Now it might perhaps be suspected that the new structures generated are very similar to the best binders from the original seed compounds. For this reason, the structural diversity of generated compounds is analyzed in the next section.

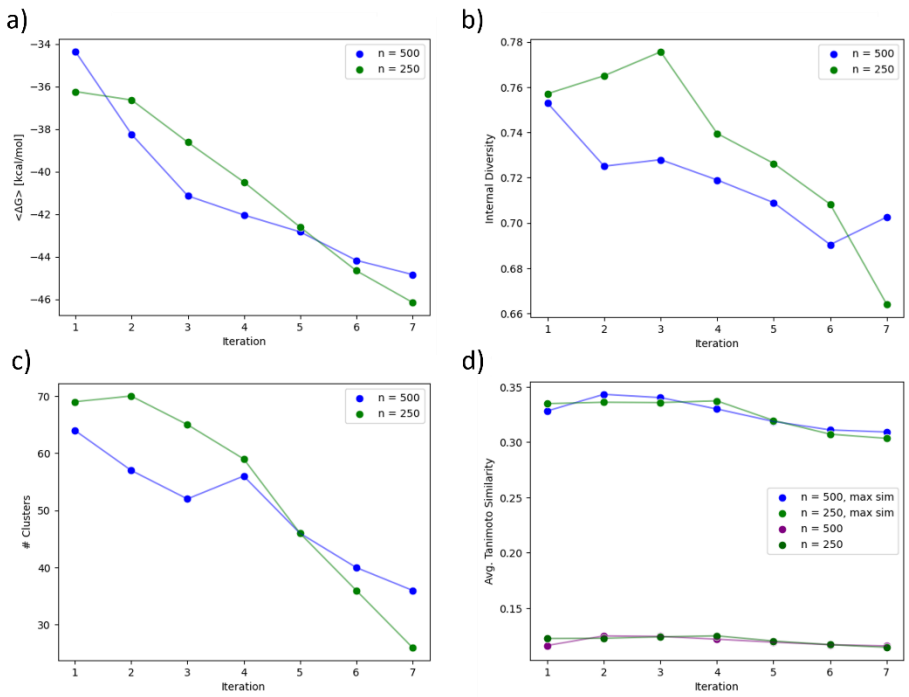

Figure 4. (a) Average $\Delta G_{ESMACS,}$ (b) internal structural compound diversity, (c) number of structure clusters and (d) average Tanimoto similarity between generated molecules and seed compounds used for initial surrogate model training for each GAL iteration for learning batch sizes 250 and 500, in green and blue, respectively, for $3CL^{pro}$. In (d) similarity was averaged over all pairs of generated molecules and seed compounds, in purple and dark green, as well as averaged over pairs of generated and their most similar seed compound, in blue and green. Only the 100 compounds with lowest ΔG_{ESMACS} were considered and taken from the accumulated pool of compounds after each iteration.

3.1.3 Structural Diversity

In Fig. 4b the average internal compound diversity of the generated 100 best binders is shown after each iteration *i*, taken from the pool of compounds accumulated throughout iterations 0 to *i*.[OBJ][OBJ]The internal diversity (Benhenda[OBJ] of a compound sample is defined a $\mathrm{ID} = 1 - < T_{ij} >$[OBJ], where T_{ij} is the Tanimoto similarity of two compounds *i* and *j*[OBJ]their two Morgan fingerprints (Morgan 1965; Rogers&Hahn[OBJ] and where the average$< T_{ij} >$[OBJ] involves all possible compound combinations *i* and *j*.

ID starts at a very high value of about 0.75, indicating that the initially generated structures are highly diverse. This value then decreases during GAL when compounds from subsequent cycles are added but a generally high value > 0.65 (i.e. smaller than an average Tanimoto similarity of 0.35) was maintained throughout. A decrease in ID during GAL is symptomatic for convergence of compound generation, where increasingly similar compounds are generated once REINVENT learns which type of compounds are good binders. With the smaller training batch size, a somewhat larger compound diversity was achieved.

To get a better understanding of the structural diversity in the structures generated, a cluster analysis was performed for the 100 best binders after each cycle using the accumulated pool of generated structures. We used the Butina clustering algorithm (Butina, 1999) with a cut-off value of 0.5 Tanimoto similarity. In Fig. 4c the total number of found clusters is shown. A strong decrease in

the number of structural clusters was observed, meaning that the top 100 binders could be grouped into fewer structural clusters with each GAL cycle. This finding is in line with the previously observed enrichment of the low energy tail of ΔG_{ESMACS} as seen in Fig. S5.

In Fig. S6 the population size of each found cluster together with its corresponding average ΔG_b value is shown after a different number of GAL cycles, again for the 100 best binders of the accumulated structures. During the GAL run, the number of structure clusters decreases, as already seen before, whereas the spread of compounds into clusters narrowed so that eventually only a few clusters were eventually significantly populated. Each cluster can be seen as a region in chemical space in which ΔG_{ESMACS} has a local minimum. The GAL process increasingly populates these regions in chemical space with more structures, while abandoning other local minima with higher ΔG_{ESMACS}.

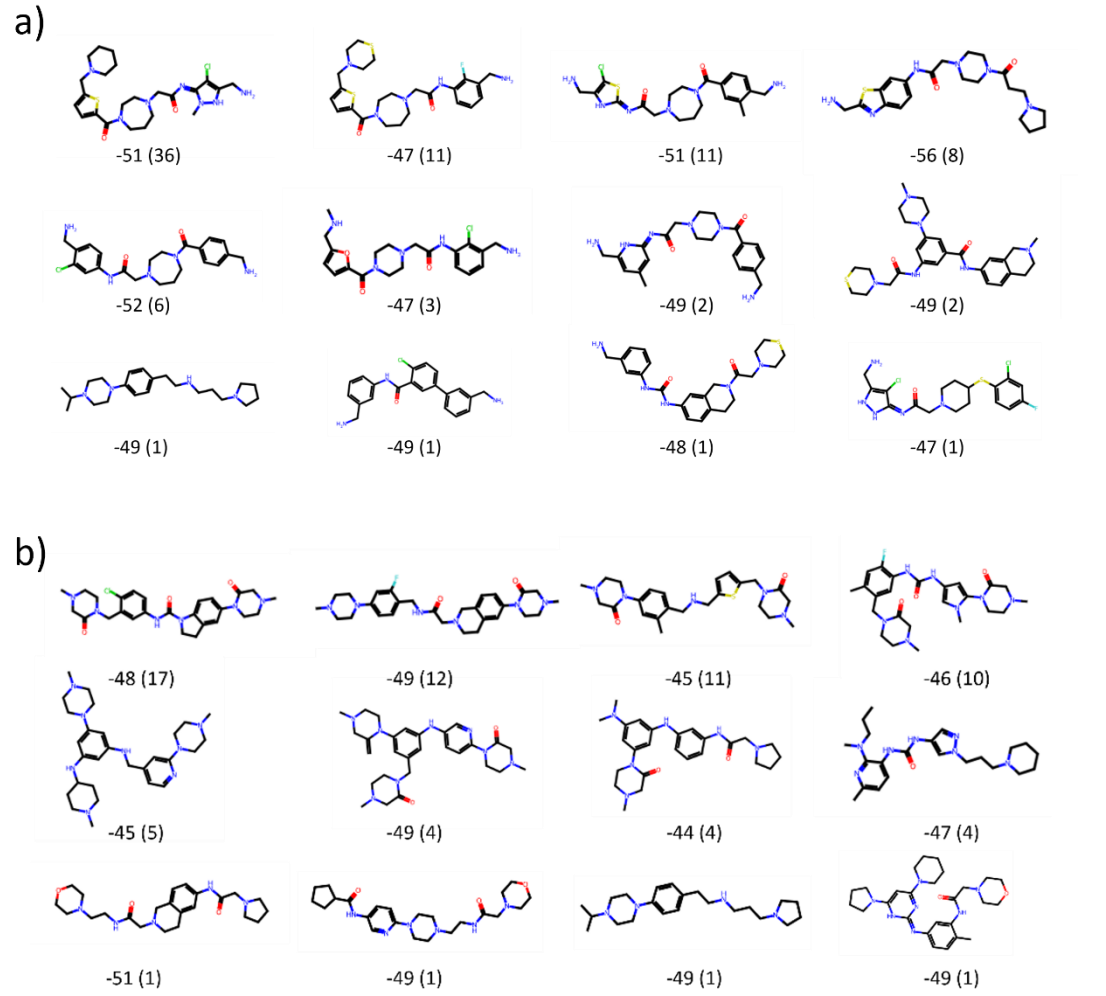

Figure 5. Representative chemical structure with lowest ΔG_{ESMAC}, for different selected structural clusters from $3CL^{pro}$. The eight most populated clusters were chosen as well as four further clusters with lowest ΔG_{ESMACS}. Cluster analysis was performed for those 100 compounds with lowest ΔG_{ESMACS} taken from the accumulated pool of compounds after each iteration for GAL training batch sizes (a) n = 250 and (b) n = 500. The energies are given in units of kcal/mol, the cluster population sizes are given in parenthesis.

To get a better understanding of the structural similarity within a cluster and between different clusters, representative chemical structures are displayed in Fig. 5. For some selected clusters the structure with lowest ΔG_{ESMACS} is displayed. Population size of each cluster is given in parenthesis as well as the maximum common substructure (MCS) for all structures in a cluster. We can see the

Butina clustering generated clusters that contain structures with varying scaffolds, even though various structural motives occur frequently. The MCSs in each cluster are of limited size, demonstrating structural diversity not only across clusters but also within them. Overall, a diverse set of good binders with different scaffolds were obtained with our GAL protocol.

The question next arises as to how similar the compounds found are to the original seed compounds that were used to train the surrogate model. Is our GAL protocol able to find genuinely new compounds or only variations of what was used to seed the process? To answer this question, the average Tanimoto similarity and average nearest neighbor Tanimoto similarity of the 100 best generated binders with the original seed compounds was evaluated and is displayed in Fig. 4d.

Very low average Tanimoto similarities of compounds generated in GAL cycles and seed compounds below 0.13 were observed. More importantly, even when considering for each generated structure only the most similar seed structure, average similarities do not exceed 0.35. This value then drops somewhat throughout the GAL run to about 0.31. This demonstrates that the best binders generated are indeed genuinely new compounds and are structurally dissimilar from the original seed compounds. Furthermore, the decreasing average similarity in Fig. 4d also indicates that, during GAL, the generated compounds drift away from the chemical space occupied by the seed compounds. To highlight this crucial aspect of GAL further, we visualize occupancy and drift of chemical space with dimensionality reduction in Fig. 6.

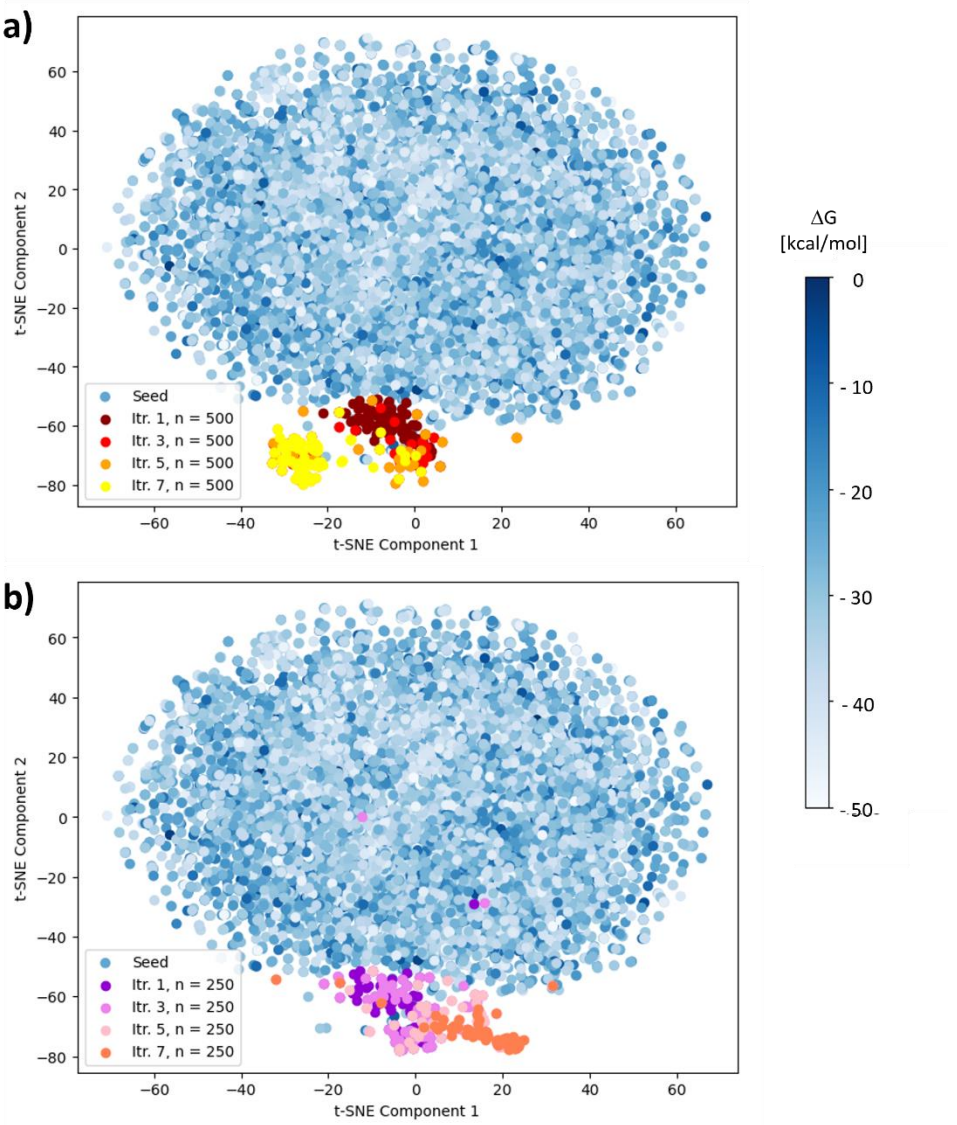

Figure 6. Morgan fingerprints of compounds projected into 2D using t-SNE, calculated for all seed compounds and generated compounds from large and small training batch sizes combined for $3CL^{pro}$. Seed compounds used for initial training of the surrogate model are shown in shades of blue according to calculated ΔG_{ESMACS}, color coded as shown in the legend on the right side. Deviations of newly generated molecules from seed compounds are shown for GAL training batch sizes (a) n = 250 and (b) n = 500 for some selected iterations, color coded as shown in the legend insets. Molecules were taken from those 100 compounds with lowest ΔG_{ESMACS} from the accumulated pool of compounds after each iteration.

We used 2D t-distributed stochastic neighbor embedding (t-SNE) (Van Der Maaten & Hinton, 2008) of the Morgan fingerprints of the 100 best generated binders as well as of the original seed compounds to visualize movement in chemical space throughout the AL process in Fig. 6. Data points in blue represent the region of chemical space occupied by the original seed compounds. We found that newly generated compounds substantially deviated from the chemical space region of the seed compounds already after the first iteration. Moreover, we also see that generated structures of good binders cover a smaller chemical space than the seed compounds, which is expected as the majority of the very diverse seed compounds also cover a wide range of binding affinities. It appears that, after about five learning cycles, GAL structures of the small and large training batch sizes moved into similar chemical space regions that are well away from the seed structure chemical space, but successively diverged into two distinct chemical space regions. In

other words, the two GAL runs converged to two different local ΔG_{ESMACS} minima in chemical space. This is not necessarily caused by the different batch sizes but more likely reflects the non-deterministic nature of the GAL protocol, which involves use of various random seeds and RL using multinomial sampling to draw samples. Most importantly, these results clearly demonstrate that the generated structures of good binders are genuinely new compounds and not just mere variations of some of the seed compounds.

Overall, we demonstrated that the AL process found genuinely new and diverse compounds with varying scaffolds with strong binding affinities exceeding those of the best binders of the original seed compounds. Moreover, after initial training of the surrogate model with 10,000 compounds, this was achieved after only 7 iteration steps with a total of 3500 and 1750 oracle calls, for large and small training batch sizes, respectively.

3.1.4 Diversity of Ligand Binding Modes

In Fig. 7 we investigated the binding modes of four selected ligands as examples for binding to the $3CL^{pro}$ target protein. These four ligands were generated during the GAL run and found to exhibit strong binding affinities with -56 kcal/mol < ΔG_b < -49 kcal/mol as well as low chemical similarity among each other.

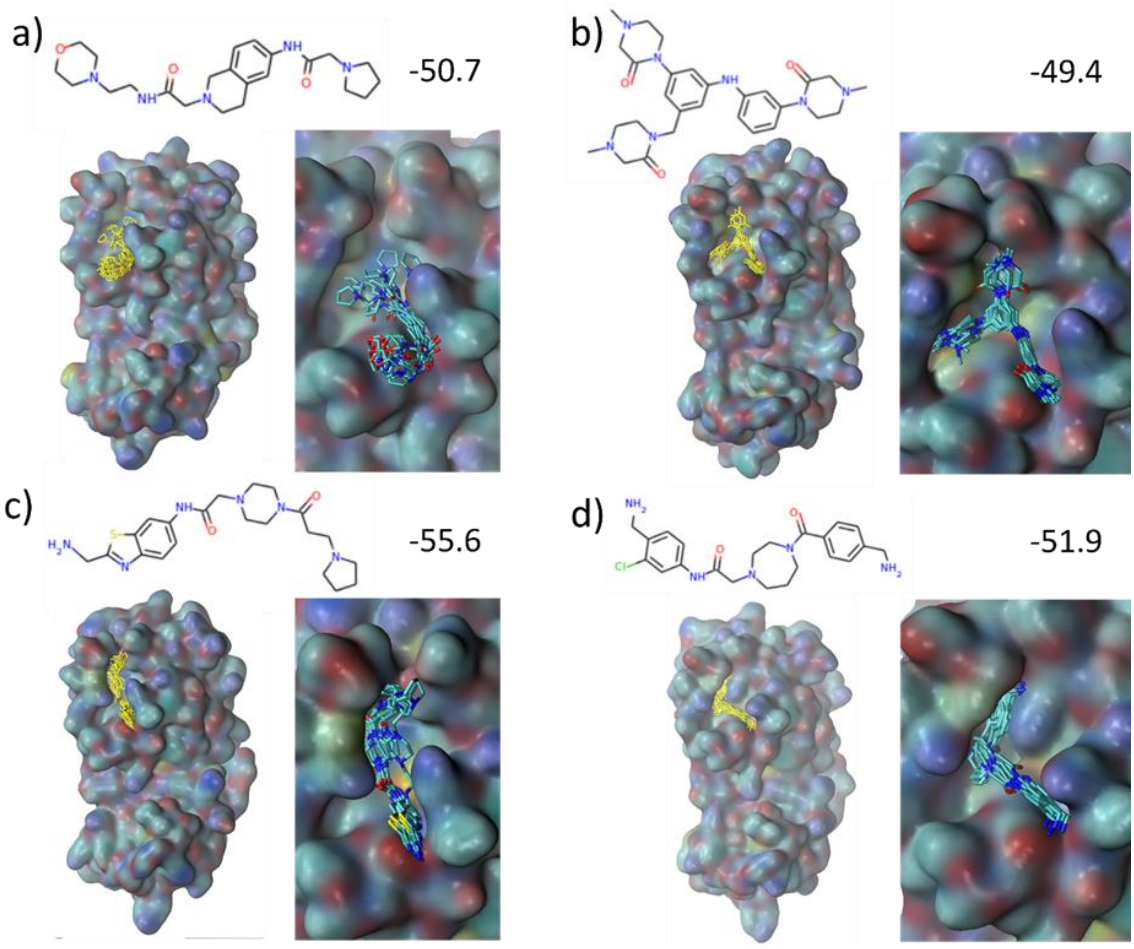

Figure 7. (a-d) Four selected ligands with predicted high target binding affinity in the $3CL^{pro}$ binding pocket. Chemical structures together with their binding free energy ΔG_{ESMACS} (in units of kcal/mol) are given on top of the figures. Shown ligand structures, highlighted in yellow on the left-hand sides, are superpositions of snapshots taken from ten ESMACS replicas after 4 ns of MD simulation. Protein surfaces are shown with coloring according to atom types of the surface atoms. The ligands together with their protein vicinity are shown in more detail on the right-hand side for each ligand.

We found different binding modes for these four ligands as demonstrated by different contact areas between ligand and protein and hence substantial differences in ligand – protein interactions. A closer inspection of these complexes showed the importance of hydrogen bonds between ligand amines, several of which are found in all generated binders with strong binding affinity, and the protein, whereas ligand carbonyl groups seem much less engaged in hydrogen bonding. Amines formed hydrogen bonds inconsistently with different protein residues when comparing across ligands. Only hydrogen bonds with anionic Glu166 and Thr25 were obtained consistently. Moreover, we found that ligands interact with different parts of the active site, where the additional ring group of the only non-linear compound shown in Fig. 7c provides specific binding to an active site region that cannot be reached by the linear compounds.

A more comprehensive analysis of binding modes is beyond the scope of the current study. However, overall, observed binding modes exhibit additional diversity that could potentially be exploited in a successive study by combining moieties found to interact with different sections of the binding pocket.

3.1.5 Computational Efficiency

Comparing the quality of GAL compounds generated with different training batch sizes, as discussed in this work, the question arises which batch size is the computationally most efficient. The computationally most demanding step by far in the GAL workflow is the derivation of ESMACS ΔG_b values for each oracle call. While larger batch sizes require more computational effort, fewer iterations are needed to find good binders as discussed before. This behavior could favor larger batch sizes where many compute nodes are available to run in parallel; however, when that is not the case, smaller batch sizes might be more efficient.

To investigate this further, we define computational efficiency η as the number of suitable structures found per made oracle call:

$$\eta = \frac{N_{\mathrm{CG},\ \Delta G\mathrm{max}}}{n_{\mathrm{oracle}}}$$

The number of oracle calls n_{oracle} is then simply the training batch size multiplied by the number of iterations. Defining when a generated compound is considered as suitable is less straightforward. Generally, a good compound should exhibit a large binding affinity and it should be sufficiently different in structure from other compounds generated so far. We defined compounds as suitable when their binding free energy $\Delta G_b < \Delta G_{max}$, where different values for the threshold ΔG_{max} were considered. Additionally, a cluster analysis was performed, using the Butina algorithm, for those structures with $\Delta G_b < \Delta G_{max}$, and counted the number of found clusters as the number of compound groups sufficiently different form each other, N_{CG}, ΔG_{max}. The similarity cut-off of the clustering algorithm, s_{cutoff}, is the similarity required for two structures to belong to the same cluster. A larger

value of s_{cutoff} therefore means that only very similar compounds are grouped into the same cluster and means that larger and fewer clusters are generated, while individual clusters cover a larger area in chemical space.

The situation where a small value of s_{cutoff} is used, i.e. where compounds are grouped together into more clusters with fewer cluster members, together with a more demanding ΔG_{max} = -40 kcal/mol corresponds to a situation where compounds are sought after that exhibit a high binding affinity but where structural diversity is less important as essentially all newly generated compounds are counted towards η, regardless of how similar they are to already generated structures. This is a typical situation akin to lead optimization in drug discovery. In contrast, in case of a more relaxed ΔG_{max} = -35 kcal/mol and larger s_{cutoff}, structural diversity and coverage of chemical space is more emphasized than binding affinity, which is typical for a more explorative hit finding scenario. Therefore, through selection of ΔG_{max} and s_{cutoff} we can evaluate GAL effectiveness for these different situations.

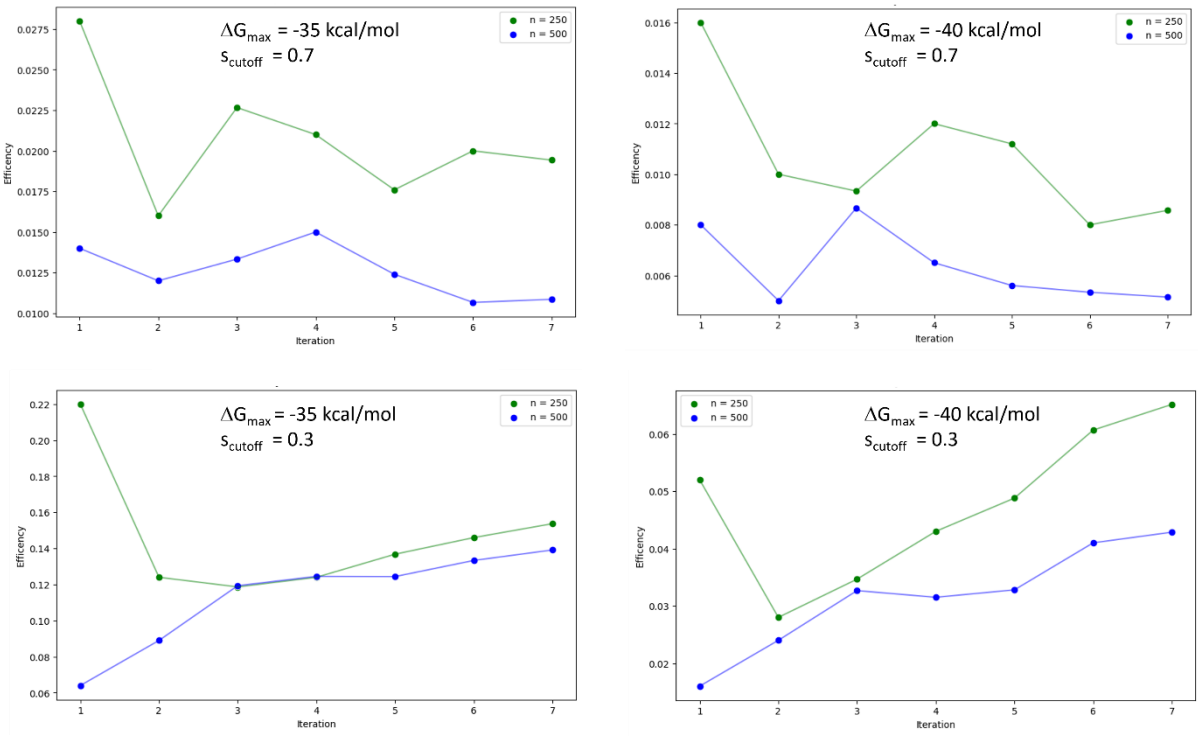

Figure 8. Efficiency of GAL process for the $3CL^{pro}$ systems using two different training batch sizes after each iteration step. Efficiency is defined as the number of structural clusters found per oracle call of all cumulatively generated structures at a given iteration step. Only ligands with $\Delta G_b < \Delta G_{max}$ were considered. Clustering was carried out using the Butina algorithm with a similarity cut-off given in respective plots.

The results for η are shown in Fig. 8 for the two different training batch sizes and two different values for ΔG_{max} and s_{cutoff} throughout the GAL process for $3CL^{Pro}$. ΔG_{max} = -35 kcal/mol corresponds approximately to the maximum of the ΔG distribution obtained after the last GAL iteration (see Fig. 3), whereas ΔG_{max} = -40 kcal/mol represents a stricter criterion to define suitable structures. In all cases, the smaller training batch size of n = 250 was found to be computationally more efficient. For s_{cutoff} = 0.3 improvements in efficiency were observed with increasing iteration steps, whereas for s_{cutoff} = 0.7 efficiency remained more or less the same throughout the GAL run.

3.2 Tankyrase-2 Target

3.2.1 Generation of First Compound Batch

Starting point of GAL for the tankyrase-2 (TNKS2) target was a congeneric series of 27 ligands with experimentally confirmed favorable binding free energies with a range from -35 to -26 kcal/mol as computed with ESMACS. Binding affinities derived with ESMACS were also compared to measured affinities, as shown in Fig. S7, where a reasonable ranking of compounds was obtained through simulations. We then proceeded to generate the first training batch of compounds with the method described in section 2.1. In the following, we will refer to the GAL step using the initially generated compounds as iteration zero. Five GAL steps were carried out with variations only in the training batch size, using sizes n = 100, 300, 500, 700 and 1000 compounds per iteration. Overall, we found that only four iterations were sufficient for results to converge, as shown below.

3.2.2 Quality of surrogate model

In Fig. 9 binding affinity predictions of the ChemProp surrogate model are compared with ΔG derived with ESMACS for a selection of batch sizes and iteration steps (Fig S8 contains this data for the entire dataset). First, we obtained a substantially improved quality of the surrogate model predictions compared to $3CL^{pro}$. Large values for Spearman ranking and R^2 coefficients above 0.7 and 0.6, respectively, were obtained. A comparison across batch sizes shows that reliable surrogate models were found after iteration 3 for all sizes, except to a lesser extent for the smallest training batch size of n = 100. We also observe that for n = 500 and larger, good surrogate models were found already after the first iteration step. For GAL it means that far more true positives, i.e. compounds with favorable binding affinities according to ESMACS, were found and a higher precision was achieved. How this affected GAL overall will be described in the following. Possible reasons behind the better surrogate model quality as compared to $3CL^{pro}$ are discussed in section 3.2.5.

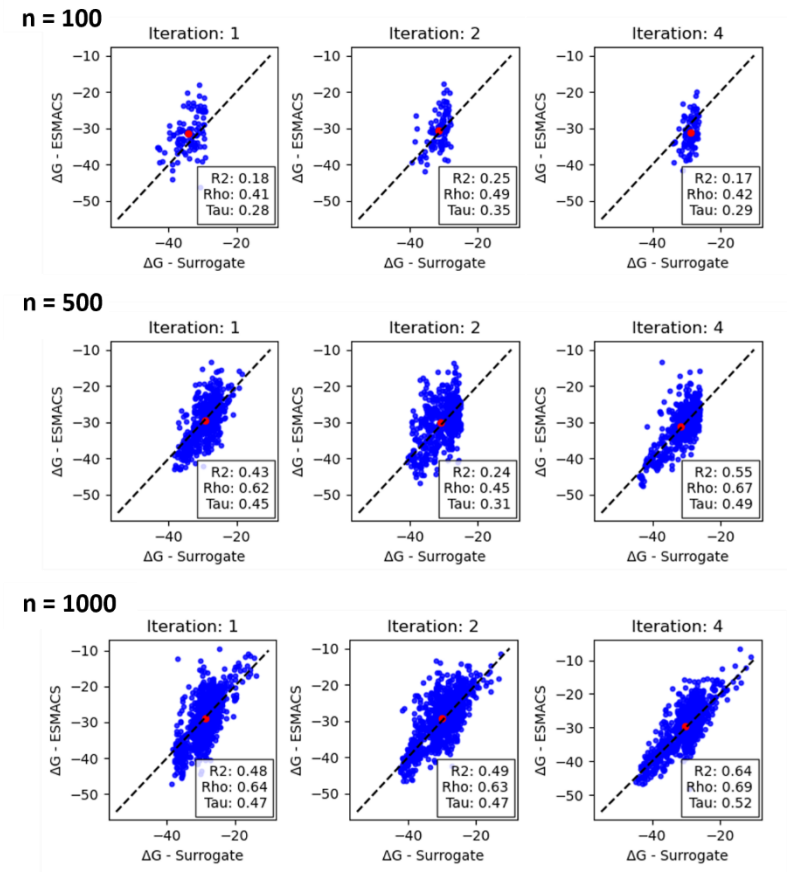

Figure 9. Comparison of surrogate model predictions of ΔG_b with calculated ESMACS values for training batch sizes between 100, 500 and 1000 molecules for selected GAL iteration steps for TNKS2. R^2-coefficient as well as Spearman and Kendall rank correlation coefficients rho and tau are given in the insets of each plot. The average ΔG_b of all surrogate model predictions and ESMACS calculations within an iteration is shown as a red circle. All energies are given in units of kcal/mol. Results for all training batch sizes and iteration steps are shown in Fig. S8.

3.2.3 ΔG_b – Distribution

In Fig 10 distributions of obtained ΔG_b values for selected iteration steps are shown (full dataset included in Fig S9). For comparison, the measured values of the congeneric series of the original 27 ligands are also shown together with the ΔG_b distribution of the compounds from iteration zero. We found that, already after the first iteration, substantial improvements were achieved compared to the initial iteration zero compound sample and that the high affinity tail of the distribution even extended beyond the binding energies of the experimentally confirmed ligands, which are already characterized by strong binding affinities. Moreover, we found that for n = 500 and larger, iterations after step 1 barely improved the energy distribution further, whereas for the smaller batch size of n = 300 improvements were still achieved with additional iterations. For n = 100 not much improvement was obtained with more iterations and the high affinity tail was less extended towards lower energies than it was the case for the other batch sizes.

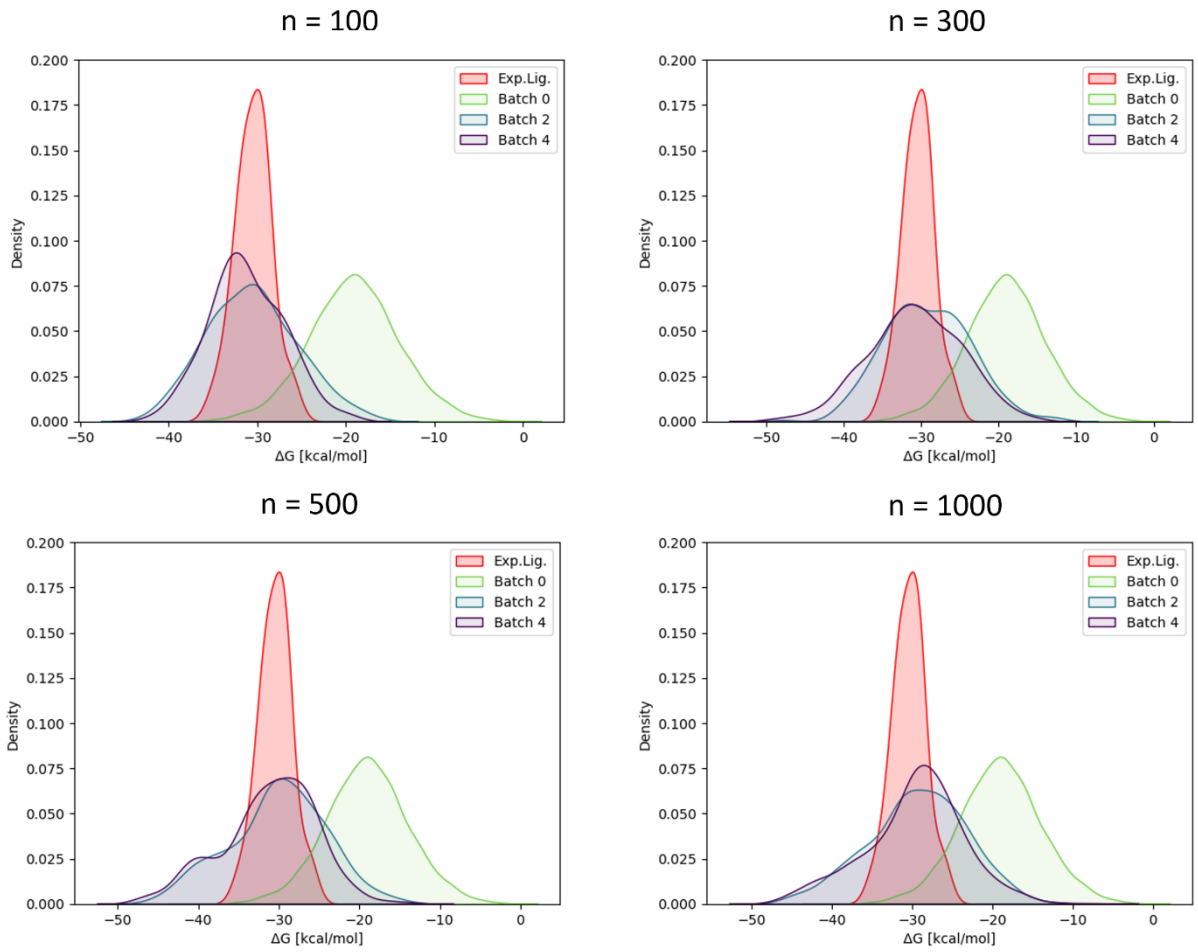

Figure 10. Distribution of calculated ΔG_{ESMACS} for selected GAL iteration using different batch sizes for TNKS2. The ΔG_{ESMACS} distribution of 10k seed compounds used to train the initial surrogate model is shown in green as batch 0. The ΔG_b distribution of 27 measured compounds is shown for comparison in red. Results for all batch sizes and iteration steps are shown in Fig. S9.

Overall, convergence of ΔG_b distributions was basically achieved rapidly within a single iteration (except for n = 300), compared to five or more iterations for $3CL^{pro}$. In Fig 11a changes in the average ΔG_b throughout the GAL run are shown for the best 100 binders of the cumulatively generated compounds: there is a trend to lower values of $<\Delta G_b>$ for larger batch sizes and a steady decrease with an increasing number of iteration steps. Here, larger batch sizes enable the development of more precise surrogate models as shown before, facilitating moderate improvements in $<\Delta G_b>$ that, however, come with the computational cost of more oracle calls. The computational efficiency of GAL is discussed in the final sub-section (3.2.6).

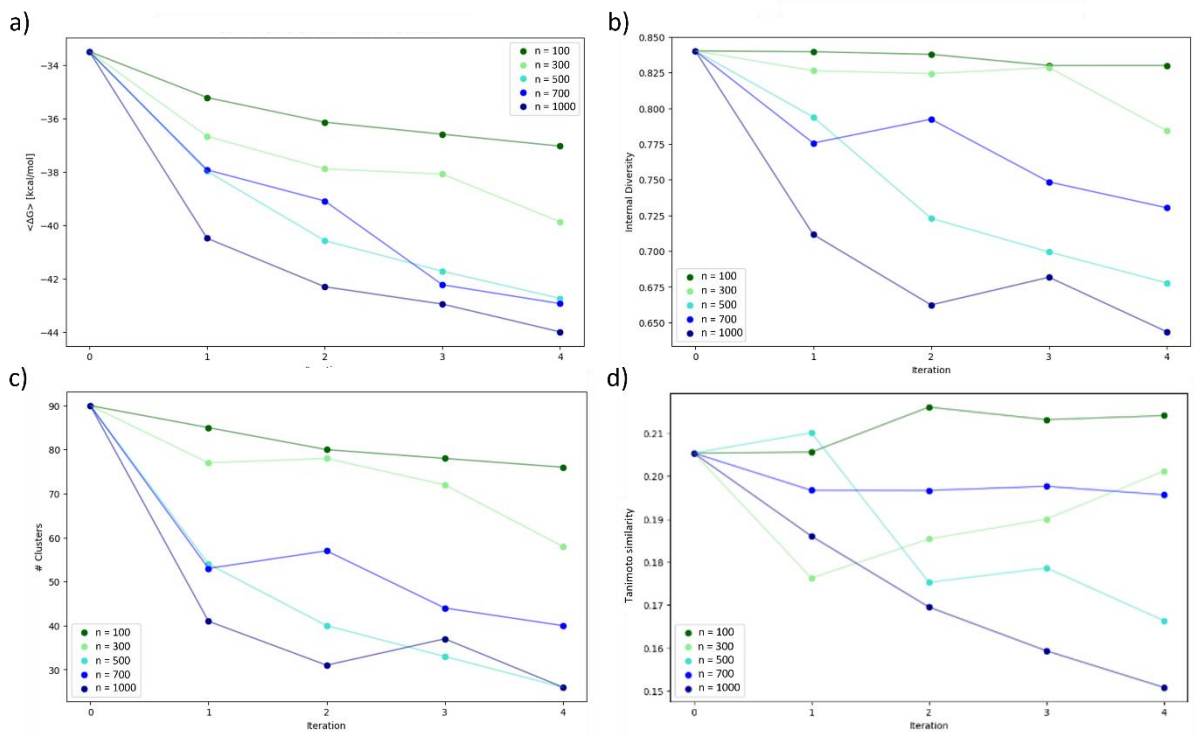

Figure 11. (a) Average $\Delta G_{ESMACS,}$ (b) internal structural compound diversity, (c) number of structure clusters and (d) average Tanimoto similarities of pairs of generated molecules and their most similar compound from the initial 10,000 compounds from iteration zero for each GAL iteration for different learning batch sizes used for TNKS2. Only the 100 compounds with lowest ΔG_{ESMACS} were considered and taken from the accumulated pool of compounds after each iteration.

We also explored if structural diversity could be improved further by augmenting the surrogate model with known TNKS2 IC_{50} - values from BindingDB (Liu et al., 2007). Structures having the same core as the 27 benchmark compounds and IC_{50} >= 1 µm were removed leaving 484 ligands. From these we randomly chose 20 compounds and added to those 80 compounds from iteration 3 to resume GAL. Resulting ΔG_b distributions are shown in Fig. S11. We found that, following minor improvements of the ΔG_b distribution after iteration step 2, no improvements were obtained after 5 iterations compared to GAL without any BindingDB infusions.

The quality of the surrogate model trained with the BindingDB infused compound samples is shown in Fig. S12. The results are comparable to the previously trained surrogate models shown in Fig. 9. Internal diversity values between 0.84 and 0.85 were obtained, which were essentially the same as what was previously observed for GAL without infusions, see Fig. 11b. Overall, it appears that the infusion of additional structures from BindingDB did not improve the quality of GAL results.

3.2.4 Structural diversity

Internal structural diversity of generated structures is shown in Fig. 11b. Generally, diversity of structures tends to be even larger than for $3CL^{Pro}$. Diversity decreased from iteration to iteration, except for training batch size n = 100, suggesting convergence in chemical space. Less diversity was obtained in larger training batch sizes, where it can be assumed that a more exhaustive sampling of chemical space near minima of ΔG_b led to more similar structures.

In Fig 11c, the number of structure clusters according to the Butina algorithm with a cut-off of 0.5 in Tanimoto similarity is shown for the 100 best binders from the cumulatively generated structures. Also, in Fig. S10 the average ΔG for each structural cluster is shown together with its population size. As for $3CL^{Pro}$ we see that the number of clusters substantially decreased during the GAL process as clusters with structures associated with low ΔG_b energy minima were increasingly populated at the expense of clusters related to higher ΔG_b minima. Also, GAL performed with larger batch sizes led to a smaller number of clusters, which once again is related to a more pronounced convergence in chemical space.

Some structures of generated compounds are exhibited in Fig. 12; this demonstrates that generated compounds with favorable binding affinities involve a large variety of different scaffolds. This is particularly noteworthy considering that the initial sample of generated compounds was based solely on a small congeneric series sharing the same quinazolinone scaffold. However, we also find quinazolinone and similar scaffolds (see Fig. 12) frequently in the predicted binders. When directly compared to the congeneric series, GAL generated structures indeed manifest low maximum Tanimoto similarity < 0.22 to any of those 27 ligands, as shown in Fig. 11d.

Moreover, we see marked differences between the structures obtained with training size n = 100, compared to n = 1000. As shown above, larger training sizes led to structures with lower ΔG_b and as we will see also to more progression towards convergence in chemical space, which led to inclusion of functional groups for n = 1000 such as nitriles and bridged cycles. In section 3.2.5 we will show how these groups contribute to a minimum in ΔG_b. Apart from improvements of ΔG_b we also observed somewhat higher QED scores, i.e. more drug-likeness, for larger batch sizes as shown in Fig. S14, despite seeing in Fig. 12b also some more uncommon compound moieties that are not found in typical drugs.

At this point, it is important to note that many generated compounds with favorable ESMACS binding affinity ΔG_b and some drug-likeness as per QED would otherwise unlikely be chosen for progression in a drug discovery project, as they may be too difficult to synthesize or lack other properties. Improvement on those issues is beyond the scope of the present work.

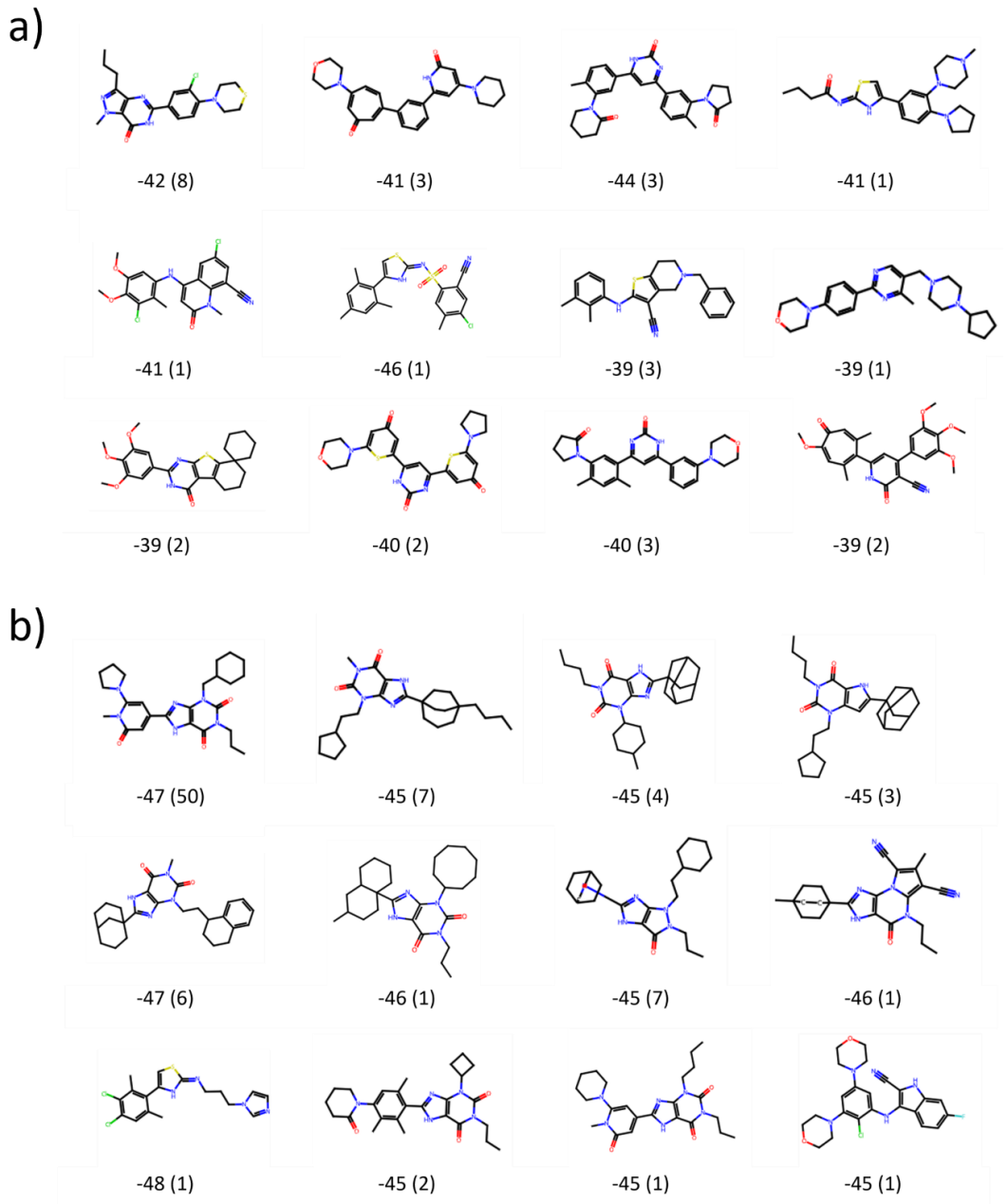

Figure 12. Chemical structure representatives with lowest ΔG_{ESMAC} for different selected structural clusters from TNKS2. The eight most populated clusters were chosen as well as four further clusters with lowest ΔG_{ESMACS}. Cluster analysis was performed for those 100 compounds with lowest ΔG_{ESMACS} taken from the accumulated pool of compounds after each iteration for GAL training batch sizes (a) n = 100 and (b) n = 1000. Cluster population sizes are given in parenthesis. Energies are given in units of kcal/mol.

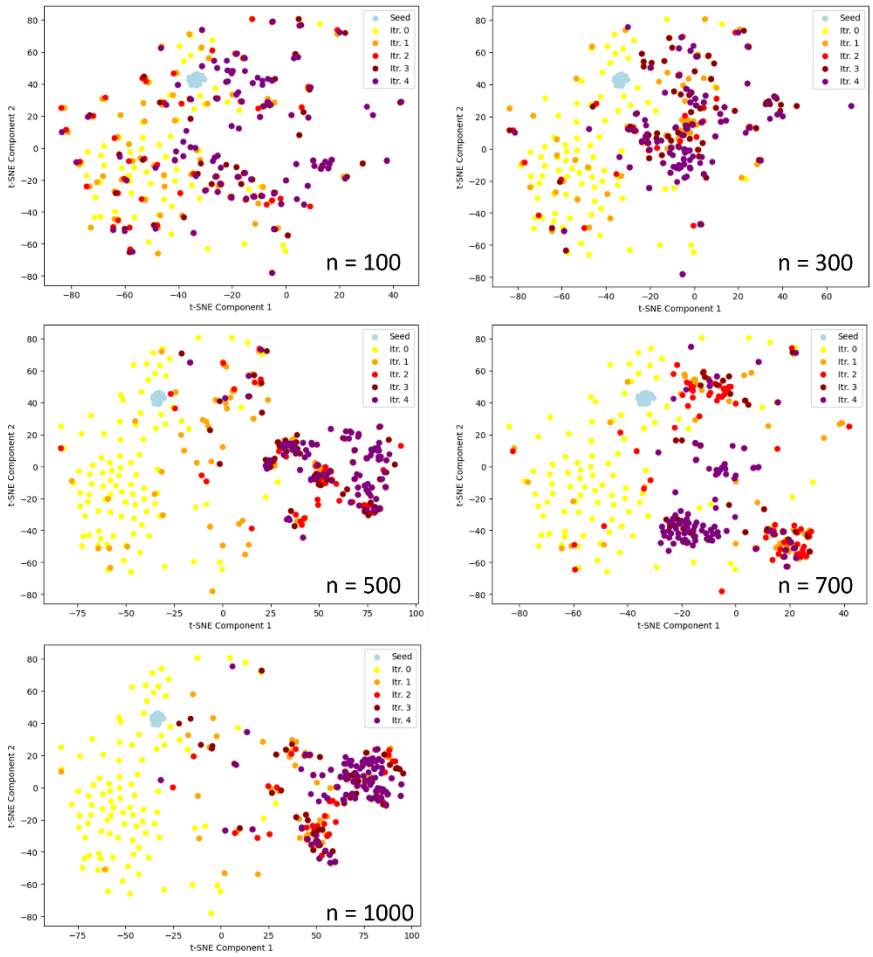

Figure 13. Morgan fingerprints for TNKS2 compounds projected into 2D space using t-SNE computed on the combined dataset from all batches shown here to enable comparison of chemical space. Compounds from iteration 0, shown in yellow, refer to molecules taken from the 10,000 initial compounds. The 27 ligands with measured binding free energies are also included and are shown in light blue. Generated molecules are color coded as per the legend insets. Shown molecules were taken from those 100 compounds with lowest ΔG_{ESMACS} from the accumulated pool of compounds after each iteration.

In Fig. 13 2D t-SNE visualizes the chemical space traversed throughout GAL for different training batch sizes. We observe that the original 27 ligands, shown as light-blue data points, are focused within a very small area, whereas the 10k generated structures, shown in yellow, cover a much larger chemical space. Throughout the GAL process, convergence into chemical spaces distinct from the space covered by the structures from iteration 0 is observed, which is more pronounced for GAL with larger compound training sizes. For n = 100 and 300 it appears that convergence did not yet occur. Furthermore, we observe that convergence was achieved for n = 500 and n = 1000 to a similar space region, whereas for n = 700 GAL converged to a somewhat different region.

Overall, the findings reported in this sub-section indicate that GAL can generate structures with high internal diversity, well distinct from the original 27 ligands. The results also suggest that for training sizes of n >= 500, according to results shown in Fig. 13 and S10, the overall 100 best binding compounds generated after the last GAL iteration could be grouped into fewer structural clusters with lower ΔG_b, i.e. the best generated binding compounds occupied narrower regions in chemical space. It also shows that GAL with smaller batch sizes is likely to benefit from more iteration steps for structural convergence and ΔG_b optimization.

3.2.5 Diversity of Binding Modes

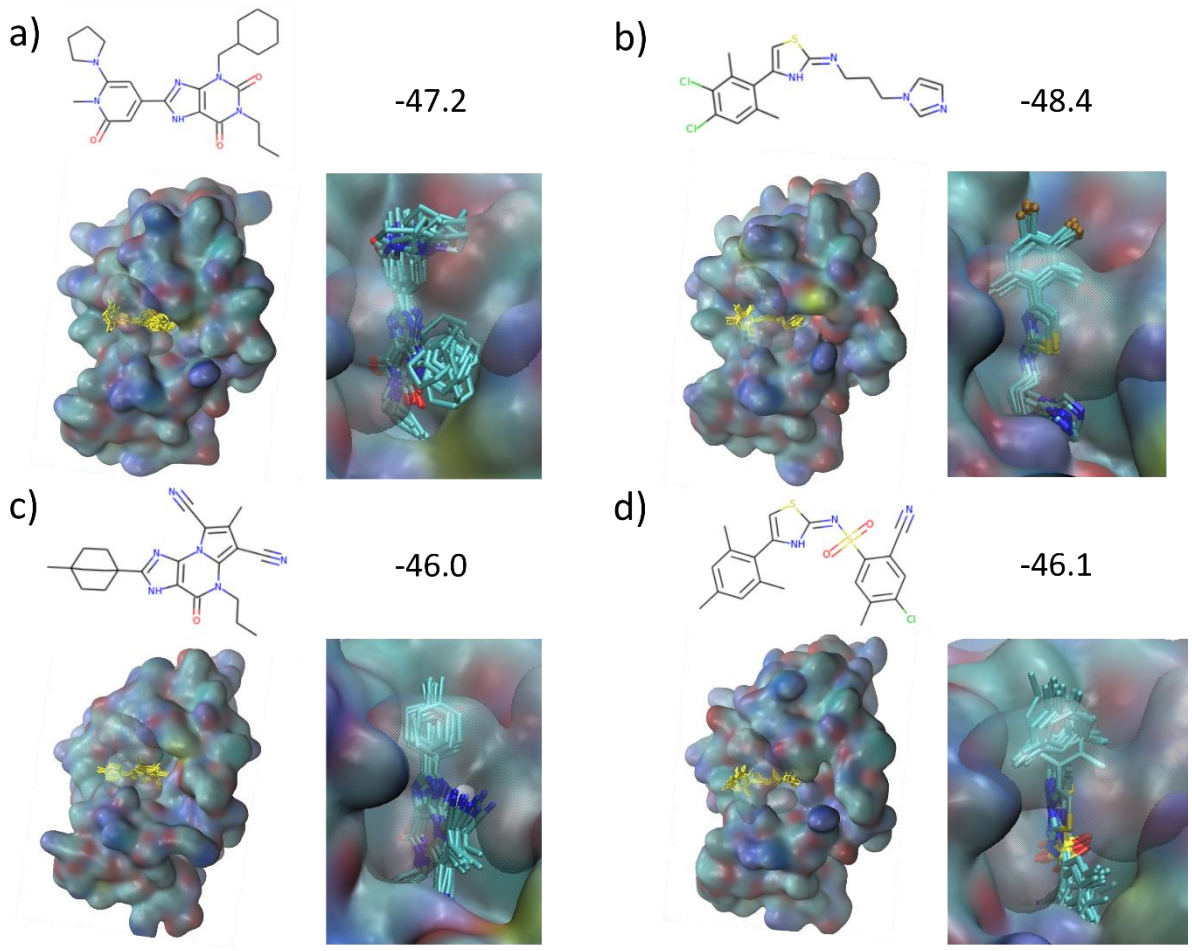

Figure 14. (a-d) Four selected ligands with predicted high binding affinity in the TNKS2 binding pocket. Chemical structures together with their binding free energy ΔG_{ESMACS} (in units of kcal/mol) are given above the images. The ligand structures shown, highlighted in yellow at the left-hand side of each image, are superpositions of snapshots taken from the ten ESMACS replicas per ensemble after 4 ns of MD simulation. Protein surfaces are shown with coloring according to atom type of the surface atoms. The ligands together with their protein vicinity are shown in more detail on the right-hand side for each ligand.

Four representative, low ΔG_b structures bound to TNKS2 are shown in Fig. 14. Compared to $3CL^{Pro}$ a qualitative difference is apparent: in $3CL^{Pro}$ a variety of different binding modes was found due to the large open binding pocket of the target, whereas in the case of TNKS2 the binding pocket is narrow and closed, which led to a well-defined binding mode and far less conformational variation of ligands in the binding pocket, despite large structural differences across ligands. A closer inspection of the ligand – protein interface reveals that some of the moieties improving ΔG_b were frequently found within the generated structures. For instance, nitriles were included into ligands to optimize the distance between the ligand and the protein surface through the triple bond. Bridged bicyclic moieties turned out to act as bespoke “plugs” that close the opening of the active site on one side, thereby maximizing van der Waals interactions with the protein, which would have been otherwise difficult to achieve with flat rings or other groups. These are instructive examples showing how REINVENT optimized the design of compounds in GAL to fit the active site of a protein target.

It is reasonable to assume that the confined space in the active site of TNKS2 and the correspondingly well-defined binding pocket structure permit a less ambiguous correspondence of the ligand structure, represented in REINVENT as a 1D SMILES code, and the 3D docking pose of the ligand inside that active site and thereby binding free energy. This could have facilitated the development of a more predictive surrogate model for GAL than in the case of $3CL^{Pro}$ (given that our surrogate model only considers ligand structures but not those of the protein, thereby does not capture protein-ligand interactions explicitly), in turn enabling the discovery of new structures with good binding affinities within a smaller number of GAL steps. This indicates that performance of the GAL process, even though successful for both selected test targets, can be dependent on the extent of ligand confinement in the active site of the target.

3.2.6. Computational Efficiency

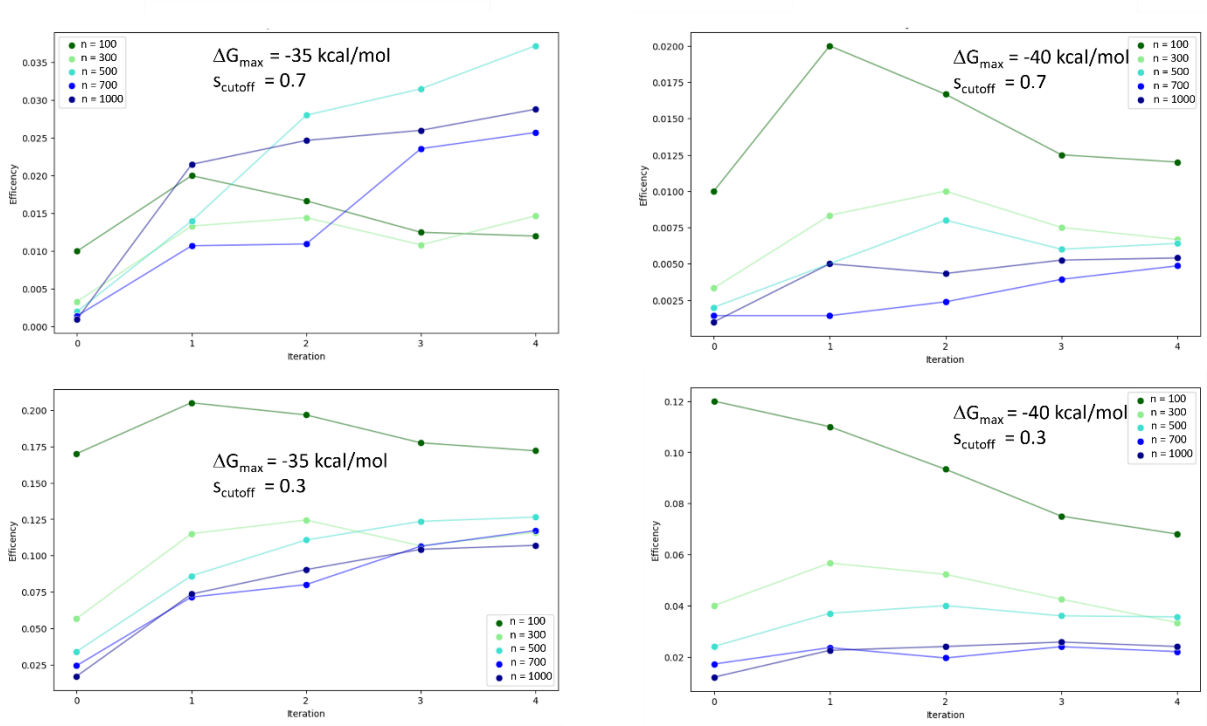

Figure 15. Efficiency of GAL for the TNKS2 systems using different training batch sizes, color coded as specified in the plot legends, after each iteration step. Efficiency is defined as the number of structural clusters found per oracle call of all cumulatively generated structures at a given iteration step. Only ligands with $\Delta G_b < \Delta G_{max}$ were considered. Clustering was carried out using the Butina algorithm with a Tanimoto similarity cut-off given in respective plots.

The results for the computational efficiency η of GAL for TNKS2 are shown in Fig. 15 for different training batch sizes, and two different values for ΔG_{max} and s_{cutoff} throughout GAL for TNKS2. A value of ΔG_{max} = -35 kcal/mol describes approximately the maximum of the ΔG_b distributions obtained after GAL, whereas a value of –40 kcal/mol was chosen as a stricter criterion for selecting suitable compounds.

According to these results we observe improving efficiency within increasing iteration steps for larger batch sizes of n = 500 and larger, whereas for n = 100 efficiency remained the same or decreased. This finding is in line with the quality of the surrogate models, as shown in Fig. 9. The more precise surrogate models for n ≥ 500 are more capable in identifying suitable new compounds,

thereby increasing efficiency during GAL, however, at the cost of more numerous oracle calls. It turns out that smaller batch sizes were more efficient in most cases, especially the smallest batch size n = 100. Only in the case of less stringent ΔG_{max} = -35 kcal/mol and s_{cutoff} = 0.7 was a batch size of n = 500 found to be most efficient, followed by n = 1000 and n = 700. As explained in subsection 3.1.5, these parameters define a η that is more suitable to describe a situation that prioritizes chemical space exploration over finding particularly high binding affinities.

These findings indicate that a more explorative GAL would benefit from larger training batch sizes, whereas exploitation would perform more efficiently with a small batch size. This finding is somewhat different from the situation in $3CL^{Pro}$ where the smaller training batch size always led to a more efficient GAL. The likely cause for this difference is that for $3CL^{Pro}$ the precision of the surrogate model remained limited in all cases compared to that for the TNKS2 case. It appears that, for a large training batch size to improve GAL efficiency, a sufficiently precise surrogate model is required. Overall, our findings indicate that in most scenarios a small batch size is more efficient and a safer choice when the quality of the surrogate model is *a priori* unknown.

4. Discussion

We have shown in this article that AL can be an effective method to improve compound optimization for a particular target. It is important to point out how the term “active learning” is being used as we follow here the convention which has been adopted in at least part of the community (Crivelli-Decker et al., 2023; Gorantla et al., 2024; Gusev et al., 2023; Khalak et al., 2022; Knight et al., 2021; Konze et al., 2019; Mohr et al., 2022; Thompson et al., 2022). However, we have also pointed out in the Introduction that we are not principally interested in creating a surrogate model for the purpose of optimally finding new labels i.e., binding affinities. In fact, in our application of AL the models (the surrogate model and the RL agent which is effectively a secondary surrogate model) are merely artefacts and the primary focus is in producing a sufficient number of good binders as estimated by ESMACS. We are not truly trying to find the extremum of a function in the strict meaning of “optimization”. Figure S1 shows when good binders are found over the course of an AL run, and we see that good binders can be found very early in any RL cycle and also in any AL step. This is important in budget-constraint situations where only a limited amount of time or money is available. As we show, GAL is perfectly capable of finding quality compounds within such regimes.

Drug discovery today is still very much a “numbers-game”, meaning that a sufficient number of suggestions need to be generated (ideation) which are then separately assessed by medicinal chemists. We also stress that we have here used only a very limited number of scoring components which are nowhere near sufficient to assess the quality of a compound as a viable drug candidate. The current setup leaves out developability of a molecule i.e., synthesisability, PK/PD, ADMET, IP, formulation, physico-chemical properties, and so on. In this sense, the current work is a proof-of-concept to show-case GAL as an effective and resource-saving tool for molecule optimization. Therefore, in future it will be necessary to strive for a more realistic target profile. The scoring components employed in this work are binding free energy prediction (the main oracle) and two scoring components (QED, chemical alerts) which are rather minimalistic models to restrain the chemistry generated. We have completely left out e.g. synthesizability which, however, is a major design goal in practical molecular design for obvious reasons. Thus, unsurprisingly, the chemical

space explored with GAL cannot principally be expected to resemble the one from actually synthesized and experimentally confirmed binders. For the same reason, generated good binders may not exhibit much similarity with previously reported binders which, however, may be beneficial to inform new molecule designs based on the newly found chemical space. In future we plan to use in-house solutions like AiZynthFinder (Genheden et al., 2020) and others to address synthesizability and biology.

We also demonstrate the need for combining a ligand-based generative model with physics-based methods. It has been shown that structure-based methods significantly improve outcomes of generative models (Papadopoulos et al., 2021; Sauer et al., 2022; Thomas et al., 2023) and it has been shown how physics-based MD descriptors are needed especially in low data regimes (Chew et al., 2024). Docking itself typically correlates rather poorly with experiment or with more accurate MD-based methods and we also observe this in the present study (see Figure S13). Hence, docking is often employed for enrichment and pre-filtering as scoring functions tend to be rather inaccurate (Li et al., 2019; Xu et al., 2018). Here we use the physics-based method ESMACS to inform an RL agent via a surrogate model and demonstrate how this can drive our GAL workflow to efficiently generate better binders.

Our work also shows that there are many opportunities to refine and improve the GAL workflow. We have built a fixed, static surrogate model *via* a deep learning model with set hyper-parameters and geometry. In principle, we would have to search all of model space to find models which are compatible with predicting binding free energies in accord with physics-based binding affinity prediction. We note here that this model is entirely ligand based and does not take protein structure or ligand-protein interactions into account. Earlier attempts to incorporate some structure-based information into AL has only shown limited utility (Khalak et al., 2022). Previous RBFE studies have made use of automatic QSAR model building (de Oliveira et al., 2023; Gusev et al., 2023) and we could make use of our own Qptuna software (Mervin et al., 2024). However, building a model with ChemProp is a rather time-consuming undertaking especially when making use of multi-fold cross-validation and ensembles. Depending on data and data size, the search for the best model may require the change of the model building algorithm during AL, e.g. we have used in this study a random tree model for the relatively small data set size of 27 ligands in the case of TNKS2 while for the actual AL runs we opted to seed with a rather large number of compounds where neural network models are much better suited (Heid et al., 2024). We have also created the surrogate model from “scratch” in every AL step i.e., starting anew from the accumulated compound dataset where weights and biases were seeded from the previous AL step’s model. But we could make also use of layer-freezing techniques (Heid et al., 2024) to accelerate surrogate model training. Likewise, the RL agent is being retrained in every AL step; we could check to see if the final or late-stage agent from the previous AL step could serve as a suitable starting point in the next GAL cycle. This would need to be gauged against concerns about model plasticity (Abbas et al., 2023; Dohare et al., 2022; Lyle et al., 2024) where covariate shift, which here comes in the form of changing chemical diversity, is one cause of concern. While AL is not continual learning this may still be an issue for even a limited number of steps of agent progression over time.

In the present work we have carried out a virtual screening or virtual hit finding exercise. The classical REINVENT prior enables *de novo* design which means that molecule generation is not limited by structural restraints but only by restraints stemming from the scoring components. In this sense we are carrying out inverse design, that is we create new molecules by describing the desired property space. For actual lead finding, in particular with AL and free energy simulation, it is typical to develop new chemistry around a common core or a limited number of cores which can be

connected via a RBFE mapping network. To facilitate this with REINVENT we could make use of the Mol2Mol model (He et al., 2024) which restrains molecule generation based on a similarity relationship of molecule pairs, e.g. using Tanimoto similarity or matched molecular pairs. This would enable us to modify a given scaffold itself depending on the tightness of the similarity criterion. If a scaffold *constraint* is desired our Libinvent model (Fialková Vendy et al., 2022) can be used and therefore generation of new molecules would occur around a fixed, pre-set core. Currently, REINVENT does not support fragment-based approaches directly i.e. our Linkinvent model (Guo et al., 2023) only supports two "warheads" (fragments) which can then be linked with a single fragment. This could be extended to allow for multiple fragments. For example, the $3CL^{Pro}$ target has a rather large, "branched" binding site (see Fig. 7) and so multiple known fragments in different binding site locations, as e.g. found in this study, could be joined together.

ESMACS is an ensemble MD method which means that multiple independent MD simulations of exactly the same molecular system are run to reinforce the statistical robustness of the simulation and its results. In the same vein, we can think about running REINVENT in an ensemble fashion. Similar to MD, the RL method is highly stochastic in nature and therefore it is advisable to carry out multiple RL runs too. There are various options available to do so. We could start multiple REINVENT runs from a single ChemProp model. We could also make use of multiple ChemProp models generated with cross-validation as several folds and also several ensembles within each fold can be made available for that (Heid et al., 2024) as we have done here. It should be noted, however, that querying multiple such models will increase inference time; in fact, the scaling is O(N) with N being the number of models. Here, the time cost for a single inference step is about 1 sec and 300-500 RL steps are carried out and would require around a day with 25 models (based on the hardware available in this study). It would therefore be more computationally efficient to infer from individual models in parallel and compute the statistics *post-hoc*. In this way we could also make use of this as uncertainty quantification (UQ). It has been shown, however, that for RBFE UQ may be of limited use at least for the system under study (Thompson et al., 2022). Alternatively, multiple independent surrogate models could be built, by varying geometries of a given network or make use of entirely different model algorithm as discussed above (resulting possibly in models of different quality/fidelity), or in a multi-fidelity fashion (Di Fiore et al., 2023; Eckmann et al., 2024; Hernandez-Garcia et al., 2023) where data of heterogeneous quality is used to create a single surrogate model.

The multi-fidelity approach has been shown to be particularly effective when a pre-trained surrogate model is trained on data that is abundantly or easily available (docking, BindingDB data) and expensive high-fidelity data (ABFE) is added to the surrogate in small volumes (Eckmann et al., 2024). The authors note, however, that this approach could be outperformed by a surrogate built on high-fidelity data only. The surrogate model in this work is based on 10,000-13,000 data points. It is of note that their generative model (Eckmann et al., 2022) appears to have been trained on the MOSES benchmark dataset of compounds (Polykovskiy et al., 2020) which is a subset of the lead-like ("rule of 3.5") ZINC 12 Clean Lead dataset (Irwin et al., 2012). REINVENT (Loeffler et al., 2024) priors are not limited to lead-like compounds. Some molecules generated in this work exhibit molecular masses larger than 500 Daltons. We also note that their generative model requires a QED weight twice as high as the ABFE component which indicates that the model struggles to produce high-quality drug-like compounds which may also reduce sample efficiency. Our weights are in the ratio 3:1 for the ABFE component to the QED component.

5. Conclusions

An AL protocol for the *de-novo* design of protein ligands was introduced by combining generative molecular AI performed by REINVENT with physics-based ABFE scoring of ligands using ensemble MD simulations and MM-PBSA. The latter uses a coarse-grained version of the ESMACS simulation protocol. The protocol, called GAL (generative active learning), was applied to two test target proteins, $3CL^{Pro}$ and TNKS2. In both cases, compounds with large binding affinities were generated that were structurally diverse and with substantial variations in their scaffolds. At the same time, these structures were found to be very dissimilar to the structures that were initially used to train the surrogate model of our GAL workflow and different regions in chemical space were efficiently explored. We have found many compounds whose binding affinities exceed those of structures generated by extensive surrogate docking models in the case of $3CL^{Pro}$ and those ligands that have been experimentally confirmed to exhibit strong binding affinities for TNKS2.

The narrowly confined binding pocket of TNKS2 led to a well-defined binding mode of ligands that translated into higher precision of GAL surrogate models for binding affinities, which in turn enabled finding structures with optimized binding affinities in fewer GAL iterations. In contrast, $3CL^{Pro}$ is characterized by a larger open binding pocket that allows variations in binding modes and thereby resulted in surrogate models with reduced precision for which more GAL iterations were required. However, GAL performed successfully in both cases: only 3 - 4 iteration steps were required in the case of TNKS2 and 5 - 7 for $3CL^{Pro}$, for generated structures to converge in chemical space. This convergence of GAL could be described as a transition from exploration of chemical space in the first iteration steps to a more exploitative regime, i.e. refinement of the structures discovered during the final iteration steps. This transition was not explicitly controlled, and all GAL steps were carried out using the same automated protocol, including the acquisition function.

Different GAL training batch sizes were tested, and qualitatively similar results were obtained. However, the larger batch sizes led in the case of TNKS2 to more precise surrogate models, accelerating convergence of GAL at the additional computational cost of more oracle calls. When comparing computational efficiency of different training batch sizes, smaller batch sizes were found to be more efficient in most cases.

Overall, even though there is clearly the potential for making further improvements, our REINVENT – ESMACS GAL protocol in its current form has been applied successfully as a generator of new compound ideas to two very different target proteins. We have demonstrated that this protocol, when integrated into a larger workflow to assess other important compound properties and to refine short-listed compounds further, is capable of *de novo* drug design to shorten the design-make-test-analysis cycle in drug discovery campaigns. It shows the immense potential of a combined AL and physics-based approach when employed reliably using ensemble simulations to control uncertainty, which is highly effective on a large supercomputer.

Acknowledgements

The authors acknowledge funding support from (i) the European Commission for EU H2020 CompBioMed2 Centre of Excellence (grant no. 823712), (ii) the UK EPSRC for the Computational Biomedicine at the Exascale (CompBioMedX) grant (EP/X019276/1) and for the Computational Biomedicine Exascale Engagement References (CompBioMedEE) grant (EP/Y008731/1), and (iii) DOE INCITE awards for years 2022-2023 and 2024-2025 providing access to computational resources on supercomputers at the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which provided access to Frontier. We thank Dilip Asthagiri and Balint Joo (ORNL) for

their help in installing various codes on Frontier (NAMD3, AmberTools, REINVENT). We are also grateful for valuable discussions with J. P. Janet.

References

Abbas, Z., Zhao, R., Modayil, J., White, A., & Machado, M. C. (2023). *Loss of Plasticity in Continual Deep Reinforcement Learning*.

Arús-Pous, J., Blaschke, T., Ulander, S., Reymond, J. L., Chen, H., & Engkvist, O. (2019). Exploring the GDB-13 chemical space using deep generative models. *Journal of Cheminformatics*, *11*(1). https://doi.org/10.1186/s13321-019-0341-z

Bajusz, D., Rácz, A., & Héberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics*, *7*(1). https://doi.org/10.1186/s13321-015-0069-3

Bellmann, L., Penner, P., Gastreich, M., & Rarey, M. (2022). Comparison of Combinatorial Fragment Spaces and Its Application to Ultralarge Make-on-Demand Compound Catalogs. *Journal of Chemical Information and Modeling*, *62*(3). https://doi.org/10.1021/acs.jcim.1c01378

Bemis, G. W., & Murcko, M. A. (1996). The properties of known drugs. 1. Molecular frameworks. *Journal of Medicinal Chemistry*, *39*(15). https://doi.org/10.1021/jm9602928

Benhenda, M. (2017). *ChemGAN challenge for drug discovery: can AI reproduce natural chemical diversity?* http://arxiv.org/abs/1708.08227

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. *Nature Chemistry*, *4*(2). https://doi.org/10.1038/nchem.1243

Blaschke, T., Arús-Pous, J., Chen, H., Margreitter, C., Tyrchan, C., Engkvist, O., Papadopoulos, K., & Patronov, A. (2020). REINVENT 2.0: An AI Tool for de Novo Drug Design. *Journal of Chemical Information and Modeling*, *60*(12). https://doi.org/10.1021/acs.jcim.0c00915

Blaschke, T., Engkvist, O., Bajorath, J., & Chen, H. (2020). Memory-assisted reinforcement learning for diverse molecular de novo design. *Journal of Cheminformatics*, *12*(1). https://doi.org/10.1186/s13321-020-00473-0

Butina, D. (1999). Unsupervised data base clustering based on daylight's fingerprint and Tanimoto similarity: A fast and automated way to cluster small and large data sets. *Journal of Chemical Information and Computer Sciences*, *39*(4). https://doi.org/10.1021/ci9803381

Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., Forouzesh, N., Giese, T. J., Götz, A. W., Gohlke, H., Izadi, S., Kasavajhala, K., Kaymak, M. C., King, E., Kurtzman, T., Lee, T. S., Li, P., Liu, J., Luchko, T., … Merz, K. M. (2023). AmberTools. *Journal of Chemical Information and Modeling*, *63*(20), 6183–6191. https://doi.org/10.1021/ACS.JCIM.3C01153/ASSET/IMAGES/LARGE/CI3C01153_0002.JPEG

Chew, A. K., Sender, M., Kaplan, Z., Chandrasekaran, A., Chief Elk, J., Browning, A. R., Kwak, H. S., Halls, M. D., & Afzal, M. A. F. (2024). Advancing material property prediction: using physics-informed machine learning models for viscosity. *Journal of Cheminformatics*, *16*(1), 31. https://doi.org/10.1186/s13321-024-00820-5

Clyde, A., Galanie, S., Kneller, D. W., Ma, H., Babuji, Y., Blaiszik, B., Brace, A., Brettin, T., Chard, K., Chard, R., Coates, L., Foster, I., Hauner, D., Kertesz, V., Kumar, N., Lee, H., Li, Z., Merzky, A., Schmidt, J. G., … Stevens, R. (2022). High-Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Noncovalent Inhibitor. *Journal of Chemical Information and Modeling*, *62*(1). https://doi.org/10.1021/acs.jcim.1c00851

Clyde, A., Liu, X., Brettin, T., Yoo, H., Partin, A., Babuji, Y., Blaiszik, B., Mohd-Yusof, J., Merzky, A., Turilli, M., Jha, S., Ramanathan, A., & Stevens, R. (2023). AI-accelerated protein-ligand docking for SARS-CoV-2 is 100-fold faster with no significant change in detection. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-28785-9

Crivelli-Decker, J. E., Beckwith, Z., Tom, G., Le, L., Khuttan, S., Salomon-Ferrer, R., Beall, J., Gómez-Bombarelli, R., & Bortolato, A. (2023). *Machine Learning Guided AQFEP: A Fast & Efficient Absolute Free Energy Perturbation Solution for Virtual Screening*.

de Oliveira, C., Leswing, K., Feng, S., Kanters, R., Abel, R., & Bhat, S. (2023). FEP Protocol Builder: Optimization of Free Energy Perturbation Protocols Using Active Learning. *Journal of Chemical Information and Modeling*, *63*(17), 5592–5603. https://doi.org/10.1021/acs.jcim.3c00681

Di Fiore, F., Nardelli, M., & Mainini, L. (2023). *Active Learning and Bayesian Optimization: a Unified Perspective to Learn with a Goal*.

Dohare, S., Sutton, R. S., & Mahmood, A. R. (2022). *Continual Backprop: Stochastic Gradient Descent with Persistent Randomness*.

Eckmann, P., Sun, K., Zhao, B., Feng, M., Gilson, M. K., & Yu, R. (2022). *LIMO: Latent Inceptionism for Targeted Molecule Generation*.

Eckmann, P., Wu, D., Heinzelmann, G., Gilson, M. K., & Yu, R. (2024). *MFBind: a Multi-Fidelity Approach for Evaluating Drug Compounds in Practical Generative Modeling*.

Fialková, V., Zhao, J., Papadopoulos, K., Engkvist, O., Bjerrum, E. J., Kogej, T., & Patronov, A. (2021). LibINVENT: Reaction-based Generative Scaffold Decoration for in Silico Library Design. *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.1c00469

Fialková Vendy, Zhao, J., Papadopoulos Kostas, Engkvist Ola, Bjerrum Esben Jannik, Kogej Thierry, & Patronov Atanas. (2022). LibINVENT: Reaction-based Generative Scaffold Decoration for in Silico Library Design. *Journal of Chemical Information and Modeling*, *62*(9), 2046–2063. https://doi.org/10.1021/acs.jcim.1c00469

Filella-Merce, I., Molina, A., Orzechowski, M., Díaz, L., Zhu, Y. M., Mor, J. V., Malo, L., Yekkirala, A. S., Ray, S., & Guallar, V. (2023). *Optimizing Drug Design by Merging Generative AI With Active Learning Frameworks*. http://arxiv.org/abs/2305.06334

Garnett, R. (2023). *Bayesian Optimization*. Cambridge University Press.

Genheden, S., Thakkar, A., Chadimová, V., Reymond, J. L., Engkvist, O., & Bjerrum, E. (2020). AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of Cheminformatics*, *12*(1). https://doi.org/10.1186/s13321-020-00472-1

Gentile, F., Agrawal, V., Hsing, M., Ton, A. T., Ban, F., Norinder, U., Gleave, M. E., & Cherkasov, A. (2020). Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. *ACS Central Science*, *6*(6). https://doi.org/10.1021/acscentsci.0c00229

Gorantla, R., Kubincová, A., Suutari, B., Cossins, B. P., & Mey, A. S. J. S. (2024). Benchmarking Active Learning Protocols for Ligand-Binding Affinity Prediction. *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.4c00220

Graff, D. E., Shakhnovich, E. I., & Coley, C. W. (2021). Accelerating high-throughput virtual screening through molecular pool-based active learning. *Chemical Science*, *12*(22). https://doi.org/10.1039/d0sc06805e

Gusev, F., Gutkin, E., Kurnikova, M. G., & Isayev, O. (2023). Active Learning Guided Drug Design Lead Optimization Based on Relative Binding Free Energy Modeling. *Journal of Chemical Information and Modeling*, *63*(2). https://doi.org/10.1021/acs.jcim.2c01052

He, J., Tibo, A., Janet, J. P., Nittinger, E., Tyrchan, C., Czechtizky, W., & Ola, E. (2024). *Evaluation of Reinforcement Learning in Transformer-based Molecular Design*. https://doi.org/10.21203/rs.3.rs-4106688/v1

Heid, E., & Green, W. H. (2021). Machine Learning of Reaction Properties via Learned Representations of the Condensed Graph of Reaction. *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.1c00975

Heid, E., Greenman, K. P., Chung, Y., Li, S.-C., Graff, D. E., Vermeire, F. H., Wu, H., Green, W. H., & McGill, C. J. (2024). Chemprop: A Machine Learning Package for Chemical Property Prediction. *Journal of Chemical Information and Modeling*, *64*(1), 9–17. https://doi.org/10.1021/acs.jcim.3c01250

Heinzelmann, G., & Gilson, M. K. (2021). Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-020-80769-1

Hernandez-Garcia, A., Saxena, N., Jain, M., Liu, C.-H., & Bengio, Y. (2023). *Multi-Fidelity Active Learning with GFlowNets*.

Homeyer, N., & Gohlke, H. (2012). Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method. *Molecular Informatics*, *31*(2), 114–122. https://doi.org/10.1002/minf.201100135

Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. (2012). ZINC: A free tool to discover chemistry for biology. In *Journal of Chemical Information and Modeling* (Vol. 52, Issue 7). https://doi.org/10.1021/ci3001277

Kearns, M., Rubinfeld, R., Mansour, Y., Schapire, R. E., Ron, D., & Univemity, H. (1994). On the learnability of discrete distributions. *Proceedings of the Annual ACM Symposium on Theory of Computing*, *Part F129502*. https://doi.org/10.1145/195058.195155

Khalak, Y., Tresadern, G., Hahn, D. F., De Groot, B. L., & Gapsys, V. (2022). Chemical Space Exploration with Active Learning and Alchemical Free Energies. *Journal of Chemical Theory and Computation*, *18*(10). https://doi.org/10.1021/acs.jctc.2c00752

Kim, M. K. (2018). Novel insight into the function of Tankyrase (Review). In *Oncology Letters* (Vol. 16, Issue 6). https://doi.org/10.3892/ol.2018.9551

Knight, J. L., Leswing, K., Bos, P. H., & Wang, L. (2021). Impacting Drug Discovery Projects with Large-Scale Enumerations, Machine Learning Strategies, and Free-Energy Predictions. In *ACS Symposium Series* (Vol. 1397). https://doi.org/10.1021/bk-2021-1397.ch008

Konze, K. D., Bos, P. H., Dahlgren, M. K., Leswing, K., Tubert-Brohman, I., Bortolato, A., Robbason, B., Abel, R., & Bhat, S. (2019). Reaction-Based Enumeration, Active Learning, and Free Energy Calculations to Rapidly Explore Synthetically Tractable Chemical Space and Optimize Potency of Cyclin-Dependent Kinase 2 Inhibitors. *Journal of Chemical Information and Modeling*, *59*(9). https://doi.org/10.1021/acs.jcim.9b00367

Li, J., Fu, A., & Zhang, L. (2019). An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. In *Interdisciplinary Sciences – Computational Life Sciences* (Vol. 11, Issue 2). https://doi.org/10.1007/s12539-019-00327-w

Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007). BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Research*, *35*(SUPPL. 1). https://doi.org/10.1093/nar/gkl999

Loeffler, H. H., He, J., Tibo, A., Janet, J. P., Voronov, A., Mervin, L. H., & Engkvist, O. (2024). Reinvent 4: Modern AI–driven generative molecule design. *Journal of Cheminformatics*, *16*(1), 20. https://doi.org/10.1186/s13321-024-00812-5

Lyle, C., Zheng, Z., Khetarpal, K., van Hasselt, H., Pascanu, R., Martens, J., & Dabney, W. (2024). *Disentangling the Causes of Plasticity Loss in Neural Networks*.

Marin, E., Kovaleva, M., Kadukova, M., Mustafin, K., Khorn, P., Rogachev, A., Mishin, A., Guskov, A., & Borshchevskiy, V. (2023). Regression-Based Active Learning for Accessible Acceleration of Ultra-Large Library Docking. *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.3c01661

McCloskey, M., & Cohen, N. J. (1989). Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. *Psychology of Learning and Motivation - Advances in Research and Theory*, *24*(C). https://doi.org/10.1016/S0079-7421(08)60536-8

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, *3*(29). https://doi.org/10.21105/joss.00861

Mervin, L., Voronov, A., Kabeshov, M., & Engkvist, O. (2024). *Qptuna: an automated QSAR modelling platform for molecular property prediction in drug design*. https://doi.org/10.26434/chemrxiv-2024-2rlk7

Mohr, B., Shmilovich, K., Kleinwächter, I. S., Schneider, D., Ferguson, A. L., & Bereau, T. (2022). Data-driven discovery of cardiolipin-selective small molecules by computational active learning. *Chemical Science*. https://doi.org/10.1039/d2sc00116k

Molecular Modeling Software | OpenEye Scientific. (2024). https://www.eyesopen.com

Olivecrona, M., Blaschke, T., Engkvist, O., & Chen, H. (2017). Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics*, *9*(1). https://doi.org/10.1186/s13321-017-0235-x

Papadopoulos, K., Giblin, K. A., Janet, J. P., Patronov, A., & Engkvist, O. (2021). De novo design with deep generative models based on 3D similarity scoring. *Bioorganic and Medicinal Chemistry*, *44*. https://doi.org/10.1016/j.bmc.2021.116308

Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., McGreevy, R., Melo, M. C. R., Radak, B. K., Skeel, R. D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z., … Tajkhorshid, E. (2020). Scalable molecular

dynamics on CPU and GPU architectures with NAMD. *Journal of Chemical Physics*, *153*(4), 44130. https://doi.org/10.1063/5.0014475/16709547/044130_1_ACCEPTED_MANUSCRIPT.PDF

Polykovskiy, D., Zhebrak, A., Sanchez-Lengeling, B., Golovanov, S., Tatanov, O., Belyaev, S., Kurbanov, R., Artamonov, A., Aladinskiy, V., Veselov, M., Kadurin, A., Johansson, S., Chen, H., Nikolenko, S., Aspuru-Guzik, A., & Zhavoronkov, A. (2020). Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models. *Frontiers in Pharmacology*, *11*. https://doi.org/10.3389/fphar.2020.565644

RDKit: Open-source cheminformatics. (n.d.).

Sadiq, S. K., Wright, D., Watson, S. J., Zasada, S. J., Stoica, I., & Coveney, P. V. (2008). Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *Molecular Simulation*, *Md*, 1909–1919. https://doi.org/10.1021/ci8000937

Sauer, S., Matter, H., Hessler, G., & Grebner, C. (2022). Optimizing interactions to protein binding sites by integrating docking-scoring strategies into generative AI methods. *Frontiers in Chemistry*, *10*. https://doi.org/10.3389/fchem.2022.1012507

Schindler, C. E. M., Baumann, H., Blum, A., Böse, D., Buchstaller, H. P., Burgdorf, L., Cappel, D., Chekler, E., Czodrowski, P., Dorsch, D., Eguida, M. K. I., Follows, B., Fuchß, T., Grädler, U., Gunera, J., Johnson, T., Lebrun, C. J., Karra, S., Klein, M., … Kuhn, D. (2020). Large-scale assessment of binding free energy calculations in active drug discovery projects. *Journal of Chemical Information and Modeling*, *60*(11). https://doi.org/10.1021/acs.jcim.0c00900

Schlander, M., Hernandez-Villafuerte, K., Cheng, C. Y., Mestre-Ferrandiz, J., & Baumann, M. (2021). How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment. In *PharmacoEconomics* (Vol. 39, Issue 11). https://doi.org/10.1007/s40273-021-01065-y

Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., Fisher, J., Jansen, J. M., Duca, J. S., Rush, T. S., Zentgraf, M., Hill, J. E., Krutoholow, E., Kohler, M., Blaney, J., Funatsu, K., Luebkemann, C., & Schneider, G. (2020). Rethinking drug design in the artificial intelligence era. In *Nature Reviews Drug Discovery* (Vol. 19, Issue 5). https://doi.org/10.1038/s41573-019-0050-3

Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackerman, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A Deep Learning Approach to Antibiotic Discovery. *Cell*, *180*(4). https://doi.org/10.1016/j.cell.2020.01.021

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2). https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Sugiyama, M., & Nakajima, S. (2009). Pool-based active learning in approximate linear regression. *Machine Learning*, *75*(3). https://doi.org/10.1007/s10994-009-5100-3

Thomas, M., Bender, A., & de Graaf, C. (2023). Integrating structure-based approaches in generative molecular design. *Current Opinion in Structural Biology*, *79*, 102559. https://doi.org/https://doi.org/10.1016/j.sbi.2023.102559

Thompson, J., Walters, W. P., Feng, J. A., Pabon, N. A., Xu, H., Goldman, B. B., Moustakas, D., Schmidt, M., & York, F. (2022). Optimizing active learning for free energy calculations. *Artificial Intelligence in the Life Sciences*, *2*. https://doi.org/10.1016/j.ailsci.2022.100050

Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*.

Wan, S., Bhati, A. P., Wade, A. D., & Coveney, P. V. (2023). Ensemble-Based Approaches Ensure Reliability and Reproducibility. *Journal of Chemical Information and Modeling*, *63*, 6959–6963. https://doi.org/10.1021/ACS.JCIM.3C01654

Wan, S., Knapp, B., Wright, D. W., Deane, C. M., & Coveney, P. V. (2015). Rapid, Precise and Reproducible Prediction of Peptide-MHC Binding Affinities from Molecular Dynamics that Correlate Well with Experiment. *Journal of Chemical Theory and Computation*, *11*, 3346–3356. https://doi.org/10.1021/acs.jctc.5b00179

Warmuth, M. K., Liao, J., Rätsch, G., Mathieson, M., Putta, S., & Lemmen, C. (2003). Active learning with support vector machines in the drug discovery process. *Journal of Chemical Information and Computer Sciences*, *43*(2). https://doi.org/10.1021/ci025620t

Wesolowski, S. S., & Brown, D. G. (2016). *The Strategies and Politics of Successful Design, Make, Test, and Analyze (DMTA) Cycles in Lead Generation*. https://doi.org/10.1002/9783527677047.ch17

WHO. (2023). *WHO chief declares end to COVID-19 as a global health emergency*. https://news.un.org/en/story/2023/05/1136367

Wouters, O. J., McKee, M., & Luyten, J. (2020). Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. In *JAMA - Journal of the American Medical Association* (Vol. 323, Issue 9). https://doi.org/10.1001/jama.2020.1166

Wu, D. (2019). Pool-Based Sequential Active Learning for Regression. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(5). https://doi.org/10.1109/TNNLS.2018.2868649

Xu, X., Huang, M., & Zou, X. (2018). Docking-based inverse virtual screening: methods, applications, and challenges. *Biophysics Reports*, *4*(1). https://doi.org/10.1007/s41048-017-0045-8

Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., Palmer, A., Settels, V., Jaakkola, T., Jensen, K., & Barzilay, R. (2019). Analyzing Learned Molecular Representations for Property Prediction. *Journal of Chemical Information and Modeling*, *59*(8). https://doi.org/10.1021/acs.jcim.9b00237