

Data-Driven Construction of Antitumor Agents with Controlled Polypharmacology

Chenxiao Da,[†] Dehui Zhang,[†] Michael Stashko,[†] Eleana Vasileiadi,[‡] Rebecca E. Parker,[‡] Katherine A. Minson,[‡] Madeline G. Huey,[‡] Justus M. Hulse,[‡] Debra Hunter,^{§,‡} Thomas S. K. Gilbert,[‡] Jacqueline Norris-Drouin,[†] Michael Miley,[‡] Laura E. Herring,[‡] Lee M. Graves,[‡] Deborah DeRyckere,[‡] H. Shelton Earp,^{§,‡} Douglas K. Graham,[‡] Stephen V. Frye,^{‡,§} Xiaodong Wang,^{*,†} and Dmitri Kireev^{*,†}

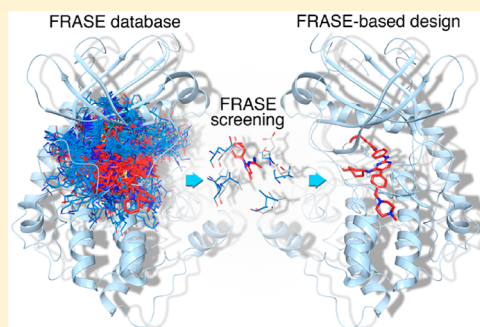
[†]Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7363, United States

[‡]Department of Pharmacology and [§]Lineberger Comprehensive Cancer Center, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[‡]Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, and Department of Pediatrics, Emory University, Atlanta, Georgia 30322, United States

Supporting Information

ABSTRACT: Controlling which particular members of a large protein family are targeted by a drug is key to achieving a desired therapeutic response. In this study, we report a rational data-driven strategy for achieving restricted polypharmacology in the design of antitumor agents selectively targeting the TYRO3, AXL, and MERTK (TAM) family tyrosine kinases. Our computational approach, based on the concept of fragments in structural environments (FRASE), distills relevant chemical information from structural and chemogenomic databases to assemble a three-dimensional inhibitor structure directly in the protein pocket. Target engagement by the inhibitors designed led to disruption of oncogenic phenotypes as demonstrated in enzymatic assays and in a panel of cancer cell lines, including acute lymphoblastic and myeloid leukemia (ALL/AML) and nonsmall cell lung cancer (NSCLC). Structural rationale underlying the approach was corroborated by X-ray crystallography. The lead compound demonstrated potent target inhibition in a pharmacodynamic study in leukemic mice.



INTRODUCTION

Polypharmacology: Embracing the Inevitable. Large families of structurally and often functionally related proteins, such as G-protein coupled receptors (GPCRs),^{1,2} protein kinases,^{3,4} and more recently HDAC⁵ or bromodomains,⁶ are the most important source of drug targets. Due to the high intrafamily similarity, especially within smaller subfamilies, a single small-molecule drug, in general, binds multiple family members, presenting both a challenge and an opportunity for drug discovery.⁷ Beyond the basic consideration that promiscuous binding to multiple targets is a major cause of adverse effects⁸ and general toxicity,⁹ it has also been shown that excessive selectivity for a single target may be a fatal safety concern,¹⁰ while binding several related targets could be a prerequisite for more efficacious drug action.¹¹ Hence, to achieve a desired therapeutic response, a drug discoverer should ideally be able to control the polypharmacology of a drug candidate, that is, to disable its promiscuous binding to unwanted targets, while enhancing potency against therapeutically relevant targets. However, controlling which particular members of a large protein family would be inhibited by a drug

is an extremely challenging endeavor. Here we describe a data-driven approach to design small-molecule drugs with fine-tuned selectivity profiles in the context of highly homologous protein targets. In our approach, 3D structures of bioactive ligands are directly “assembled” from smaller building blocks in the protein binding pocket.

From Virtual Screener to Virtual Chemist. Previous computational efforts to design multitarget drugs focused on extending bioinformatics models or scaling-up cheminformatics techniques. In particular, various flavors of binding-pocket similarity have been used to make predictions based on the principle “similar targets bind similar ligands”.^{12–14} Alternatively, broad panels of computational structure–activity relationship (SAR) models were used to predict targets for ligand libraries.^{15–18} Most practical applications focused on the two largest drug-target families: G-protein coupled receptors (GPCRs)¹⁹ and protein kinases.²⁰ Although highly diverse in terms of underlying principles and assumptions, all previously

Received: August 10, 2019

Published: September 9, 2019

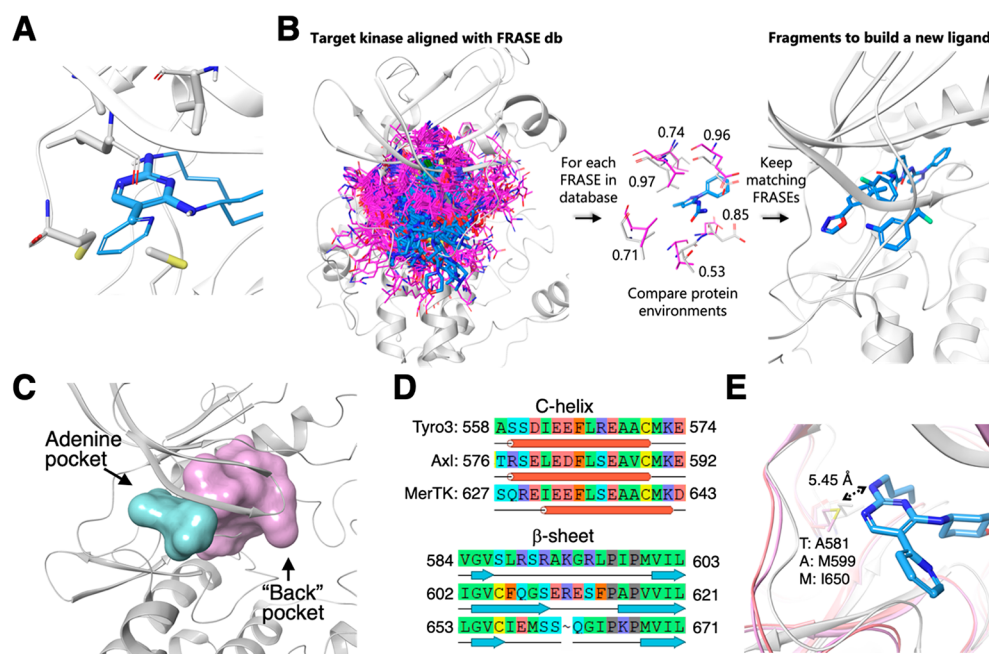


Figure 1. FRASE-based approach to “assemble” inhibitors selective for specific TAM family members. (A) Example of a typical FRASE: the constituent fragments are rendered as thick sticks (gray for the protein portion of the FRASE, blue for the ligand portion). (B) Simplified representation of the FRASE-based design: (left to right) a target kinase (gray cartoon) with projected reference FRASEs from the database (magenta and blue sticks). Evaluation of how individual components of a reference protein environment (magenta sticks) match the nearby target residues (gray sticks) (the numbers represent per-residue Tscore values, for a total Tscore of 0.79). Ligand fragments from the matching FRASEs (blue sticks) that can be used for design. (C) Adenine (cyan) and back (pink) pockets in a kinase enzyme. (D) Aligned sequences of residues lining the back pocket show 44 pairwise differences. (E) The only amino acid position (rendered as sticks) showing variance in the adenine pocket of TAM kinases and the X-ray structure for the template inhibitor **1** in complex with MERTK.

reported approaches share one common function: they act as a “virtual screener”, that is, receive a ligand structure as the input and report its activity against a given target as the output. “Virtual screeners” are highly efficient when applied to a small panel of well-studied targets with large SAR data sets. However, they are suboptimal for dealing with larger panels of predominantly novel targets: first, because there are no SAR data to develop predictive SAR models, and, second, a “virtual screener” still needs a human chemist to feed it with newly designed compounds. Yet, traditional SAR-based medicinal chemistry is difficult to scale up to multitarget discovery. Intuitively, the design of chemically novel ligands for understudied targets may benefit from (i) thousands of experimental structures of protein–ligand complexes, (ii) hundreds of thousands of biological activities for those ligands in public databases, and (iii) an intricate network of evolutionary relationships within large protein families. Tools are needed to jointly exploit disparate pieces of information from heterogeneous sources. The approach to polypharmacological ligand design, that we introduce, is based on the concept of a fragment in structural environment (FRASE). A single FRASE extracted from a high-affinity ligand explicitly encapsulates 3D structural information and, implicitly, SAR and sequence alignment data. Several FRASEs extracted from different protein–ligand complexes can be readily combined into a novel ligand for an orphan protein target. Hence, FRASE-based design represents a first step in the evolution of computer-aided drug design from “virtual screener” to “virtual chemist”.

Targeting Tumor-Survival Promoters TYRO3, AXL, and MERTK. Here we demonstrate the potential of the FRASE-based approach by using it to design ligands that can

interfere with the function of the TYRO3, AXL, and MERTK (TAM) family of receptor tyrosine kinases.²¹ These receptors play a physiologic role in dampening inflammatory responses, a role that is subverted in cancer-associated myeloid cells in the immunosuppressive tumor microenvironment.^{22,23} TAM kinases, in particular MERTK, have been shown to upregulate expression of the T cell immune checkpoint molecule PD-L1 and to promote phosphatidylserine-dependent efferocytosis and AKT-mediated chemoresistance in tumor cells.^{22,24} In addition, TAM kinases are ectopically expressed in many human cancers, including both hematologic malignancies and solid tumors, where they function to promote tumor cell survival, chemoresistance, and motility.²³ Consequently, targeted inhibition of TAM family members could reverse the immunosuppressive microenvironment and directly target tumor cells in a wide range of cancers.

The therapeutic value of TAM inhibitors may be largely dependent on their respective polypharmacologic profiles, both within the TAM family and with respect to other protein kinases. Therefore, it is important to design chemical probes selective either for a single TAM family member or for any combination thereof. However, controlling the intrafamily selectivity is challenging because TAM kinases were among the latest to evolve, and hence, they have highly similar kinase-domain structures and near-identical ATP pockets, providing almost no opportunity for developing intrafamily selectivity. Here, we made use of FRASE-based design to develop inhibitors selective for specific TAM family members and explore their potential as anticancer agents.

■ RESULTS

FRASE-Based Design. For a long time, ligand- and structure-based strategies in computer-aided drug design evolved separately, one toward embracing bigger data for more predictive and comprehensive activity models and the other toward faster and more accurate prediction of ligand–protein interactions.^{25,26} However, despite significant progress, both strategies conserve their essential limitations. Ligand-based approaches cannot predict activities for “orphan” targets, and structure-based techniques still have a long way to go before it might be possible to profile hundreds of ligands against hundreds of targets with a suitable accuracy in a reasonable time. FRASEs are structural descriptors that are intended to merge the chemical and protein structure spaces to overcome both of the above limitations. FRASE-based design enables processing of thousands of 3D protein–ligand complexes and tens of thousands of bioactivity data points to find the most appropriate building blocks for new active ligands. Moreover, the suggested building blocks are shown as 3D protein-bound poses and can be easily shuffled to yield synthetically tractable combinations with diverse family-wide polypharmacology profiles. A typical FRASE (Figure 1A) includes a chemically sound fragment of a 3D protein-bound ligand pose, e.g., a cycle and/or functional group, as well as the nearby protein fragments (in this study, full residue structures). There are two primary assumptions behind the FRASE-based design: (a) a ligand has high affinity to a given protein due to an “endorsement” by every fragment it consists of and (b) a fragment is “good” due to its favorable interactions with its immediate protein environment. By combining (a) and (b) we obtain (c): if a fragment is “good” in one protein, it will also be “good” in any other protein that features an identical structural environment.

Conceptually, design of a new ligand for a given kinase involves (i) identification of structural environments (stored in the FRASE database) that match those in the binding pocket of the kinase of interest and (ii) inspect locations and orientations of the ligand fragments belonging to the matching FRASEs (Figure 1B) and combine them into synthetically tractable compounds. The algorithmic implementation of step (i) is straightforward. First, it must be noted that all kinase structures (either those used to create the FRASE database or those for which ligands are to be identified) are aligned in 3D space, thus enabling an instantaneous projection of all FRASEs on the target structure (Figure 1B). Hence, the i -th protein residue belonging to a structural environment from the FRASE database (reference environment) can be directly compared to the closest (i' -th) residue in the target protein, and the match between them can be quantified as a function of their chemical similarity and proximity in 3D space,

$$\text{Tscore}_i = \exp\left(-2\left(\frac{d_{ii'}}{4}\right)^2\right) T_{ii'},$$
 where $d_{ii'}$ is the distance between the residue centers, and $T_{ii'}$ is the Tanimoto²⁷ similarity between them. The total score for a reference environment is an average over all residues in the reference protein environment, $\text{Tscore} = \langle \text{Tscore}_i \rangle$ (Tscore would be equal to one if all reference residues are chemically identical to and spatially aligned with their target counterparts). Figure 1B exemplifies evaluation of a single reference environment with Tscore_i shown for each individual residue. The total Tscore value in this example is 0.79, which is high enough and would warrant the selection of the respective ligand fragment as a

potential building block for design. The database search for matching FRASEs can be significantly restrained by setting a particular location for the FRASE's ligand fragment. This would allow one to build a new ligand sequentially, by finding the next ligand fragment close to the attachment site of the previous one (which is the way ligands were designed in this study). The final selection of ligand fragments is closely associated with step (ii), that is, assembling fragments into synthetically tractable compounds. More than one factor is taken into consideration. First, Tscore is used to narrow down the choice. As in any screening process, the quantitative end point only serves to select a manageable number of hits that can be evaluated using alternative, more time- and resource-demanding criteria (a few thousand FRASEs are usually selected based on Tscore). There is no clearly defined, optimal Tscore threshold. Certainly, it should not be set significantly lower than 0.5, but most of the times the threshold is determined by how many FRASE hits one can process. The next step is triage of the initial fragment hits. One notable aspect of ligand fragments in the FRASE database is their significant redundancy due to the fact that a fragment can be cut out of a ligand in multiple overlapping ways (this redundancy was intentionally allowed to maximize the sensitivity of the database search). We remove this redundancy by clustering on chemical similarity of the ligand fragments and the location of their geometrical centers. Finally, binding poses of the remaining fragments (irrespective of their Tscores) are visually inspected, and their relevance is assessed based on the synthetic tractability, availability of reagents, as well as the opportunity of connecting the selected fragment (to other fragments or the pre-existing structural template) by a common minimal linker, such as a single bond, ether, amide, or urea. Sometimes a fragment can be slightly modified to facilitate its fusion with the structural template.

The FRASE database, the key prerequisite for the design, integrates the information from 3D structures of protein–ligand complexes, ligand structures, and their inhibitory potencies. First, a cross-search in the Uniprot,²⁸ PDB,²⁹ and PDBbind³⁰ databases was performed to yield 2800 3D structures of high-affinity complexes (involving 2100 unique ligands and 230 kinases). Next, each ligand in every protein–ligand complex was broken into fragments (cycles and acyclic functional groups). Finally, each ligand fragment from the previous step is combined into a single structure with the nearby protein fragments (here, whole residues were used as protein fragments). A total of ~230 000 FRASEs were identified. We observed that, on average, a single structural environment can be found in ~30 kinases, that is, ~6% of the kinome. This appears to be an acceptable balance between uniqueness and redundancy. Indeed, if a structural environment occurs in every kinase, it would not be useful for designing selective inhibitors, while if it occurs in only one kinase it would only be useful for designing an inhibitor for this particular kinase (and none other). The selectivity profile of a FRASE-based inhibitor would be an intersection of kinase lists that each individual FRASE is “good” for. Technical details on the development of the database and related tools are given in the Methods section.

Kinase Back Pocket Offers Structural Variance for Selective TAM-Member Inhibitors. Our strategy consists of taking a type I TAM inhibitor and expanding it to new protein pockets featuring more significant structural differences in the vicinity of the ligand, thus providing a better opportunity to

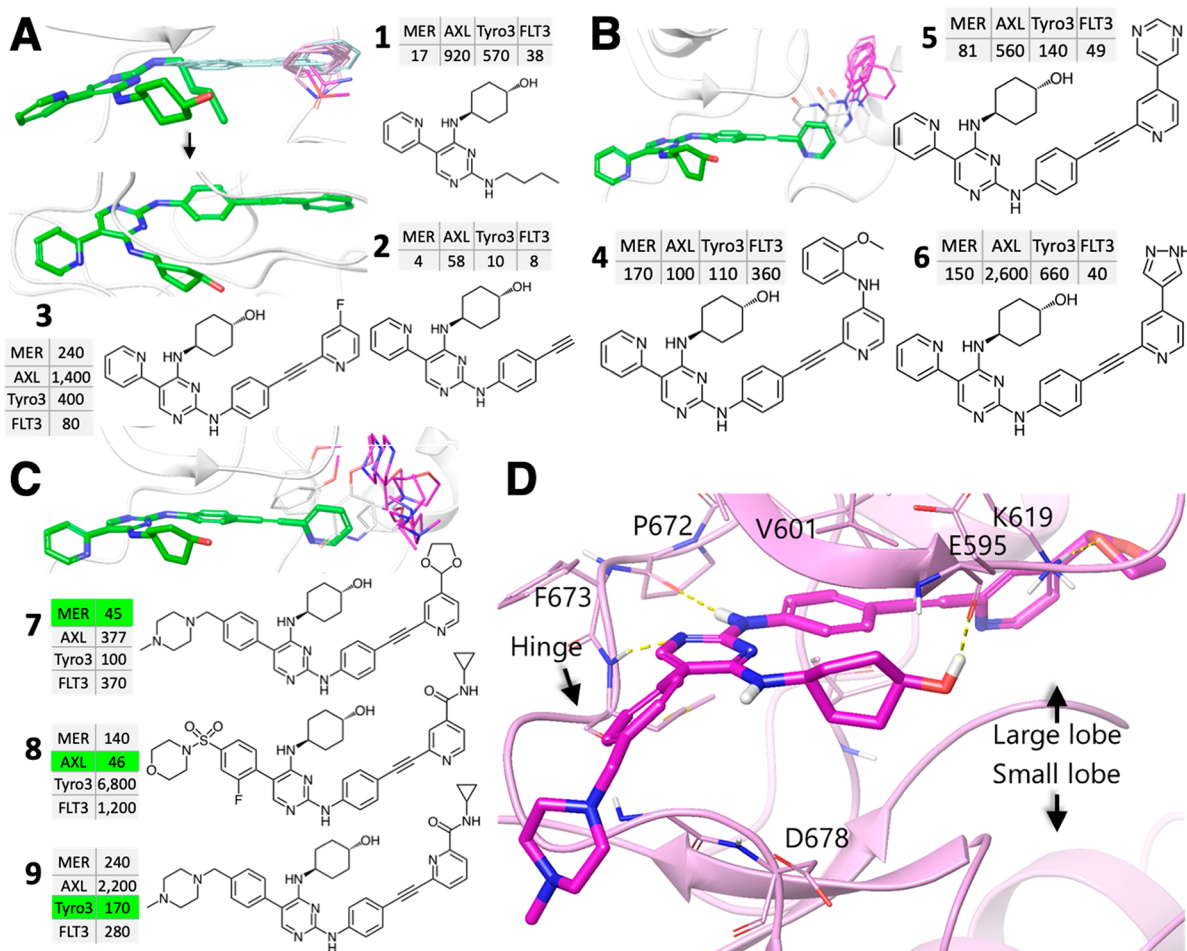


Figure 2. Steps toward potent TAM back-pocket binders. (A) Growing the aryl-alkyne-aryl spacer: template compound **1** (green), aryl-alkyne (cyan) and aryl (magenta) FRASEs, and TAM/FLT3 profiles of inhibitors (IC_{50} values given in nM) with a partial (**2**) and full (**3**) spacers. (B) Filling the back pocket by aromatic groups: template compound (green), source FRASEs (magenta), and sample inhibitors **4–6** not showing sufficient potency and selectivity. (C) Filling the back pocket with smaller aliphatic groups: template compound (green), source FRASEs (magenta), and sample inhibitors **7** and **8** showing high potency and promising selectivity. (D) Co-crystal structure of **7** in complex with MERTK and its major interactions with the adenine pocket (Pro672, Phe673, and Glu595) and back pocket (Lys619).

tune the inhibitor selectivity for any given TAM kinase. Structural analysis of all TAM members suggests that the “back pocket” (Figure 1C), located between the *c*-helix and the adjacent antiparallel β -sheet, displays significant structural diversity and hence provides multiple opportunities for enhancing inhibitors’ selectivity. The three TAM sequences of *c*-helices (the rear wall of the pocket) and adjacent antiparallel β -sheets (the ceiling) display 44 pairwise-differing residues (Figure 1D), whose side chains may either directly interact with the ligand or induce changes in relative orientations of these structural modules. This structural diversity in the back pocket is in sharp contrast with the high conservation of the adenine pocket, to which typical type I kinase inhibitors, such as our pyridine-pyrimidine **1**,³¹ bind. In this pocket, there is only one amino acid position showing variation across the TAM family: MERTK:Ile650 vs AXL:Met599 vs TYRO3:Ala581 (Figure 1E). In all three cases, the differing residues do not have close contacts or make high-energy interactions with the ligand. Hence, we hypothesized that the back pocket (Figure 1C,E) would be a more suitable target for pyridine-pyrimidine expansion. A sequential FRASE-based strategy was applied to design TAM ligands interacting with the back pocket. Inhibitor **1** was used as a

structural template with its 2-amino nitrogen on the pyrimidine ring as an attachment site for a new fragment. Then, the position of each further fragment was constrained by the proximity of the previous one and lying on an imaginary line pointing toward the back pocket.

Opening the Back Door and Reaching the Back Pocket. A first step in expanding the type I inhibitor **1** toward the back pocket was to find a spacer that would “deliver” a functional group of choice to the back pocket while avoiding bad contacts with the gatekeeper residue (MERTK:Leu671/AXL:Leu621/TYRO3:Leu603). To this end, we have extracted from our FRASE database those FRASEs that would possess the required spacer properties and match the respective structural environments of TAM family members (1786 FRASEs from 73 unique PDB structures representing 30 kinases). The triage process yielded 203 nonredundant fragments from which several most minimalistic fragments were retained as relevant for the design (Figure 2A). They were extracted from the following structures: AKT2 (PDB: 3D0E; Tscore: 0.74), CHK1 (4FST/0.53), and MELK (4UMP/0.52, 4UMT/0.57, 4UMU/0.56) kinases. The respective TAM inhibitor **2** was readily inferred, synthesized, and tested in an enzymatic microcapillary electrophoresis (MCE)

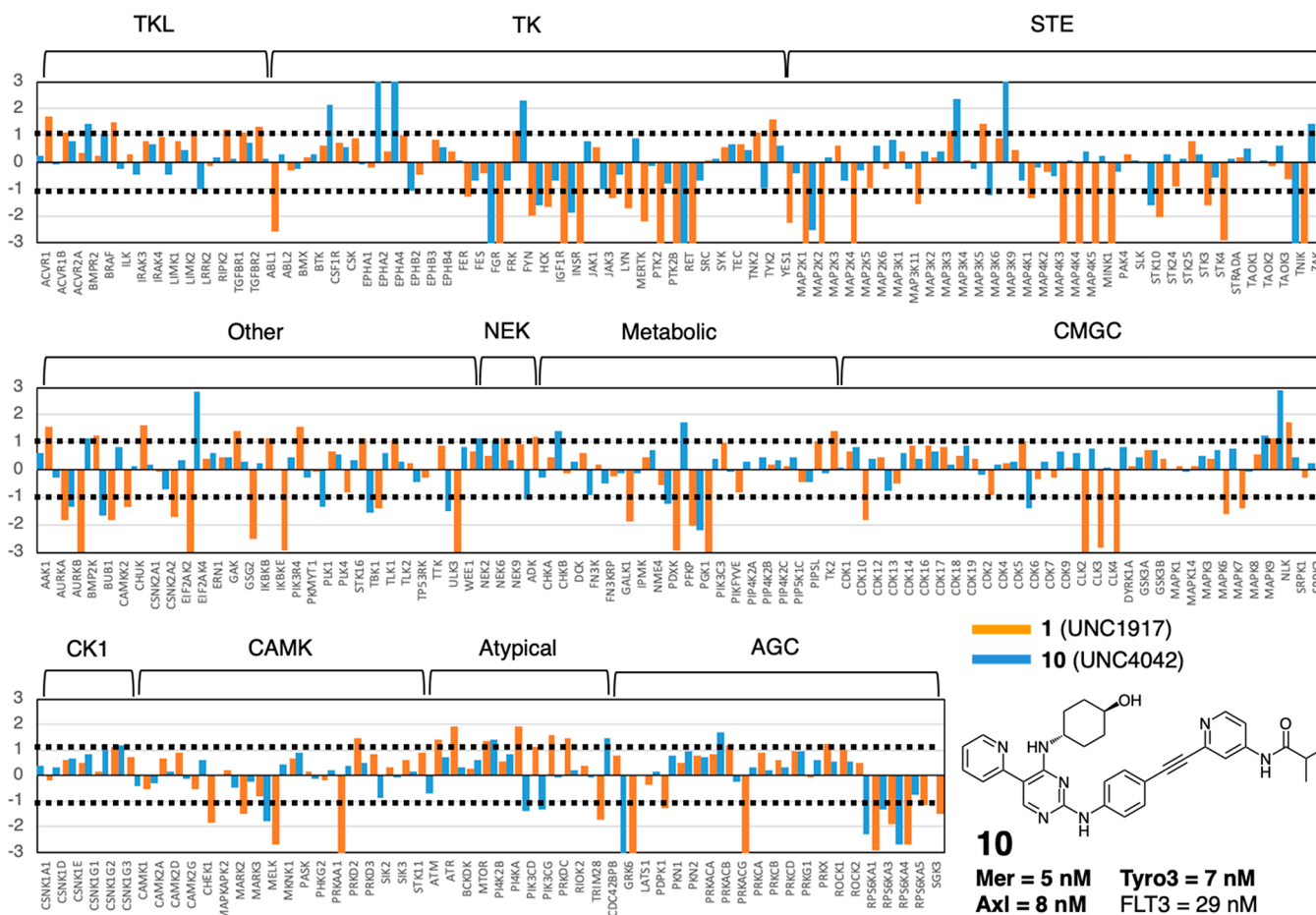


Figure 3. Kinome selectivity profiles. MIB-MS kinome profiles for inhibitor **10** (see structure and TAM/FLT3 IC₅₀ values) and the initial template **1**. Each kinase features an orange (inhibitor **1**) and a blue (inhibitor **10**) bar, with bar heights corresponding to log₂ of the inhibitor/control ratio. Thresholds of −1 and 1 (shown in dotted lines) specify significant enrichment or decrease of kinases on the beads with respect to control. The brackets with labels on the top of the charts indicate the respective kinase family.

assay³² (Figure 2A). It had low nanomolar IC₅₀ values against MERTK and TYRO3, and its potency against AXL was approximately one order of magnitude weaker (Figure 2A and Supporting Information Table S1 for a full account of IC₅₀ values in this study). The new spacer is expected to efficiently deliver any functional group to the back pocket, and any substitution at the alkyne-bound methyl would be past the gatekeeper residue. The next step was to identify a fragment that would penetrate into the back pocket. One of the relevant FRASEs, donated by CHK1 (4FST), features an alkyne–aryl fragment that aligns with the aryl–alkyne group of **2**. We also searched for fragments and features located at the entry to the back pocket and identified 151 such fragments (from 9 unique PDB structures). The most frequent feature at this location was a hydrogen bond (H-bond) acceptor pointing toward the large lobe (Figure 2B, top), which suggested using a 2-pyridine ring as an aromatic scaffold within the back pocket. The respective structural environments come from 6 PDB structures from 4 unique kinases (VEGFR2, c-Met, and p38α) with T-scores ranging from 0.50 to 0.84. Combining this 2-pyridine ring with **2** resulted in **3** (Figure 2A). Although it demonstrated reduced potencies against all TAM members (Figure 2A), inhibitor **3** provides a promising template to explore the back-pocket SAR.

Small Chemical Modifications in the Back Pocket Alter Intrafamily Selectivity. To explore the chemical space

to fill the back pocket, we used two one-point queries, with the radius of 2 Å, that mostly cover its volume. One of the queries yielded 1837 fragment hits (119 nonredundant, 30 unique PDB, 7 kinases) and the other 155 hits (21 nonredundant, 11 unique PDB, 8 kinases). The fragments from both queries can be split into two broad categories: rigid, flat aromatic groups, and more flexible and bulky aliphatic ones (a few fragments featured both aromatic and aliphatic moieties). Our first step was to investigate the impact of aromatic groups that could be readily attached to template **3** (such groups were observed in several p38α structures at the previous step of design). In the course of further visual analysis, a set of aromatic fillings were retained to inspire our ligand design. These fragments were “donated” by eight diverse kinases including BRAF1 (PDB: 3NNW, Tscore: 0.98), c-MET (3EFK/0.54, 2RFN/0.53), FAK2 (3FZT, 0.53), HCK (2C0O/0.69), JNK2 (3NPC/0.99), JNK3 (3E92/0.92, 3D7Z/0.89, 3DA6/0.65), p38α (2BAJ/0.93, 3BYS/0.91, 3HV3/0.87, 3NNV/0.87, 3BYU/0.84, 3HV5/0.84, 3HV4/0.84, 3GCU/0.84, 3GCV/0.83, 1W82/0.83, 2OFV/0.80, 3NNX/0.78, 3HV6/0.78, 3HV7/0.72, 3GCQ/0.80, 1KV2/0.69, 3NNU/0.68), and PLK (3DB6/0.70). A series of such aromatic substituents were attached to **3** through a minimal linker (see examples in Figure 2B). The respective inhibitors **4–6** have slightly improved potencies against TAM members showing, respectively, pan-TAM (**4**), MERTK/TYRO3/FLT3 (**5**), and MERTK/FLT3 (**6**) profiles.

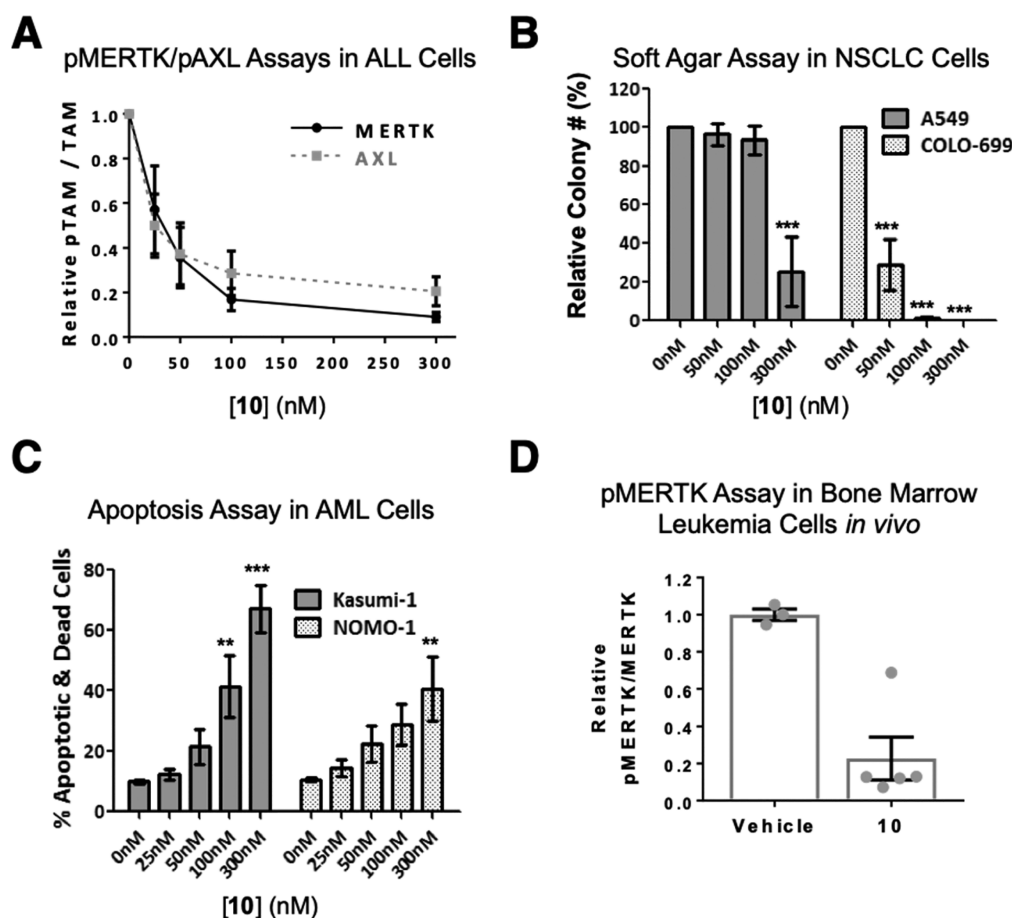


Figure 4. Target engagement in cellular and animal tumor models. Compound **10** inhibits MERTK and AXL phosphorylation, decreases colony forming potential, and induces cell death in tumor cell lines. (A) 697 ALL cells (pMERTK assay) and A549 NSCLC cells (pAXL assay) were cultured with **10** or vehicle for 1 h and then treated with pervanadate phosphatase inhibitor to stabilize phosphorylated proteins. MERTK and AXL were immunoprecipitated from cell lysates and phosphorylated, and total MERTK or AXL proteins were detected by an immunoblot. Phosphorylated and total proteins were quantitated by densitometry. (B) NSCLC (A549 and COLO-699) cells were cultured in soft agar overlaid with medium containing **10** or vehicle only for 14 days. Media and **10** were refreshed every 3 days. (C) Kasumi-1 and NOMO-1 AML cells were cultured with **10** or a vehicle for 72 h. Then cells were stained with PoPro-1-iodide (PoPro) and propidium iodide (PI) dyes, and apoptotic (PoPro +, PI-) and dead (PI+) cells were detected by flow cytometry. Mean values and standard errors from 3 to 4 independent experiments are shown. Statistically significant differences relative to the vehicle were determined using 1-way ANOVA (** $p < 0.01$, *** $p < 0.001$). (D) Compound **10** inhibits MERTK phosphorylation in bone marrow leukemia cells *in vivo*. NSG mice were transplanted with 697 human leukemia cells, and leukemic mice were treated with a single dose of 60 mg/kg of **10** or saline vehicle (VEH) by intraperitoneal injection. Bone marrow cells were isolated from femurs 2 h later and incubated with a pervanadate phosphatase inhibitor for 10 min to stabilize phosphoproteins. Phosphorylated and total human MERTK proteins were quantitated by densitometry. The ratio of phosphorylated and total MERTK proteins was determined for each sample, and values relative to the mean for vehicle-treated mice in each experiment were determined. Mean values and standard errors derived from two independent experiments are shown (** $p = 0.0026$, student's unpaired t test).

Unfortunately, the potency improvement achieved for **4–6** is not sufficient to warrant further progression of these compounds toward *in vivo* application. Hence, we explored flexible, aliphatic back-pocket-filling fragments located in potentially more diverse structural environments. They originate from 5 kinases, including BRAF1 (PDB: 3NNW, Tscore: 0.98), FAK2 (3FZT, 0.53), JNK2 (3DS6/0.66), p38 α (3F82/0.84, 3BYS/0.91, 3NNV/0.87, 3CTQ/0.85, 3BYU/0.84, 3E93/0.80, 2OFV/0.80, 3NNX/0.78, 2OG8/0.71, 3P7A/0.69, 1W83/0.69, 3P7C/0.67, 3GCQ/0.80, 3P7B/0.75, 3P79/0.75, 1WBS/0.71, 3NNU/0.68), and TIE2 (2OO8, 0.72). Incorporation of these aliphatic fragments in the template compound was less straightforward synthetically than in the previous steps of the design. We eventually developed a large series of modified inhibitors whose “decoration” in the back pocket closely matches the respective

FRASEs. Three of them, inhibitors **7–9**, are exemplified in Figure 2C and show significant improvement in potency as well as a higher diversity of selectivity profiles compared to **4–6**. In particular, compound **7** (UNC3437A) is primarily selective for MERTK with slightly lower potency against FLT3; **8** is clearly AXL-biased, a remarkable result since none of our previously published inhibitors were selective for AXL; and **9** is primarily a TYRO3-biased inhibitor with slightly lower MERTK potency. The sensitivity of the TAM/FLT3 polypharmacology profile to small chemical changes demonstrates the high potential of FRASE-based design to provide a rational method of optimization. To corroborate the structural rationale underlying the approach, we solved the cocrystal structure of MERTK in complex with inhibitor **7** (Figure 3d and Supporting Information Table S2). The structure confirmed that indeed the pyridine-dioxolane group is bound

to the back pocket extended toward the small lobe, consistent with the respective FRASE locations. On the other hand, the pyridine–pyrimidine core binding to the adenine pocket featured the same interactions with hinge residues Pro672, Phe673, and Glu595 or Asp678 as the template pyridine–pyrimidine compound **1**. Overall, the X-ray structure confirmed the assumptions underlying the design of the novel series and, more generally, the utility of the FRASE-based approach for modulating TAM/FLT3 selectivity profiles.

Kinome Profiling and *in Vivo* Pharmacokinetics. We selected a pan-TAM inhibitor **10** (UNC4042A) (Figure 3) as a lead compound for more comprehensive biological characterization. At this initial step of the series evaluation, a probe with a broad TAM potency spectrum was intentionally chosen to test whether the *in vitro* inhibitory potency translates into target engagement in a broad range of cellular and *in vivo* assays.

We next sought to know whether, and how, extending type I inhibitors toward the back pocket would affect the inhibitors' kinome-wide polypharmacology. To this end, we compared the selectivity profile of inhibitor **10** to that of inhibitor **1** (Figure 1), the initial chemical template for FRASE-based design. Multiplexed inhibitor bead kinome enrichment coupled to quantitative mass spectrometry analysis (MIB-MS) was used to determine specificity of the compounds in an unbiased manner. In MIB-MS, mixtures of Sepharose beads with covalently immobilized, linker-adapted kinase inhibitors are used to capture a large portion of the kinome present in the analyzed cell lysate.³³ The identity of the kinases captured is then established by means of MS analysis. In a competition MIB-MS assay, an inhibitor is preincubated with the lysate prior to the analysis, so that the inhibitor-bound kinases would not be captured by the beads. The kinases captured with and without the inhibitor present in lysate can then be compared to determine which kinases were down- or upregulated by the inhibitor.

In this study, NOMO-1 human AML cells were treated with DMSO, **1**, or **10** (both at 100 nM) for 1 h, and then, kinases were enriched using MIBs and analyzed by LC-MS/MS. A set of 222 kinases were quantified across the samples using a label-free approach. Inhibitor **1** showed significantly lower kinome selectivity compared to **10** (Figure 3); 62 kinases showed lower enrichment on the beads upon treatment with **1** compared to DMSO control, whereas 27 kinases were decreased upon treatment with **10**. Of these 27 kinases, 8 were significantly decreased ($p < 0.05$ and \log_2 ratio **10**/control ≥ 2). In addition to higher selectivity, the kinases decreased by **10** are not a mere subset of the 62 kinases decreased by **1**, suggesting a different mode of action. Detailed MIB-MS data, including results for inhibitors **1** and **10** at 1 μ M, are given in the Supporting Information "FRASE-Proteomics.xlsx".

We also assessed whether the lead compound is suitable for *in vivo* studies. To this end, pharmacokinetic (PK) studies in mice have been performed with compound **10** administered via intravenous (i.v.), intraperitoneal (i.p.), and oral (p.o.) routes. As can be seen from the evaluation of drug concentration in plasma over time, **10** has a PK profile suitable for *in vivo* application in murine models (Supporting Information Figure S1, $C_{\max} = 435$ ng/mL; $AUC_{po} = 1661$ ng h/mL, and $T_{1/2} = 4.97$ h via the i.p. route). The weak oral bioavailability of compound **10** precludes the p.o. route. However, oral administration is not crucial for animal studies

and can be addressed by further chemical optimization. Details of the *in vivo* PK study are described in the Supporting Information.

Target Engagement in Cellular and Animal Models.

To confirm potent on-target inhibition and tumor suppression activity mediated by the lead compound **10**, we evaluated the effects of treatment with **10** in solid tumor and hematologic malignancy models. First, we determined target engagement by monitoring the level of active, phosphorylated MERTK, and AXL in cancer cells as a function of inhibitor concentration. As expected, **10** inhibited both MERTK ($IC_{50} = 28$ nM, 95% confidence interval = 16–52 nM) and AXL ($IC_{50} = 32$ nM, 95% confidence interval = 18–57 nM) at nanomolar concentrations (Figure 4A and Supporting Information Figure S2A,B). Near complete inhibition of pMERTK and pAXL was observed in cells treated with higher concentrations of **10**. This assay provides a potency level reference for pharmacologic effects of **10** that might be observed in cellular tumor models and be attributable to MERTK or AXL inhibition.

Consistent with previous data demonstrating oncogenic roles for MERTK in AML cells and for both MERTK and AXL in NSCLC cells,^{34,35} treatment with **10** mediated potent inhibition of colony formation in soft agar cultures of both AML and NSCLC cell lines (Figure 4B and Supporting Information Figure S2C). More specifically, both OCI-AML5 AML cells and COLO-699 NSCLC cells were very sensitive to **10**, with concentrations as low as 50 nM mediating a 70% reduction in colony number and complete abrogation of colony formation in response to treatment with 300 nM **10**. The NOMO-1 AML and A549 NSCLC cell lines were sensitive to **10** as well, although with reduced potency compared to the other cell lines. In these cases, a 300 nM concentration was required to affect a significant 60–75% reduction in colony number, perhaps indicating a requirement for more complete target inhibition. In addition, treatment with 300 nM **10** was sufficient to induce tumor cell death in 67% and 41% of Kasumi-1 and NOMO-1 AML cells, respectively (Figure 4C).

Consistent with the favorable PK profile for **10** in mice, treatment with **10** also inhibited MERTK phosphorylation *in vivo*. Immune compromised NOD/SCID/IL2Rgamma $^{-/-}$ (NSG) mice were inoculated intravenously with human B-cell ALL cells to generate orthotopic xenografts, and mice with advanced leukemia were treated with a single 60 mg/kg dose of **10** administered by i.p. injection. Bone marrow cells were collected 2 h post-treatment and phosphorylated, and total human MERTK proteins were detected by immunoblot. MERTK phosphoprotein was significantly decreased in leukemia cells collected from mice treated with **10** relative to mice treated with a vehicle (0.23 ± 0.116 versus 1.00 ± 0.031 , $p = 0.0026$) (Figure 4D and Supporting Information Figure S2D). These pharmacodynamic data confirm inhibition of MERTK mediated by **10** *in vivo* and demonstrate the utility of **10** for translational studies utilizing murine models.

DISCUSSION AND CONCLUSIONS

FRASE-Based Series of *in Vivo* Antitumor Agents. The enzymatic, structural, cellular, and *in vivo* data obtained in this study suggest that the pyridine–pyrimidine–alkyne series represent useful *in vivo* chemical probes with readily adjustable intra-TAM/FLT3 selectivity profiles and the potential to significantly affect their kinome-wide polypharmacology profiles. This is exemplified by the compound **10**, which

shows an improved kinome selectivity profile in cells compared to the initial chemical template 1, cell-based activity consistent with the presumed pharmacology of the target, abrogation of oncogenic phenotypes and induction of cytotoxicity in cell-based assays, chemical stability, and a SAR profile suggestive of a specific interaction with the target.³⁶ We have shown that small chemical changes, suggested by FRASE analysis, induce significantly different selectivity profiles within the TAM family. Hence, probes with varying TAM and kinome-wide selectivity can be readily derived using additional FRASEs. Development of a more comprehensive probe set would enable a more specific analysis of the role of the TAM tyrosine kinase family and individual members in their physiologic and pathologic functions, including prevention of chronic inflammatory responses and autoimmunity. In particular, availability of more selective kinase inhibitors with specific TAM profiles may facilitate a better understanding of the extent to which inhibition of individual or all TAM kinases may relieve the deleterious immunosuppressive action of tumor-associated macrophages and myeloid-derived suppressor cells. Selective and potent inhibition would also reverse the overexpressed TAM kinase action to drive tumor cell survival and chemoresistance. Remarkably, these probes have been designed with the help of a “virtual medicinal chemist”, a cheminformatics technique that is able to “assemble” new bioactive compounds based on available 3D structures and chemogenomic data.

Interaction–Activity Relationships. We have chosen the polypharmacology of TAM kinase inhibitors as an opportune test case, given the potential importance of TAM inhibitors for cancer therapy.²³ While the focus of this study was on protein kinases, we expect the FRASE approach to be applicable to any other protein family, provided there is a sufficient amount of 3D and SAR data, and even across protein families. Indeed, FRASE could potentially serve as an informational key for data integration across multiple protein families, so that structures and data abundant in one family, such as protein kinases, would benefit protein classes with sparse structure–activity information. Such extrapolation of SAR beyond a given protein, or even a protein family, should be possible due to the malleable redundancy of FRASE descriptors. That is, the size of ligand and protein fragments constituting a single FRASE can be set small enough for similar structural environments to occur in multiple structurally diverse proteins. Hence, FRASE-based design might overcome some fundamental limitations of SAR analysis and structure-based design. The traditional SAR analysis is target-blind: no structural or sequence similarity between the target proteins is being considered. The activity (or the target) is identified by a label, and if the labels in two SAR data sets are different, the sets cannot be cross-linked irrespective of how similar the respective target proteins are. New compounds are prioritized for synthesis based on a formal or informal examination of chemical similarity between active and inactive compounds. At the same time, structure-based design is historically SAR-blind. That is, a new compound is prioritized for synthesis based on a computed or an *ad hoc* assessment of its binding affinity to the protein of interest. There is no formal algorithm that allows direct involvement of available SAR data in this decision. FRASEs, par excellence, jointly exploit explicit representations of both ligand and protein structures, thus converting conventional SAR into interaction–activity relationships to guide construction and prioritization of new compounds for

synthesis. Here, we have shown that even unweighted FRASEs, i.e., those whose “goodness” was inferred from a mere occurrence in a high-affinity complex, can be effectively used as building blocks for biologically active compounds. Future studies will focus on development and application of algorithms to identify FRASEs that are more likely than others to be “good” and to quantify the degree of “goodness”. This can potentially be achieved through a combination of machine learning techniques with a proven efficiency in chemoinformatics^{37,38} and specialized affinity assessment techniques developed for structure-based design.^{39–41}

METHODS

Methods and any associated references are available in the [Supporting Information](#).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b08660. The full FRASE database used in this study was made available to public as a shared archive at: https://drive.google.com/open?id=1xkIIdMDz_XMzdBBVikmivr8RVRoX9Q6X.

Hitlists including fragment structures and environmental information (PDF)

Source pdb ids and Tcores (ZIP)

AUTHOR INFORMATION

Corresponding Authors

*xiaodonw@unc.edu

*dmitri.kireev@unc.edu

ORCID

Thomas S. K. Gilbert: 0000-0002-6833-6788

Laura E. Herring: 0000-0003-4496-7312

Lee M. Graves: 0000-0002-4736-9855

Stephen V. Frye: 0000-0002-2415-2215

Xiaodong Wang: 0000-0003-3293-7538

Dmitri Kireev: 0000-0001-8479-8555

Notes

The authors declare the following competing financial interest(s): D.D., D.K.G., H.S.E., S.V.F., X.W., and D.K. have stock in Meryx, Inc.

ACKNOWLEDGMENTS

This work was supported by the University Cancer Research Fund (UCRF), Federal Funds from the National Cancer Institute, National Institute of Health, under Contract No. HHSN261200800001E, and a philanthropic gift to the CICBDD. We thank Dr. Brenda Temple for her help in depositing the X-ray crystallographic structure to the Protein Data Bank (PDB). This research is based in part upon work conducted using the UNC Proteomics Core Facility, which is supported by P30 CA016086 Cancer Center Core Support Grant to the UNC Lineberger Comprehensive Cancer Center.

REFERENCES

- (1) Sriram, K.; Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **2018**, 93 (4), 251–8.
- (2) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR drug discovery: new agents,

targets and indications. *Nat. Rev. Drug Discovery* **2017**, *16* (12), 829–42.

(3) Cohen, P. A.; Alessi, D. R. Kinase Drug Discovery – What's Next in the Field? *ACS Chem. Biol.* **2013**, *8* (1), 96–104.

(4) Wu, P.; Nielsen, T. E.; Clausen, M. H. Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. *Drug Discovery Today* **2016**, *21* (1), 5–10.

(5) Li, Y.; Seto, E. HDACs and HDAC Inhibitors in Cancer Development and Therapy. *Cold Spring Harbor Perspect. Med.* **2016**, *6* (10), No. a026831.

(6) Theodoulou, N. H.; Tomkinson, N. C.; Prinjha, R. K.; Humphreys, P. G. Clinical progress and pharmacology of small molecule bromodomain inhibitors. *Curr. Opin. Chem. Biol.* **2016**, *33*, 58–66.

(7) Barnash, K. D.; James, L. I.; Frye, S. V. Target class drug discovery. *Nat. Chem. Biol.* **2017**, *13* (10), 1053–6.

(8) Edwards, I. R.; Aronson, J. K. Adverse drug reactions: definitions, diagnosis, and management. *Lancet* **2000**, *356* (9237), 1255–9.

(9) Smith, G. F. Designing Drugs to Avoid Toxicity. *Prog. Med. Chem.* **2011**, *50*, 1–47.

(10) Grosser, T.; Fries, S.; FitzGerald, G. A. Biological basis for the cardiovascular consequences of COX-2 inhibition: therapeutic challenges and opportunities. *J. Clin. Invest.* **2005**, *116* (1), 4–15.

(11) Roth, B. L.; Sheffler, D. J.; Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discovery* **2004**, *3* (4), 353–9.

(12) Zavodszky, M. I.; Rohatgi, A.; Van Voorst, J. R.; Yan, H.; Kuhn, L. A. Scoring ligand similarity in structure-based virtual screening. *J. Mol. Recognit.* **2009**, *22* (4), 280–292.

(13) Vulpetti, A.; Kalliokoski, T.; Milletti, F. Chemogenomics in drug discovery: computational methods based on the comparison of binding sites. *Future Med. Chem.* **2012**, *4* (15), 1971–9.

(14) Glinca, S.; Klebe, G. Cavities Tell More than Sequences: Exploring Functional Relationships of Proteases via Binding Pockets. *J. Chem. Inf. Model.* **2013**, *53* (8), 2082–92.

(15) Keiser, M. J.; Setola, V.; Irwin, J. J.; Laggner, C.; Abbas, A. I.; Hufeisen, S. J.; Jensen, N. H.; Kuijter, M. B.; Matos, R. C.; Tran, T. B.; Whaley, R.; Glennon, R. A.; Hert, J.; Thomas, K. L. H.; Edwards, D. D.; Shoichet, B. K.; Roth, B. L. Predicting new molecular targets for known drugs. *Nature* **2009**, *462* (7270), 175–81.

(16) Lounkine, E.; Keiser, M. J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J. L.; Lavan, P.; Weber, E.; Doak, A. K.; Cote, S.; Shoichet, B. K.; Urban, L. Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **2012**, *486* (7403), 361–7.

(17) Cheng, F.; Zhou, Y.; Li, J.; Li, W.; Liu, G.; Tang, Y. Prediction of chemical–protein interactions: multitarget-QSAR versus computational chemogenomic methods. *Mol. Biosyst.* **2012**, *8* (9), 2373.

(18) Martin, E.; Mukherjee, P.; Sullivan, D.; Jansen, J. Profile-QSAR: A Novel meta-QSAR Method that Combines Activities across the Kinase Family To Accurately Predict Affinity, Selectivity, and Cellular Activity. *J. Chem. Inf. Model.* **2011**, *51* (8), 1942–56.

(19) Besnard, J.; Ruda, G. F.; Setola, V.; Abecassis, K.; Rodriguez, R. M.; Huang, X.-P.; Norval, S.; Sassano, M. F.; Shin, A. I.; Webster, L. A.; Simeons, F. R. C.; Stojanovski, L.; Prat, A.; Seidah, N. G.; Constam, D. B.; Bickerton, G. R.; Read, K. D.; Wetsel, W. C.; Gilbert, I. H.; Roth, B. L.; Hopkins, A. L. Automated design of ligands to polypharmacological profiles. *Nature* **2012**, *492* (7428), 215–20.

(20) Erickson, J. A.; Mader, M. M.; Watson, I. A.; Webster, Y. W.; Higgs, R. E.; Bell, M. A.; Vieth, M. Structure-guided expansion of kinase fragment libraries driven by support vector machine models. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804* (3), 642–52.

(21) Linger, R. M. A.; Keating, A. K.; Earp, H. S.; Graham, D. K. TAM Receptor Tyrosine Kinases: Biologic Functions, Signaling, and Potential Therapeutic Targeting in Human Cancer. In *Adv. Cancer Res.*; George, F. V. W., George, K., Eds.; Academic Press; 2008; pp 35–83.

(22) Kasikara, C.; Kumar, S.; Kimani, S.; Tsou, W.-I.; Geng, K.; Davra, V.; Sriram, G.; Devoe, C.; Nguyen, K.-Q.N.; Antes, A.; Krantz,

A.; Rymarczyk, G.; Wilczynski, A.; Empig, C.; Freimark, B.; Gray, M.; Schlunegger, K.; Hutchins, J.; Kutenko, S. V.; Birge, R. B. Phosphatidylserine Sensing by TAM Receptors Regulates AKT-Dependent Chemoresistance and PD-L1 Expression. *Mol. Cancer Res.* **2017**, *15* (6), 753–64.

(23) Graham, D. K.; DeRyckere, D.; Davies, K. D.; Earp, H. S. The TAM family: phosphatidylserine-sensing receptor tyrosine kinases gone awry in cancer. *Nat. Rev. Cancer* **2014**, *14* (12), 769–85.

(24) Nguyen, K.-Q.N.; Tsou, W.-I.; Calareso, D. A.; Kimani, S. G.; Singh, S.; Hsieh, S.; Liu, Y.; Lu, B.; Wu, Y.; Garforth, S. J.; Almo, S. C.; Kutenko, S. V.; Birge, R. B. Overexpression of MERTK receptor tyrosine kinase in epithelial cancer cells drives efferocytosis in a gain-of-function capacity. *J. Biol. Chem.* **2014**, *289* (37), 25737–49.

(25) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'Min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57* (12), 4977–5010.

(26) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137* (7), 2695–703.

(27) Nikolova, N.; Jaworska, J. Approaches to Measure Chemical Similarity—A Review. *QSAR Comb. Sci.* **2003**, *22* (910), 1006–26.

(28) Apweiler, R.; Bairoch, A.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2004**, *32*, 115–119.

(29) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **1999**/12/11. **2000**, *28* (1), 235–42.

(30) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004**, *47* (12), 2977–80.

(31) Zhang, W.; Zhang, D.; Stashko, M. A.; DeRyckere, D.; Hunter, D.; Kireev, D.; Miley, M. J.; Cummings, C.; Lee, M.; Norris-Drouin, J.; Stewart, W. M.; Sather, S.; Zhou, Y.; Kirkpatrick, G.; Machius, M.; Janzen, W. P.; Earp, H. S.; Graham, D. K.; Frye, S. V.; Wang, X. Pseudo-cyclization through intramolecular hydrogen bond enables discovery of pyridine substituted pyrimidines as new mer kinase inhibitors. *J. Med. Chem.* **2013**, *56* (23), 9683–92.

(32) Zhang, W.; DeRyckere, D.; Hunter, D.; Liu, J.; Stashko, M. A.; Minson, K. A.; Cummings, C. T.; Lee, M.; Glaros, T. G.; Newton, D. L.; Sather, S.; Zhang, D.; Kireev, D.; Janzen, W. P.; Earp, H. S.; Graham, D. K.; Frye, S. V.; Wang, X. UNC2025, a potent and orally bioavailable MER/FLT3 dual inhibitor. *J. Med. Chem.* **2014**, *57* (16), 7031–41.

(33) Duncan, J. S.; Whittle, M. C.; Nakamura, K.; Abell, A. N.; Midland, A. A.; Zawistowski, J. S.; Johnson, N. L.; Granger, D. A.; Jordan, N. V.; Darr, D. B.; Usary, J.; Kuan, P. F.; Smalley, D. M.; Major, B.; He, X.; Hoadley, K. A.; Zhou, B.; Sharpless, N. E.; Perou, C. M.; Kim, W. Y.; Gomez, S. M.; Chen, X.; Jin, J.; Frye, S. V.; Earp, H. S.; Graves, L. M.; Johnson, G. L. Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* **2012**, *149* (2), 307–21.

(34) Lee-Sherick, A. B.; Eisenman, K. M.; Sather, S.; McGranahan, A.; Armistead, P. M.; McGary, C. S.; Hunsucker, S. A.; Schlegel, J.; Martinson, H.; Cannon, C.; Keating, A. K.; Earp, H. S.; Liang, X.; DeRyckere, D.; Graham, D. K. Aberrant Mer receptor tyrosine kinase expression contributes to leukemogenesis in acute myeloid leukemia. *Oncogene* **2013**, *32* (46), 5359–68.

- (35) Linger, R. M. A.; Cohen, R. A.; Cummings, C. T.; Sather, S.; Migdall-Wilson, J.; Middleton, D. H. G.; Lu, X.; Barón, A. E.; Franklin, W. A.; Merrick, D. T.; et al. Mer or Axl receptor tyrosine kinase inhibition promotes apoptosis, blocks growth and enhances chemosensitivity of human non-small cell lung cancer. *Oncogene* **2013**, *32* (29), 3420–3431.
- (36) Frye, S. V. The art of the chemical probe. *Nat. Chem. Biol.* **2010**, *6* (3), 159–61.
- (37) Varnek, A.; Baskin, I. Machine Learning Methods for Property Prediction in Chemoinformatics: Quo Vadis? *J. Chem. Inf. Model.* **2012**, *52* (6), 1413–37.
- (38) Mitchell, J. B. O. Machine learning methods in chemoinformatics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (5), 468–81.
- (39) Lindstroem, A.; Edvinsson, L.; Johansson, A.; Andersson, C. D.; Andersson, I. E.; Raubacher, F.; Linusson, A. Postprocessing of Docked Protein-Ligand Complexes Using Implicit Solvation Models. *J. Chem. Inf. Model.* **2011**, *51* (2), 267–82.
- (40) Jain, A. N. Scoring functions for protein-ligand docking. *Curr. Protein Pept. Sci.* **2006**, *7* (5), 407–420.
- (41) Mobley, D. L.; Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annu. Rev. Biophys.* **2017**, *46* (1), 531–58.