

Prospective evaluation and success of a machine learning hit-to-lead drug development program against phosphatidylinositol 3-kinase α

Thomas M. Kaiser,^a Zackery W. Dentmon,^b Pieter B. Burger,^c Qi Shi,^b James P. Snyder,^b Yuhong Du,^d Haian Fu,^d and Dennis C. Liotta^{*b}

^aSt. Peter's College, University of Oxford, Oxford, U.K. OX1 2DL; ^cDepartment of Drug Discovery & Biomedical Sciences, College of Pharmacy, Medical College of South Carolina, 70 President St., Charleston, SC 29425 USA;

^bDepartments of Chemistry and ^dPharmacology & Chemical Biology, Emory University, 1521 Dickey Dr., Atlanta, GA 30322, USA

Email: dliotta@emory.edu

This is dedicated to the successful career of Professor Lanny Liebeskind
and to the memory of Dr. James P. Snyder

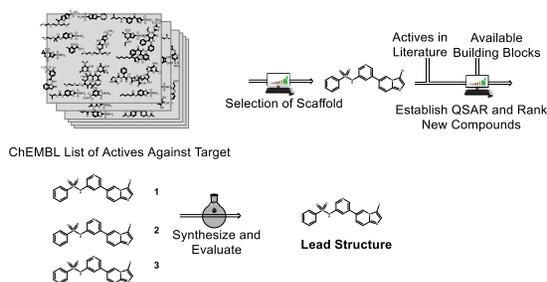
Received 07-16-2020

Accepted 09-22-2020

Published on line 10-06-2020

Abstract

As a result of the rapidly increasing cost of drug development, efficient methods for early identification of compounds with a high probability of clinical success are needed. Herein, we describe a cheminformatics protocol which dramatically increases quality candidate identification and should reduce the attrition rate of compounds entering the clinic, increasing the cost-effectiveness of drug development. Against the oncology target phosphatidylinositol 3-kinase α , all five compounds synthesized from the protocol were found to have low nanomolar activity. We therefore propose that our protocol can be used as a tool for reducing the synthetic burden required for hit-to-lead optimization.



Keywords: Machine learning, hit-to-lead optimization, PI3K α inhibitors, chemotherapy

Introduction

At more than 10^{60} , the sheer number of possible structural variations for a drug-like molecule is staggering.¹ Identifying bioactive molecules has, more often than not, relied on serendipitous findings resulting from the large chemical space available to explore.² Also, when this is coupled with rising drug development program costs, it becomes clear that innovative, efficient methods for identifying high quality drug candidates are needed to provide a pipeline for new medicines.^{3,4} Virtual screening (VS) is frequently employed as a filter where large compound libraries ($>10^7$) are processed to give an enriched list (tens to hundreds of compounds) capable of being examined in biological assays.⁵ Modern VS strategies often employ multiple protocols that are then combined to yield an overall ranking in order to increase the true positive to false positive ratio.⁶⁻⁸ There are now abundant virtual screening methods working from purely 2D descriptors like extended connectivity fingerprints and molecular graphs, hybrid 2D/3D methods of virtual screening and purely 3D methods of screening both from a machine learning perspective and from a structural biology perspective.⁹⁻¹⁶

As part of our initial investigation of new virtual screening methods, we describe here a second prospective evaluation of our original 2D/3D hybrid cheminformatics protocol termed FRESH.¹¹ We have now demonstrated across two distinct biological targets and chemical spaces that FRESH can dramatically increase the ease of identifying commercially unavailable, quality lead compound candidates and should, as a consequence, reduce the attrition rate of compounds entering the clinic and increase the cost-effectiveness of drug development.^{17,18,19}

Results and Discussion

A successful medicinal chemistry program requires the balancing of target and pharmacological properties, intellectual property issues and synthetic considerations. Achieving this balance can consume considerable time and labor costs.²⁰ FRESH (FRagment-based Exploitation of modular Synthesis by vHTS) is a computational workflow designed to prioritize molecules to be synthesized during the drug lead-optimization stage with the aim of increasing efficiency in the hit-to-lead process.¹¹ After a hit has been identified in a general drug discovery program, medicinal chemistry is undertaken to develop a structure-activity relationship (SAR) around the hit structure. FRESH takes advantage of the often hundreds of compounds and their biological data obtained from that SAR to suggest new structures for further development.

FRESH is constructed primarily on the cheminformatics platform Pipeline Pilot (PLP, BIOVIA, Dassault).^{11, 21} PLP performs the task of developing an initial virtual enumeration compound library based on a core structure with user-marked attachment points. The workflow of FRESH involves four key steps. First, the program establishes quantitative structure activity relationships (QSARs) based on the available data to evaluate the biological activity of a molecule of interest. The input/training set for this can be derived from published data, like ChEMBL, or existing unpublished experimental data within a lab.²² The results from other programs, like docking or free energy of binding calculations, can also be included for evaluating potency, when appropriate. Since there is no universally applicable QSAR method, we used at least two independent QSAR evaluations in order to increase the stringency and lessen the molecular case dependency of the QSAR filter.²³

A virtual molecule library is constructed during the second step. As stated previously, this step is strictly based on established synthetic schemes and can be regarded as an *in silico* synthesis, mirroring a wet-lab synthesis.¹¹ The variation of the R groups to be explored is derived from the various corresponding building block structures in a virtual compound library. The source for the building block may be a commercial

electronic database, like Zinc, Maybridge, *etc.*, or a laboratory's proprietary inventory. Fragments with potential liability, stability and reactivity concerns are removed. Also, building blocks violating the fragment rule of three are filtered out.^{11,24} This strategy of varying one site of the core structure works well because, in many cases, a large portion of the hit structure is conserved during hit-to-lead optimization.²⁵

In the third step, the molecule library is filtered using widely accepted rules, like the Lipinski Rules of Five and the Jorgensen Rules of Three,^{26,27} and ADME parameters from the QikProp package.²⁸ The previously established QSAR model is also applied in this step to prioritize molecules, which are likely to have improved potency. Unlike many traditional QSAR studies, the case presented below first uses Bayesian statistics to construct a classification QSAR using a leave-one-out cross-validation in Pipeline Pilot v7.0.²¹ Rather than giving a direct evaluation of the biological activity, as in the case with regression models which can be challenging and prone to failure, the Bayesian classifier method only gives a score which corresponds to the likelihood of whether the compound in question is active or not active (user-defined).^{29,30} Additionally, the Bayesian classifier has the additional property of being highly noise tolerant as compared to other machine learning methods.³¹ Another advantage of using the Bayesian classifier applied in FRESH is that it uses 2D extended connectivity fingerprints (ECFP) making it suitable for use on large data sets due to its rapid speed.³² The Bayesian classifier is therefore applied ahead of more computationally expensive methods like docking and binding free energy calculations in the task-stream of FRESH. Altogether, the Bayesian classifier is more applicable to FRESH early on in the program, as it fits exactly the purpose of prioritizing rather than precisely evaluating compounds in a timely fashion. The final part of the FRESH protocol is the processing and merging of the docking (GLIDE) and free binding energy calculation (MM-GBSA) results. These results (docking, binding energy calculation, and Bayesian classification) are then used to rank the compounds in descending order of most probable to be active followed by visual inspection of the final structures of interest.

We used five criteria to select targets for a FRESH analysis. The first was that the protein target has a therapeutic potential. We then only worked with biological targets that had recent data (the past 15 years). Also, the core scaffold must have been modular and not a potent ligand itself ($IC_{50} > 100$ nM). Finally, the literature must have contained at least one highly potent molecule ($IC_{50} < 10$ nM) that was not used in the QSAR model built by FRESH. With these criteria, we expected the data to be sufficiently stringent for FRESH to work.

In an effort to push beyond traditional limitations of virtual screen method evaluation,²³ we decided to test FRESH prospectively in two cases that would mimic a lead optimization program. Our prospective study was therefore designed to function in a 'real world' setting where the most synthetically accessible compounds with the highest output ranking would be evaluated first. For each case (carbonic anhydrase II and PI3K α), we set a limit that each test set would have five compounds selected from it. If synthetic difficulties were encountered with a given compound, we would move to the next attractive target on the list. Our work on carbonic anhydrase II (CAII) has been published previously;¹³ we now wish to report our findings on PI3K α .

Phosphatidylinositol 3-kinase α (PI3K α)

The phosphatidylinositol 3-kinase (PI3K) family of enzymes are involved in a number of cell signaling pathways and serve as regulators of cell proliferation, survival, chemotaxis and differentiation.³³ Upregulation of PI3K α promotes angiogenesis and is associated with the development of solid tumors. Also, dysregulated PI3K α has been found in 30% of human solid tumors.³⁴ Additionally, since the pathway has been implicated in the development of resistance to conventional therapies, the use of PI3K inhibitors in combination with standard chemotherapies represents a potentially attractive strategy for treating a variety of intractable solid tumors.

An excellent review of the progression of PI3K inhibitors in clinical trials was written by Yang *et al.* in 2019.³⁵ There are currently four anti-PI3K drugs which have received FDA approval (Figure 1).

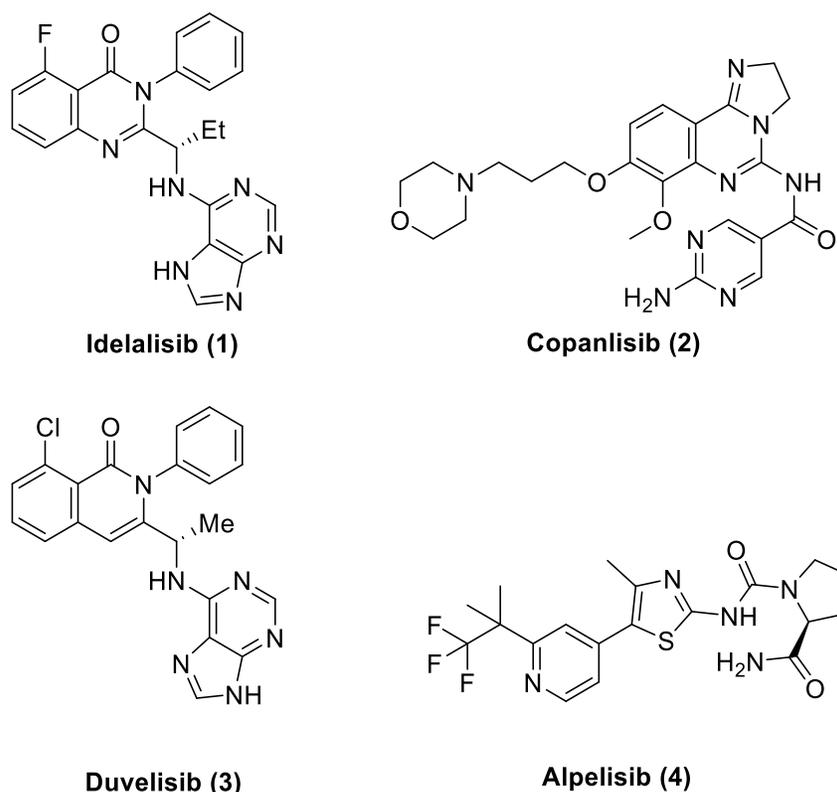


Figure 1. FDA-approved PI3K inhibitors.

PI3K α FRESH Construction

The prospective PI3K-based case study originates from the work of Kim *et al.* in 2011.³⁶ The initial hit compound (**5**) which emerged from a compound screen, exhibits an IC₅₀ of 360 nM against PI3K α (Figure 2). We first evaluated the ability of FRESH to identify PI3K α inhibitors in a retrospective fashion. Specifically, PI3K α inhibitors present in the ChEMBL database were employed to construct the training and test sets for the QSAR in the FRESH program. This QSAR is one of the screens used to predict potencies of novel chemical structures. In order to accurately test the FRESH QSAR against PI3K α at the time of the Kim paper, all compounds with either the same structure as those in the Kim paper or those that appeared in the literature after 2010 were excluded.

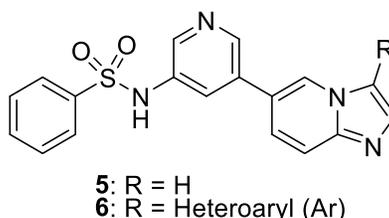


Figure 2. Initial PI3K α hit scaffold (**5**) and proposed site-modified analogs (**6**).

A 10 nM cutoff value was used to partition compounds between “active” and “inactive” analogs for constructing the ligand-based Bayesian model, and ECFP_4 was used as the molecular descriptor. The selected ChEMBL database compounds were randomly divided into a 2:1 ratio using PLP for training and test sets, respectively, to build the QSAR. The training set contained ~500 compounds with ~70 actives and the test set contained ~250 compounds with ~30 actives. A Bayesian model was generated from the training set data using a leave-one-out cross-validation, and the performance of the Bayesian model was evaluated by the receiver operating characteristic (ROC) curve.³⁷ The area under the curve (AUC) of the ROC curve was used to evaluate the ROC curve with an AUC close to 1 being good and an AUC under 0.6 being unacceptable. The test set was subjected to the above-generated Bayesian model and resulted in an AUC of 0.93. This AUC demonstrates that the Bayesian model has a high capacity for predicting activity. With this performance well validated, the Bayesian score with ECFP_4 as the descriptor was used as one of the filters for evaluating activity.

At the time of this project, there were no crystal structures of PI3K α available. However, Kim *et al.* used a homology model based on PI3K γ to model the interaction between small molecule ligands and the PI3K α kinase domain. A homology model using the same template (PDB code: 1E8Z) was constructed and used for receptor-based docking analysis.³⁸ The ChEMBL database compounds were prepared by LigPrep in the Maestro package and processed for both Glide and MM-GBSA scores.^{28, 39} The corresponding AUC values of ROC curves are 0.64 and 0.72 (acceptable), so both the Glide and MM-GBSA scores were used as criteria for potency evaluation.

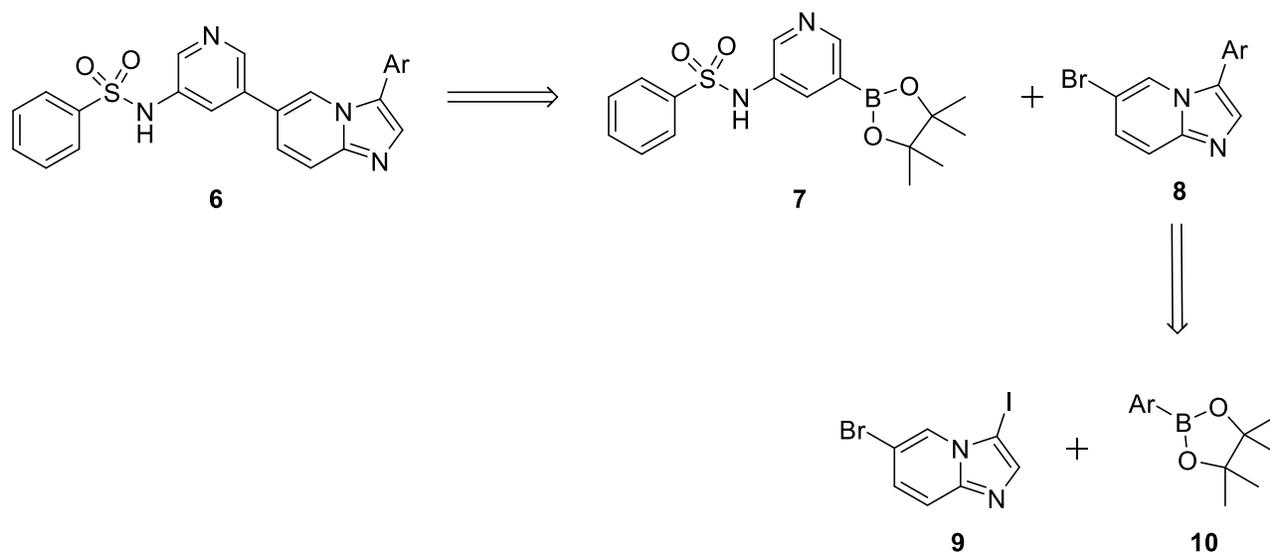
According to the synthesis route from Kim *et al.*, the –Ar group is attached to the core structure via Suzuki coupling (**6**, Figure 2). To obtain the fragment library for enumerating virtual structure libraries, the building blocks (aryl bromides and aryl boronic esters) were queried against the “Zinc bb now” commercial compound database. This resulted in ~44,000 fragments. The –Ar fragment was first filtered by fragment rule of three and groups with potential liability and reactivity concerns were taken into account. The surviving fragments were then covalently connected to the core structure at the –Ar position (Figure 2) to generate structures of possible PI3K α inhibitors. Subsequently, a series of widely-accepted, drug-like filters (Lipinski Rule of Five, Jorgensen Rule of Three, *etc.*) was applied and structures with desirable drug-like properties were established for further processing.^{26,27} Finally, the Bayesian, Glide and MM-GBSA scores were obtained for the filtered structures. Using the corresponding scores of the hit compound (**5**, 360 nM) as a reference, structures with all three scores better than the hit compound were included in the final list for this round. After all screens were applied, approximately 40 structures remained from this round of selection.

Due to the ROC curve of the Bayesian score having the highest AUC, the ligand-based Bayesian model was used to rank and prioritize the remaining ~40 structures. The prioritized list obtained was further filtered for structures already extant in PubMed, thus removing the Kim *et al.* compounds. Similar to the CAII prospective analysis,¹³ we again pursued only those compounds that were quick to synthesize; those that had a high synthetic burden were discarded. General considerations for synthetic candidates could be summarized as follows: 1) the compound must be synthetically accessible from inexpensive building blocks; 2) the compound must have few to no stereocenters; 3) the compound must be outside the established IP; 4) the compound should be drug-like in appearance; 5) the compound should be made within two months or it is abandoned, and the next lower compound should be attempted.

Synthetic Chemistry and Biological Evaluation

The compounds predicted by FRESH revolved around the synthesis employed by Hong *et al.* with minor modifications (Scheme 1).³⁶ The synthesis was highly modular in that **6** could be accessed by a Suzuki coupling

between **7** and **8** with variation at the Ar position of **8**. The biaryls of type **8** were expected to come from the known compound **9** and an array of various aryl pinacolboronic esters **10**.



Scheme 1. Retrosynthesis of **6**.

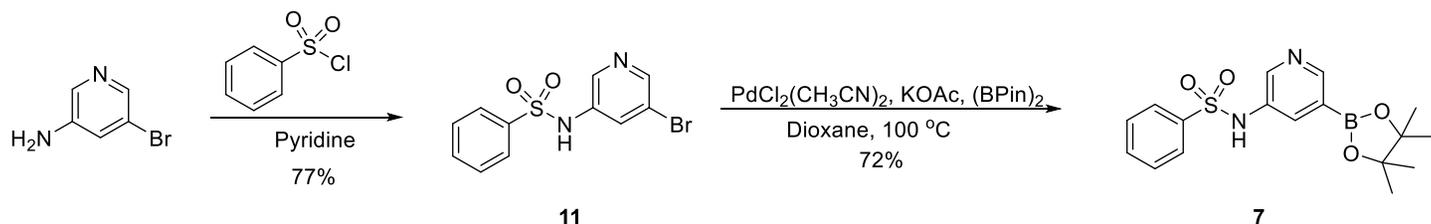
Again, the molecules that were selected for synthesis were chosen because they represented the least synthetic effort to access. The heteroaryl components of the candidates are shown below in Table 1.

Table 1. Candidates chosen from FRESH list

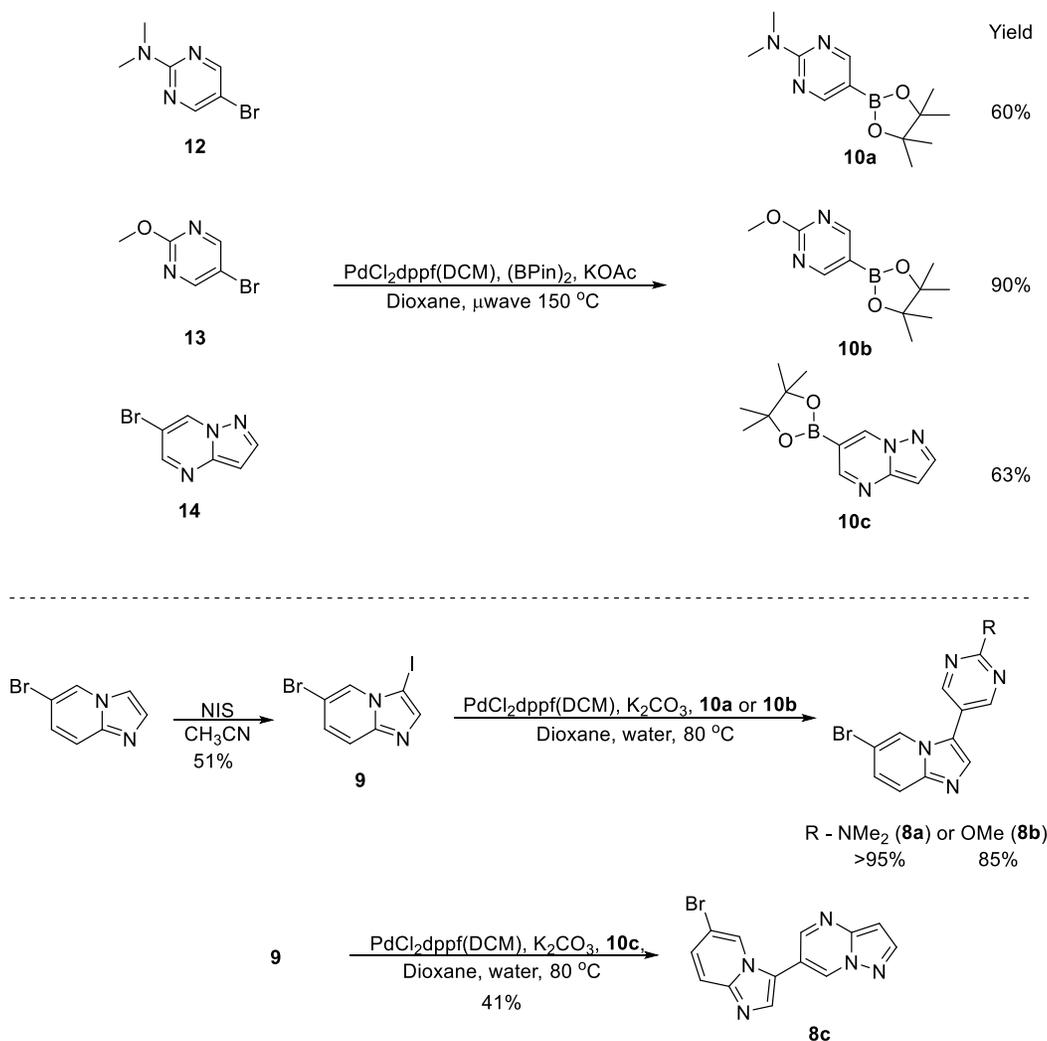
Heteroarene (Ar)	Rank in FRESH	Heteroarene (Ar)	Rank in FRESH
	1		4
	5		8
	19		22
	26		39

To initiate the synthetic campaign, compound **7** was synthesized by reacting 5-bromopyridin-3-amine with benzenesulfonyl chloride to give **11** in 77% yield (Scheme 2). The preparation of **7** was completed by a

palladium-mediated reaction to install the pinacolboronate in 72% yield. Next, the synthesis of the variations of **8** for the final targets began by iodinating 6-bromoimidazo[1,2-*a*]pyridine with NIS to give **9** in 51% yield (Scheme 3). The key step of this synthesis was the selective coupling of aryl boronic esters to the iodo handle of **9**. Therefore, the corresponding aryl halides **12**, **13** and **14** were used to generate the desired boronic esters **10a**, **10b** and **10c** in 60, 90 and 63% yields, respectively. These boronic esters were then coupled with **9** via Suzuki reaction to give the desired **8a**, **8b** and **8c** in 41 to >95% yields.

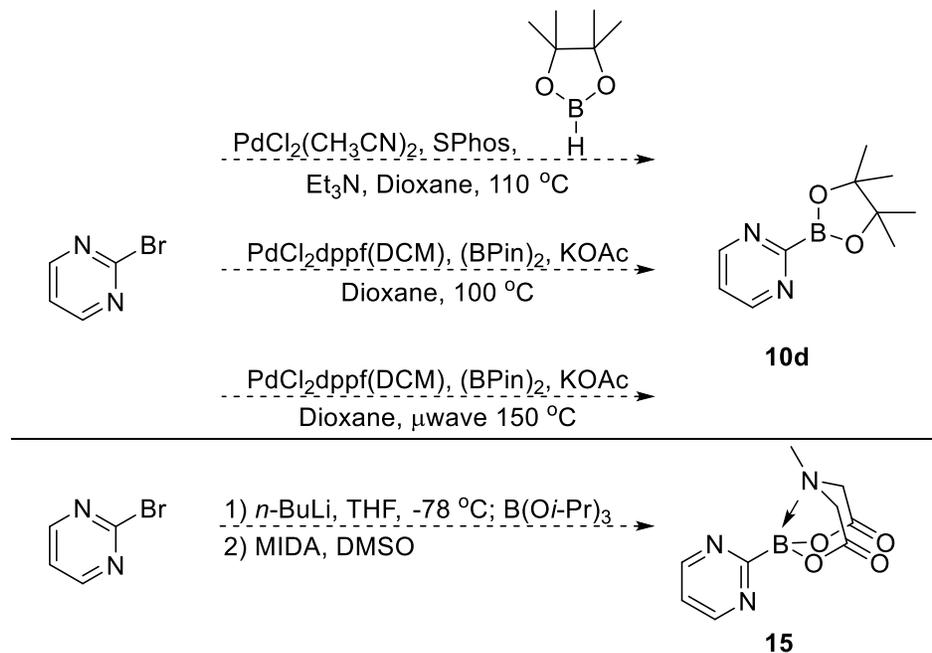


Scheme 2. Synthesis of boronic ester **7**.



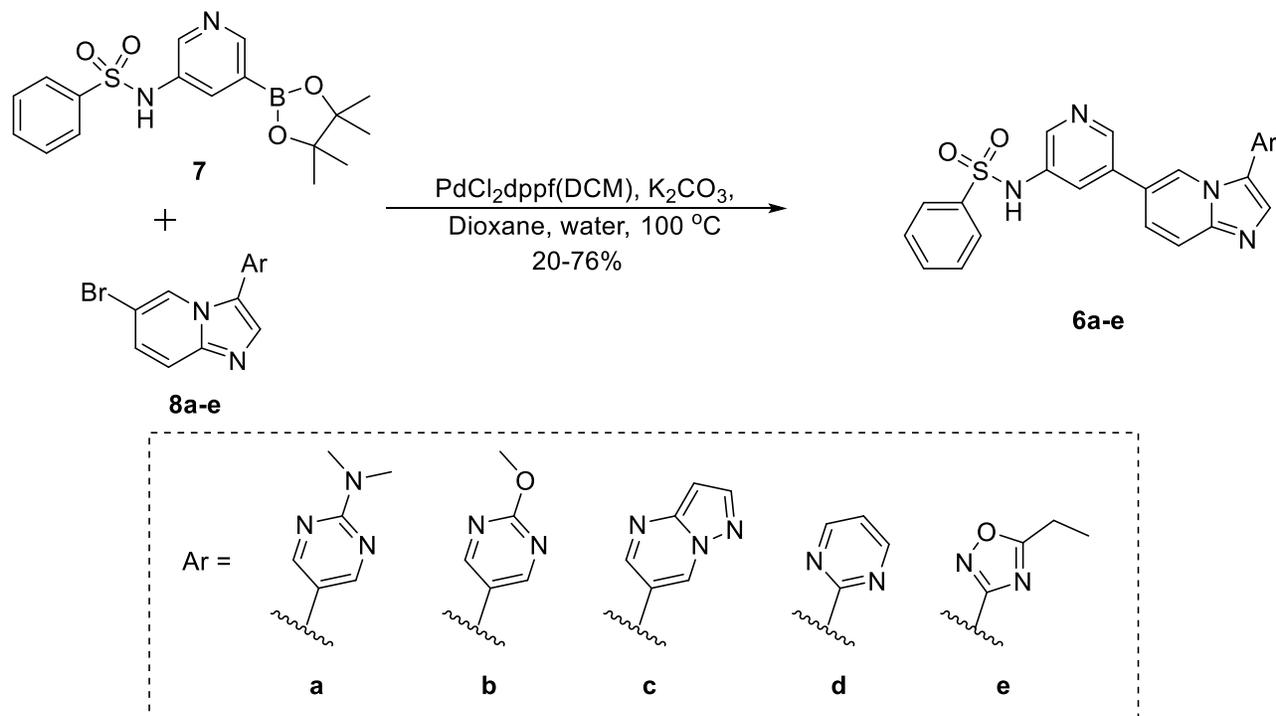
Scheme 3. Preparation of **10a-c** to access **8a-c**.

However, our approach for accessing **10a**, **10b** or **10c** did not prove to be general. Reaction of 2-bromopyrimidine under the boronylation conditions did not afford any desired **10d** (Scheme 4). Probing several different methods reported in the literature did indeed afford small amounts of **10d**, observable by NMR, but the material could not be purified due to protodeboronylation on silica gel. Attempts to cross-couple the crude pinacolboronate also did not afford the desired **8d**. The 2-pyridyl problem, as reported by Burke and co-workers, is due to the instability of 2-boro-pyridine species and, presumably, 2-pyrimidyl boron species would be even more labile due to the greater electrophilicity of the pyrimidyl system over the pyridyl system.⁴⁰ In spite of the MIDA boronate solution to the 2-pyridyl problem by Burke, attempts at using the MIDA boronate of 2-bromopyrimidine (**15**) met with failure (Scheme 4).



Scheme 4. Failed attempts to access 2-pyrimidylboronates.

The method of Buchwald and co-workers, where very active palladacycle precatalysts of type **16** are used to cross-couple electrophilic heteroaryl rings with aryl nucleophiles, was tried as a means of directly synthesizing **8d** (Scheme 5).⁴¹ Dihalogenated **9** was selectively lithiated at the iodo position and then transmetalated to the zincate. Zincates have served as coupling partners in the literature for other electrophilic aryl halides, yet the desired molecule **8d** was not produced.^{42,43} Higher catalyst loading or making boronates *in situ* only produced **8d** in trace amounts. Therefore, a suitable method to install a 2-pyrimidyl ring onto substrates was still needed.



Scheme 7. Synthesis of target final compounds **6a-6e**.

All compounds predicted by FRESH were found to have nanomolar potency against PI3K α using the ADP-Glo Lipid Kinase System (Table 2). Also, of all the PI3K α inhibitors in the literature at the time of our study (2,149), only 44% had a IC₅₀ less than or equal to 329 nM (954). Furthermore, only 23% (489) of all PI3K α inhibitors at the time of our study had an activity equal to or less than 43 nM. Additionally, the compounds were further evaluated in an anti-proliferative assay against MCF-7 breast cancer cell lines housing either wild-type or E545K mutant PI3K α . In each case, the compounds demonstrated increased potency to the mutant relative to wild-type, with single-digit to sub-micromolar potencies. We also calculated the Tanimoto score for each of our compounds against the 30 most potent compounds against PI3K α . Again, the highest similarity coefficient for **6b-6e** was only 0.17822, which indicates that the new compounds proposed by FRESH are chemically distinct from the input training set.⁴⁴ Taken together, the results of FRESH against PI3K α and CAII¹³ strongly suggest that FRESH has the ability to identify compounds with a high likelihood of potent activity.

Table 2. Enzymatic PI3K α and anti-proliferative potencies^a

Compound ID	Enzymatic PI3K α IC ₅₀ (nM) \pm SD ^b	MCF-7 WT IC ₅₀ (μ M) \pm SD	MCF-7 E545K IC ₅₀ (μ M) \pm SD	Rank in FRESH
6a	23.9 \pm 8.70	>10	0.94 \pm 0.19	3
6b	11.5 \pm 1.40	3.35 \pm 0.80	0.78 \pm 0.10	4
6c	47.4 \pm 16.0	>10	4.08 \pm 0.67	22
6d	7.40 \pm 4.90	3.41 \pm 0.55	0.81 \pm 0.17	26
6e	6.10 \pm 3.60	1.27 \pm 0.28	0.46 \pm 0.12	1

^a Data were collected in triplicate and IC₅₀ values were calculated from n = 3 enzymatic and n = 2 cellular experiments; ^b SD: standard deviation

Conclusions

We have evaluated the possibility of using FRESH as a lead development tool for reducing the intensity of synthetic preparation required in the optimization process. In both CAII¹³ and PI3K α , FRESH was capable of generating tens of thousands of synthesis candidates and reducing that pool to tens of compounds predicted to be highly active and have good pharmacological parameters. In the case of PI3K α , all of our synthesized compounds were found to have low nanomolar activity against the enzyme and low-to-submicromolar activity against the E545K mutant MCF-7 cell line. The sites of variation were chosen via consultation with synthetic chemists such that diversity and activity could both be accessed quickly through chemical synthesis. FRESH made use of structural information and ligand-based methods, and also proved useful when only limited structural information was available. The possibility to create a list of proposed active compounds using only a machine learning/Bayesian model without the input of structural biology represents an exciting and successful feature of FRESH^{11,13,45} as well as an area for continued investigation. One of the limitations of FRESH is the requirement of a preexisting and broad activity profile to give predictive capability to the models. However, in a lead development program, this limitation is not present and FRESH can take advantage of substantial SAR to make a high-priority target list efficiently, for further developmental investigation.

Experimental Section

Chemistry. Thin-layer chromatography was carried out on TLC Silica gel 60 F254 commercial plates (Merck KGaA). ¹H NMR spectra were recorded on Varian Inova 400 (400 MHz) spectrometer. All spectra are referenced to the residual solvent peak (2.50 ppm for *d*₆-DMSO; 7.26 for CDCl₃). The chemical shift (δ) of each signal is reported in parts per million (ppm) and all coupling constants (*J*) are reported in Hertz (Hz). ¹³C NMR spectra were recorded on Varian Inova 400 (100 MHz) spectrometer. All spectra are referenced to the residual solvent peak. The chemical shift (δ) of each signal is reported in parts per million (ppm). The Emory University Mass Spectrometry Center recorded high resolution mass spectra. All compounds were analyzed by Agilent 1200 series LCMS or Agilent 1100 series HPLC and found to be greater than 95% pure.

General procedure for the synthesis of 6a-e. Either of **8a-e** (1 equiv) was combined with **7** (1 equiv), PdCl₂dppf(CH₂Cl₂) (5 mol%) and K₂CO₃ (3.8 equiv) in a dried flask and placed under argon. Dioxane and water (3:1, 0.16 M) were added, and the reaction was sealed and heated to 100 °C overnight. The reaction slurry was then filtered through a Celite pad, and the pad was washed with ~125 mL of EtOAc. The organics were then concentrated and purified via silica gel flash column chromatography (0 to 5% MeOH in CH₂Cl₂) to afford the corresponding desired product **6a-e**.

***N*-(5-(3-(2-(Dimethylamino)pyrimidin-5-yl)imidazo[1,2-*a*]pyridin-6-yl)pyridin-3-yl)benzenesulfonamide (6a).**

The material was isolated as a tan residue (0.031 g, 20% yield). ¹H NMR (400 MHz, CDCl₃) δ 8.51 (d, *J* 2.0 Hz, 1H), 8.48 (s, 2H), 8.36 (d, *J* 2.4 Hz, 1H), 8.19 (s, 1H), 7.83-7.78 (m, 4H), 7.69 (s, 1H), 7.58-7.52 (m, 1H), 7.45 (t, *J* 8.0 Hz, 2H), 7.36 (dd, *J* 1.6, 9.2 Hz, 1H), 3.27 (s, 6H). ¹³C NMR (100 MHz, CDCl₃) δ 161.8, 157.7, 145.5, 144.2, 141.8, 139.1, 134.5, 133.8, 133.5, 133.3, 129.5, 127.3, 126.6, 124.6, 123.6, 121.6, 121.0, 118.8, 110.1, 37.4. HRMS (EI+) *m/z* calculated for C₂₄H₂₂N₇O₂S [M+H]⁺: 472.1550, found: 472.15446. The purity was determined to be 99.16 % by HPLC at 254 nm (Chiralpak OD-RH, 75:25 ACN:water, 0.5 mL/min, 5.071 min).

***N*-(5-(3-(2-Methoxypyrimidin-5-yl)imidazo[1,2-*a*]pyridin-6-yl)pyridin-3-yl)benzenesulfonamide (6b).** The material isolated as a brown residue (0.0913 g, 76%). ¹H NMR (400 MHz, CDCl₃) δ 8.78 (s, 2H), 8.52 (d, *J* 2.0

Hz, 1H), 8.42 (d, *J* 2.4 Hz, 1H), 8.27 (s, 1H), 7.85-7.79 (m, 5H), 7.54 (t, *J* 7.2 Hz, 1H) 7.46-7.42 (m, 3H), 4.09 (s, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 165.5, 159.0, 146.0, 144.0, 141.7, 139.1, 134.7, 134.1, 133.5, 133.4, 129.4, 127.2, 126.4, 125.3, 124.2, 120.8, 119.8, 119.0, 117.3, 55.6. HRMS (EI+) *m/z* calculated for C₂₃H₁₉N₆O₃S [M+H]⁺: 459.1234, found: 459.12305. The purity was determined to be 99.99 % by HPLC at 254 nm (Chiralpak ODRH, 75:25 ACN:water, 0.5 mL/min, 5.194 min).

N-(5-(3-(Pyrazolo[1,5-a]pyrimidin-6-yl)imidazo[1,2-a]pyridin-6-yl)pyridin-3-yl)benzenesulfonamide (6c). The material was isolated as a beige solid (0.0536 g, 63% yield). ¹H NMR (400 MHz, *d*₆-DMSO) δ 10.75 (brs, 1H), 9.62-9.92 (m, 1H), 8.85-8.84 (m, 2H), 8.70 (d, *J* 2.0 Hz, 1H), 8.33 (d, *J* 2.4 Hz, 1H), 8.28 (d, *J* 2.4 Hz, 1H), 7.97 (s, 1H), 7.83-7.80 (m, 3H), 7.77 (t, *J* 2.4 Hz, 1H), 7.62-7.52 (m, 4H), 6.85 (dd, *J* 0.8, 2.4 Hz, 1H). ¹³C NMR (100 MHz, *d*₆-DMSO) δ 150.3, 147.2, 145.6, 145.2, 144.0, 140.6, 139.0, 134.6, 134.5, 133.3, 133.0, 129.5, 126.8, 125.6, 124.8, 123., 122.2, 119.6, 117.8, 110.6, 96.6. HRMS (EI+) *m/z* calculated for C₂₄H₁₈N₇O₂S [M+H]⁺: 468.1237, found: 468.12361. The purity was determined to be 95.3135 % by HPLC at 254 nm (Chiralpak OD-RH, 75:25 ACN:water, 0.5 mL/min, 5.053 min).

N-(5-(3-(Pyrimidin-2-yl)imidazo[1,2-a]pyridin-6-yl)pyridin-3-yl)benzenesulfonamide (6d). The material was isolated as a beige solid (0.0625 g, 63% yield). ¹H NMR (400 MHz, CDCl₃/*d*₆-DMSO) δ 10.78 (brs, 1H), 10.15 (s, 1H), 8.89 (d, *J* 4.8 Hz, 2H), 8.61 (d, *J* 2.0 Hz, 1H), 8.52 (s, 1H), 8.34 (d, *J* 2.4 Hz, 1H), 7.86-7.82 (m, 4H), 7.67 (dd, *J* 9.3, 1.9 Hz, 1H) 7.63-7.53 (m, 3H), 7.33 (t, *J* 5.2 Hz, 1H). ¹³C NMR (100 MHz, CDCl₃/*d*₆-DMSO) δ 158.8, 157.5, 147.3, 143.3, 141.0, 139.7, 139.5, 133.4, 133.3, 129.5, 127.0, 126.3, 125.9, 125.0, 123.7, 118.4, 118.1, 110.0. HRMS (EI+) *m/z* calculated for C₂₂H₁₇N₆O₂S [M+H]⁺: 429.1128, found: 429.11221. The purity was determined to be 96.90 % by HPLC at 254 nm (Chiralpak OD-RH, 75:25 ACN:water, 0.5 mL/min, 6.844 min).

N-(5-(3-(Pyrimidin-2-yl)imidazo[1,2-a]pyridin-6-yl)pyridin-3-yl)benzenesulfonamide (6e). The material was isolated as a beige solid (0.1203 g, 40% yield). ¹H NMR (400 MHz, *d*₆-DMSO) δ 10.89 (brs, 1H), 9.13 (s, 1H), 8.59 (d, *J* 2.0 Hz, 1H), 8.35 (d, *J* 2.4 Hz, 1H), 8.29 (s, 1H), 7.90-7.85 (m, 3H), 7.81 (t, *J* 2.0 Hz, 1H), 7.73 (dd, *J* 1.6, 9.2 Hz, 1H), 7.66-7.58 (m, 3H), 3.04 (q, *J* 7.6 Hz, 2H), 1.36 (t, *J* 7.6 Hz, 3H). ¹³C NMR (100 MHz, D₆DMSO) δ 180.3, 160.4, 146.6, 143.2, 140.8, 138.9, 138.2, 134.8, 133.4, 132.3, 129.6, 126.8, 126.2, 125.0, 124.2, 123.9, 118.0, 113.7, 19.5, 10.4. HRMS (EI+) *m/z* calculated for C₂₂H₁₉N₆O₃S [M+H]⁺: 447.1234, found: 447.12277. The purity was determined to be 98.67 % by HPLC at 254 nm (Chiralpak OD-RH, 75:25 ACN:water, 0.5 mL/min, 8.012 min).

N-(5-Bromopyridin-3-yl)benzenesulfonamide (11). To a solution of 5-bromopyridine-3-amine (1.0 g, 5.78 mmol) in pyridine (4 mL) was added benzenesulfonyl chloride (0.82 mL, 6.36 mmol) drop-wise. The reaction mixture was stirred at r.t. for 16 h. After this time, water (10 mL) was added to the reaction mixture. The resulting crystalline solid was collected by filtration, washed with water and dried by rotary evaporation (2x) with 20 mL methanol, 2x with 20 mL 1:1 methanol: CH₂Cl₂ and 1x with CH₂Cl₂. The material was isolated as a beige solid (1.47 g, 81% yield). ¹H NMR (400 MHz, CDCl₃/*d*₆-DMSO) δ 10.52 (brs, 1H), 8.23 (s, 1H), 8.20 (s, 1H), 7.73 (d, *J* 7.2 Hz, 2H), 7.63 (s, 1H), 7.51 (t, *J* 7.2 Hz, 1H), 7.43 (t, *J* 7.6 Hz, 2H). ¹³C NMR (100 MHz, CDCl₃/*d*₆-DMSO) δ 144.7, 138.9, 138.3, 134.8, 132.3, 128.34, 128.26, 125.9, 119.2. This material had spectra that were indistinguishable from literature values.

N-(5-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)pyridin-3-yl)benzenesulfonamide (7). Compound **11** (1.4388 g, 4.59 mmol), 4,4,4',4',5,5,5'-octamethyl-2,2'-bi(1,3,2-dioxaborolane) (1.4 g, 5.51 mmol), PdCl₂(dppf)- CH₂Cl₂ (150 mg, 0.18 mmol), and KOAc (1.804 g, 18.38 mmol) were combined in an oven-dried flask and placed under Ar. Dioxane (23 mL) was added and the flask was sealed and heated to 100 °C and followed by TLC. The reaction was run for approximately 4 h. The reaction mixture was concentrated and dissolved in EtOAc. Celite was added and the mixture was concentrated to yield a clumpy solid. This material was purified by Combiflash at 60-100 % EtOAc in hexanes to afford a yellow solid. The solid was triturated with

3x15 mL of hexanes to remove the yellow color and yielded the product as a white solid (1.19 g, 72%). ^1H NMR (400 MHz, CDCl_3) δ 8.69 (d, J 1.2 Hz, 1H), 8.36 (d, J 2.8 Hz, 1H), 7.92-7.91 (m, 1H), 7.78-7.75 (m, 2H), 7.54 (t, J 7.6 Hz, 1H), 7.43 (t, J 8.0 Hz, 2H), 1.33 (s, 12H). This material had spectra that were indistinguishable from the literature values.

***N,N*-Dimethyl-5-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)pyrimidin-2-amine (10a)**. A mixture of 5-bromo-*N,N*-dimethylpyrimidin-2-amine (0.966 g, 4.78 mmol), 4,4,4',4',5,5,5',5'-octamethyl-2,2'-bi(1,3,2-dioxaborolane) (1.46 g, 5.74 mmol), $\text{PdCl}_2\text{dppf}(\text{CH}_2\text{Cl}_2)$ (0.114 g, 0.14 mmol) and KOAc (1.407 g, 14.34 mmol) were placed under Ar in a 20 mL microwave flask. Anhydrous 1,2-dimethoxyethane (16 mL) was added and the flask was irradiated at 150 °C for 15 minutes. The reaction was filtered through Celite, concentrated and slurried in EtOAc. The reaction was filtered through Celite again and the organics were concentrated and purified by column chromatography 0 to 30% EtOAc in Hexanes. The material was isolated as a light teal solid (0.718 g, 60% yield). ^1H NMR (400 MHz, CDCl_3) δ 8.59 (s, 2H), 3.21 (s, 6H), 1.31 (s, 12H). ^{13}C NMR (100 MHz, CDCl_3) δ 164.1, 163.0, 83.7, 77.4, 37.2, 24.9. HRMS (EI+) m/z calculated for $\text{C}_{12}\text{H}_{12}\text{BN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$: 250.1721, found: 250.17189.

2-Methoxy-5-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)pyrimidine (10b). A mixture of 5-bromo-2-methoxypyrimidine (0.903 g, 4.78 mmol), 4,4,4',4',5,5,5',5'-octamethyl-2,2'-bi(1,3,2-dioxaborolane) (1.46 g, 5.74 mmol), $\text{PdCl}_2\text{dppf}(\text{CH}_2\text{Cl}_2)$ (0.114 g, 0.14 mmol) and KOAc (1.407 g, 14.34 mmol) were placed under Ar in a 20 mL microwave flask. Anhydrous 1,2-dimethoxyethane (16 mL) was added and the flask was irradiated at 150 °C for 15 minutes. Note: starting material and product are indistinguishable by TLC. The material was concentrated and slurried in EtOAc. The material was then filtered through Celite and purified by column chromatography 0 to 60% EtOAc in hexanes. The compound was isolated as a white solid (1.01 g, 90% yield). ^1H NMR (400 MHz, CDCl_3) δ 8.97 (s, 2H), 4.20 (s, 3H), 1.50 (s, 12H). ^{13}C NMR (100 MHz, CDCl_3) δ 167.0, 165.7, 84.5, 77.4, 55.1, 25.0. HRMS (EI+) m/z calculated for $\text{C}_{11}\text{H}_{17}\text{BN}_2\text{O}_3$ $[\text{M}+\text{H}]^+$: 237.1405, found: 237.14008.

6-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)pyrazolo[1,5-*a*]pyrimidine (10c). A mixture of 6-bromopyrazolo[1,5-*a*]pyrimidine (0.947 g, 4.78 mmol), 4,4,4',4',5,5,5',5'-octamethyl-2,2'-bi(1,3,2-dioxaborolane) (1.46 g, 5.74 mmol), $\text{PdCl}_2\text{dppf}(\text{CH}_2\text{Cl}_2)$ (0.114 g, 0.14 mmol) and KOAc (1.407 g, 14.34 mmol) were placed under Ar in a 20 mL microwave flask. Anhydrous 1,2-dimethoxyethane (16 mL) was added and the flask was irradiated in a microwave reactor at 150 °C for 15 minutes. (Note: The reaction vessel took 9 min. to reach reaction temperature). The material was concentrated and slurried in EtOAc. The material was then filtered through Celite and purified by column chromatography (0 to 80% EtOAc in Hexanes). The material was isolated as a cream colored solid (0.7386 g, 63% yield). NMR (400 MHz, CDCl_3) δ 8.99 (m, 1H), 8.69 (d, J 1.6 Hz, 1H), 8.16 (d, J 2.0 Hz, 1H), 6.68 (dd, J 0.8, 2.0 Hz, 1H), 1.67 (s, 12 H). ^{13}C NMR (100 MHz, CDCl_3) δ 153.6, 149.2, 146.2, 141.9, 97.2, 84.8, 77.4, 25.0. HRMS (EI+) m/z calculated for $\text{C}_{12}\text{H}_{17}\text{BN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$: 246.1408, found: 246.14066.

6-Bromo-3-iodoimidazo[1,2-*a*]pyridine (9). To a solution of 6-bromoimidazo[1,2-*a*]pyridine (5 g, 25.4 mmol) in CH_3CN (139 mL) was added NIS (6.85 g, 30.5 mmol). The reaction was stirred at r.t. for 1 h and the reaction was concentrated. The residue was diluted with CH_2Cl_2 (140 mL) and washed with 100 mL of 10% NaOH, 100 mL of saturated thiosulfate solution, and then 100 mL of water. The organic layer was dried over MgSO_4 , filtered and concentrated. The material was isolated as a beige solid (4.163 g, 51% yield). ^1H NMR (400 MHz, CDCl_3) δ 8.28 – 8.27 (m, 1H), 7.69 (s, 1H), 7.51 (dd, J 1.2, 12.8 Hz, 1H), 7.29 (dd, J 2.8, 12.8 Hz, 1H). This material had spectra that were indistinguishable from the literature.

5-(6-Bromoimidazo[1,2-*a*]pyridin-3-yl)-*N,N*-dimethylpyrimidin-2-amine (8a). **9** (0.1 g, 0.31 mmol), **10a** (0.073 g, 0.31 mmol), $\text{PdCl}_2\text{dppf}(\text{CH}_2\text{Cl}_2)$ (13 mg, 0.015 mmol) and K_2CO_3 (0.162 g, 1.174 mmol) were combined in a dried flask and placed under Ar. Dioxane (1.4 mL) and water (0.47 mL) were added and the reaction was

heated to 80 °C and followed by TLC. The reaction was stirred for 3.5 h. The reaction was diluted with EtOAc and filtered through a Celite plug. The plug was washed with ~125 mL of EtOAc and concentrated. The material was then purified by combiflash (0 to 10% MeOH in CH₂Cl₂) to give a beige solid (0.1033g, >95% yield, the material could not be separated from pinacol and was used in the next step 14.5% contamination). ¹H NMR (400 MHz, CDCl₃) δ 8.42 (s, 2H), 8.19 (s, 1H), 7.59 (s, 1H), 7.53 (d, *J* 9.6 Hz, 1H), 7.23 (dd, *J* 2.0, 9.6 Hz, 1H), 3.25 (s, 6H). ¹³C NMR (100 MHz, CDCl₃) δ 161.8, 157.6, 155.0, 144.7, 133.3, 127.8, 123.3, 121.1, 119.0, 110.0, 107.8, 37.3. HRMS (EI+) *m/z* calculated for C₁₃H₁₃BrN₅ [M+H]⁺: 318.0349, found: 318.03431.

5-(6-Bromoimidazo[1,2-*a*]pyridin-3-yl)-*N,N*-dimethylpyrimidin-2-amine (8b). **9** (0.2 g, 0.62 mmol), **10b** (0.146 g, 0.62 mmol), PdCl₂dppf(CH₂Cl₂) (26 mg, 0.031 mmol) and K₂CO₃ (0.324 g, 2.35 mmol) were combined in a dried flask and placed under Ar. Dioxane (2.8mL) and water (0.93 mL) were added and the reaction was heated to 80 °C and followed by TLC. The reaction was diluted with EtOAc and filtered through a Celite plug. The plug was washed with ~125 mL of EtOAc and concentrated. The material was then purified by combiflash (0 to 10% MeOH in CH₂Cl₂) to give a beige solid (0.1219 g, 65%, the material could not be separated from pinacol and was used in the next step 7.3% contamination). ¹H NMR (400 MHz, CDCl₃) δ 8.69 (s, 2H), 8.22 (m, 1H), 7.69 (s, 1H), 7.59 (dd, *J* 0.8, 9.6 Hz, 1H), 7.30 (dd, *J* 2.0, 9.6 Hz, 1H), 4.10 (s, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 165.6, 159.0, 145.2, 134.2, 128.5, 123.0, 119.4, 119.3, 119.2, 117.1, 108.3, 55.5. HRMS (EI+) *m/z* calculated for C₁₂H₁₀BrN₄O [M+H]⁺: 305.0032, found: 305.00286.

6-(6-Bromoimidazo[1,2-*a*]pyridin-3-yl)pyrazolo[1,5-*a*]pyrimidine (8c). Compound **9** (0.1 g, 0.31 mmol), **10c** (0.076 g, 0.31 mmol), PdCl₂dppf(CH₂Cl₂) (13 mg, 0.015 mmol) and K₂CO₃ (0.162 g, 1.174 mmol) were combined in a flask and placed under Ar. Dioxane (1.4 mL) and water (0.47 mL) were added and the reaction was heated to 80 °C and followed by TLC. The reaction was diluted with EtOAc and filtered through a Celite plug. The plug was washed with ~125 mL of EtOAc and concentrated. The material was then purified by combiflash (0 to 10% MeOH in CH₂Cl₂) to give a beige solid (0.042 g, 41% yield). NMR (400 MHz, CD₃OD/CDCl₃) δ 8.99 (d, *J* 1.6 Hz, 1H), 8.62 (d, *J* 2.4 Hz, 1H), 8.39 (d, *J* 0.8 Hz, 1H), 8.22 (d, *J* 2.4 Hz, 1H), 7.76 (s, 1H), 7.59 (d, *J* 9.6 Hz, 1H), 7.42-7.39 (m, 1H), 6.81 (s, 1H). ¹³C NMR (100 MHz, CD₃OD/CDCl₃) δ 186.7, 149.2, 147.4, 146.1, 144.8, 134.2, 133.8, 129.5, 123.2, 118.3, 110.0, 108.7, 97.6. HRMS (EI+) *m/z* calculated for C₁₃H₉BrN₅ [M+H]⁺: 314.0036, found: 314.00299.

6-Bromoimidazo[1,2-*a*]pyridine-3-carbonitrile (18). **Step 1.** To a solution of 5-bromopyridin-2-amine (1.7 g, 9.83 mmol) in MeOH (10 mL) was added 1,1-dimethoxy-*N,N*-dimethylmethanamine (1.6 mL, 11.9 mmol). The reaction was heated at 70 °C for 6 h. The reaction was concentrated and the product was recrystallized from ~20 mL of hexanes heated to reflux with a heat gun to afford crystals of *N'*-(5-bromo-2-pyridinyl)-*N,N*-dimethylimidoforamide as a yellow solid (1.8318 g, 82%). ¹H NMR (400 MHz, CDCl₃) δ 8.38 (s, 1H), 8.25 (dd, *J* 0.8, 2.8 Hz, 1H), 7.60 (dd, *J* 2.8, 8.8 Hz, 1H), 6.83 (dd, *J* 0.4, 8.4 Hz, 1H), 3.09 (s, 3H), 3.07 (2, 3H). Spectrum was indistinguishable from the literature values.

Step 2. To a mixture of the product from Step 1 (0.9274 g, 4.07 mmol) in isopropanol (18.5 mL) was added bromoacetonitrile (0.334 mL, 4.80 mmol) followed by NaHCO₃ (0.59 g, 5.90 mmol) in a sealable reaction tube. The reaction was purged with Ar, sealed and heated to 100 °C for 12 h via microwave irradiation. The solvent had evaporated from the reaction after overnight heating. The reaction was concentrated and the material was brought up in 100 mL of CH₂Cl₂ and 50 mL of water. The organics were removed and the aqueous phase was extracted with another 100 mL of CH₂Cl₂. The combined organics were then dried over Na₂SO₄, filtered and concentrated. The residual brown solid was then recrystallized from EtOH to give the product as tan crystals (0.2065 g, 23% yield). ¹H NMR (400 MHz, CDCl₃) δ 8.52 (d, *J* 2.4 Hz, 1H), 8.15 (s, 1H), 7.68 (dd, *J* 0.8, 9.6 Hz, 1H), 7.53 (dd, *J* 2.0 Hz, 9.6 Hz, 1H). ¹³C NMR (100 MHz, CDCl₃) δ 145.8, 142.8, 132.1, 125.9, 119.3, 110.6, 110.2. Spectra were indistinguishable from the literature.

6-Bromo-3-(pyrimidin-2-yl)imidazo[1,2-*a*]pyridine (8d). 18 (0.229, 1.033 mmol) and 2,2,2-trifluoroacetamide (0.304 g, 2.69 mmol) were placed in a microwave vial, and EtOH (3.04 mL) and (dimethylamino)acrylaldehyde (0.153 mL, 1.529 mmol) were added. The vial was sealed and the reaction was irradiated to 125 °C for 2 h. The reaction was opened, and trifluoroacetamide (0.152 g, 1.35 mmol) and (dimethylamino)acrylaldehyde (0.077 mL, 0.765 mmol) were added. The reaction was placed under Ar and irradiated to 125 °C for 3 h. The reaction was transferred with CH₂Cl₂ and then concentrated on rotovap. A combiflash column was run 0 to 70 % EtOAc in Hexanes. Material isolated as a light yellow oil (0.0634 g, 24% yield, >95% BRSM). ¹H NMR (400 MHz, CDCl₃) δ 10.14-10.13 (m, 1H), 8.79 (d, *J* 5.2 Hz, 2H), 8.60 (s, 1H), 7.63 (d, *J* 9.6 Hz, 1H), 7.43 (dd, *J* 2.0, 9.6 Hz, 1H), 7.13 (t, *J* 4.8 Hz, 1H). ¹³C NMR (100 MHz, CDCl₃) δ 159.0, 157.1, 139.8, 129.6, 129.5, 128.8, 118.5, 118.0, 108.4. HRMS (EI+) *m/z* calculated for C₁₁H₈BrN₄ [M+H]⁺: 274.9927, found: 274.99294.

3-(6-Bromoimidazo[1,2-*a*]pyridin-3-yl)-5-ethyl-1,2,4-oxadiazole (8e). To a solution of **18** (217 mg, 0.976 mmol) in EtOH (3.9 mL) were added hydroxylamine hydrochloride (0.081 g, 1.17 mmol) and Et₃N (0.272 mL, 2.0 mmol). The reaction was sealed and stirred at ambient temperature for 19 h and then concentrated. The reaction was concentrated via rotary evaporation and placed on high vacuum for 2 h. The material was clean by TLC and was then placed under Ar to be used directly. Dioxane (3.9 mL) was added followed by propionic anhydride (0.375 mL, 2.93 mmol). The flask was sealed, heated to 120 °C and followed by TLC. The reaction was concentrated and the material was resuspended in CH₂Cl₂. The mixture was then partitioned between 25 mL of CH₂Cl₂ and 25 mL of saturated NaHCO₃. The layers were shaken and separated. The aqueous phase was then extracted with CH₂Cl₂ (2x 25 mL). The combined organics were dried over Na₂SO₄, filtered and concentrated. A combiflash was run 0 to 60% EtOAc in hexanes. The material was isolated as a light yellow solid (0.2279 g, 80% yield). ¹H NMR (400 MHz, CDCl₃) δ 9.30 (dd, *J* 0.9, 1.9 Hz, 1H), 8.33 (s, 1H), 7.63 (dd, *J* 0.8, 9.6 Hz, 1H), 7.43 (dd, *J* 1.6, 9.2 Hz, 1H), 3.00 (q, *J* 7.6 Hz, 2H), 1.47 (t, *J* 7.6 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 180.2, 160.9, 146.4, 138.5, 130.0, 128.0, 118.6, 114.2, 109.0, 20.3, 10.9. HRMS (EI+) *m/z* calculated for C₁₁H₁₀BrN₄O [M+H]⁺: 293.0032, found: 293.00313.

Biological evaluation: ADP-Glo™ Kinase Assay (Promega). The effect of compounds on PI3K (p110α/p85α) activity was determined using ADP-Glo Lipid Kinase systems (Promega, Madison, WI). We first optimized the kinase assay by titrating enzyme in a 384-well plate format for high throughput screening (uHTS). Briefly, the enzyme was diluted in 2.5X kinase buffer. 4 μl/well of increasing concentrations of kinase was then added into a 384-well low volume white solid round bottom plate (Corning Costar, Cat# 4513). The substrate (4 μl/well) reconstituted in 2.5X lipid buffer was added into each. 2 μl/well of ATP (100 μM final concentration) was added into the wells to start the kinase reaction. After incubating at rt for 1 h, 10 μl/well of ADP-Glo reagent was added and incubated at rt for 40 min. 20 μl/well of kinase detection buffer was then added and the luminescence signal was measured using Envision multilabel plate reader (PerkinElmer).

Based on the results of enzyme titration, we selected 0.2 μg/ml of enzyme for compound evaluation. The kinase was diluted in 2.5X kinase buffer and mixed with substrate in 2.5X lipid buffer. 13 μl of enzyme/substrate was dispensed into 384-well low volume plate using Multidrop combi dispenser (Thermo-fisher Scientific). 0.5 μl of compound diluted in DMSO was added using 384-channel pipette of Beckman NX automated liquid handling system (Beckman Coulter, Danaher Corporation). 2 μl of ATP (100 μM final concentration) was then dispensed into wells and incubated for 40 min followed by kinase detection as described above. The effect of compound on kinase activity is expressed as % of control and calculated according to the following equation:

$$\% \text{ of Control} = (\text{Lum}_{\text{compound}} - \text{Lum}_{\text{blank}}) / (\text{Lum}_{\text{DMSO}} - \text{Lum}_{\text{blank}}) \times 100\%$$

where Lum_{blank} is the average signal from wells in the presence of the kinase and substrate without ATP, which is the background signal; Lum_{DMSO} is the average signal from well with the Kinase, substrate and ATP with

DMSO control without compound, which defines the highest signal; Lum_{compound} is the signal from wells with both kinase, substrate, ATP and compound.

Finally, the IC₅₀ was calculated using GraphPad Prism software (GraphPad software LLC).

Cell viability assay

The effect of compounds on cell viability was determined by CellTiter Blue cell viability assay as previously described.^{46,47} MCF-7 E545K cells and MCF-7 WT cells⁴⁸ were kindly provided by Dr. Josh Lauring M.D., Ph.D. (the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins). 1000 cells in 45 µl of cell culture medium were seeded in a 384-well cell culture black wall clear bottom plate (Corning #3724). After overnight incubation at 37 °C, 0.5 µl of each test compound serially diluted in DMSO was added to the cells. The plates were centrifuged at 800 rpm for 5 minutes and incubated for 3 days in a cell culture incubator. After 3 days, 5 µl CellTiter Blue (Promega) was added to each well. After incubating for 4 h, the fluorescence intensity (FI), which is correlated with the number of viable cells, was measured using Envision Multilabel plate reader (PerkinElmer) with excitation at 545 nm and emission at 615 nm. Cells containing medium alone without cells were used as blank control for background correction.

The effect of compound on cell viability was expressed as % of control and calculated using the following equation:

$$\% \text{ of Control} = (FI_{\text{compound}} - FI_{\text{blank}}) / (FI_{\text{DMSO}} - FI_{\text{blank}}) \times 100\%$$

Acknowledgements

The authors wish to thank Dr. John DiRaddo and Prof. Stephen Traynelis for their assistance in the preliminary evaluation of the compounds, as well as Dr. Josh Lauring for his provision of the MCF-7 cells. Additionally, the authors acknowledge the use of shared instrumentation maintained by the Emory University NMR Research Center, directed by Dr. Shoaxiong Wu and Dr. Bing Wang, as well as the Emory University Mass Spectrometry Center directed by Dr. Fred Strobel.

Supplementary Material

NMR spectra are available as supplementary material to this article.

References

1. Bohacek, R. S.; McMartin, C.; Guida, W. C. *Med. Res. Rev.* **1996**, *16*, 3-50.
[https://doi.org/10.1002/\(SICI\)1098-1128\(199601\)16:1<3::AID-MED1>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1098-1128(199601)16:1<3::AID-MED1>3.0.CO;2-6)
2. Dobson, C. M. *Nature* **2004**, *432*, 824-828.
<https://doi.org/10.1038/nature03192>
3. Khanna, I. *Drug Discovery Today* **2012**, *17*, 1088-1102.
<https://doi.org/10.1016/j.drudis.2012.05.007>

4. Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L. *Nat. Rev. Drug Discovery* **2010**, *9*, 203-214.
<https://doi.org/10.1038/nrd3078>
5. Tanrikulu, Y.; Kruger, B.; Proschak, E. *Drug Discovery Today* **2013**, *18*, 358-364.
<https://doi.org/10.1016/j.drudis.2013.01.007>
6. Holliday, J. D.; Kanoulas, E.; Malim, N.; Willett, P. *J. Cheminf.* **2011**, *3*, 1-15.
<https://doi.org/10.1186/1758-2946-3-29>
7. da Silva Rocha, S. F. L.; Olanda, C. G.; Fokoue, H. H.; Sant'Anna, C. M. R. *Curr. Top. Med. Chem.* **2019**, *19*, 1751-1767.
<https://doi.org/10.2174/1568026619666190816101948>
8. Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. *Drug Discovery Today* **2018**, *23*, 1241-1250.
<https://doi.org/10.1016/j.drudis.2018.01.039>
9. Feinberg, E. N.; Sur, D.; Wu, Z.; Husic, B. E.; Mai, H.; Li, Y.; Sun, S.; Yang, J.; Ramsundar, B.; Pande, V. S. *ACS Central Sci.* **2018**, *4*, 1520-1530.
<https://doi.org/10.1021/acscentsci.8b00507>
10. Wallach, I. D., Michael; Heifets, Abraham. AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery. 2015, arXiv:1510.02855. arXiv.org e-Print archive.
<https://arxiv.org/abs/1510.02855> (accessed July 6, 2020).
11. Shi, Q.; Kaiser, T. M.; Dentmon, Z. W.; Ceruso, M.; Vullo, D.; Supuran, C. T.; Snyder, J. P. *ACS Med. Chem. Lett.* **2015**, *6*, 518-522.
<https://doi.org/10.1021/acsmchemlett.5b00062>
12. Kaiser, T. M.; Burger, P. B.; Butch, C. J.; Pelly, S. C.; Liotta, D. C. *J. Chem. Inf. Model.* **2018**, *58*, 1544-1552.
<https://doi.org/10.1021/acs.jcim.7b00475>
13. Kaiser, T. M.; Dentmon, Z. W.; Dalloul, C. E.; Sharma, S. K.; Liotta, D. C. *ACS Med. Chem. Lett.* **2020**, *11*, 491-496.
<https://doi.org/10.1021/acsmchemlett.9b00612>
14. Stokes, J. M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N. M.; MacNair, C. R.; French, S.; Carfrae, L. A.; Bloom-Ackermann, Z.; Tran, V. M.; Chiappino-Pepe, A.; Badran, A. H.; Andrews, I. W.; Chory, E. J.; Church, G. M.; Brown, E. D.; Jaakkola, T. S.; Barzilay, R.; Collins, J. J. *Cell* **2020**, *180*, 688-702 e13.
<https://doi.org/10.1016/j.cell.2020.01.021>
15. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; Zhao, S. *Nat. Rev. Drug Discovery* **2019**, *18*, 463-477.
<https://doi.org/10.1038/s41573-019-0024-5>
16. Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. *Acc. Chem. Res.* **2017**, *50*, 1625-1632.
<https://doi.org/10.1021/acs.accounts.7b00083>
17. Empfield, J. R.; Clark, M. P., Attrition: The Biggest Enemy of the Pharmaceutical Industry. In *Reducing Drug Attrition*, Empfield, J. R.; Clark, M. P., Eds. Springer-Verlag: New York, 2014; p vii.
<https://doi.org/10.1007/978-3-662-43914-2>
18. Ekins, S.; Waller, C. L.; Swaan, P. W.; Cruciani, G.; Wrighton, S. A.; Wikel, J. H. *J. Pharmacol. Tox. Met.* **2000**, *44*, 251-272.
[https://doi.org/10.1016/S1056-8719\(00\)00109-X](https://doi.org/10.1016/S1056-8719(00)00109-X)
19. Gleeson, M. P.; Hersey, A.; Montanari, D.; Overington, J. *Nat. Rev. Drug Discovery* **2011**, *10*, 107-208.
<https://doi.org/10.1038/nrd3367>

20. Jorgensen, W. L. *Angew. Chem. Int. Ed.* **2012**, *51*, 11680-11684
<https://doi.org/10.1002/anie.201204625>
21. BIOVIA <http://accelrys.com/products/collaborative-science/biovia-pipeline-pilot/>.
22. P., B. A.; Gaulton, A.; Hersey, A.; Bellis, L. J.; Chambers, J.; Davies, M.; Kruger, F. A.; Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J. P. *Nucleic Acids Res.* **2014**, *42*, 1083-1090.
<https://doi.org/10.1093/nar/gkt1031>
23. Strumpfe, D.; Bajorath, J., Critical Assessment of Virtual Screening for Hit Identification. In *Chemoinformatics for Drug Discovery*, Bajorath, J., Ed. John Wiley and Sons, Inc.: Hoboken, New Jersey, 2014; p 115.
<https://doi.org/10.1002/9781118742785.ch6>
24. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A. *Drug Discovery Today* **2003**, *8*, 876-877.
[https://doi.org/10.1016/S1359-6446\(03\)02831-9](https://doi.org/10.1016/S1359-6446(03)02831-9)
25. Kerns, E. H.; Di, L., *Drug-like Properties: Concepts, Structure Design and Methods*. Elsevier Inc.: New York, 2008.
<https://doi.org/10.1016/B978-012369520-8.50025-5>
26. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3-26.
27. Jorgensen, W. L. *Acc. Chem. Res.* **2009**, *42*, 724-733.
<https://doi.org/10.1016/B978-012369520-8.50025-5>
28. *Small-Molecule Drug Discovery Suite 2014-1: Glide*, 6.2; Schrödinger, LLC: New York, NY, 2014.
29. Rogers, D.; Hahn, M. J. *Chem. Inf. Model.* **2010**, *50*, 742-754.
<https://doi.org/10.1021/ci100050t>
30. Mitchell, J. B. O. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 468-481.
<https://doi.org/10.1002/wcms.1183>
31. Kaiser, T. M.; Burger, P. B. *Molecules* **2019**, *24*, 2115.
<https://doi.org/10.3390/molecules24112115>
32. Rogers, D.; Brown, R. D.; Hahn, M. J. *Biomol. Screening* **2005**, *10*, 682-686.
<https://doi.org/10.1177/1087057105281365>
33. Cain, R. J.; ARidley, A. J. *Biol. Cell* **2009**, *101*, 13-29.
<https://doi.org/10.1042/BC20080079>
34. Akinleye, A.; Avvaru, P.; Furqan, M.; Song, Y.; Liu, D. J. *Hematol. Oncol.* **2013**, *6*, 88-104.
<https://doi.org/10.1186/1756-8722-6-88>
35. Yang, J.; Nie, J.; Ma, X.; Wei, Y.; Peng, Y.; Wei, X. *Mol. Cancer* **2019**, *18*, 26.
<https://doi.org/10.1186/s12943-019-0954-x>
36. Kim, O.; Jeong, Y.; Lee, H.; Hong, S.-S.; Hong, S. J. *Med. Chem.* **2011**, *54*, 2455-2466.
<https://doi.org/10.1021/jm101582z>
37. Swets, J. A., *Signal detection theory and ROC analysis in psychology and diagnostics: collected papers* Lawrence Erlbaum Associates: Mahwah, N.J., 1996.
38. Walker, E. H.; Pacold, M. E.; Perisic, O.; Stephens, L.; Hawkins, P. T.; Wymann, M. P.; Williams, R. L. *Mol. Cell* **2000**, *6*, 909-919.
[https://doi.org/10.1016/S1097-2765\(05\)00089-4](https://doi.org/10.1016/S1097-2765(05)00089-4)
39. *Schrödinger Release 2014-1: Prime*, 3.5; Schrödinger, LLC: New York, NY, 2014.
40. Dick, G. R.; Woerly, E. M.; Burke, M. D. *Angew. Chem. Int. Ed.* **2012**, *51*, 2667-2672.
<https://doi.org/10.1002/anie.201108608>

41. Milne, J. E.; Buckwald, S. L. *J. Am. Chem. Soc.* **2004**, *126*, 13028-13032.
<https://doi.org/10.1021/ja0474493>
42. Chanteau, S. H.; Tour, J. M. *J. Org. Chem.* **2003**, *68*, 8750-8766.
<https://doi.org/10.1021/jo0349227>
43. Yang, Y.; Oldenhuis, N. J.; Buchwald, S. L. *Angew. Chem. Int. Ed.* **2013**, *2013*, 615-619.
<https://doi.org/10.1002/anie.201207750>
44. Maggiora, G.; Vogt, M.; Stumpfe, D.; Bajorath, J. *J. Med. Chem.* **2014**, *57*, 3186-3244.
<https://doi.org/10.1021/jm401411z>
45. Pribut, N.; Kaiser, T. M.; Wilson, R. J.; Jecs, E.; Dentmon, Z. W.; Pelly, S. C.; Sharma, S.; Bartsch III, P. W.; Burger, P. B.; Hwang, S. S.; Le, T.; Sourimant, J.; Yoon, J. J.; Plemper, R. K.; Liotta, D. C. *ACS Infect. Dis.* **2020**, *6*, 922-929.
<https://doi.org/10.1021/acsinfecdis.9b00524>
46. Webber, P. J.; Park, C.; Qui, M.; Ramalingam, S. S.; Khuri, F. R.; Fu, H.; Du, Y. *Oncoscience* **2015**, *2*, 765-776.
<https://doi.org/10.18632/oncoscience.245>
47. Mo, X.; Tang, C.; Niu, Q.; Ma, T.; Du, Y.; Fu, H. *Cell Chem. Biol.* **2019**, *26*, 331-339 e3.
<https://doi.org/10.1016/j.chembiol.2018.11.011>
48. Beaver, J. A.; Gustin, J. P.; Yi, K. H.; Rajpurohit, A.; Thomas, M.; Gilbert, S. F.; Rosen, D. M.; Ho Park, B.; Lauring, J. *Clin. Cancer Res.* **2013**, *19*, 5413-22.
<https://doi.org/10.1158/1078-0432.CCR-13-0884>

This paper is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>)