Journal of Medicinal Chemistry

pubs.acs.org/jmc

Article

# Novel Efficient Multistage Lead Optimization Pipeline Experimentally Validated for DYRK1B Selective Inhibitors

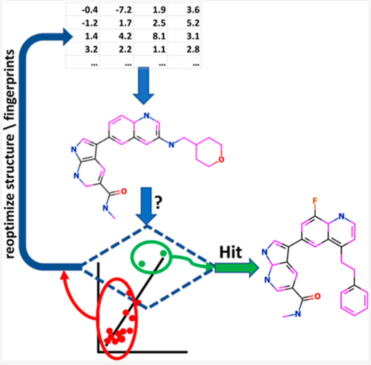
Vadim Alexandrov, Maria Vilenchik, Omar Kantidze, Nika Tsutskiridze, Daviti Kharchilava, Pema Lhewa, Aleksandr Shishkin, Yuriy Gankin,*,# and Alexander Kirpich#

Cite This: J. Med. Chem. 2022, 65, 13784−13792

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

**ABSTRACT:** In addition to general challenges in drug discovery such as the identification of lead compounds in time- and cost-effective ways, specific challenges also exist. Particularly, it is necessary to develop pharmacological inhibitors that effectively discriminate between closely related molecular targets. DYRK1B kinase is considered a valuable target for cancer-specific mono- or combination chemotherapy; however, the inhibition of its closely related DYRK1A kinase is not beneficial. Existing inhibitors target both kinases with essentially the same efficiency, and the unavailability of the DYRK1B crystal structure makes the discovery of DYRK1B-specific inhibitors even more challenging. Here, we propose a novel multi-stage compound discovery pipeline aimed at *in silico* identification of both potent and selective small molecules from a large set of initial candidates. The method uses structure-based docking and ligand-based quantitative structure−activity relationship modeling. This approach allowed us to identify lead and runner-up small-molecule compounds targeting DYRK1B with high efficiency and specificity.



## ■ INTRODUCTION

One of the major goals in drug discovery is shortening the overall lead compound identification and optimization synthesis cycle. It is critical to force a relatively low number of iterations and the number of compounds on each iteration to correctly identify the most potent compounds from patentable scaffolds experimentally found to contain marginally effective small molecules for a target of interest. In this context, potency indicates a quantitative experimental measure of a compound's ability to selectively inhibit *in vitro* either a molecular target or biological process as well as additional pharmacokinetic properties such as high solubility, low toxicity, high microsomal stability, and so forth.

Although the preliminary choice of desired scaffolds limits the size of the initial compound set, the final number of small molecules can be in hundreds or more.[1] Because the sequential generation and versatile *in vitro* evaluation of each small molecule is time- and cost-consuming, the capability to accurately narrow down the initial set of compounds for experimental evaluation is of paramount importance. Thus, a reliable pipeline that can help to decrease *in silico* the number of compounds for synthesis and evaluation would be a great saver of time and resources. Such a pipeline (i) should be efficient computationally (the algorithm should be easy to run within a reasonable time frame) and (ii) should reliably determine the optimal compound in the fewest number of iterations with the fewest number of compounds in each prediction−synthesis−evaluation iteration. When scaffolds cannot be easily enumerated, the *in silico* optimization procedure should be able to suggest structurally similar *de novo* compounds, which would be predicted to have higher potency and good druggability.

The first QSAR iteration (ligand-based modeling) usually starts in its training set with the compounds that already have some experimental information on their efficiency. Otherwise, such initial compounds would also have to be identified *in silico* by approximating their potency by binding affinity to the target and selecting ones with the highest binding affinity. The selected small initial subset with the highest computed affinity values is then evaluated *in vitro* for selective potency (sP); these experimentally measured potency values would be the actual values to be modeled on the subsequent QSAR iteration(s). It is worth noting that the calculated binding affinities can serve only as very crude surrogate predictors for experimental compound potencies,[2] so one should not expect high predictive power at this step. However, once the actual measured potency values enter the QSAR model, higher predictive power is usually achieved. For all the subsequent QSAR iterations, the compounds from the screening universe with the highest predicted potencies are selected for experimental validation, and after their validation the training

**Received:** June 21, 2022
**Published:** October 14, 2022

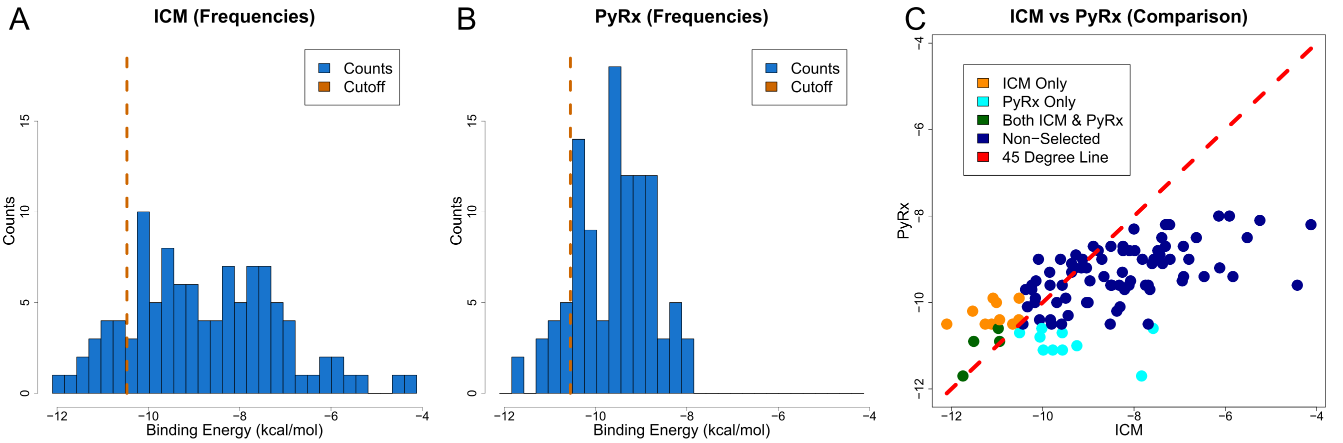


**Figure 1.** Histogram of the binding affinities for the ligands docked by ICM (panel A) and PyRx (panel B) software to the DYRK-1B homology model. The sets of compounds with the highest predicted affinities were 15 (for ICM) and 14 (for PyRx), which were separated by vertical bars. Those sets were selected for the first QSAR iteration. Panel C displays the comparison between the sets of compounds selected by ICM (orange and green) and PyRx (green and cyan), respectively, as well as the correspondence between binding affinities returned by each method.

set for the next iteration is updated with these compounds and their experimental values. The procedure is stopped when the algorithm fails to produce higher predicted potency values for the rest of the screening universe. For validation purposes, one can still evaluate the rest of the universe to see if the most potent compounds were indeed selected by such an optimization procedure.

With these requirements in mind, we designed the QSAR pipeline, an open-source multi-stage compound discovery algorithm. The initial screening stage is a structure-based docking performed on a large set of candidates based on their selected scaffolds and desired binding properties. The selected small subset with the highest computed affinity values is validated *in vitro* for potency and serves as an input for the first ligand-based QSAR iteration. The measured sP is a response (modeled) variable, and fingerprint values for each corresponding compound are the input features (dependent variables) for the machine learning-based QSAR algorithm. The optimized QSAR model is applied to predict the potency of the remaining (not tested *in vitro*) compounds from the original compound universe. At every subsequent QSAR iteration, the small subset of compounds with the highest potency predicted by the algorithm is selected, experimentally evaluated, and then added to the training set for the next iteration.

This procedure has been applied to optimize the dual-specificity tyrosine phosphorylation-regulated kinase 1B (DYRK1B) inhibitors, which was challenging due to the unavailability of the DYRK1B crystal structure and its structural similarity with DYRK1A kinase.[3] DYRK1B is overexpressed in several cancer types and maintains cellular quiescence.[4] Moreover, DYRK1B can enhance cancer cell survival by upregulating antioxidant gene expression and reducing intracellular levels of reactive oxygen species.[5,6] It is well documented that either genetic depletion of DYRK1B or its pharmacological inhibition leads to the cell cycle re-entry and apoptosis of DYRK1B-expressing quiescent cancer cells,[7−9] which brings DYRK1B inhibitors in focus of novel highly potent cancer therapies. At the same time, the inhibition of the DYRK1A kinase, which has high homology to DYRK1B in its active site yet different functions and tissue distribution, bears risks of adverse drug reactions because DYRK1A is expressed in multiple tissues and is involved in a plethora of housekeeping cellular processes.[10] Here, using the QSAR pipeline, we identified the lead and two runner-up compounds in just one docking-based screening, followed by two QSAR iterations with less than 15 compounds selected on each step, bringing the total number of experimentally evaluated compounds to less than 50. To verify the procedure, the entire initial data set of candidate inhibitors was evaluated *in vitro* for DYRK1B/DYRK1A potency and selectivity, and the results were compared with the predictions from the pipeline. It turned out that the most potent selective inhibitor (lead) was correctly identified as well as the two runner-ups. Moreover, the rest of the compounds from the tested universe were less potent than the top compounds correctly identified by our procedure, proving the correct stoppage point.

## ■ RESULTS

**Selecting the Initial Set of DYRK1B Ligands for Subsequent QSAR Modeling.** For the very first iteration, the goal was to select 10−15 compounds (<10% of the initial scaffold set containing 164 small molecules) with the highest selective binding affinities. The exact number of selected compounds in the batch has been determined based on the resulting distribution of the binding energies and our medicinal chemist's experience and recommendation. Retrospectively, the sensitivity analysis showed that the final number of tested compounds and QSAR iterations was quite robust for this initial selection and had a detrimental effect only if the number of the initially selected compounds was less than 10. The binding affinities were produced by PyRx and ICM (Figure 1A,B). A comparison of the computed binding affinities obtained by PyRx and ICM for each compound showed that the results for the two methods have been very similar (Figure 1C). Moreover, a steep drop in the histogram of affinities helped to determine the cutoff point making the compound selection process easier and to produce the desired sets of 15 (ICM) and 14 (PyRx) compounds.

The subsequent *in vitro* evaluation revealed that less than 50% of selected compounds produced by either PyRx or ICM were active. There was not a single meaningfully potent leading compound in the subset identified by homology modeling/docking at this step. Such results were quite expected by the authors because (i) the binding affinities are usually poorly translated to the actual selective inhibition potencies and (ii) DYRK1B potencies were obtained by homology modeling, which introduced an additional source of error. Nevertheless, these initial compound sets formed good starting points for QSAR model training. Their 3D structural features helped to avoid selecting decoys (non-active counterpart compounds) in the follow-up iterations. It is also worth noting that both ICM and PyRx starting points, which were somewhat different (Figure 1C), eventually led to the same number of QSAR iterations and approximately the same number of finally selected compounds (44 and 47, respectively) in the last algorithm iteration.

**Iterative Selection of Compounds to Identify the Lead.** The appropriate number of compounds in each iteration was chosen to be 10−15, which was supposed (i) to be sufficient to instruct the algorithm on what improves the desired potency and (ii) to keep the algorithm's cost-effectiveness. While the chosen number of compounds was similar to the earlier iterative QSAR studies,[11] the number of iterations themselves used here was much less compared to the 10 iterations needed to identify the lead compound in the above-mentioned study. Our approach also agrees with earlier findings where the set of 10 compounds was found to be the minimal meaningful input for model refinement.[12]

**Importance of Shape-Based Fingerprints.** We found that excluding shape-based descriptors (fingerprints) from the set of features would add one more iteration to the overall optimization cycle, which is undesired. On the contrary, when fingerprints were included, the resulting percentage of the ODDT features, that is, the prevalence of the shape-based descriptors on the first QSAR iteration was 24% (6 out of 25) and 43% (3 out of 7), respectively. This was much larger than would be expected under random selection (15 out of 338 resulting in 4%). In summary, the usage of shape-based descriptors implied fewer cross-validation (CV) iterations, which improved computational efficiency, and demonstrated that 3D features were more informative and important because their selection proportion was much higher than the one for the non-3D features.

**Compound Summaries from Each QSAR Iteration.** From the initial (first) iteration, the 15 compounds with the highest ICM-predicted binding affinity were selected to seed the QSAR model (Table 1). The corresponding inhibition potencies for DYRK1A and DYRK1B of the synthesized compounds, which were measured *in vitro* and calculated according to formula 1 are also provided in Table 1. These measured potencies were used as the QSAR model training inputs. There is a discrepancy between the PyRx/ICM-predicted energies and evaluated *in vitro* potencies, which indicates that the initial set of compounds can be selected either by PyRx or ICM. This, however, does not guarantee the potent set of initial compounds and can serve only as the algorithm's starting point. Overall, the results for the ICM and PyRx predictions were essentially the same—some compounds with high binding energies were potent and some were not.

For the second iteration, the fitted "optimal" model potency predictions from the first iteration out of the entire 164-compound universe of potential inhibitors are summarized in Table 2. In the same way, as for the first iteration, the top 15 compounds with the highest predicted potencies were experimentally evaluated for their actual potencies (Table 2). The actual potencies from the first and second iterations were used as the modeled response values in the training set for the third iteration.

Table 1. *First iteration* Summaries: the Choice of Compounds with the Highest Predicted Selective Affinity[a]

| ID | E(DYRK1B)-E(DYRK1A) | measured sP |
|---|---|---|
| Compound9163 | −1.2 | −0.03 |
| Compound9069 | 0.1 | −0.55 |
| Compound8814 | 0.2 | 0.3 |
| Compound9160 | 0.2 | −1.44 |
| Compound9630 | 0.2 | −1 |
| Compound9965 | 0.3 | −1.41 |
| Compound9716 | 0.3 | 0.4 |
| Compound10731 | 0.3 | −1.18 |
| Compound9402 | 0.4 | −0.75 |
| Compound3702 | 0.4 | 0.49 |
| Compound9070 | 0.4 | −2.16 |
| Compound9162 | 0.4 | −0.47 |
| Compound9411 | 0.4 | −0.85 |

[a] Here, deltaEnergy is defined as the ligand's predicted binding affinity (DYRK1B) minus predicted binding affinity (DYRK1A). Potency is a sP as defined in formulas 1 and 2. The results are presented for ICM, while the results for PyRx are similar.

Table 2. Second Iteration Summaries: Selection of Compounds with the Highest Predicted sP[a]

| ID | measured potency | predicted potency |
|---|---|---|
| Compound9447 | 0.88 | 0.88 |
| Compound10289 | 0.4 | 0.52 |
| Compound9716 | 0.4 | 0.4 |
| Compound9422 | −0.59 | 0.32 |
| Compound8814 | 0.3 | 0.3 |
| Compound9445 | 0.65 | 0.29 |
| Compound9790 | −0.91 | 0.27 |
| Compound9394 | −0.86 | 0.27 |
| Compound9449 | 0.6 | 0.26 |
| Compound3702 | 0.49 | 0.26 |
| Compound3421 | 0 | 0.21 |
| Compound9397 | 0.49 | 0.21 |

[a] The experimentally measured ground truth calculated according to eqs 1 and 2 is provided in the sP column. The cutoff was made at the 12th compound to make the total number of lab-tested compounds less than or equal to 25.

The results of the third (most recent) iteration are summarized in Table 3. Because the model on the third iteration did not identify any additional compounds with higher potency than those already present in the training set from previous iterations, the resulting model was declared to be final. To validate such a statement, the rest of the compounds in the 164-compound universe, which had not been evaluated during the above-mentioned iterations, were synthesized and experimentally evaluated to test the predictive power of the model.

In summary, all "top hits" from the universe, that is, compounds with the highest actual selective potencies have been identified correctly during the final iteration. For a visual illustration, the experimentally confirmed top hits from the final iteration are presented in Figure 2A−C. One can see that they all belong to the same 7-azaindole-quinoline scaffold, containing a carboxylic amide substituent at position five of the azaindole moiety and substituents at positions 3, 4, and/or eight of the quinoline. The rest of the initial universe not

**Table 3. Third (Final) Iteration Summaries: the Resulting Set Has Only Two Inactive Compounds, While all Potent Compounds within the Universe of 164 Were Extracted**[a]

| ID | measured potency | predicted potency |
|---|---|---|
| Compound9447 | 0.88 | 0.88 |
| Compound9445 | 0.65 | 0.65 |
| Compound9449 | 0.6 | 0.6 |
| Compound9857 | 0.85 | 0.55 |
| Compound9799 | 1 | 0.49 |
| Compound9397 | 0.49 | 0.49 |
| Compound9401 | 0.65 | 0.41 |
| Compound10289 | 0.4 | 0.4 |
| Compound9716 | 0.4 | 0.4 |
| Compound3702 | 0.49 | 0.35 |
| Compound8814 | 0.3 | 0.3 |
| Compound9446 | 0.54 | 0.26 |
| Compound3457 | 0 | 0.25 |
| Compound8796 | −1.7 | 0.11 |
| Compound9465 | 0.88 | 0.07 |
| Compound9659 | 0.89 | 0.04 |
| Compound3421 | 0 | 0 |
| Compound9466 | 0.78 | 0 |
| Compound9163 | −0.03 | −0.03 |
| Compound9549 | 1.06 | −0.05 |
| Compound10315 | 0.78 | −0.1 |

[a] If no validation is involved, the total number of tested compounds is 45. The most potent set of 32 is presented.

selected by the model for processing in the lab had more than 70% of all considered compounds, which substantially reduced the synthesis effort and number of experimental validations. Compounds rejected by the model were indeed found experimentally to be way less potent than the selected ones with an average measured sP of −0.1 in comparison to an average of 0.3 for the compounds selected on the final iteration.

## DISCUSSION

Despite the promising results, the proposed pipeline has some inherited limitations, which are coming from multiple sources and which are discussed in this section.

**Structure-Based Screening Stage and Homology Modeling Limitations.** The usual culprits limiting the reliability of the in-silico binding scores are as follows: the forcefield and the scoring function,[13] uncertainties resulting from modeling potential water-mediated interactions[14,15] and of course, the use of a homology model itself, which would conventionally be more suitable for virtual screening rather than lead optimization.[13,16,17] Homology modeling can have multiple inherited issues from modeling errors to difficulties in translating predicted binding affinities to the observed inhibition concentrations.[18] Moreover, the above-mentioned issues are expected to be more frequent because our modeled response variable is composed of the two predicted components (binding affinities from docking calculations for both DYRK1A and DYRK1B for the initial set and then the predicted pIC50 values as a combination of individual pIC50 values for both DYRK1A and DYRK1B). Because the structures of both targets are also very similar, it greatly influences the accuracy of the resulting "difference" model.[19]

**Feature Selection Potential Biases.** Feature selection can have great influence on the model performance[20] while in general, a comprehensive evaluation of all possible feature combinations is prohibitively expensive. However, for the methods that we present in this study (SVR, ensembles of decision trees, and Gaussian processes), robust heuristics exist that produce reasonable results in most cases.[21] Nevertheless, if the data set dimensions grow significantly larger than in the present study, the sampling of feature space can become either a computational bottleneck or can result in suboptimal model selection.

**Small Sample Size Limitations and Potential for Overfit.** The desire to perform lead optimization in the most efficient manner, that is, with as few iterations and as few compounds in each iteration as possible, inevitably leads to a small training set size and potential for an overfit, given that the number of features exceeds the number of samples in the model. Keeping the number of selected features fewer than the number of samples along with repeated extensive CV of the resulting models can attenuate this problem to some extent but cannot make the setup, which will be universally generalizable to every imaginable lead optimization program. As a consequence, for the foreseeable future, human input (e.g., from a trained medicinal chemist who is an expert in the field regarding selection of patentable scaffolds and molecular building blocks) will still be needed.

**Use of AlphaFold and Similar AI Approaches at the First (Structure-Based Screening) Iteration.** In light of the recent AI-based structure prediction developments (e.g., AlphaFold[22]), we envision that our lead optimization pipeline could indeed benefit in those situations when the target's crystal structure is not available and the closest homology model would have a sequence similarity of 50% or even lower (esp. when there is large uncertainty in the target's binding site region) to fully utilize the predictive and generalization power of AlphaFold.[23] In the present work, however, our target structure (DYRK1B) had 84% identity with a DYRK1A template in the N-terminus and the catalytic domain, so we

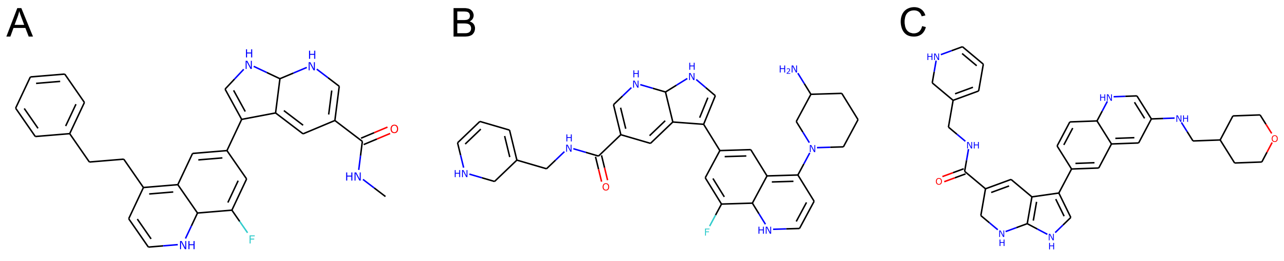


Figure 2. Experimentally confirmed top hits from the *final iteration* with identifiers are presented for compound #8548 (panel A), for compound #8658 (panel B), and for compound #8464 (panel C).

would not expect to gain much benefit in this particular situation from using a pure AI approach. Note also that the essence of the presented pipeline here is based upon the QSAR power on the subsequent ligand-only-based iterations and that structure-based methods alone in general are more suitable for screening rather than optimization as mentioned above.

## ■ CONCLUSIONS

The proposed pipeline has shortened the optimization cycle to only three iterations on subsets of size <10% each and has reduced the required number of compounds to synthesize by 70%. The method has addressed both potency and selectivity and has been successfully illustrated for a challenging DYRK1A/DYRK1B problem, where DYRK1B crystal structures were not available and there was a need to refer to homology modeling. Overall, thorough feature selection and model optimization has been deemed essential and the algorithm resulted in a synthesis of new potent and selective DYRK1B inhibitor compounds in a relatively small number of iterations. It turned out that our method was able to select the top hit and the top five most potent compounds in the entire evaluated universe, proving that it would shorten the lead optimization effort by more than 300%. The entire proposed universe of compounds has been evaluated in the lab to validate the pipeline output. Moreover, the proposed pipeline is not bounded by the presented universe of compounds used for current evaluation and can be used to suggest compounds outside of this set for future hit and lead optimization efforts.

## ■ EXPERIMENTAL SECTION

**Algorithm Inputs: Evaluation of the DYRK1B Inhibitor Binding Affinities.** When the crystal structure of the encoded protein is known, the evaluation of docking properties and binding affinities is a straightforward task.[24,25] Unfortunately, because the target protein crystal structure for DYRK1B is unknown,[19] the evaluation of binding affinities also becomes challenging from the computational chemistry perspective.

The workaround for DYRK1B is to refer to similar and known crystal structures of DYRK1A target protein and to use a homology modeling approach to "predict" the DYRK1B target protein structure.[26] The proposed homology modeling approach is plausible because the resulting binding pockets of DYRK1B and DYRK1A are (expected) to be very similar.

The homology modeling was performed *via* ICM.[27] The resulting DYRK1B structure was derived *via* an ICM homology modeling protocol with a reference 4yll structure from a protein data bank,[28] representing DYRK1A as the starting point. DYRK1A is the closest sequence to DYRK1B according to the results of a blast search with a similarity of 85%. DYRK1A (4yll) PDB entry required pre-processing as the deposited structure has bonds to ligand it was crystallized with, 10-bromo-2-iodo-11H-indolo[3,2-*c*]quinoline-6-carboxylic acid (transferase inhibitor). The ligand was separated from the remaining target structure, which was further optimized for the next step. Alignment was made between the DYRK1A sequence (from the actual PDB 4yll record) and the DYRK1B uniprot Q9Y463 sequence.[29] The gaps between amino acids were blocked to prevent the structure from any insertions and homology modeling with the full model builder tool was used to construct the DYRK1B molecule from alignment results (the insert margin value = 2, deletion margin value = 3).

For docking purposes, PyRx[30] and also ICM (for consistency comparison) were applied for the derived DYRK1B structure, which resulted from homology modeling. The obtained docking results from two software suites were subsequently compared for consistency. PyRx is a convenient user interface, which utilizes AutoDock Vina, as a docking engine.[31] Autodock Vina is an open-source docking engine software routinely used for docking tasks and benchmark comparisons for more than a decade.[32−35] In the meantime, ICM is a direct commercial competitor of the free AutoDock software.[36]

In the ICM docking procedure, we used the standard all-atom vacuum force field ECEPP/3 with appended terms to account for solvation free energy and entropic contribution. Ligand conformational sampling was performed based on the ICM's default biased probability Monte Carlo (BPMC) procedure as described in the literature.[37]

For AutoDock Vina, we used the standard Vina force-field and AutoDock4.2 scoring function described and cross-validated in detail in the literature[38] in a standard rigid-receptor setup. All ligands were treated as flexible.

The docking results defined the optimization pipeline's starting point and provided a good input for the first QSAR iteration. In particular, the small subset (<10% of the initial set) with the highest predicted binding affinities to DYRK1B and low affinities to DYRK1A were evaluated in the lab for their selective potencies to inhibit DYRK1B and DYRK1A, and these experimental measures were subsequently used as a modeled response on the first QSAR iteration (with compounds' fingerprints as predictor values, respectively).

**Determination of the Inhibitory Activity *In Vitro*.** The protocol for the determination of the inhibitory activity *in vitro* has been used earlier and the details have been both published[19] and submitted for a patent.[39] In brief, the compounds of the present invention were tested for their inhibitory activity against both DYRK1A and DYRK1B. Multiple measurements were taken to increase the precision of measurements. Later, the average (mean) value of the measurements was taken.

The DYRK-inhibitory activity of the compounds was tested using the ADP-Glo assay.[40] In particular, the procedure for determining the IC50 values with the ADP-Glo assay *in vitro* kinase assays consisted of two sequential parts: (i) kinase reaction performed under optimized conditions and (ii) detection of ADP as a product of the reaction using the ADP-Glo system (Promega). The tested compounds listed in Tables 4 and 5 are dissolved in dimethyl sulfoxide (DMSO), then transferred to the V-bottom PP plate to perform 9 serial dilutions (in order to obtain dose−response curves) in 25% DMSO.

**Table 4. Optimized Conditions for Measuring DYRK1A *In Vitro* Kinase Activity**

| reagent/condition | final concentration |
|---|---|
| buffer | 50 mM Tris, pH 7.5 |
| $MgCl_2$ | 10 mM |
| NaCl | 25 mM |
| DTT | 0,1 mM |
| ATP ($K_m$) (ultrapure, from ADP-Glo ™ kit) | 70 μM |
| substrate ($K_m$): RRRFRPASPLRGPPK (Lipopharm) | 3 μM |
| enzyme—DYRK1A (Carna Bioscience) catalogue no. 04-130 | 2 nM or 0.7 nM |
| time of reaction | 30 min |
| temperature of reaction | rt |

The protocol followed the patent.[39,41] Two mixes were prepared on ice, where Mix 1 contained substrate, ATP and reaction buffer while Mix 2 contained reaction buffer and the kinase. In particular, 15 μL per well of Mix 1 was transferred to wells of a 96-well plate. Next, 2.5 μL of the pre-diluted tested compound was added to Mix 1, followed by the addition of 12.5 μL of Mix 2. The total reaction volume was 30 μL per well. The experiment was duplicated twice for each data point being examined. The estimated dose−response curve for the positive control was carried out on each assay plate by adding the reference inhibitory compound staurosporine. In addition to this, three controls were also included for each test: (i) 30 μL of the reaction mixture containing the reaction buffer, ATP, kinase, and DMSO without substrate (quasi-positive control); (ii) 30 μL of the reaction mixture containing the reaction buffer, substrate, ATP, and DMSO without

Table 5. Optimized Conditions for Measuring DYRK1B *In Vitro* Kinase Activity

| reagent/condition | final concentration |
|---|---|
| Buffer | 5 mM MOPS, pH 7.5 |
| $MgCl_2$ | 5 mM |
| EDTA | 0.4 mM |
| DTT | 1 mM |
| ATP ($K_m$) (ultrapure, from ADP-Glo $^{TM}$ kit) | 15 $\mu$M |
| substrate ($K_m$): RRRFRPASPLRGPPK (Lipopharm) | 7 $\mu$M |
| enzyme—DYRK1B (Carna Bioscience) catalogue no. 04-131 | 1 nM or 0.3 nM |
| time of reaction | 1 h |
| temperature of reaction | Rt |

kinase (background control); and (iii) 30 $\mu$L of the reaction mixture containing the reaction buffer, substrate, kinase, ATP, and DMSO (vehicle control). The final concentration of DMSO in the reaction was 2%. To detect the ADP amount produced during the kinase reaction, the commercially available kit ADP-Glo Kinase assay (Promega, cat. No # V9103) was used. The protocol used for the detection was based on the Technical Bulletin of the ADP-Glo Kinase assay (Promega) and was adapted to a 96-well plate containing a 30 $\mu$L reaction mixture.

More precisely, 30 $\mu$L of ADP-Glo reagent was added to each well of a 96-well plate containing 30 $\mu$L of reaction mixture to terminate the kinase reaction and to deplete the remaining ATP. The plate was incubated for 60 min on a shaker at room temperature (RT). Then, the 60 $\mu$L of kinase detection solution was added to each well of 96-well plate containing 60 $\mu$L of the solution to convert ADP to ATP and to allow the newly synthesized ATP to be measured using a luciferase/luciferin reaction (ratio of kinase reaction volume to ADP Glo Reagent volume to kinase detection solution volume was maintained at 1:1:2). The plate was incubated for 40 min on a shaker at RT while protected from light. Luminescence was measured in the plate reader, wherein the luminescent signal is proportional to the ADP concentration produced and thus directly correlated with the kinase activity.

The IC50 values were determined, using the GraphPad Prism 6.0 as log(agonist) *versus* normalized response minus the variable slope after data normalization to controls (complete reaction mix and no-substrate control) after putting the test compound concentrations on a natural logarithmic scale.

**Quantitative Metric for sP.** At every pipeline iteration, the targeted small subset (<10% of the initial set) of compounds is evaluated for sP in the lab. Such a set is either proposed by (i) docking for binding affinity estimation as an input for the first QSAR iteration or (ii) QSAR prediction for the rest of the yet (experimentally) unevaluated compound universe during subsequent iterations.

The desired compound has to be both selective and potent. From the earlier experimental experience,[19] our team could not identify chemical scaffolds, which would be both potent (inhibit DYRK1B at very low ligand concentrations) and selective (inhibit DYRK1B but not so much DYRK1A). Therefore, the metric of interest has to incorporate both potency and selectivity. In particular, DYRK1B inhibition potency should be encouraged while there should also be some penalty within the metric for high DYRK1A inhibition potency. Therefore, the utilized combined potency metric (sP) is defined as follows

$$\mathrm{sP} = \mathrm{pIC50(DYRK1B)} - \mathrm{pIC50(DYRK1A)} \quad \text{if } \mathrm{pIC50(DYRK1B)} > 0 \tag{1}$$

$$\mathrm{sP} = \mathrm{pIC50(DYRK1B)} - \mathrm{abs(pIC50(DYRK1A))} \quad \text{if } \mathrm{pIC50(DYRK1B)} < 0 \tag{2}$$

where pIC50 = −log(IC50) is a log transformation for more uniform distribution (and thus easier modeling) of the response variable (sP). Conditions 1 and 2 penalize the total sP values for potency to DYRK1A, that is, to account for non-potent and/or non-selective compounds.

**Combined Docking and Ligand-Based Machine-Learning Approach for QSAR.** The proposed combined (docking and ligand-based) approach has been encouraged by the literature.[42] This published work justifies the use of combined structure-based docking and ligand-based QSAR (in particular, structural docked pose descriptors) to reduce the rate of false positives in compound optimization.

On the initial iteration, the compounds were selected solely based on their docked binding energies. These initial compounds were included in the following QSAR interactions after their experimental selective potencies were evaluated in the lab. Because 3D and non-3D descriptors (features) naturally contribute to explaining the compounds' activity, the standard StarDrop Automodeler[43] set of features has been augmented with the 3D shape descriptors from open drug discovery toolkit (ODDT),[44] which contributed to additional 15 real-valued features per conformer (docked pose). The very purpose of shape descriptors is to describe the 3D molecular conformation in the most accurate and concise way and in the situations, when the number of training instances is small (our case), the compact feature representation of the training set would provide an additional protection against overtraining and add accuracy. That is exactly what was found in the recent research: Bonanno and Ebejer[45] report a mean enrichment factor improvement of 430% when shape-based descriptors are used in machine learning models rather than in traditional Tanimoto-based virtual screening. In our recent work,[46] we show that it is specifically 3D features that helped us identify compounds with similar biologic activity that would be missed otherwise if only 2D descriptors are used. In this work, we also found that shape descriptors were significantly enriched in the selected feature sets for our QSAR models, which are described in the Discussion section.

Because both ICM and PyRx packages optimize ligands and select the optimal conformation for docking that docked conformation can be saved in the 3D structure data file (SDF) format and fingerprinted. ODDT allows reading molecules in multiple industry accepted formats, including 3D SDF, which consists of multiple chemical table (CTAB) units separated by quadruple dollar signs. This format was developed by molecular design limited (MDL) information systems, Inc that is now a part of Dassault Systems. ODDT allows reading CTAB with one line of Python code

$$\text{molecule\_object} = \text{oddt. toolkit. readstring('sdf', CTAB)} \tag{3}$$

generating an internal "molecule" object from the CTAB string. After that, one can calculate the electroshape object using the following instruction

$$\text{shape} = \text{oddt. shape. electroshape(molecule\_object)} \tag{4}$$

where shape would be just a Python list of 15 floating point numbers.

There were multiple reasons to use ODDT to augment the initial set of 3D features. In particular, shape and electrostatic charge along the surface are missing in most conventional non-3D fingerprints since the exact pose ,that is, docked to the target is often unknown in the ligand-based-only QSAR modeling. This is not acceptable for DYRK1A/DYRK1B studies because both targets' binding pockets and their ligands have similar shapes and the homology model contributes a lot of uncertainty to the binding-only sP estimation. However, once the actually measured pIC50 values are revealed for each binding prediction, the structural features would be able to provide valuable information on the next iteration to the QSAR model regarding what structural features were not contributing to the modeled potency and selectivity. Therefore, we generated 3D shape fingerprints using the standard ODDT package first and added them to the StarDrop Automodeler set, which then collectively empowered the feature selection algorithm to do the selection of only those fingerprint components that would bring the most predictive power to the QSAR

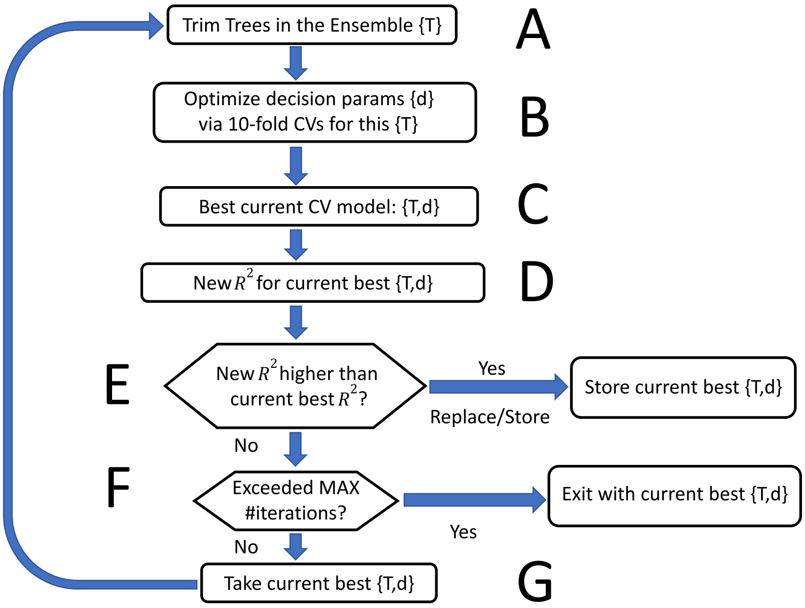

Trim Trees in the Ensemble {T}
A
Optimize decision params {d} via 10-fold CVs for this {T}
B
Best current CV model: {T,d}
C
New $R^2$ for current best {T,d}
D
E
New $R^2$ higher than current best $R^2$?
Yes
Replace/Store
Store current best {T,d}
No
F
Exceeded MAX #iterations?
Exit with current best {T,d}
Yes
No
Take current best {T,d}
G

**Figure 3.** Summary diagram of the QSAR algorithm for each iteration. The diagram summarizes the repeated optimization of the decision tree ensemble model using CV (for parameters tuneup) and tree size trimming *via* feature selection. The features and the size of the trees on each iteration are selected as implemented in the Caret package.[21] The diagram summarizes every algorithm step: (A) specification of decision tree ensemble, (B) optimization of parameters for the trees, (C) selection of the best model based on the CV approach, (D) computation of $R^2$ for the current algorithm iteration, (E) computation of $R^2$ for the current algorithm iteration and comparison with the previous iteration, (F) evaluation of the algorithm stopping conditions, and (G) move to the next iteration if the stopping conditions are not met.

model. As a result, a combination of enriched conformer-specific 3D fingerprints with the initial set of non-binary general-purpose 3D fingerprints, through proper feature selection and extensive CV resulted in the optimal model, from the $R^2$ metric perspective and robust informative sets of features that are paramount to use for small sample sizes.

Because the desired goal of the proposed pipeline is to minimize the number of tested compounds (*i.e.*, the number of samples in the respective training set) during each iteration, the preliminary step is to find the shortest numerical representation of each compound shape to avoid over-training problems. Typically, the lengths of a fingerprint can range from 128 to 4096 bits; however, much shorter alternatives are available. In particular, the ElectroShape method,[47] which is implemented in the ODDT package, is based on the algorithm that incorporates shape, chirality, and electrostatics, and represents each conformer *via* a fixed-length vector of only 15 real-valued numbers. This short representation, however, is quite powerful and predictive, which has been confirmed by recent ligand-based screening studies.[45] Therefore, this 15-dimensional 3D shape representation as implemented in the ODDT package has been chosen for the proposed pipeline as the most condensed one. The proposed method, however, is not restricted by such choice and if longer fingerprint-length alternatives are desired they can be used. The example of such an alternative includes the E3FP package, which also utilizes an alignment-invariant 3D representation of molecular conformers as a fixed-length (2048) binary vector for each conformer[48] as well as 48-dimensional spectrophores.[49]

**Feature Selection and Cross Validation for Building Robust Predictive Models.** The proposed QSAR machine learning-based optimization procedure uses the augmented set of features and automatically selects a subset from the set to build the best predictive model based on the sP metrics (1) and (2). The "best" model in this context is determined by predictive power in the $R^2$ sense and is based on the extensive CV procedure to prevent overfitting. The summaries of the machine learning algorithm are outlined in Figure 3A−G.

Overfitting may be a problem for situations when the number of features (337-dimensional fingerprint vector) far exceeds the number of compounds in the training set (164 total, while only up to 50 compounds in any given QSAR model). To prevent overfitting, we followed the repeated grid-search CV methodology described in the literature.[50] More precisely, the data set has been divided randomly into $K$ folds, and the model has been refitted $K$ times with the data points in each fold withheld in turn from the training set (non-stratified CV). No stratification was used because Breiman and Spector[51] reported no improvement of stratified CV *versus* non-stratified CV in a regression setting. Also, ref 50 has found that with a large number of repeated CVs, the issue of stratification becomes redundant.

Parameter tuning has been performed following the repeated grid-search approach also described in ref 50. CV has been used to select both the optimal number of parameters and their values. On each optimization step, feature selection has been performed as implemented in the Caret package.[52,53] The grid search CV optimization procedure was repeated until the $R^2$ quality metric[54] had converged. The $R^2$ was considered converged if the observed change in the value during iterations was less than $10^{-4}$. The optimization procedure is schematically depicted in Figure 3A−G.

According to our knowledge, the value and the importance of repeated CV on a grid search has not been extensively explored and discussed in the literature. We believe that this can be partially explained by the associated high computational costs, which are expected during implementation. Those computational costs, however, are "manageable" provided the size of the investigated data set is relatively small (164 compounds, 388 features per compound), which is exactly the case here because we are performing lead optimization on a purposefully small number of compounds.

## ■ COMPETING FINANCIAL INTERESTS

The protocol has been submitted for a patent.[39,41] The identified compounds have also been patented. The patented products can be a source of income for a commercial for-profit company Felicitex Therapeutics, Inc.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.2c00988.

> Homology model files which include the DYRK1A/DYRK1B sequences alignment file (extension: aln) and the DYRK1B predicted structure file (extension: pdb), SDF structure files for the lead compounds (extension: sdf), and SMILES structure files for the lead compounds (extension: csv)[55] (ZIP)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Yuriy Gankin** − *Quantori LLC, Cambridge, Massachusetts 02142, United States;* orcid.org/0000-0003-0046-1037; Email: yuriy.gankin@quantori.com

### Authors

**Vadim Alexandrov** − *Liquid Algo LLC, Hopewell Junction, New York 12533, United States*

**Maria Vilenchik** − *Felicitex Therapeutics, Natick, Massachusetts 01760, United States*

**Omar Kantidze** − *Quantori LLC, Cambridge, Massachusetts 02142, United States;* orcid.org/0000-0002-7507-7307

**Nika Tsutskiridze** − *Quantori LLC, Cambridge, Massachusetts 02142, United States; Tbilisi State Medical University, Tbilisi 0186, Georgia*

**Daviti Kharchilava** − *Quantori LLC, Cambridge, Massachusetts 02142, United States; Tbilisi State Medical University, Tbilisi 0186, Georgia*

**Pema Lhewa** − *Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia 30303, United States*

**Aleksandr Shishkin** − *Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia 30303, United States*

**Alexander Kirpich** − *Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia 30303, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jmedchem.2c00988

### Author Contributions

#Y.G. and A.K. contributed equally.

### Notes

The authors declare the following competing financial interest(s): The protocol has been submitted for a patent. The identified compounds have also been patented. The patented products can be a source of income for a commercial for-profit company Felicitex Therapeutics, Inc.

## ■ ACKNOWLEDGMENTS

The authors would like to thank Michael Frid for his assistance with compound synthesis and evaluation as well as for his help with the manuscript preparation.

## ■ ABBREVIATIONS USED

AI, artificial intelligence; BPMC, biased probability Monte Carlo; CTAB, chemical table; CV, cross validation; DYRK1A, dual specificity tyrosine-phosphorylation-regulated kinase 1A; DYRK1B, dual specificity tyrosine-phosphorylation-regulated kinase 1B; E3FP, Extended 3-Dimensional FingerPrints; IC50, half-maximal inhibitory concentration; MDL, molecular design limited information systems; ODDT, open drug discovery toolkit; PDB, Protein Data Bank; QSAR, quantitative structure−activity relationships; SDF, structure-data file

## ■ REFERENCES

(1) Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162*, 1239.

(2) Jansson-Löfmark, R.; Hjorth, S.; Gabrielsson, J. Does In Vitro Potency Predict Clinically Efficacious Concentrations? *Clin. Pharmacol. Ther.* **2020**, *108*, 298−305.

(3) Becker, W. A wake-up call to quiescent cancer cells - potential use ofDYRK1B inhibitors in cancer therapy. *FEBS J.* **2018**, *285*, 1203−1211.

(4) Friedman, E. Mirk/Dyrk1B in Cancer. *J. Cell. Biochem.* **2007**, *102*, 274−279.

(5) Deng, X.; Ewton, D. Z.; Friedman, E. Mirk/Dyrk1B Maintains the Viability of Quiescent Pancreatic Cancer Cells by Reducing Levels of Reactive Oxygen Species. *Cancer Res.* **2009**, *69*, 3317−3324.

(6) Deng, X.; Mercer, S. E.; Sun, C.-Y.; Friedman, E. The Normal Function of the Cancer Kinase Mirk/dyrk1B Is to Reduce Reactive Oxygen Species. *Genes Cancer* **2014**, *5*, 22−30.

(7) Hu, J.; Deng, H.; Friedman, E. A. Ovarian Cancer Cells, Not Normal Cells, Are Damaged by Mirk/Dyrk1B Kinase Inhibition. *Int. J. Cancer* **2013**, *132*, 2258−2269.

(8) Chen, H.; Shen, J.; Choy, E.; Hornicek, F. J.; Shan, A.; Duan, Z. Targeting DYRK1B Suppresses the Proliferation and Migration of Liposarcoma Cells. *Oncotarget* **2017**, *9*, 13154−13166.

(9) Vilenchik, M.; Kuznetsova, A.; Frid, M.; Duey, M.; Damiani, A.; De Leon, L.; Law, J.; Golbin, D. A.; Shishkina, L. V.; Potapova, O. Implication of DYRK1B Kinase in Dormant Glioblastoma Cancers and Utilization of DYRK1B Inhibitors as a Novel Therapeutic Strategy for Glioblastoma. *J. Clin. Oncol.* **2019**, *37*, No. e14670.

(10) Soppa, U.; Becker, W. DYRK Protein Kinases. *Curr. Biol.* **2015**, *25*, R488−R489.

(11) Pickett, S. D.; Green, D. V. S.; Hunt, D. L.; Pardoe, D. A.; Hughes, I. Automated Lead Optimization of MMP-12 Inhibitors Using a Genetic Algorithm. *ACS Med. Chem. Lett.* **2011**, *2*, 28.

(12) Varela, R.; Walters, W.; Goldman, B. B.; Jain, A. N. Iterative Refinement of a Binding Pocket Model: Active Computational Steering of Lead Optimization. *J. Med. Chem.* **2012**, *55*, 8926−8942.

(13) Yunta, M. J. R. Docking and Ligand Binding Affinity: Uses and Pitfalls. *Am. J. Model. Optim.* **2016**, *4*, 74−114.

(14) Shahroz, M. M.; Sharma, H. K.; Altamimi, A. S. A.; Alamri, M. A.; Ali, A.; Ali, A.; Alqahtani, S.; Altharawi, A.; Alabbas, A. B.; Alossaimi, M. A.; Riadi, Y.; Firoz, A.; Afzal, O. Novel and Potential Small Molecule Scaffolds as DYRK1A Inhibitors by Integrated Molecular Docking-Based Virtual Screening and Dynamics Simulation Study. *Molecules* **2022**, *27*, 1159.

(15) Henderson, S. H.; Sorrell, F.; Bennett, J.; Fedorov, O.; Hanley, M. T.; Godoi, P. H.; Ruela de Sousa, R. R.; Robinson, S.; Ashall-Kelly, A.; Hopkins Navratilova, I. H.; Walter, D. S.; Elkins, J. M.; Ward, S. E. Discovery and Characterization of Selective and Ligand-Efficient DYRK Inhibitors. *J. Med. Chem.* **2021**, *64*, 11709.

(16) Chen, Y. C. Beware of Docking! *Trends Pharmacol. Sci.* **2015**, *36*, 78.

(17) Gupta, M.; Sharma, R.; Kumar, A. Docking Techniques in Pharmacology: How Much Promising? *Comput. Biol. Chem.* **2018**, *76*, 210.

(18) Salahudeen, P. S. N.; Nishtala, P. S. An Overview of Pharmacodynamic Modelling, Ligand-Binding Approach and Its Application in Clinical Practice. *Saudi Pharm. J.* **2017**, *25*, 165.

(19) Szamborska-Gbur, A.; Rutkowska, E.; Dreas, A.; Frid, M.; Vilenchik, M.; Milik, M.; Brzózka, K.; Król, M. How to Design Potent and Selective DYRK1B Inhibitors? Molecular Modeling Study. *J. Mol. Model.* **2019**, *25*, 41.

(20) Khan, P. M.; Roy, K. Current Approaches for Choosing Feature Selection and Learning Algorithms in Quantitative Structure-Activity Relationships (QSAR). *Expert Opin. Drug Discovery* **2018**, *13*, 1075.

(21) Kuhn, M.The caret Package. https://topepo.github.io/caret/ (accessed Oct 05, 2022).

(22) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.

(23) David, A.; Islam, S.; Tankhilevich, E.; Sternberg, M. J. E. The AlphaFold Database of Protein Structures: A Biologist's Guide. *J. Mol. Biol.* **2022**, *434*, 167336.

(24) 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*, 1923.

(25) Gorgulla, C.; Boeszoermenyi, A.; Wang, Z.-F.; Fischer, P. D.; Coote, P.; Padmanabha Das, K. M.; Malets, Y. S.; Radchenko, D. S.; Moroz, Y. S.; Scott, D. A.; Fackeldey, K.; Hoffmann, M.; Iavniuk, I.; Wagner, G.; Arthanari, H. An Open-Source Drug Discovery Platform Enables Ultra-Large Virtual Screens. *Nature* **2020**, *580*, 663.

(26) Muhammed, M. T.; Aki-Yalcin, E. Homology Modeling in Drug Discovery: Overview, Current Applications, and Future Perspectives. *Chem. Biol. Drug Des.* **2019**, *93*, 12.

(27) Neves, M. A. C.; Totrov, M.; Abagyan, R. Docking and Scoring with ICM: The Benchmarking Results and Strategies for Improvement. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 675.

(28) Pdb. R. RCSB PDB: Homepage. https://www.rcsb.org/ (accessed Oct 5, 2022).

(29) UniProt. UniProt. https://www.uniprot.org/ (accessed Oct 5, 2022).

(30) Dallakyan, S.; Olson, A. J. Small-Molecule Library Screening by Docking with PyRx. *Methods Mol. Biol.* **2015**, *1263*, 243.

(31) Autodock Vina. AutoDock Vina. https://vina.scripps.edu/ (accessed Oct 5, 2022).

(32) Trott, A. J. O. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455.

(33) Gaillard, T. Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark. *J. Chem. Inf. Model.* **2018**, *58*, 1697−1706.

(34) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964−12975.

(35) Pham, T. N. H.; Nguyen, T. H.; Tam, N. M.; Y. Vu, T.; Pham, N. T.; Huy, N. T.; Mai, B. K.; Tung, N. T.; Pham, M. Q.; Vu, V.; Ngo, S. T. Improving ligand-ranking of AutoDock Vina by changing the empirical parameters. *J. Comput. Chem.* **2022**, *43*, 160−169.

(36) Pagadala, N. S.; Syed, K.; Tuszynski, J. Software for Molecular Docking: A Review. *Biophys. Rev.* **2017**, *9*, 91.

(37) Abagyan, R.; Totrov, M. Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *J. Mol. Biol.* **1994**, *235*, 983.

(38) Eberhardt, J.; Santos-Martins, D.; Tillack, A.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, *61*, 3891.

(39) Felicitex Therapeutics, Inc.Derivatives of quinoline as inhibitors of DYRK1A and/or DYRK1B kinases. https://patents.justia.com/patent/10577365 (accessed Oct 5, 2022).

(40) Zegzouti, H.; Zdanovskaia, M.; Hsiao, K.; Goueli, S. A. ADP-Glo: A Bioluminescent and Homogeneous ADP Monitoring Assay for Kinases. *Assay Drug Dev. Technol.* **2009**, *7*, 560−572.

(41) Dreas, A.; Fabritius, C.-H.; Dzienia, A.; Buda, A.; Galezowski, M.; Kachkovskyi, G.; Kulesza, U.; Kucwaj-Brysz, K.; Szamborska-Gbur, A.; Czardybon, W.; Vilenchik, M.; Frid, M.; Kuznetsova, A.Derivatives of Quinoline as Inhibitors of DYRK1A And/or DYRK1B Kinases. U.S. Patent 10,577,365 B2, March 3, 2020.

(42) Adeshina, Y. O.; Deeds, E. J.; Karanicolas, J. Machine Learning Classification Can Reduce False Positives in Structure-Based Virtual Screening. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 18477−18488.

(43) ADME QSAR. Optibrium. https://www.optibrium.com/project/adme-qsar/ (accessed Oct 5, 2022).

(44) Wójcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): A New Open-Source Player in the Drug Discovery Field. *J. Cheminform.* **2015**, *7*, 26.

(45) Bonanno, E.; Ebejer, J.-P. Applying Machine Learning to Ultrafast Shape Recognition in Ligand-Based Virtual Screening. *Front. Pharmacol.* **2020**, *10*, 1675.

(46) Polyakov, V. R.; Alexandrov, V.; Maderna, A.; Bajjuri, K.; Li, X.; Zhou, S. Indexing Ultrafast Shape-Based Descriptors in MongoDB to Identify TLR4 Pathway Agonists. *J. Chem. Inf. Model.* **2022**, *62*, 2446−2455.

(47) Armstrong, M. S.; Morris, G. M.; Finn, P. W.; Sharma, R.; Moretti, L.; Cooper, R. I.; Richards, W. G. ElectroShape: Fast Molecular Similarity Calculations Incorporating Shape, Chirality and Electrostatics. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 789−801.

(48) Axen, S. D.; Huang, X.-P.; Cáceres, E. L.; Gendelev, L.; Roth, B. L.; Keiser, M. J. A Simple Representation of Three-Dimensional Molecular Structure. *J. Med. Chem.* **2017**, *60*, 7393−7409.

(49) Gladysz, R.; Dos Santos, F. M.; Langenaeker, W.; Thijs, G.; Augustyns, K.; De Winter, H. Spectrophores as One-Dimensional Descriptors Calculated from Three-Dimensional Atomic Properties: Applications Ranging from Scaffold Hopping to Multi-Target Virtual Screening. *J. Cheminform.* **2018**, *10*, 9.

(50) Krstajic, D.; Buturovic, L. J.; Leahy, D. E.; Thomas, S. Cross-Validation Pitfalls When Selecting and Assessing Regression and Classification Models. *J. Cheminform.* **2014**, *6*, 10−15.

(51) Breiman, L.; Spector, P. Submodel Selection and Evaluation in Regression. The X-Random Case. *Int. Stat. Rev.* **1992**, *60*, 291.

(52) Kuhn, M.20 Recursive Feature Elimination. https://topepo.github.io/caret/recursive-feature-elimination.html (accessed Oct 05, 2022).

(53) Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1−26.

(54) Alexander, D. L. J.; Tropsha, A.; Winkler, D. A. Beware of R2: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *J. Chem. Inf. Model.* **2015**, *55*, 1316.

(55) GitHub - Quantori, https://github.com/quantori/DYRK1B (accessed Oct 5, 2022).

## ■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on October 14, 2022, with the images for Figure 1 and Figure 2 transposed. The corrected version was reposted on October 17, 2022.