# Combining Cloud-Based Free Energy Calculations, Synthetically Aware Enumerations and Goal-Directed Generative Machine Learning for Rapid Large-Scale Chemical Exploration and Optimization

Phani Ghanakota[‡], Pieter H. Bos[‡], Kyle D. Konze, Joshua Staker, Gabriel Marques, Kyle Marshall, Karl Leswing, Robert Abel, Sathesh Bhat*

[‡]These authors contributed equally to this work

*satesh.bhat@schrodinger.com

Schrödinger Inc, 120 West 45th St, 17th floor, New York, New York, 10036

## Abstract

The hit identification process usually involves the profiling of millions to more recently billions of compounds either via traditional experimental high throughput screens (HTS) or computational virtual high throughput screens (vHTS). We have previously demonstrated that by coupling reaction-based enumeration, active learning and free energy calculations, a similarly large-scale exploration of chemical space can be extended to the hit-to-lead process. In this work, we augment that approach by coupling large scale enumeration and cloud-based FEP profiling with goal-directed generative machine learning, which results in a higher enrichment of potent ideas compared to large scale enumeration alone, while simultaneously staying within the bounds of a predefined drug-like property space. We are able to achieve this by building the molecular distribution for generative machine learning from the PathFinder rules-based enumeration and optimizing for a weighted sum QSAR based multi-parameter optimization function. We examine the utility of this combined approach by designing potent inhibitors of cyclin-dependent kinase 2 (CDK2) and demonstrate a coupled workflow that can: (1) provide a 6.4 fold enrichment

improvement in identifying < 10nM compounds over random selection, and a 1.5 fold enrichment in identifying < 10nM compounds over our previous method (2) rapidly explore relevant chemical space outside the bounds of commercial reagents, (3) use generative ML approaches to "learn" the SAR from large scale *in silico* enumerations and generate novel idea molecules for a flexible receptor site that are both potent and within relevant physicochemical space and (4) produce over 3,000,000 idea molecules and run 2153 FEP simulations, identifying 69 ideas with a predicted $IC_{50}$ < 10nM and 358 ideas with a predicted $IC_{50}$ <100 nM. The reported data suggest combining both reaction-based and generative machine learning for ideation results in a higher enrichment of potent compounds over previously described approaches, and can rapidly accelerate the discovery of novel chemical matter within a predefined potency and property space.

**Introduction**

Computational chemistry and molecular modeling have recently seen a resurgence in preclinical drug discovery. This is likely due to the increased computational power of modern CPUs and GPUs, and advances in computational methods. For example, virtual screening and pharmacophore modeling are regularly used as hit finding strategies,[1–4] and physics-based methods like free energy perturbations (FEP) for small molecule potency prediction are heavily used in the hit-to-lead and lead optimization phases of drug discovery.[5–10] During a typical virtual screen, millions of compounds are profiled *in silico* for the target of interest. However, as a hit is progressed into a lead, and subsequently lead optimization, the number of compounds considered is significantly smaller in scale than was profiled in the virtual screen.[11] The lead optimization phase of drug discovery tasks medicinal chemists with the job of designing and synthesizing analogs of a lead compound to alter the physicochemical properties which affect multiple endpoints of interest. These include, but are far from limited to: maintaining or increasing on-target potency, reducing off-target potency, increasing solubility, increasing metabolic stability,

increasing permeability, and reducing clearance. Oftentimes, this is accomplished by modulating common physicochemical properties such as tPSA, logP and MW. For example, a compound that has high lipophilicity may have aqueous solubility issues, so the chemist may look for opportunities to incorporate polarity by replacing a phenyl with a more polar heterocycle or replacing a methyl group with a methoxy group. Sometimes the modifications are less obvious, like incorporating a bicyclopentane to replace a phenyl for the purpose of increasing solubility.[12] It is up to the medicinal chemists to use their intuition, past expertise, and known SAR data to design the next series of analogs. This process, which drives the design cycle, is limited by human capability; one cannot expect a chemist to remember every chemical modification that was previously utilized, nor have a full recollection of the plethora of past transformations described in the literature. On the computational front, due to the limit on how many compounds can be synthesized and assayed within a traditional lead optimization timeframe there can often be insufficient data available to appropriately train generative machine learning models, which in literature have required hundreds of thousands to millions of relevant exemplar compounds [13,14].

Historically, three main factors have contributed to the smaller scale of exploration during the lead optimization phase: (1) a lack of computational tools accessible to the computational or medicinal chemist to easily create large sets of drug-like and synthetically tractable idea molecules, (2) a lack of accurate methods to prospectively optimize multiple parameters simultaneously, and (3) a lack of computational power to perform the necessary calculations on a timescale that can impact the design cycle of a drug discovery program. Further, more than half of the costs of preclinical drug discovery are incurred during the lead optimization phase, highlighting the need for more sophisticated approaches to support this phase of drug discovery.[15] We believe that providing the necessary computational tools to quickly explore a large portion of chemical space will allow drug discovery programs to find better compounds more efficiently, and will reduce the

number of compounds needed to be synthesized and assayed before entering the clinic; saving both time and money in the process.

Currently, there is significant interest in utilizing computational methods for the de novo design of compounds in drug discovery.[16–19] Some examples include: (1) generating ligands from known building blocks using simulated organic steps, (2) generative machine learning design utilizing reaction vectors,[20] (3) fragment replacement using rule-based fragmentation in combination with a pharmacophore,[21] (4) using Reaction-MQL encoded reactions,[22] and (5) generative machine learning approaches.[13,23–27] A number of tools exist to aid synthetic route design as well: (1) LHASA, pioneered by EJ Corey in the 1960s,[15,28] (2) network algorithms that mine the literature for applicable reactions,[20] (3) automated retrosynthetic rule generation based on pattern recognition,[21] and (4) neural-symbolic approaches to retrosynthesis and reaction planning.[21,22] Previously, we described PathFinder,[29] a reaction-based enumeration tool that enables rapid, facile generation of synthetically tractable drug-like ligands from commercially available building blocks using practical chemical reactions. To illustrate the impact this technology can have on drug discovery, we combined PathFinder, multi-parameter optimization, docking, machine learning, and cloud-based FEP simulations to design and predict the potency of CDK2 inhibitors.[29] The PathFinder workflow allows for the rapid exploration of a large portion of chemical space, but there are several limitations: (1) the large number of ligands generated are nevertheless bounded by commercially available reagents and represent only a small fraction of total chemical space,[29,30] (2) the ligands are generated using building blocks containing specific functional groups compatible with the synthetic chemistry encoded in PathFinder. Consequently, on-target potency and other properties are not explicitly considered during ligand generation, and (3) as we have previously demonstrated[29] an enumeration workflow can generate upwards of thousands of virtual SAR data points per week. These virtual datasets quickly become prohibitively difficult for a medicinal chemistry team to analyze in order to tease out SAR trends for the next round of ideation, which could increase the number of design-test cycles to arrive at optimal compounds.

To address the aforementioned shortcomings, we aim to couple PathFinder with cloud-based FEP and generative machine learning methods to increase the relevant chemical space explored during lead optimization. Particularly, we show the superior performance of this coupled workflow in enriching for more potent compounds that fit a predefined drug-like space (Table 1).

## Methods

### PathFinder R-group enumeration

For the R-group enumeration we performed a retrosynthetic analysis of the starting ligand **1a** (Figure 1) using PathFinder, followed by enumeration of all possible reaction routes while keeping the core of the molecule constant. In each route the reagent that contains the immutable region is kept fixed and the other reagent is systematically varied. After the exhaustive enumeration of all possible synthetic routes the output was combined and duplicate products were removed resulting in a total of 1.36 million compounds. This set of ligands was filtered using a proprietary set of ~600 SMARTS patterns to remove known and potential chemical liabilities and PAINS offenders. This step reduced the number of compounds from 1.36 million to 607,357. Next, we removed all ligands that would be considered outside reasonable drug-like space using a physicochemical property filter criteria based on the "rule of five"[31]: molecular weight ≤ 500, a polar surface area between 50 - 130 $Å^2$, LogP between -1 and 4, maximum of 5 hydrogen bond donors (HBD), maximum of 10 hydrogen bond acceptors (HBA), maximum of 6 rotatable bonds, maximum of 5 aromatic rings, and no chiral centers. After filtering based on these physicochemical properties, the library size was reduced to 253,372 compounds. The set of ligands was prepared for docking using LigPrep[32] and only neutral ligands with a total and absolute charge equal to zero were retained. The resulting library of 196,290 structures was docked in the CDK2 receptor (PDB: 1H1S) using Glide SP,[33,34] while employing both a hinge

constraint and a core constraint (RMSD of 0.1 Å compared to the reference ligand (Figure 1, **1b**, $R_1$ = H). All ligands that fit in the binding pocket and satisfied the constraints were retained and resulting ligands were filtered once more by their physicochemical properties as the 2D to 3D conversion stage introduced certain stereocenters or tautomeric variants outside of the desired property space. The final set of ligands from the R-group enumeration that passed the filtering protocol detailed above contained 52,932 unique ligands. A random subset of 935 ligands was selected for FEP calculations to train the ML-FEP model ($QSAR_{FEP}$).

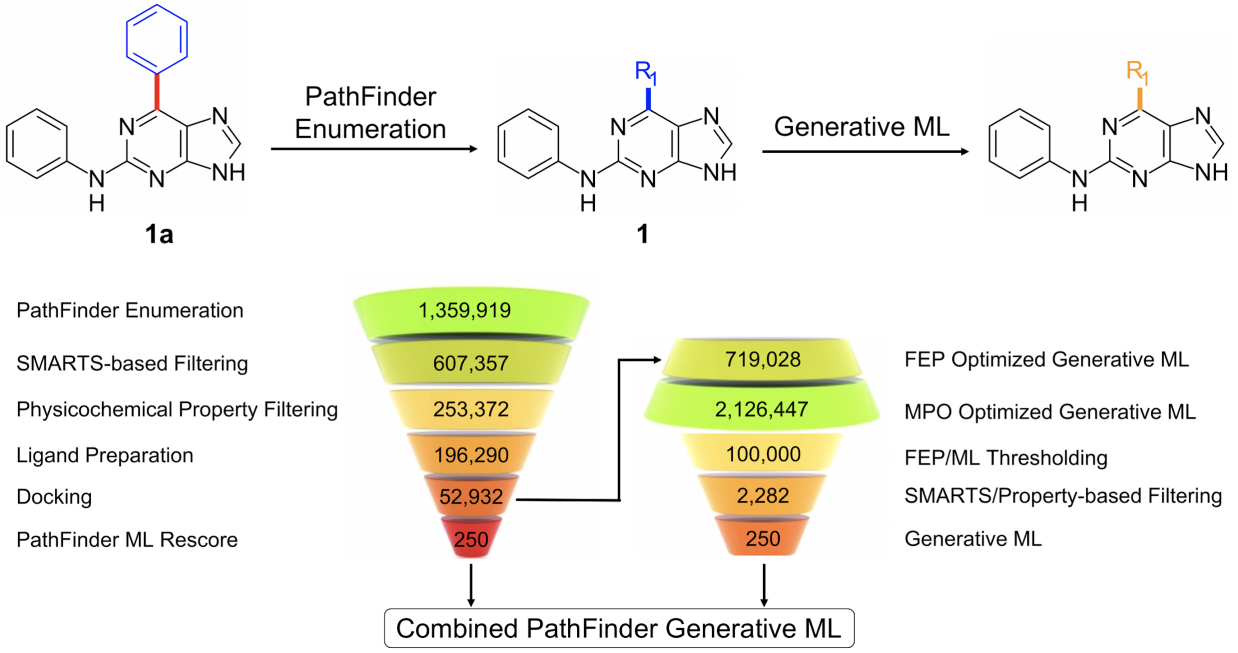



**Figure 1.** PathFinder Generative ML workflow: PathFinder R-group enumeration followed by ML-FEP and Generative ML.

### Free Energy Perturbation Calculations

Relative binding free energy calculations were performed using the FEP+ implementation within the Schrodinger 2019-3 software release.[35] The CDK2 protein system was part of the original set of test systems validated with the FEP+ methodology.[5] Further validation was also reported in our

recent work using CDK2.[29] The ligand forcefield torsions were parameterized with the approach described in the earlier cited work using OPLS3e forcefield.[36] In order to run FEP, the core of ligand was predefined using a SMARTS pattern based on the reference ligand **1b** (see Table 2) and the input pose for FEP was obtained from GLIDE docking described previously. The FEP simulations were run using a single edge from the reference for 1 nanosecond (ns) on cloud computing GPU resources. The perturbations were run using the default 12 λ windows. Representative snapshots from the FEP simulations were auto-generated for each FEP edge from clustering the λ=0 window trajectory frames. The clustering was done based on the protein-ligand interactions and the centroid of the largest cluster is referred to as the representative frame in this work. The representative frames from all the FEP simulations were used to highlight the interactions of the ligands with the protein and for further analysis including SiteMap[37,38] calculations.

**Goal-Directed Generative Machine Learning**

Generative methods have been shown to work when optimizing in general chemical space[39] but in an early lead optimization effort, a drug discovery team is usually modifying only a single substituent at a time while holding the remainder of the molecule constant, primarily in order to elucidate SAR. To augment that design process computationally, here we describe a protocol for applying REINVENT[14] in order to generate millions of unique compounds at a single R-group location on a single core.

The REINVENT algorithm is a two-stage process -- first training a prior network on a population of chemical matter, and second shifting the distribution to also perform well on a utility function. In the original work the REINVENT algorithm was initially trained on a population of molecules selected from the ChEMBL database. In this work we trained REINVENT on the PathFinder

generated molecules. Because of this, 96% of valid molecules generated by our agent networks had the correct core and relevant R-group location for lead optimization in this study, without explicitly putting a term for this substructure in our utility function.

In order to run REINVENT we need a utility function whose gradient trends with interesting molecular matter for a drug discovery project. We estimate that for it to run in a reasonable amount of time to impact a standard drug discovery project, this utility function has to be evaluated on the order of milliseconds per molecule. In this study we explicitly optimized molecules for two QSAR models trained with AUTOQSAR/DC.[40,41] The first QSAR model, a regression model, $QSAR_{FEP,}$ was trained on FEP results from a random 935 molecules selected from a PathFinder enumeration. The second QSAR model, a classification model, $QSAR_{property}$, was trained on property and substructure filters from the PathFinder workflow. To get the compounds used to train $QSAR_{property}$ we ran a round of $QSAR_{FEP}$ generative machine learning optimization. We then down-sampled the generated molecules to balance the $QSAR_{property}$ training dataset between compounds that pass the property and substructure filters and compounds that fail the property or substructure filters.

To combine these two QSAR models into a single utility function we used EQ 2. EQ 2 is a weighted sum of a linear transformation of the QSAR models. We sample over possible weightings to approximate the Pareto optimal frontier. The goal of the linear transform equation is to normalize the dynamic range of an arbitrary function between negative infinity and one. y_min and y_max are the minimum and maximum value of the QSAR model's training data, target is the goal value of the QSAR model for optimization. This transformation will scale linearly as the distance from the target in this newly normalized space. For this study we set the target to -0.2, 20% of the dynamic range lower than the tightest binding molecule in the training set. We set the property

model target to 0, that is the model reports a 0% chance the molecule would fail our property or SMARTS filters.

$$LT(x, target, y_{min}, y_{max}) = 1 - \left| \frac{x - y_{min}}{y_{max} - y_{min}} - target \right|$$

**EQ 1.**

$$U_i = \begin{cases} w_0 * LT(QSAR_{fep}(m_i)) + w_1 * LT(QSAR_{property}(m_i)) \\ -1 \end{cases} \quad \text{if } m_i \text{ is an invalid molecule}$$

**EQ 2. $\Sigma w_j = 1$, $w_1 \in \{0.0, 0.05, 0.09, 0.17, 0.23\}$**

After the second round of MPO driven optimization we took the top 100,000 compounds generated with the lowest predicted FEP value and ran PathFinder structure and property filters resulting in 500 unique molecules being prioritized.

Alongside the Generative ML design protocol we repeated the "Round 1 Protocol" from our previous work with CDK2[29] where ligands that passed docking were reprioritized based on the same AutoQSAR/DC model used for generative machine learning design ($QSAR_{FEP}$).

**Matched Molecular Pairs Analysis**

Matched molecular pairs (MMP) were generated by fragmentation of the molecules from the combined dataset of the PathFinder ML Rescore and Generative ML molecules using a max heavy atom difference of 8 atoms and a maximum of 2 cuts.[42] Matched molecular pairs were grouped by their MMP constant structure and each group was internally sorted to maximize disparity of FEP score between PathFinder ML Rescore and Generative ML populations. Groups

were then ranked by their internal FEP disparity score in order to analyze the change in predicted potency between similar PathFinder and Generative ML designs.

**Physicochemical Descriptors, Chemical Diversity, and Clustering**

All physicochemical descriptors were calculated using RDKit[43] other than LogP which was computed using Canvas[44]. Lipophilic ligand efficiency (LLE)[45–47] was calculated as $pIC_{50}$ - LogP.

To gain further insights into the diversity and similarity of all the ideas profiled using FEP from each of the three subsets of ligands, we initially carried out a principal component analysis (PCA), which revealed the top two components only accounted for 10% of the variance in the data. Thus to provide a more accurate description of the diversity of the dataset we turned to a t-distributed stochastic neighbor embedding (tSNE) analysis that would account for a greater proportion of the variance in the data.[48] The top 50 PCA components were selected to reduce the dimensionality of the dataset followed by a tSNE analysis.

**SiteMap Analysis**

SiteMap is a binding site druggability analysis tool that takes into account the hydrophobic/hydrophilic character along with the curvature of binding sites.[37,38] The fast and accurate nature of these calculations allowed us to profile representative frames from each of the FEP calculations run and analyze the dynamics of the binding pocket in the presence of chemical matter. SiteMap descriptors were calculated for the binding pocket within 6Å of the ligand.

## Results and Discussion

The rate-limiting step in most large scale *in silico* workflows is generally the most expensive calculation, both from a cost and timescale perspective. In this case the FEP calculations represent the most computationally expensive step, so in real-world drug discovery applications, we would ideally prefer to advance only those compounds to FEP that have the highest probability of having good predicted activity while simultaneously not violating a predefined property space. We examine the performance of three workflows to accomplish this task: (1) selecting a random selection of ideas from the PathFinder enumeration for FEP (PathFinder Random), (2) using a ML model trained on FEP data to prioritize compounds from the PathFinder enumeration for FEP (PathFinder Rescore) (3) prioritize compounds from Generative ML for FEP (Generative ML) and (4) prioritize an equal amount of the top compounds from PathFinder Rescore and the Generative ML for FEP calculations (PathFinder Generative ML).

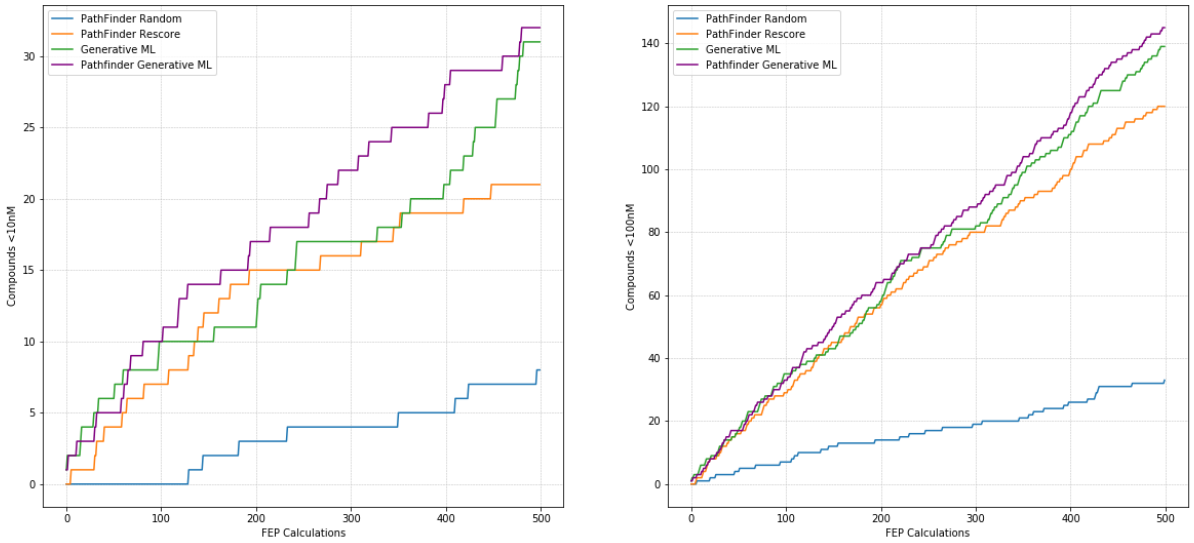



**Figure 2.** a) Cumulative Compounds Identified < 10nM Given A Number Of FEP Calculations By Prioritization Method b) Cumulative Compounds Identified < 100nM Given A Number Of FEP Calculations By Prioritization Method.

**Table1.** Cumulative Predicted potencies of Pathfinder Random, Pathfinder Rescore, Generative ML and Pathfinder Generative ML Designs.

| Prioritization Approach | Number of Ideas Profiled by FEP | <10nM | <100nM | <1μM | >1μM |
|---|---|---|---|---|---|
| PathFinder Random | 935 | 9 (1.0%) | 56 (5.9%) | 205 (22%) | 730 (77%) |
| PathFinder Rescore | 500 | 21 (4.2%) | 120 (24%) | 322 (64%) | 178 (36%) |
| Generative ML | 500 | 31 (6.2%) | 139 (28%) | 314 (63%) | 186 (37%) |
| PathFinder Generative ML | 500 | **32 (6.4%)** | **145 (29%)** | **341 (68%)** | **159 (32%)** |

As outlined in Table 1, the random ideas selected from the PathFinder enumeration mainly consist of relatively weak binding compounds that would be of little interest to most lead optimization projects (77% > 1μM), suggesting that random selection would be a non-optimal approach since a majority of the FEP budget would be allocated to compounds with a low probability of meeting the desired potency criteria. The PathFinder Rescore approach is able to rapidly prioritize enumerated compounds that are predicted to be more potent (24% < 100nM) and many fewer weak binders (36% > 1uM), which agrees with previous results that demonstrated this approach is superior to random selection[29]. Picking only compounds from the Generative ML set outperforms PathFinder Rescore identifying 19 more compounds < 100nM (+16%) and 10 more compounds < 10 nM (+47%). PathFinder Generative ML slightly improves these results further by identifying one more compound < 10nM (+1%), and six more compounds < 100nM (+4%). The mixed strategy of PathFinder Generative ML is conceptually robust because we are replacing compounds that are stack ranked lower from one method, with compounds that are stack ranked higher from the other. If both prioritized pools of compounds have approximately the same hit rate distribution and the distribution skews towards enrichment over random the mixed strategy will have better results. While Table 1 discusses results given a budget of an additional 500 FEP calculations we can identify tight binders with far fewer FEP calculations. Using the PathFinder

Generative ML strategy, we can identify 5 compounds < 10nM and 17 compounds < 100nM given 50 additional FEP calculations. Figure 2 depicts how many compounds can be identified at different potency thresholds for increasing numbers of additional FEP calculations.

This suggests that the generative machine learning method is an effective approach to utilize data from large-scale *in silico* FEP enumerations to rapidly generate additional novel, potent chemical matter. While potency is an important driver in drug discovery programs, it is also of interest to examine how diverse the chemical matter is from the three approaches. This is shown in the tSNE plots in Figure 2. From examining the number of clusters, there is a drop in diversity going from the PathFinder Random set (Figure 2b) to the Pathfinder ML Rescore set (Figure 2c). This is expected since the PathFinder Rescore will select for ligands based solely on potency, which here results in focusing on specific clusters of chemical space. Interestingly, there is an additional drop in diversity in the Generative ML set (Figure 2c) suggesting that the Generative ML set is further focusing on specific areas of chemical space that meet both the property space and potency requirements (Figure 2a).

Comparing the property space of the PathFinder Rescore and PathFinder Generative ML datasets show that both sets of ideas generally occupy similar property space over the majority of molecular descriptors (see Supplementary Figure S1). This suggests that the Generative ML approach is able to produce novel ideas that fit within a predefined multi-dimensional property space, while at the same time maintaining or improving potency, which is often a key objective in drug discovery campaigns. In Figure 4a we focus on three common descriptors of interest to drug discovery projects, molecular weight (MW), the logarithm of the octanol/water partition coefficient (LogP) and total polar surface area (TPSA). On average the following trends were observed for the Generative ML ligands: 1) Reduction in LogP; 2) Increase in TPSA; 3) Slight increase in MW. This is in-line with the observed SAR of the majority of Generative ML ideas: functional groups

are added that are able to form additional polar interactions with the p-loop while simultaneously staying within the bounds of the other property space descriptors (Figure 4a, and 7).

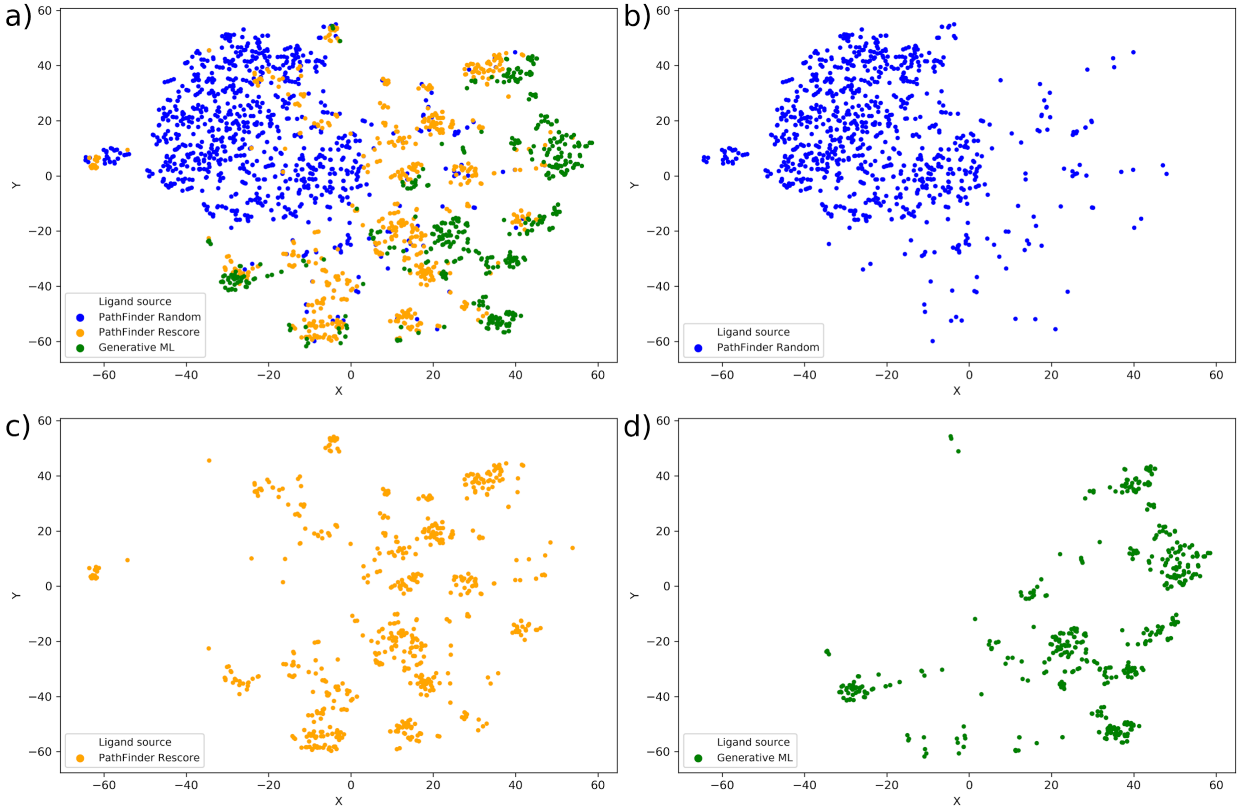



**Figure 3.** tSNE plots of various ligand datasets generated from the PathFinder and Generative ML sets. a) tSNE of all chemical space of all R-groups (from PathFinder Random, PathFinder Rescore and Generative ML); b) tSNE of R-groups from the PathFinder Random set; c) tSNE of R-groups from the PathFinder Rescore set; d) tSNE of R-groups from the Generative ML set.

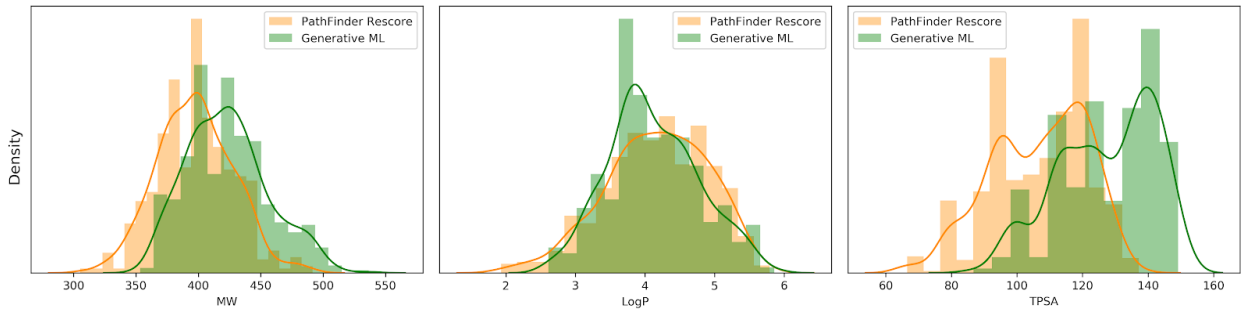



**Figure 4a.** Smoothed histogram of selected property space (MW, LogP and TPSA) for Generative ML and Pathfinder Rescore ideas.

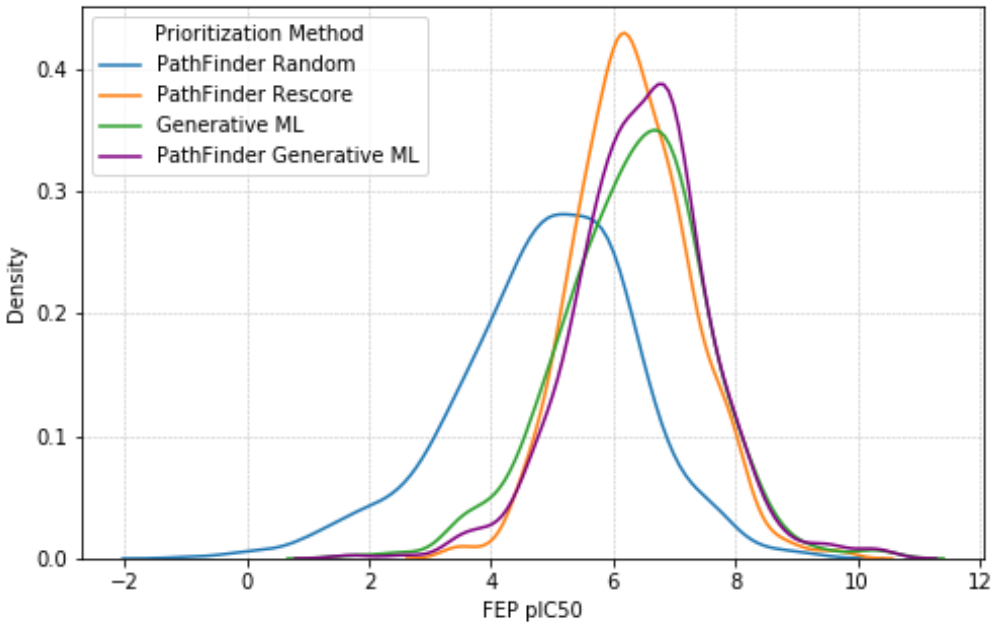



**Figure 4b.** Smoothed histogram comparing the FEP potencies for the four different selection schemes.

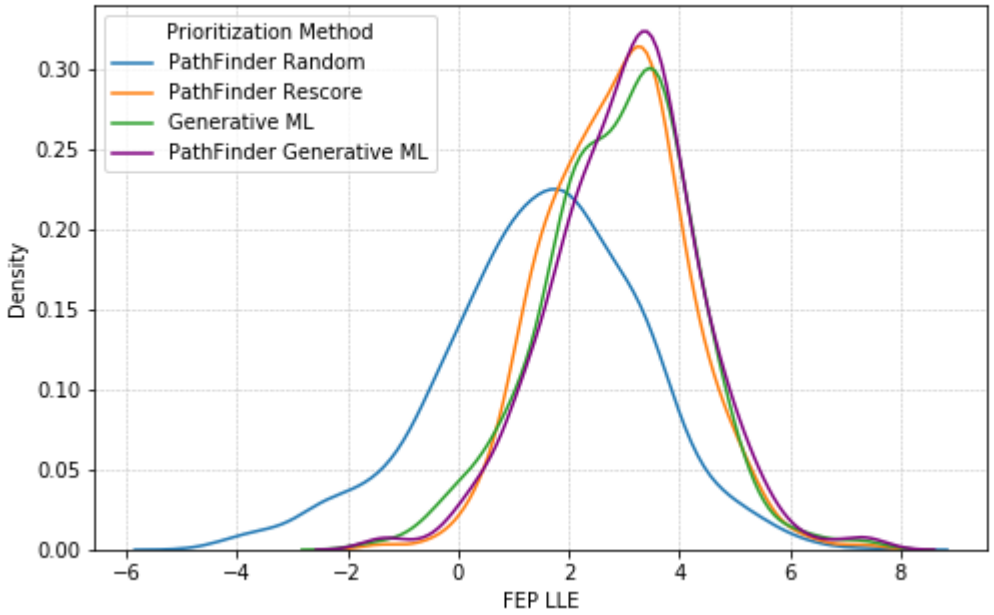



**Figure 4c.** Smoothed histogram comparing FEP Lipophilic Ligand Efficiency (LLE) for the four different selection schemes.

Looking at the overall potency distributions (Figure 4b), we observe that the Generative ML approach shows a shift towards more potent compounds relative to both the PathFinder Random and PathFinder Rescore approaches. The PathFinder Generative ML approach shows a further shift towards more potent compounds, although not as pronounced as the previous comparison. Additionally, the PathFinder Generative ML set contains fewer weak binders than the other approaches. The PathFinder Rescore, Generative ML and PathFinder Generative ML set show similar distributions in LLE (Figure 4c). A predicted LLE > 6 is considered of high quality in a drug discovery project.[47,49,50] Compared to the PathFinder Random set, we see an improved distribution in LLE for all three approaches, indicating that the compounds are occupying a more desirable druglike space. While not trained on LLE explicitly, it is encouraging to observe that the Generative ML model is able to produce compounds within the same LLE space as the top compounds from the other approaches.

Figure 5 plots the change in FEP predicted potency (pIC50) vs change in LLE for all matched molecular pairs between the PathFinder Rescore and Generative ML enumeration. The purpose of generating this data was to examine if the Generative ML approach was improving potency simply by adding lipophilicity to existing ideas, which is generally not an ideal design approach in most drug discovery campaigns [47]. It is encouraging to observe that potency increases (delta pIC50 > 0) are for the most part accompanied by an improvement in LLE (45% of ideas fit these criteria), particularly for the most potent ideas. This suggests that the generative machine learning approach, when integrated with free energy calculation and conventional enumeration technologies, is able to generate ideas with improved potency without sacrificing key physicochemical properties, which is a desirable paradigm in lead optimization efforts.

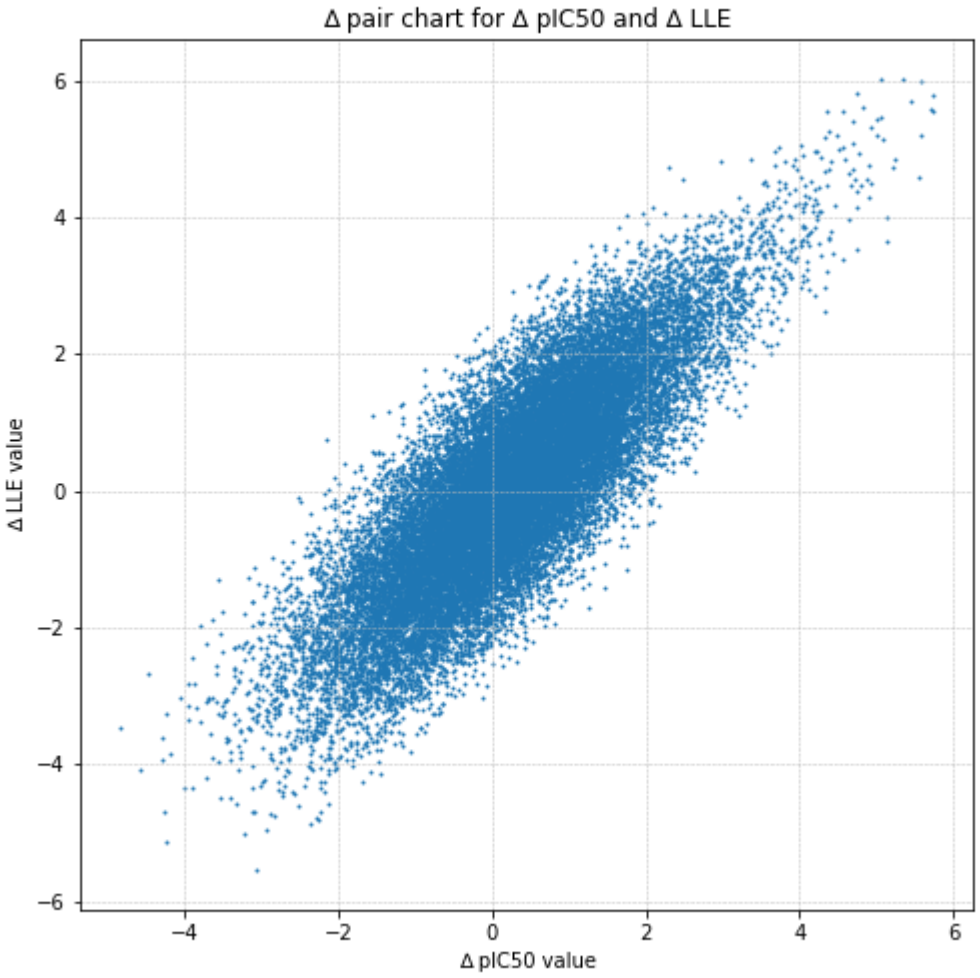



**Figure 5.** Change in LLE vs change in predicted pIC50 potency for all matched molecular pairs between PathFinder Rescore ideas and Generative ML ideas.

**SAR Analysis of PathFinder and Generative ML Designed Compounds**

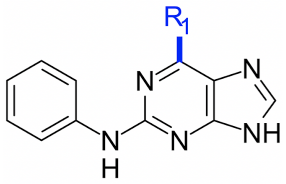
A representative set of potent R-groups that were identified by the PathFinder Rescore approach is shown in Table 2a. All of the potent R-groups displayed in Table 2a contain a phenyl ring coupled to the central purine core, are substituted at the meta position, and contain at least one hydrogen bond donor and/or acceptor. In some cases the phenyl ring was part of a fused ring system (**3**). Four of the R-group examples shown in Table 2a contain an amide linker (**2**, **3**, **4**, and **8**), and four examples have a hydroxyl group (**4**, **5**, **7**, and **8**).

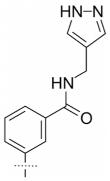
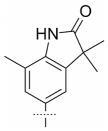
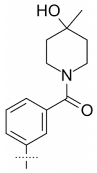
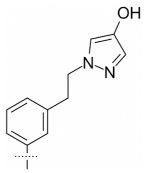
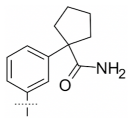
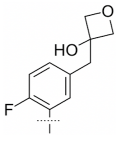
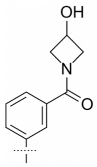
Table 2b illustrates a representative selection of potent ligands designed using Generative ML. Similar to the PathFinder Rescore ligands, the examples shown contain a 6-membered ring attached to the purine core. The main difference is that a number of compounds contain aryl replacements such as a pyridine (**10**, **13**, and **14**) or a pyridin-2(1*H*)-one (**12**). In line with the examples from Table 2a, all compounds highlighted are substituted at the meta position and contain at least one hydrogen bond donor and/or acceptor. As illustrated earlier, the LLE of both the Generative ML and the ML Rescore set are comparable and a number of potent compounds are quite attractive based on their lipophilic ligand efficiency relative to the starting ligand (LLE ≈ 6).

Previous experimental studies described the necessity of adding a primary sulfonamide at the 4-position of the aniline in order to improve potency to $pIC_{50}$ ≈ 7-8.[51] Both the PathFinder Rescore and Generative ML were able to identify a number of compounds with a potency in that desired range without the need for additional r-groups at the aniline position. By increasing the potency in the absence of the sulfonamide, we avoid the need to carry the potential physicochemical complications (high TPSA, high MW, etc) which can sometimes lead to poor properties for sulfonamide containing compounds (low permeability, lower cellular potency, *etc*.). This also underscores the design flexibility this workflow affords during the lead optimization process,

allowing for the rapid replacement of potentially problematic functional groups without sacrifices to potency or physicochemical properties.

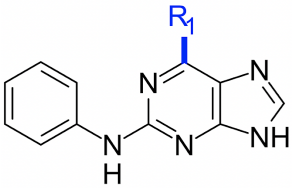
**Table 2a.** Selection of Potent R-Groups Identified by PathFinder ML Rescore

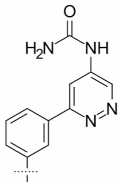
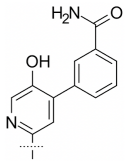
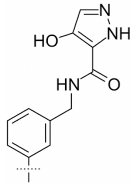
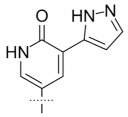
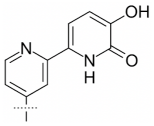
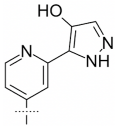


| | 1b[a] | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $R_1$ | H |  |  |  |  |  |  |  |
| **FEP $pIC_{50}$** | 4.87 (4.21) | 9.58 | 8.99 | 8.55 | 8.23 | 8.19 | 8.15 | 7.95 |
| **FEP $IC_{50}$ (nM)** | 13,500 (61,000) | 0.26 | 1.02 | 2.82 | 5.89 | 6.46 | 7.08 | 11.22 |
| **$LLE_{FEP}$** | 2.77 | 7.46 | 5.15 | 6.14 | 5.06 | 4.41 | 5.3 | 5.87 |

[a]Experimental values in parentheses, see Coxon *et al.*[51]

**Table 2b.** Selection of Potent R-Groups Identified by Generative ML



| | 1b[a] | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| $R_1$ | H |  |  |  |  |  |  |
| **FEP $pIC_{50}$** | 4.87 (4.21) | 10.39 | 8.69 | 8.46 | 7.89 | 7.28 | 7.12 |
| **FEP $IC_{50}$ (nM)** | 13,500 (61,000) | 0.041 | 2.04 | 3.47 | 12.88 | 52.48 | 75.86 |
| **$LLE_{FEP}$** | 2.77 | 7.35 | 5.83 | 6.18 | 6.07 | 4.42 | 3.17 |

[a]Experimental values in parentheses, see Coxon *et al.*[51]

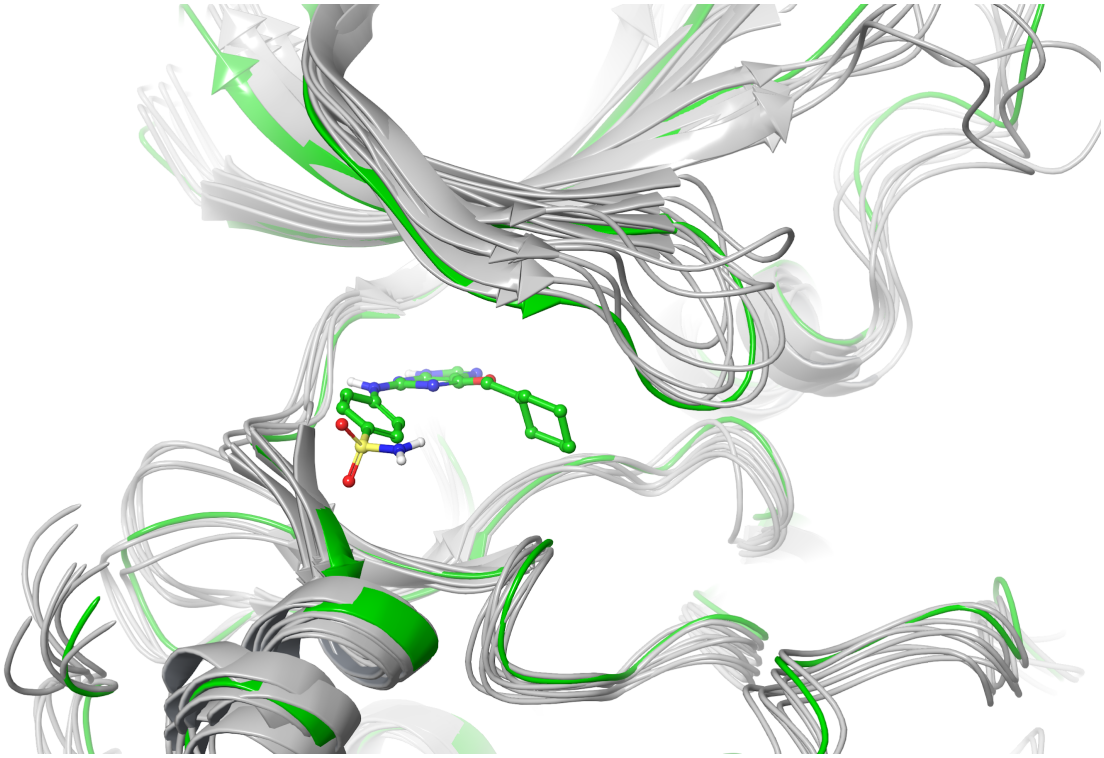


**Figure 6a**. Overlay of representative snapshots from high scoring ligands. The crystal structure of CDK2 used to run FEP is show in green, snapshots in grey (PDB ID: 1H1S)

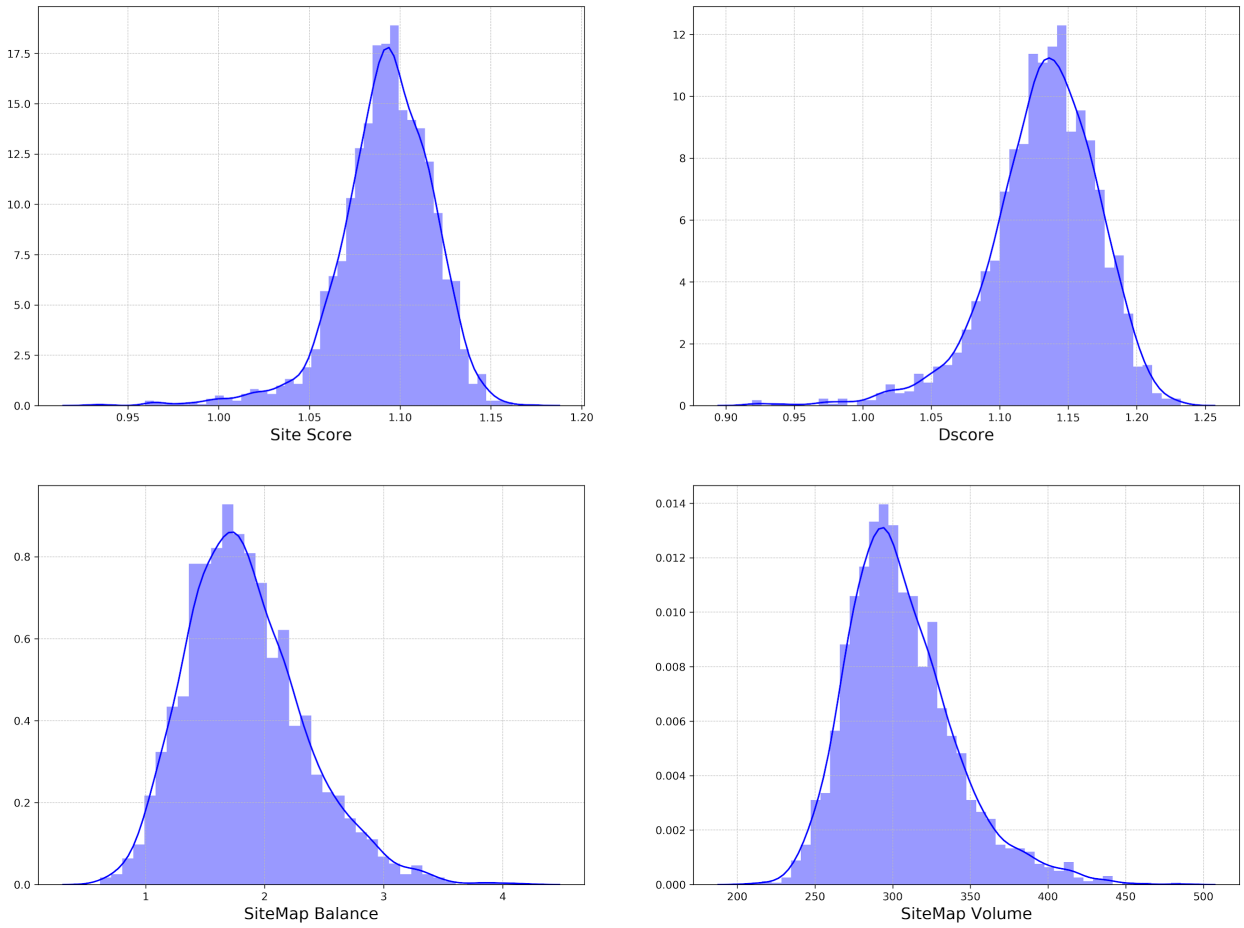



**Figure 6b.** Smoothed histograms of various descriptive properties from SiteMap calculations on FEP representative frames across all idea sets.

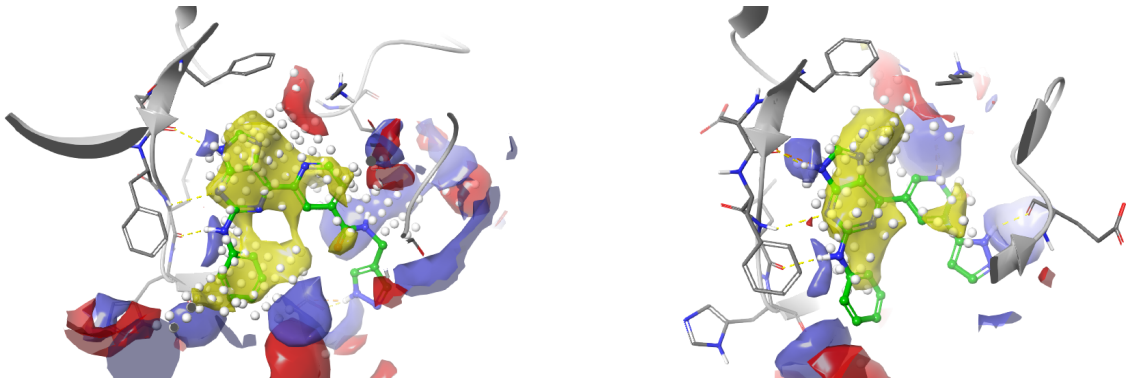


**Figure 6c.** A top down view of SiteMap results with the kinase hinge binding motif positioned to the left for compound **16** ($pIC_{50}$ = 9.58) and compound **12** ($pIC_{50}$ = 7.89). The SiteMap points used to define the volume are in white spheres. The hydrophobic component of the binding site is colored yellow, hydrogen bond donor region colored blue and the hydrogen bond acceptor region color red.

Figure 6a is an overlay of representative CDK2 receptor FEP snapshots from high scoring compounds. Interestingly, we observe significant movement of the kinase P-loop away from the starting static crystallographic conformation. The p-loop appears to adopt a variety of distinct conformations depending on the R-group interacting with that region. This suggests that designing ligand functional groups interacting with this region via traditional structure-based drug design (SBDD) could be non-trivial since the conformation of the loop could be directly affected by the functional group introduced. This is further substantiated by Figure 6b, which shows four histograms of the distribution of different SiteMap metrics across those representative FEP+ snapshots. In particular the SiteMap Balance and SiteMap Volume plots suggest significant variability in the polarity and shape of the site. This is further substantiated by Figure 6c which illustrates the high variability of the binding pocket depending on the nature of the R-group, with the more potent ligand resulting in a receptor conformation that exposes many more polar residues and hydrophobic volume. By incorporating a large number of system-specific FEP simulation predictions into $QSAR_{FEP}$, we are implicitly including this information in the generative machine learning process. This can help circumvent the difficult nature of manually designing potent ligands against a flexible pocket while maintaining a favorable property space.

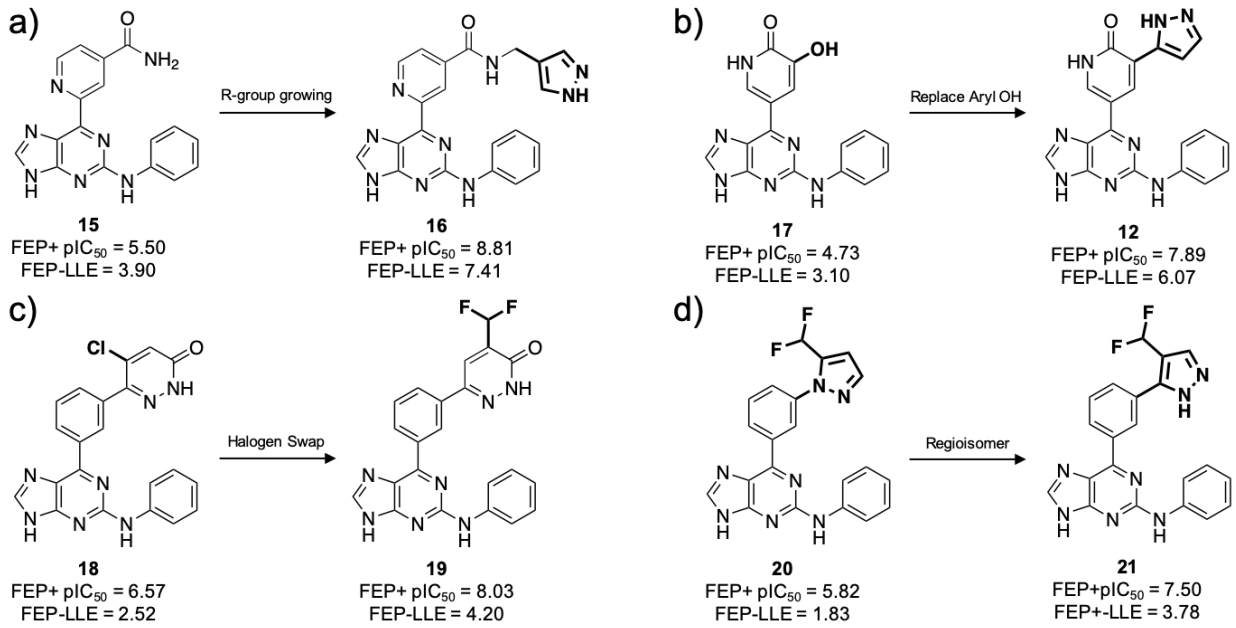



**Figure 7.** Selected matched molecular pairs highlighting R-group modifications designed by Generative ML. a) Pyrazole incorporation; b) Replacement of aryl hydroxyl; c) Substitution of chlorine for dichloromethane; d) Altering pyrazole regioisomer to increase potency. (*Reported $pIC_{50}$ values are from FEP predictions)

It is of further interest to examine the modifications Generative ML makes to commercially available reagents to improve potency while maintaining favorable drug-like properties. Figure 7 highlights some examples of matched molecular pairs that were generated by Generative ML from PathFinder enumerated ligands. Interestingly, we observe small to relatively large alterations to portions of the parent ligand. For example, Compound **15** is a weakly potent inhibitor of CDK2 based on the FEP predicted potency (Figure 7a; FEP+ $pIC_{50}$ = 5.50). With the goal of increasing potency in mind, a common technique employed a medicinal chemist may be to grow off the amide, because it is a synthetically feasible modification that can be installed with simple chemistry.[52] Generative ML created a similar ligand by growing a pyrazole from the amide via a methylene linker (**16**), and this modification is predicted to increase the potency significantly (>3 $pIC_{50}$ in FEP+; Figure 7a). Aryl-hydroxy containing compounds are often avoided during lead optimization due to potential ADME issues[53]. Figure 7b shows that Generative ML designed away from the 3-hydroxy-pyrid-2-one of **17** and replaced it with a 3-pyrazolo-pyrid-2-one in **12**, which

maintains the hydrogen-bonding character of the hydroxyl group, and is predicted to be significantly more potent than **17** (>3 $pIC_{50}$ in FEP+; Figure 7b). Another common technique in lead optimization is to move a methyl ("methyl walk") or halogen around a ring with the goal of determining the best position to substitute for on-target potency.[53,54] An example of this effort is displayed in Figure 7c, where Generative ML designed a difluoromethyl containing compound (**19**) from the halogen (Cl) containing parent (**18**). Further, the $CF_2$ group was installed at the 6-position of the pyridazinone ring for potency, rather than the 5-position like the chloro; FEP+ predicted a significant boost in potency of (~1.5 $pIC_{50}$; Figure 7c). Lastly, Generative ML demonstrated the ability to make very minor modifications to the parent chemical matter that are predicted to significantly increase the potency in FEP simulations. For example, Figure 7d shows that simply changing the regioisomer of the 5-difluoromethyl-pyrazole of **20** to the 4-difluoromethyl-pyrazole of **21** is predicted to significantly increase the potency (~1.5 $pIC_{50}$). Figure 8 highlights the additional hydrogen bonding interactions made by the Generative ML ideas with the CDK2 receptor. Collectively, the modifications in Figure 7 and Figure 8 demonstrate that Generative ML is able to "learn" the FEP SAR and produce ligands similar to those we would expect to be designed by medicinal chemists.

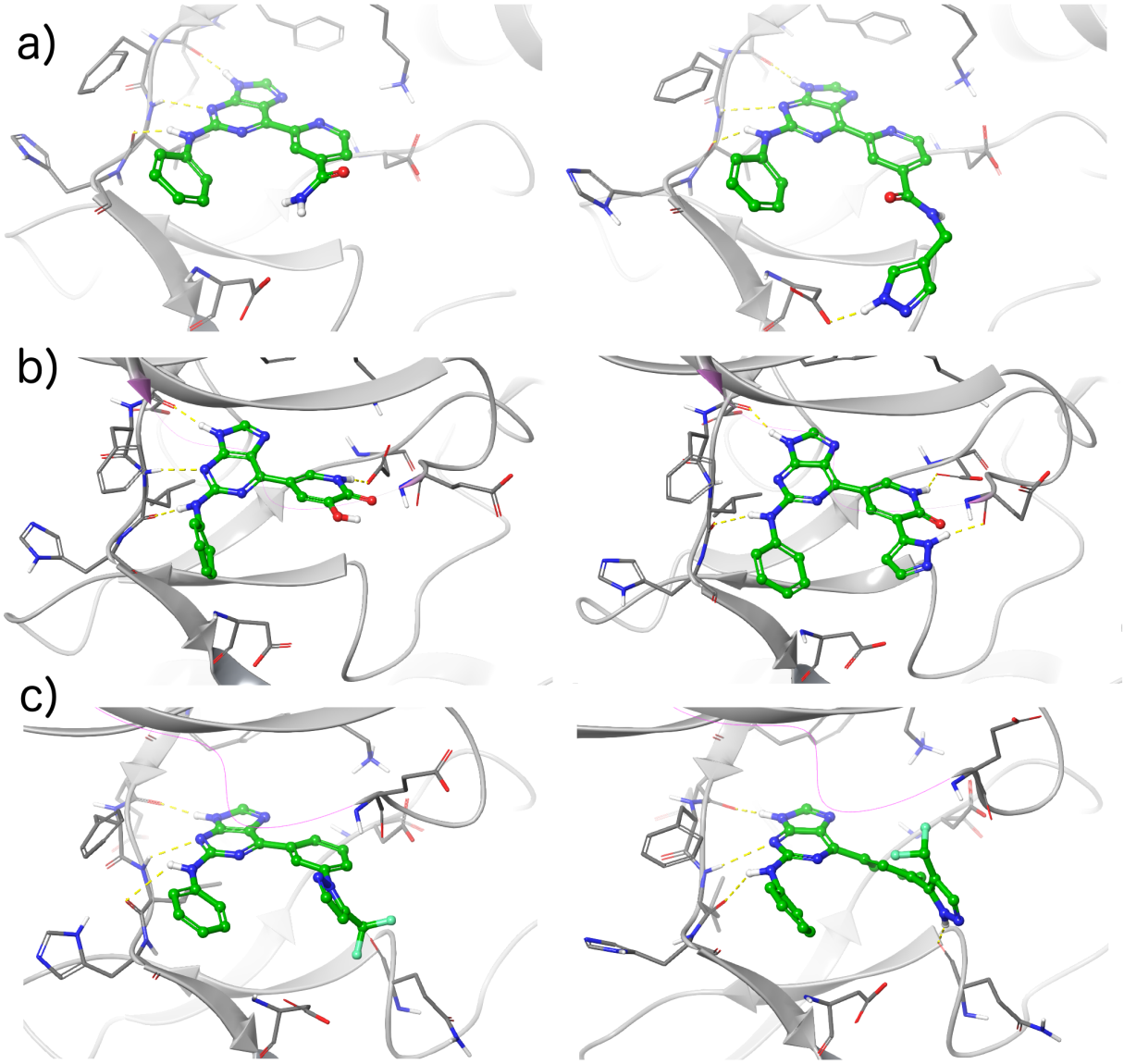



**Figure 8.** Examples of new ligand-receptor interactions formed by Generative ML designed compounds with improved potency. The PathFinder Rescore ligand is shown on the left and the corresponding Generative ML MMP with increased potency on the right. The ligands are all oriented to show the interactions of the hinge binding motif to the left. A portion of the P-loop is rendered as a thin pink wire, to better display the hinge binding interactions of the core. a) MMP of compound **15** (left) and compound **16** (right) showing increased potency with the pyrazole able to interact with Asp86. b) MMP of compound **17** (left) and compound **12** (right) in which addition of the pyrazole results in a gain in potency through a different hydrogen bond interaction with the Glu12 residue in the P-loop. c) MMP of compound **20** (left) and compound **21** (right) in which the Generative ML designed compound forms a hydrogen bond interaction with the backbone of Gln131.

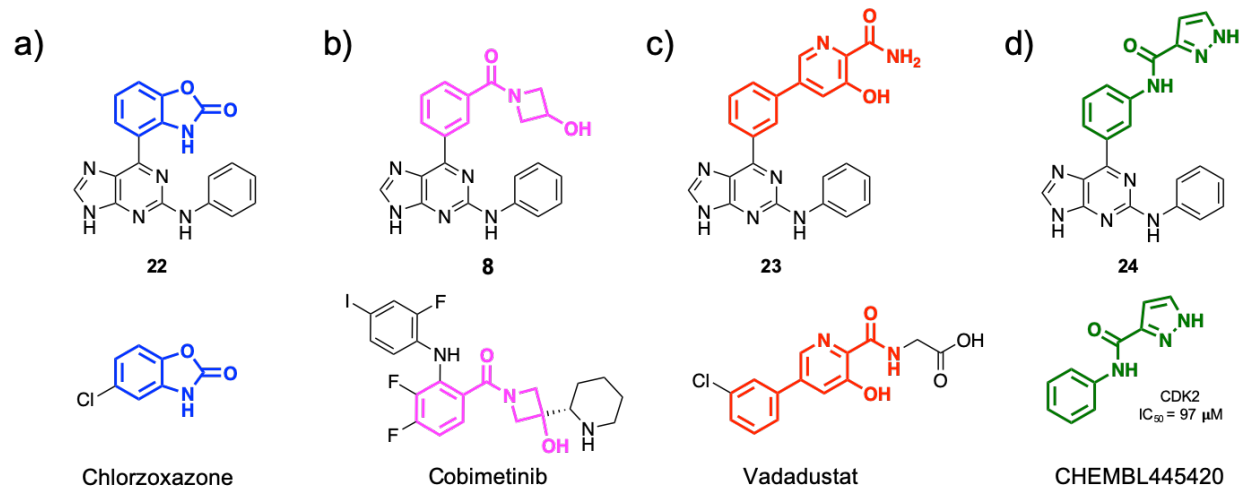



**Figure 9.** Drug-like fragments generated by the Generative ML approach. a) Fragment of **22** identified in Chlorzoxazone. b) Fragment of **8** identified in Cobimetinib. c) Fragment of **23** identified in Vadadustat. d) The identical structure of the fragment of **24** is CHEMBL445420, a very weak inhibitor of CDK2.

To further examine if the generative machine learning approach creates drug-like ligands, we queried ChEMBL[55] against a variety of the R-groups generated by the Generative ML design process. It is important to note that the Generative ML approach has not been trained on the ChEMBL dataset. Interestingly, a number of the fragments from the Generative ML approach are present in marketed drugs (Figure 9). The 2-benzoxazolinone of **22** is present with a chlorine at the 5-position of the benzoxazolinone ring in the marketed muscle relaxant Chlorzoxazone (Figure 9a). Compound **8** contains a 3-hydroxy-azetidine amide at the meta position of the phenyl ring, and a similar group is found in Cobimetinib, an approved MEK inhibitor used to treat some melanomas (Figure 9b). Compound **23** contains a 2-formamide-3-hydroxy-pyridine, which is very similar to a portion of Vadadustat, a HIF prolyl-hydroxylase inhibitor that is currently in phase 3 clinical trials for the treatment of anemia and chronic kidney disease (Figure 9c). Finally, the ChEMBL search also identified a fragment that was previously screened against CDK2 (Figure 9d). CHEMBL445420 is a very weak inhibitor of CDK2 ($IC_{50}$ = 97 µM) and is the exact structure of the fragment designed by Generative ML in compound **24**. Many other fragments from the ligands in the Generative ML set were also found in a variety of compounds that were progressed

into the clinic on drug discovery programs across a wide variety of targets such as VEGFR2, COX-1, COX-2, c-Met, Factor Xa, c-Kit, CDK5, ABL, BRCA1, and DPP4 (not shown), which further demonstrates that the ideation approach outlined here is of general interest to lead optimization drug discovery projects.

## Conclusion

We have previously shown how combining cloud-based FEP calculations with reaction-based enumeration and active learning can rapidly generate novel potent compounds.[29] In this work, we augment that approach by coupling the workflow with goal-directed machine learning generative design (PathFinder Generative ML). This new workflow provides a 6.4 fold improvement in identifying <10nM compounds over random selection, and a 1.5 fold enrichment in identifying <10nM compounds over our previous method. We also demonstrate how this approach can rapidly explore relevant chemical space outside the bounds of commercial reagents, which was a limitation of our previous workflow. Upon examination of FEP simulation snapshots of the CDK2 receptor and sitemap analysis, we observe significant mobility in the binding pocket suggesting that this a non-trivial SBDD problem. We were able to generate 3,000,000 idea molecules and run 2153 FEP simulations, identifying 69 ideas predicted $IC_{50}$ <10nM and 358 ideas with a predicted $IC_{50}$ <100 nM. The Generative ML approach is able to "learn" the SAR from large scale *in silico* enumerations and generate novel idea molecules that are potent and within relevant physicochemical space. In particular, we observe that the Generative ML ideas are able to improve potency while improving the LLE which is a desirable strategy in a lead optimization campaign. Upon examination of the MMPs between the original PathFinder ideas and their Generative ML counterparts, we observe that the generative machine learning algorithm is able to generate modifications similar to those found in traditional medicinal chemistry, and that several of the Generative ML R-groups can be found in market drugs. This study compares PathFinder

Rescore with PathFinder Generative ML for a single round of prioritization, future work can be done showing the effects of multiple rounds of prioritization. We expect that coupling FEP calculations with both reaction-based and generative machine learning approaches for enumeration will have significant utility in accelerating the drug discovery process.

## Supporting Information

All ligands from the PathFinder and Generative ML calculations with FEP and QSAR predictions and all ligands from PathFinder that passed the docking stage are available in supplementary material in SMILES format.

## References


(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* **2004**, *3* (11), 935–949.

(2) Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1923–1938.

(3) Schneider, G. Virtual Screening: An Endless Staircase? *Nature Reviews Drug Discovery*. 2010, pp 273–276. https://doi.org/10.1038/nrd3139.

(4) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; et al. Ultra-Large Library Docking for Discovering New Chemotypes. *Nature* **2019**, *566* (7743), 224–229.

(5) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *Journal of the American Chemical Society*. 2015, pp 2695–2703. https://doi.org/10.1021/ja512751q.

(6) Rombouts, F. J. R.; Tresadern, G.; Buijnsters, P.; Langlois, X.; Tovar, F.; Steinbrecher, T. B.; Vanhoof, G.; Somers, M.; Andrés, J.-I.; Trabanco, A. A. Pyrido[4,3-e][1,2,4]triazolo[4,3-A]pyrazines as Selective, Brain Penetrant Phosphodiesterase 2 (PDE2) Inhibitors. *ACS Med. Chem. Lett.* **2015**, *6* (3), 282–286.

(7) Lovering, F.; Aevazelis, C.; Chang, J.; Dehnhardt, C.; Fitz, L.; Han, S.; Janz, K.; Lee, J.; Kaila, N.; McDonald, J.; et al. Imidazotriazines: Spleen Tyrosine Kinase (Syk) Inhibitors Identified by Free-Energy Perturbation (FEP). *ChemMedChem* **2016**, *11* (2), 217–233.

(8) Christ, C. D.; Fox, T. Accuracy Assessment and Automation of Free Energy Calculations for Drug Design. *Journal of Chemical Information and Modeling*. 2014, pp 108–120. https://doi.org/10.1021/ci4004199.

(9) Keränen, H.; Pérez-Benito, L.; Ciordia, M.; Delgado, F.; Steinbrecher, T. B.; Oehlrich, D.; van Vlijmen, H. W. T.; Trabanco, A. A.; Tresadern, G. Acylguanidine Beta Secretase 1 Inhibitors: A Combined Experimental and Free Energy Perturbation Study. *Journal of Chemical Theory and Computation*. 2017, pp 1439–1453. https://doi.org/10.1021/acs.jctc.6b01141.

(10) Schindler, C.; Rippmann, F.; Kuhn, D. Relative Binding Affinity Prediction of Farnesoid X

Receptor in the D3R Grand Challenge 2 Using FEP. *Journal of Computer-Aided Molecular Design*. 2018, pp 265–272. https://doi.org/10.1007/s10822-017-0064-z.
(11) Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162* (6), 1239–1249.
(12) Stepan, A. F.; Subramanyam, C.; Efremov, I. V.; Dutra, J. K.; O'Sullivan, T. J.; DiRico, K. J.; McDonald, W. S.; Won, A.; Dorff, P. H.; Nolan, C. E.; et al. Application of the bicyclo[1.1.1]pentane Motif as a Nonclassical Phenyl Ring Bioisostere in the Design of a Potent and Orally Active γ-Secretase Inhibitor. *J. Med. Chem.* **2012**, *55* (7), 3414–3424.
(13) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science*. 2018, pp 268–276. https://doi.org/10.1021/acscentsci.7b00572.
(14) Olivecrona, M.; Blaschke, T.; Engkvist, O.; Chen, H. Molecular de-Novo Design through Deep Reinforcement Learning. *J. Cheminform.* **2017**, *9* (1), 48.
(15) Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L. How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge. *Nature Reviews Drug Discovery*. 2010, pp 203–214. https://doi.org/10.1038/nrd3078.
(16) Schneider, G.; Fechner, U. Computer-Based de Novo Design of Drug-like Molecules. *Nat. Rev. Drug Discov.* **2005**, *4* (8), 649–663.
(17) Hartenfeller, M.; Schneider, G. Enabling Future Drug Discovery by de Novo Design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2011, pp 742–759. https://doi.org/10.1002/wcms.49.
(18) Dean, P. M.; Lloyd, D. G.; Todorov, N. P. De Novo Drug Design: Integration of Structure-Based and Ligand-Based Methods. *Curr. Opin. Drug Discov. Devel.* **2004**, *7* (3), 347–353.
(19) Kutchukian, P. S.; Shakhnovich, E. I. De Novo Design: Balancing Novelty and Confined Chemical Space. *Expert Opin. Drug Discov.* **2010**, *5* (8), 789–812.
(20) Patel, H.; Bodkin, M. J.; Chen, B.; Gillet, V. J. Knowledge-Based Approach to de Novo Design Using Reaction Vectors. *J. Chem. Inf. Model.* **2009**, *49* (5), 1163–1184.
(21) Firth, N. C.; Atrash, B.; Brown, N.; Blagg, J. MOARF, an Integrated Workflow for Multiobjective Optimization: Implementation, Synthesis, and Biological Evaluation. *Journal of Chemical Information and Modeling*. 2015, pp 1169–1180. https://doi.org/10.1021/acs.jcim.5b00073.
(22) Hartenfeller, M.; Zettl, H.; Walter, M.; Rupp, M.; Reisen, F.; Proschak, E.; Weggen, S.; Stark, H.; Schneider, G. DOGS: Reaction-Driven de Novo Design of Bioactive Compounds. *PLoS Comput. Biol.* **2012**, *8* (2), e1002380.
(23) Gupta, A.; Müller, A. T.; Huisman, B. J. H.; Fuchs, J. A.; Schneider, P.; Schneider, G. Generative Recurrent Networks for De Novo Drug Design. *Molecular Informatics*. 2018, p 1700111. https://doi.org/10.1002/minf.201700111.
(24) Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. The Rise of Deep Learning in Drug Discovery. *Drug Discov. Today* **2018**, *23* (6), 1241–1250.
(25) Kadurin, A.; Nikolenko, S.; Khrabrov, K.; Aliper, A.; Zhavoronkov, A. druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico. *Molecular Pharmaceutics*. 2017, pp 3098–3104. https://doi.org/10.1021/acs.molpharmaceut.7b00346.
(26) Blaschke, T.; Olivecrona, M.; Engkvist, O.; Bajorath, J.; Chen, H. Application of Generative Autoencoder inDe NovoMolecular Design. *Molecular Informatics*. 2018, p 1700123. https://doi.org/10.1002/minf.201700123.
(27) Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Central Science*. 2018,

pp 120–131. https://doi.org/10.1021/acscentsci.7b00512.
(28) Pensak, D. A.; Corey, E. J. LHASA—Logic and Heuristics Applied to Synthetic Analysis. *ACS Symposium Series*. 1977, pp 1–32. https://doi.org/10.1021/bk-1977-0061.ch001.
(29) Konze, K. D.; Bos, P. H.; Dahlgren, M. K.; Leswing, K.; Tubert-Brohman, I.; Bortolato, A.; Robbason, B.; Abel, R.; Bhat, S. 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. *J. Chem. Inf. Model.* **2019**, *59* (9), 3782–3793.
(30) eMolecules. *Choice Reviews Online*. 2006, p 43Sup – 0288. https://doi.org/10.5860/choice.43sup-0288.
(31) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews*. 1997, pp 3–25. https://doi.org/10.1016/s0169-409x(96)00423-1.
(32) *Schrödinger Release 2019-4: LigPrep, Schrödinger, LLC, New York, NY, 2019*.
(33) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*. 2004, pp 1739–1749. https://doi.org/10.1021/jm0306430.
(34) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Thomas Pollard, W.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*. 2004, pp 1750–1759. https://doi.org/10.1021/jm030644s.
(35) Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Acc. Chem. Res.* **2017**, *50* (7), 1625–1632.
(36) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; et al. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1863–1874.
(37) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49* (2), 377–389.
(38) Halgren, T. New Method for Fast and Accurate Binding-Site Identification and Analysis. *Chem. Biol. Drug Des.* **2007**, *69* (2), 146–148.
(39) Merk, D.; Friedrich, L.; Grisoni, F.; Schneider, G. De Novo Design of Bioactive Small Molecules by Artificial Intelligence. *Mol. Inform.* **2018**, *37* (1-2). https://doi.org/10.1002/minf.201700153.
(40) Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.; Pande, V. MoleculeNet: A Benchmark for Molecular Machine Learning. *Chem. Sci.* **2018**, *9* (2), 513–530.
(41) Dixon, S. L.; Duan, J.; Smith, E.; Von Bargen, C. D.; Sherman, W.; Repasky, M. P. AutoQSAR: An Automated Machine Learning Tool for Best-Practice Quantitative Structure–activity Relationship Modeling. *Future Med. Chem.* **2016**, *8* (15), 1825–1839.
(42) Hussain, J.; Rea, C. Computationally Efficient Algorithm to Identify Matched Molecular Pairs (MMPs) in Large Data Sets. *Journal of Chemical Information and Modeling*. 2010, pp 339–348. https://doi.org/10.1021/ci900450m.
(43) RDKit: Open-source cheminformatics http://www.rdkit.org.
(44) *Schrödinger Release 2019-4: Canvas, Schrödinger, LLC, New York, NY, 2019*.
(45) Ryckmans, T.; Edwards, M. P.; Horne, V. A.; Correia, A. M.; Owen, D. R.; Thompson, L. R.; Tran, I.; Tutt, M. F.; Young, T. Rapid Assessment of a Novel Series of Selective CB(2) Agonists Using Parallel Synthesis Protocols: A Lipophilic Efficiency (LipE) Analysis. *Bioorg. Med. Chem. Lett.* **2009**, *19* (15), 4406–4409.

(46) Edwards, M. P.; Price, D. A. Role of Physicochemical Properties and Ligand Lipophilicity Efficiency in Addressing Drug Safety Risks; Annual Reports in Medicinal Chemistry; Elsevier, 2010; Vol. 45, pp 380–391.
(47) Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* **2007**, *6* (11), 881–890.
(48) Priam, R. Symmetric Generative Methods and tSNE: A Short Survey. *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. 2018. https://doi.org/10.5220/0006684303560363.
(49) Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The Role of Ligand Efficiency Metrics in Drug Discovery. *Nature Reviews Drug Discovery*. 2014, pp 105–121. https://doi.org/10.1038/nrd4163.
(50) Leeson, P. D.; Young, R. J. Molecular Property Design: Does Everyone Get It? *ACS Medicinal Chemistry Letters*. 2015, pp 722–725. https://doi.org/10.1021/acsmedchemlett.5b00157.
(51) Coxon, C. R.; Anscombe, E.; Harnor, S. J.; Martin, M. P.; Carbain, B.; Golding, B. T.; Hardcastle, I. R.; Harlow, L. K.; Korolchuk, S.; Matheson, C. J.; et al. Cyclin-Dependent Kinase (CDK) Inhibitors: Structure–Activity Relationships and Insights into the CDK-2 Selectivity of 6-Substituted 2-Arylaminopurines. *Journal of Medicinal Chemistry*. 2017, pp 1746–1767. https://doi.org/10.1021/acs.jmedchem.6b01254.
(52) Flick, A. C.; Leverett, C. A.; Ding, H. X.; McInturff, E.; Fink, S. J.; Helal, C. J.; O'Donnell, C. J. Synthetic Approaches to the New Drugs Approved During 2017. *J. Med. Chem.* **2019**, *62* (16), 7340–7382.
(53) Baell, J.; Walters, M. A. Chemistry: Chemical Con Artists Foil Drug Discovery. *Nature* **2014**, *513* (7519), 481–483.
(54) Schönherr, H.; Cernak, T. Profound Methyl Effects in Drug Discovery and a Call for New C-H Methylation Reactions. *Angewandte Chemie International Edition*. 2013, pp 12256–12267. https://doi.org/10.1002/anie.201303207.
(55) ChEMBL Database https://www.ebi.ac.uk/chembl/ (accessed Jan 28, 2020).