Protein Enrichment Strategies for Advanced Proteomic Studies

Jennifer L. Brigham

A dissertation
submitted in partial fulfillments of the
requirements for the degree of

Doctor of Philosophy
University of Washington
2013

Reading Committee:
Dustin J. Maly, Chair
Michael Gelb
Champak Chatterjee

Program Authorized to Offer Degree:
Department of Chemistry

University of Washington

Abstract

Protein Enrichment Strategies for Advanced Proteomic Studies

Jennifer L. Brigham

Chair of the Supervisory Committee:
Associate Professor Dustin J. Maly
Department of Chemistry

The intracellular environment is complex, with thousands of different proteins alongside many other types of molecules. Such complexity makes the study of individual proteins very difficult without some form of enrichment. While current methods for protein enrichment have been used successfully in many different applications, they lack the specificity and robustness needed for advanced proteomic studies.

The first chapter of this work describes a new protein catch-and-release system that allows for the enrichment and selective release of target proteins under mild conditions. This system was used to enrich target proteins of therapeutic interest, which were identified by mass spectrometry. Chapter two reports the application of this enrichment strategy to find new protein targets for the treatment of Human African Trypanosomiasis (HAT) and details the validation of these targets using biochemical, genetic, and orthogonal proteomic methods.

Table of Contents

LIST OF ABBREVIATIONS

ABL	Ableson protein tyrosine kinase
ABPP	Activity-based protein profiling
α-His6	anti-His6
ASH	SNAP-tag(AGT)/SUMO/HaloTag
Asp	Aspartate
α-SRC	anti-SRC
ATP	Adenosine-5'-triphosphate
Boc	Tert-Butyloxycarbonyl
BODIPY	Boron-dipyrromethene
BSA	Bovine serum albumin
BTK	Bruton's tyrosine kinase
CAMK	Calcium/Calmodulin dependent kinase
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
CDK	Cyclin-dependent kinase
CLP	Chloropyrimidine
CMGC	CDK MAPK GSK3 CLK
CSK	C-Src kinase
DFG	Asp-Phe-Gly
DhaA	Haloalkane dehalogenase
DIEA	*N*,*N*-Diisopropylethylamine
DME	1,2 Dimethoxyethane
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSA7	Dual Src Abl inhibitor 7
DTT	Dithiothreitol
EC_{50}	Half maximal effector concentration
EDC	1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride
EDCI	*N*-[3-Dimethylaminopropyl]-*N*′-ethylcarbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
EIF2AK2	Eukaryotic translation initiation factor 2-alpha kinase 2
EPHA	Ephrin type-A receptor
equiv.	Equivalents
EtOH	Ethanol
Ex/Em	Excitation/emission

FRK	Fyn-related kinase
Glu	Glutamate
HAT	Human African Trypanosomiasis
HCK	Hematopoietic cell kinase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
His6	Hexahistidine
HLP	Hexylchloride-labeled protein
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	1 - Hydroxybenzotriazole hydrate
HPLC	High performance liquid chromatography
IC_{50}	Half maximal inhibitory concentration
IMAC	Immobilized metal affinity chromatography
IMDM	Iscove's Modified Dulbecco's Medium
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRAK4	Interleukin-1 receptor-associated kinase 4
KD	Kinase domain
K_d	Dissociation constant
kDNA	Kinetoplast DNA
LB	Lysogeny broth
LC/MS	Liquid chromatography / Mass spectrometry
LCK	Lymphocyte-specific protein tyrosine kinase
LDL	Low-density lipoprotein
LTQ	Linear ion trap quadrupole
m/z	Charge-to-mass ratio
MAPK	Mitogen-activated protein kinase
MBP	Myelin basic protien
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NEK	NimA-related kinase
NHS	N-hydroxysulfosuccinimide
NMR	Nuclear magnetic resonance
NP-40	Nonidet P-40
NTA	Nitrilotriacetic acid
OD	Optical density
ODC	Ornithine decarboxylase
P2	Purine transporter 2
PAGE	Polyacrylamide gel electrophoresis
PAK	p21 activated kinase
PBS	Phosphate buffered saline

PDB	Protein Data Bank
Phe	Phenylalanine
PMSF	phenylmethanesulfonylfluoride
RB_{50}	Half maximal residual binding
RNA	Ribonucleic acid
RNAi	RNA interference
rpm	revolutions per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
SH	Src homology
STE	Yeast sterile
SUMO	Small ubiquitin-like modifier
T. b. gambiense	*Trypanosoma brucei gambiense*
T. b. rhodesiense	*Trypanosoma brucei rhodesiense*
T. brucei	*Trypanosoma brucei*
TAP	Tandem affinity purification
TbMRPA	*T. brucei* multidrug resistance-associated protein A
TCEP	Tris[2-carboxyethyl] phosphine
tet	Tetracycline
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
Tyr	Tyrosine
Ulp1	Ubiquitin-like Protein 1
UV	Ultraviolet
WB	Western blot
WT	Wild-type

Units

°	Degree
µ	Micro
Å	Angstrom
C	Celsius
Ci	Curie
Da	Dalton
g	Gram
h	Hour
k	Kilo

L	Liter
m	Milli; meter
M	Molar
min	Minute
mol	Mole
n	Nano
p	Pico
s	Second

NMR

d	Doublet
dd	Doublet of doublets
Hz	Hertz
J	Coupling constant in Hz
m	Multiplet
MHz	Megahertz
ppm	Parts per million
s	Singlet
t	Triplet
δ	Chemical shift in parts per million

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my appreciation to my advisor Dustin Maly for providing me with the opportunity to work on this project – one that I have found unendingly interesting. I am grateful for the technical guidance and support that he has freely given. It has made me a better scientist and helped propel my research forward when it was sputtering.

Gayani Perera originally synthesized many of the probes and inhibitors included in this work. She also worked many long hours with me performing the *T. brucei* kinase enrichment experiments. Ours was a truly enjoyable collaboration and I will count her as a lifelong friend.

The work presented in chapter two would not have been possible without the collaboration of the Van Woorhis Lab, Stuart Lab, and Novartis TIP team. Rob Gillespie and K.K. Ojo generated the growth inhibition curves and provided the *T. brucei* lysate for proteomic experiments. Chris Merritt performed all of the genetic experiments included in this work and the Novartis TIP team completed the kinase affinity resin experiments.

For Sanjay

That's why it's called "research."
If you only had to do the experiment once,
it would be called "search."

VITA

2002.....Depew High School, Depew, NY

2006.....B.A. Chemistry, B.A. English,
Wittenberg University, Springfield, OH

2008 – present.....Graduate student, University of
Washington, Seattle, WA

Publications (in reverse chronological order):

Brigham J. L.; Perera, B. G.; Maly, D. J. A Hexyl-chloride Based Catch-and-Release System for Chemical Proteomic Applications. *ACS Chem. Biol.* **2013**, *8* (4), 691-199.

Krishnamurty R; Brigham J. L.; Leonard, S.E.; Ranjitkar, P.; Larson, E. T.; Dale, E. J.; Merritt, E. A,; Maly, D. J. Active site profiling reveals coupling between domains in SRC-family kinases. *Nat Chem Biol.* **2013**, *9* (1), 43-50.

Ellison, M. D.; Morris, S. T.; Sender, M. R.; Brigham, J.; Padgett, N. E., Infrared and computational studies of the adsorption of methanol and ethanol on single-walled carbon nanotubes. *J. Phys. Chem. C* **2007,** *111* (49), 18127-18134.

Chapter 1: A Hexylchloride-Based Catch-and-Release System for Chemical Proteomic Applications[1]

I. INTRODUCTION

Derivatized small molecule probes are valuable reagents for studying biology. Support-bound small molecule ligands have facilitated the enrichment of specific classes of low abundance proteins, such as protein kinases.[1-4] Furthermore, immobilized analogs of small molecules that show interesting properties in phenotypic screens are useful for identifying the intracellular targets of bioactive molecules.[5-7] Fluorophore- and biotin-modified derivatives of small molecule probes that covalently modify the active sites of their binding partners have served as effective tools for profiling the activities of various enzyme families. These activity-based protein profiling (ABPP) probes have allowed the discovery of enzymatic activities that are misregulated in various disease models and for the selectivity profiling of inhibitors in physiologically relevant contexts.[8]

The development of a number of robust bioorthogonal reactions has revolutionized the design and use of small molecule probes. These reactions allow the use of small molecule probes that contain an inert chemical handle that minimally perturbs their solubility, cell permeability, and binding properties. Examples of bioorthogonal reactions that have been successfully used for conjugation include Diels-Alder cycloadditions,[9,10] nucleophile additions to carbonyl groups,[11] Michael additions,[12] thiol-ene reactions,[13] Staudinger ligations,[14] and alkyne-azide cycloaddition reactions.[15] Bioorthogonal reactions, in particular cycloaddition reactions utilizing alkyne and azide tags, have found widespread use in chemical proteomic studies. For example, azide and

[1] Reprinted (adapted) with permission from Brigham, J. L., Perera, B. G. K. & Maly, D. J. A Hexylchloride-Based Catch-and-Release System for Chemical Proteomic Applications. *ACS Chemical Biology* **8**, 691-699. Copyright 2013 American Chemical Society.

alkyne tags have been incorporated into ABPP probes and used to examine large families of enzymes.[16-18]

Many chemical proteomic studies rely on selectively enriching covalently or non-covalently bound proteins for subsequent identification and quantification. For small molecule probes that contain a bioorthogonal chemical handle, this is usually accomplished through selective conjugation to biotin, followed by the enrichment of probe-bound proteins with an immobilized protein (avidin or streptavidin) that recognizes biotin. While this two-step enrichment procedure has been successfully used in a number of proteomic applications, there are several drawbacks to its implementation. The bioorthogonal reactions used to conjugate biotin are not always quantitative and in some cases can lead to irreversible protein aggregation and precipitation from solution.[19,20] In addition, endogenously biotinylated proteins and proteins that bind non-specifically to the affinity matrix can lead to an increase in the complexity of the sample being analyzed.[21] Furthermore, the harsh elution conditions required to elute captured proteins do not allow differentiation of specifically versus non-specifically bound proteins. While a number of biotin analogs that contain releasable linkers have been developed to overcome this limitation,[22] the use of these reagents adds an additional non-quantitative handling step to proteomic analyses. Therefore, new bioorthogonal tags that circumvent the use of biotin-streptavidin are needed.

Here, we present a new catch-and-release strategy that utilizes a hexylchloride group as a bioorthogonal chemical handle. The hexylchloride tag is unique because it allows chemo-selective and direct conjugation to a self-labeling protein through a covalent bond. By incorporating a hexylchloride tag into a small molecule probe of interest, probe-bound proteins can be enriched with an immobilized version of HaloTag, which is an engineered form of

Rhodococcus dehalogenase that undergoes a self-labeling reaction with alkylchlorides (Figure 1-1).[23] Furthermore, by using a HaloTag fusion protein that contains a protease cleavage site, captured proteins can be selectively released under mild conditions. To demonstrate the overall utility of this strategy, we show that our hexylchloride/HaloTag catch-and-release system can be used to enrich proteins that are either covalently or non-covalently bound to kinase-directed probes.

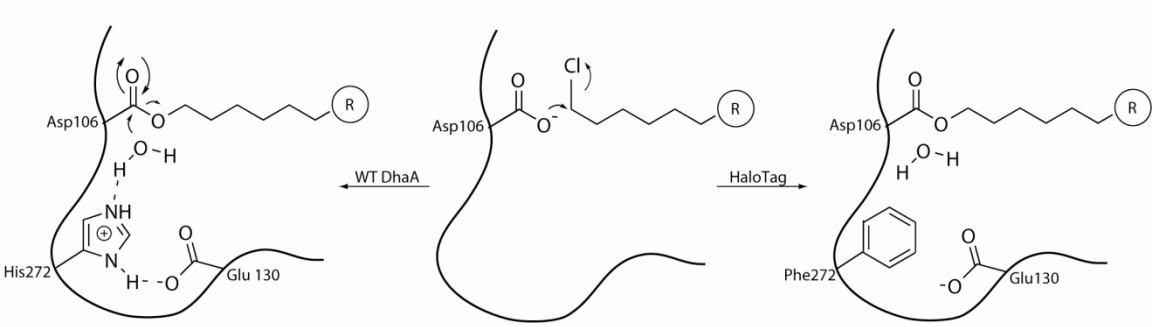


Figure 1- 1. Reaction mechanism of WT DhaA and mutant HaloTag. DhaA catalyzes the nucleophilic displacement of chlorine from alkyl chains forming an ester intermediate. In the WT dehalogenase, His272 catalyzes the hydrolysis of the ester, regenerating the enzyme. In the mutant protein, HaloTag, Phe272 cannot activate water and the enzyme is irreversibly bound to its substrate.[23]

II. RESULTS AND DISCUSSION

A. Design of a hexylchloride-based catch-and-release system

Our strategy for designing a hexylchloride-based catch-and-release system relies on the selective and rapid reaction between alkylchloride-labeled molecules and an immobilized version of the self-labeling protein HaloTag. In order to exploit this bioorthogonal reaction for proteomic studies, HaloTag must be able to be immobilized on a solid support without loss of catalytic activity. Furthermore, a method for the selective release of captured proteins is required. Towards this end, we envisioned generating a fusion protein that contains HaloTag linked through a protease cleavage site to a domain that allows immobilization to a solid support

(Figure 1-2A). The self-labeling protein SNAP-tag (also referred to as AGT), which is a mutant of O^6–alkylguanine-DNA alkyltransferase that undergoes an efficient self-labeling reaction with benzylguanine derivatives, was chosen as the immobilization domain.[24-26] SNAP-tag is an attractive immobilization option because it has already been demonstrated that fusion proteins containing this self-labeling domain can be displayed from surfaces derivatized with benzylguanine analogues. Furthermore, it has been shown that both enzymes maintain their catalytic activity in the context of a SNAP-tag/HaloTag fusion protein.[27] We next investigated the use of TEV protease, which is used in the purification of TAP-tagged proteins,[28] to selectively release immobilized HaloTag. To test the efficiency of this protease in our catch-and-release system, a fluorescently-labeled SNAP-tag/HaloTag fusion protein that contains a TEV cleavage site was generated and tested (Figure 1-2B). As expected, the fluorescently-labeled SNAP-tag/HaloTag fusion protein was efficiently captured (>90%) by an agarose resin displaying the benzylguanine derivative chloropyrimidine (CLP) (Figure 1-2B). After immobilization of the fusion protein and washing of the resin, TEV protease's efficiency in releasing the HaloTag portion of the SNAP-tag/HaloTag fusion protein was determined. Disappointingly, the ability of this protease to cleave the TEV site of this immobilized fusion protein was modest. Even at a 1:2 ratio of TEV protease to immobilized HaloTag protein, a low cleavage efficiency (~20%) was observed (Figure 1-2B). Therefore, the TEV cleavage site of the fusion protein was replaced with a SUMO (small ubiquitin-like modifier) tag, which when fused to proteins has been shown to increase their solubility and expression levels.[29] Furthermore, the SUMO Protease Ulp1 is able to cleave SUMO-tagged fusion proteins directly after the C-terminal glycine residue in SUMO with a high catalytic efficiency. This is most likely because Ulp1 recognizes the tertiary structure of SUMO, rather than a short peptide motif like the one

recognized by TEV protease. A fluorescently-labeled SNAP-tag(AGT)/SUMO/HaloTag (ASH) fusion protein was generated to test the immobilization and release of this construct. Like the fusion protein containing a TEV cleavage site, >90% of fluorescently-labeled ASH was captured by the resin. Gratifyingly, the SUMO Protease Ulp1 was able to rapidly and efficiently release the HaloTag portion of the immobilized fusion construct. A 1:80 (w/w) ratio of Ulp1: immobilized ASH allowed >85% of fluorescently-labeled HaloTag to be released from the resin (Figure 1-2B).

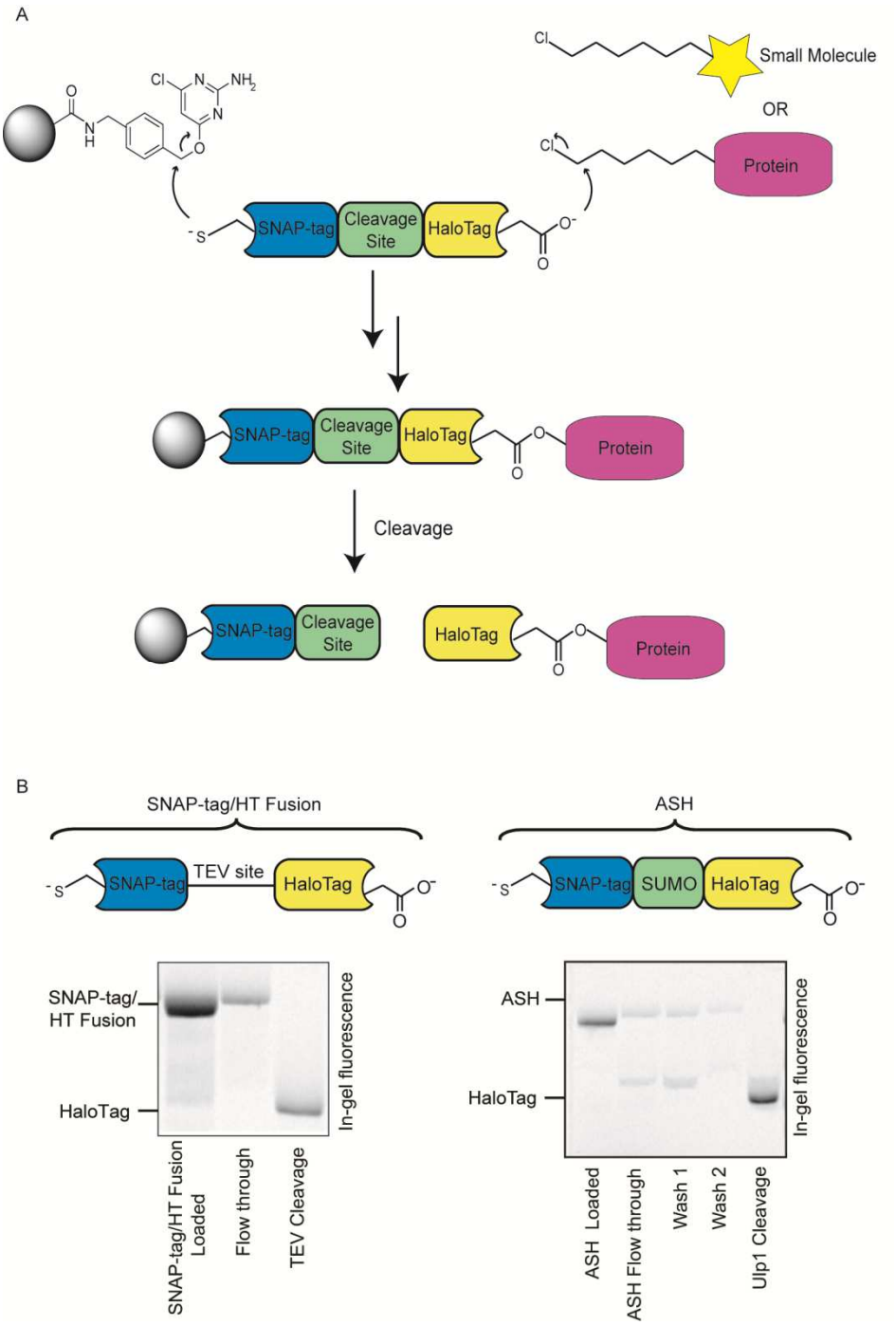


Figure 1- 2. The hexylchloride-based catch-and-release strategy. (A) Overall strategy for the catch-and-release of hexylchloride-labeled small molecules probes and proteins. A protein that contains SNAP-tag fused to HaloTag through a selectively cleavable site serves as a linker between a solid support and captured small molecule probes or proteins. The selective reaction between SNAP-tag and a chloropyrimidine (CLP)-derivatized resin allows immobilization of the bi-functional fusion protein. The support-bound HaloTag enzyme is then able to capture hexylchloride-labeled probes or proteins. Captured proteins are selectively released with a protease that cleaves the linker between SNAP-tag and HaloTag. (B) Comparison of the release efficiencies for SNAP-tag/HaloTag fusions that contain either a TEV or SUMO protease cleavage site. A fluorescently-labeled SNAP-tag/HaloTag fusion protein is captured with a CLP-derivatized resin and the fluorophore-modified HaloTag protein is released with a protease. Quantitation of the fluorescence intensities of the flow through and elution fractions allows the efficiency of the capture and release steps to be determined.

To demonstrate the stringency of the washes that can be performed with the immobilized ASH constructs, this fusion protein was immobilized on chloropyrimidine resin and then subjected to various buffers, detergents, and salts. Following these washes, HaloTag was cleaved from the resin with Ulp1 to determine the overall cleavage efficiency. The ASH catch-and-release system was found to be compatible with wash buffers containing 0.1% Tween 20, 0.1% NP-40, 0.5 mM EDTA, 1 M NaCl, and at pH 5-9 (Table 1-1). Subjecting immobilized ASH to high concentrations of detergents (> 1%), 8M urea, 6M guanidine, or SDS (> 0.1%) resulted in a dramatic decrease in the cleavage efficiency of Ulp1. This is most likely due to the partial unfolding of SUMO under these conditions.

Wash Conditions	Percent Cleavage
1 M NaCl	77%
0.5 mM EDTA	87%
pH 5	96%
pH 7	89%
pH 9	72%
6M urea	10%
0.1% SDS	1%
6M guanidine	6%
0.1% NP40	74%
0.1% Tween®20	87%

Table 1- 1. Percent elution of fluorescently labeled HaloTag protein from CLP resin after stringent wash conditions. All wash conditions were performed in 50 mM Tris pH 7.5 and 100 mM NaCl unless otherwise noted.

B. Catch-and-release of hexylchloride-labeled protein in cell lysate

Encouraged by the high cleavage efficiency of HaloTag from resin by Ulp1, we investigated the ability of the fusion protein ASH to capture proteins labeled with a hexylchloride tag. To test this, His6-SNAP-tag was covalently labeled with the hexylchloride derivative **2 (**Figure 1-3), resulting in a protein containing a single hexylchloride moiety. This

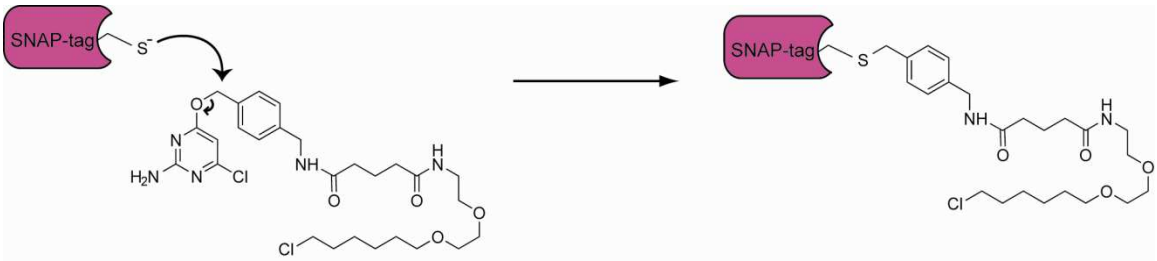


Figure 1- 3. Generation of a hexylchloride-labeled protein. The self-labeling protein SNAP-tag (AGT) undergoes a selective reaction with the chloropyrimidine moiety of compound 2. The result is a singly-labeled protein that displays the hexylchloride tag from its active site. The hexylchloride tag is directed out of the active site and is available for reaction with the enzyme HaloTag.

hexylchloride-labeled protein (HLP) was then added to mammalian cell lysate and incubated with resin displaying ASH. After a series of washes, HaloTag and the HLP/HaloTag conjugate were cleaved from the resin using Ulp1 (Figure 1-4A). Western blot analysis showed that >90% of the hexylchloride-labeled protein was captured by pre-immobilized HaloTag (Figure 1-4B). Gratifyingly, >90% of the captured HLP was released from the resin as an HLP/HaloTag conjugate following incubation with Ulp1. Only three proteins (Ulp1, unmodified HaloTag, and the HLP/HaloTag conjugate) could be detected in a silver stain analysis of the elution fraction, demonstrating the selectivity of the Ulp1 release strategy (Figure 1-4C). After optimizing the catch-and-release strategy with a pre-immobilized ASH protein, we next explored whether it would be possible to first capture hexylchloride-labeled proteins with soluble ASH in a cell lysate and then to subsequently capture this covalent complex with beads displaying a SNAP-tag substrate. Gratifyingly, this strategy provided a comparable catch-and-release efficiency to the pre-immobilization strategy (Figure 1-5).

.

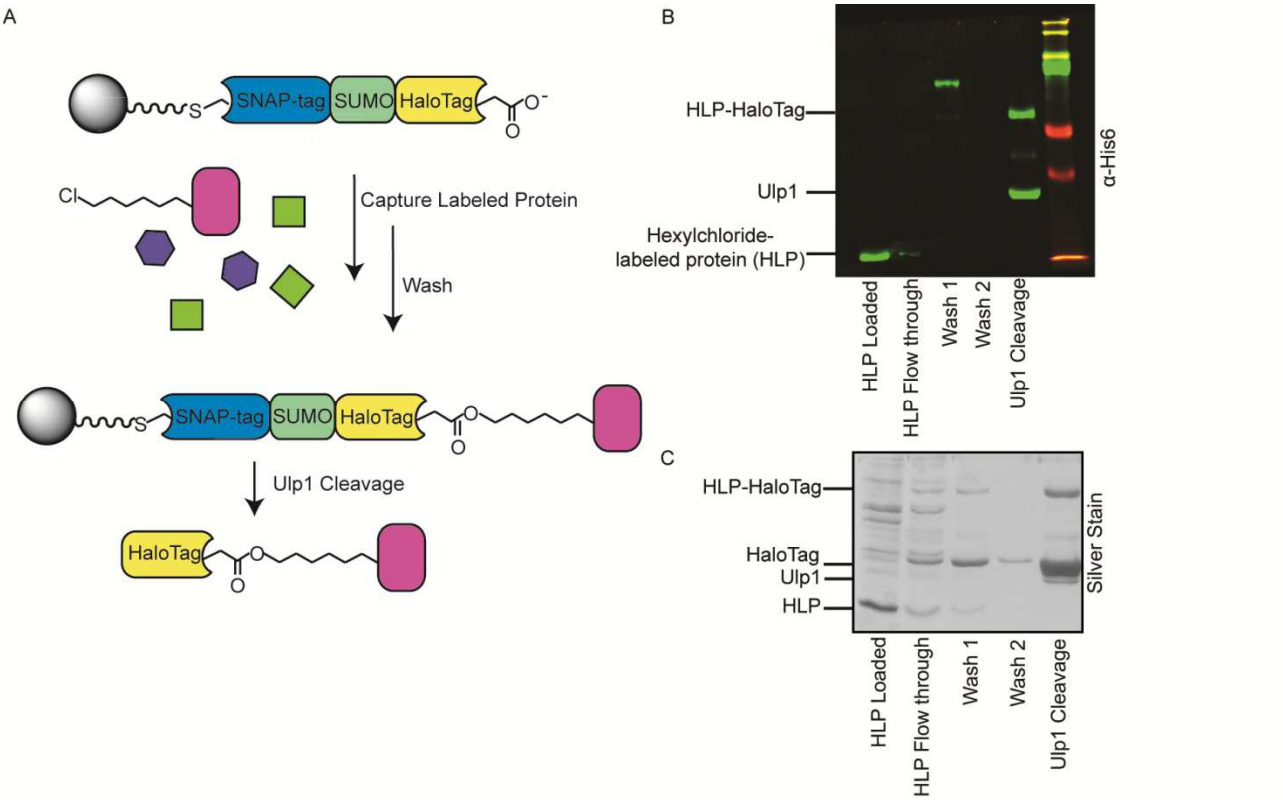


Figure 1- 4. Selective catch-and-release of a hexylchloride-labeled protein in HeLa lysate. (A) Schematic for the selective catch-and-release of hexylchloride-labeled protein (HLPs). ASH is immobilized on resin and then incubated with mammalian lysate supplemented with a protein, HLP, which contains a single hexylchloride tag (shown in pink). HaloTag selectively captures the HLP and the SUMO protease Ulp1 releases HaloTag (and any proteins that HaloTag has captured) from the beads. (B) Western Blot analysis (anti-His6) of the catch-and-release experiment described in (A). (C) Silver Stain analysis of the catch-and-release experiment described in (A).

While performing our catch-and-release experiments in cell lysate, we noticed that a small percentage of HaloTag was cleaved from the resin during the wash steps. We felt that this was likely due to the presence of endogenous SUMO proteases. To prevent the premature cleavage of HaloTag from the resin by native mammalian SUMO proteases, various protease inhibitors and inhibitor cocktails were employed. Unfortunately, none of the protease inhibitors tested was able to prevent premature cleavage of the ASH protein (data not shown). Therefore, we explored the use of an orthogonal SUMO/SUMO protease pair.[30] Mutation of two amino acid residues in the SUMO domain prevents recognition by native SUMO proteases. This SUMO

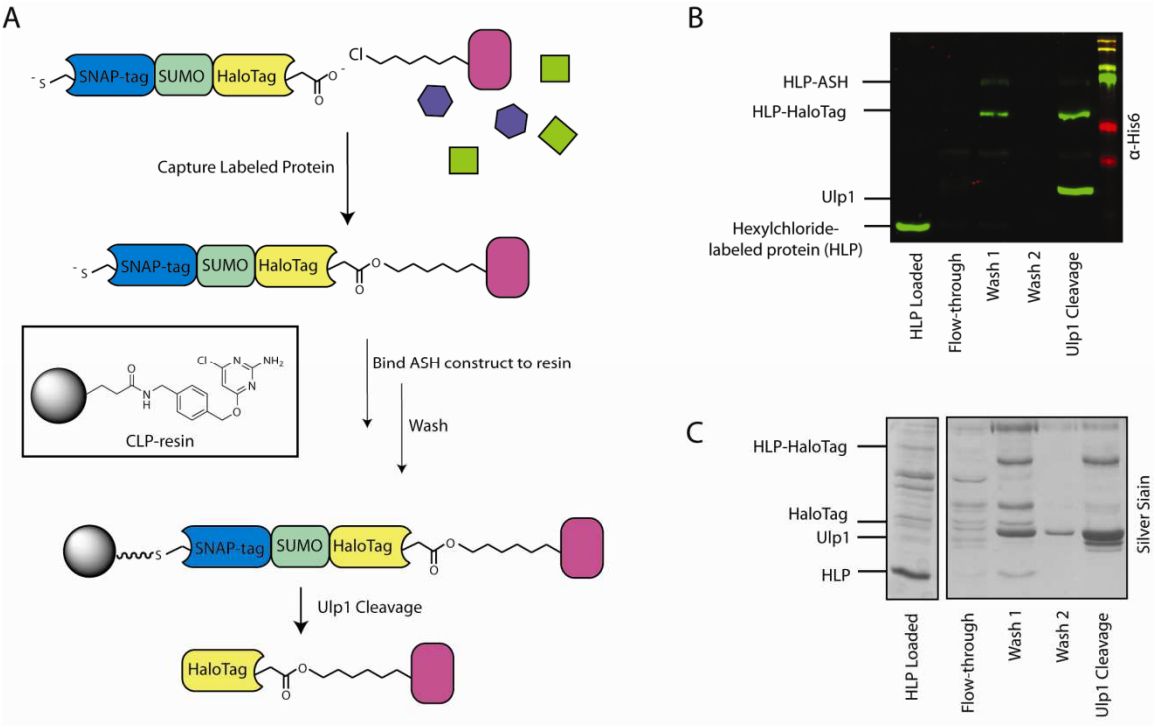


Figure 1- 5. Selective catch-and-release of a hexylchloride-labeled protein in HeLa lysate. (A) Schematic for the catch-and-release of hexylchloride-labeled protein (HLPs). ASH is incubated with mammalian lysate supplemented with a protein, HLP, which contains a single hexylchloride tag (shown in pink). HaloTag selectively captures the HLP and then is incubated with CLP-resin. The SUMO protease, Ulp1, releases HaloTag (and any proteins captured by HaloTag) from the resin. (B) Western Blot analysis (anti-His6) of the catch-and-release experiment described in (A). Greater than 90% of the HLP is captured by HaloTag and immobilized on the resin and greater than 90% of the captured protein is released from the resin by Ulp1. (C) Silver stain analysis of the catch-and-release experiment described in (A). Only three protein bands, corresponding to the molecular weights of Ulp1, HaloTag, and the HLP/HaloTag conjugate, are present in the elution.

variant, called SUMO*, can be cleaved by an engineered form of Ulp1, Ulp1*. Like Ulp1, Ulp1* cleaves the C-terminal glycine residue of SUMO*. Catch-and-release experiments in mammalian lysate were repeated with a fluorescently-labeled variant of ASH, ASH*, that contains SUMO*. As expected, endogenous SUMO proteases present in mammalian lysate do not cleave immobilized ASH*. However, Ulp1*'s efficiency in releasing HaloTag from resin (> 85%) is similar to that of Ulp1's ability to cleave ASH (Figure 1-6). Therefore, the Ulp1*/ASH* pair was used in all subsequent experiments in mammalian cell lysate.

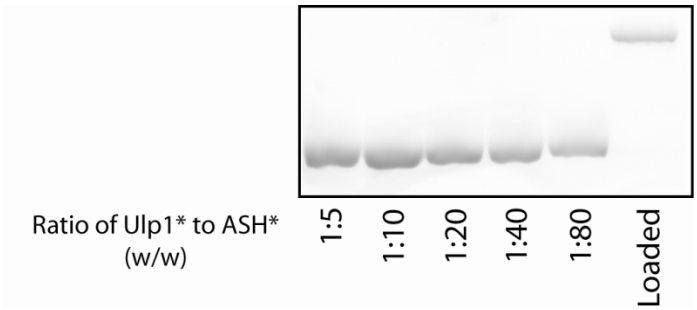


Figure 1- 6. Selective cleavage of ASH* with Ulp1*. Greater than 85% of immobilized HaloTag was released from the resin using a 1:20 (w/w) ratio of Ulp1*: immobilized ASH*.

We next investigated whether this system could be used to enrich a protein, HLP, that had been labeled with a hexylchloride tag in live cells (Figure 1-7). COS-7 cells were transfected with a protein, His6-SNAP-tag, which can be labeled with a single hexylchloride tag in cells. Transfected cells were then incubated with the hexylchloride-labeling reagent **2**. In parallel, transfected control cells were incubated with TMRstar, a rhodamine derivative that labels the model protein but does not contain a hexylchloride tag, instead of hexylchloride-labeling reagent **2**. Both sets of cells were then washed, lysed, and subjected to our optimized catch-and-release conditions with the ASH*/Ulp1* pair (Figure 1-7). Similar to the lysate pull-downs described above, only three proteins (HaloTag, Ulp1*, and the HLP/HaloTag conjugate) were detected in the elution from the catch-and-release experiments performed with cells that contain HLP (Figure 1-7B). As expected, no band corresponding to the molecular weight of HLP/HaloTag was observed in the elution from the control experiment. Therefore, the ASH*/Ulp1* pair is capable of selectively enriching hexylchloride-labeled proteins in the presence of a full complement of cellular components.

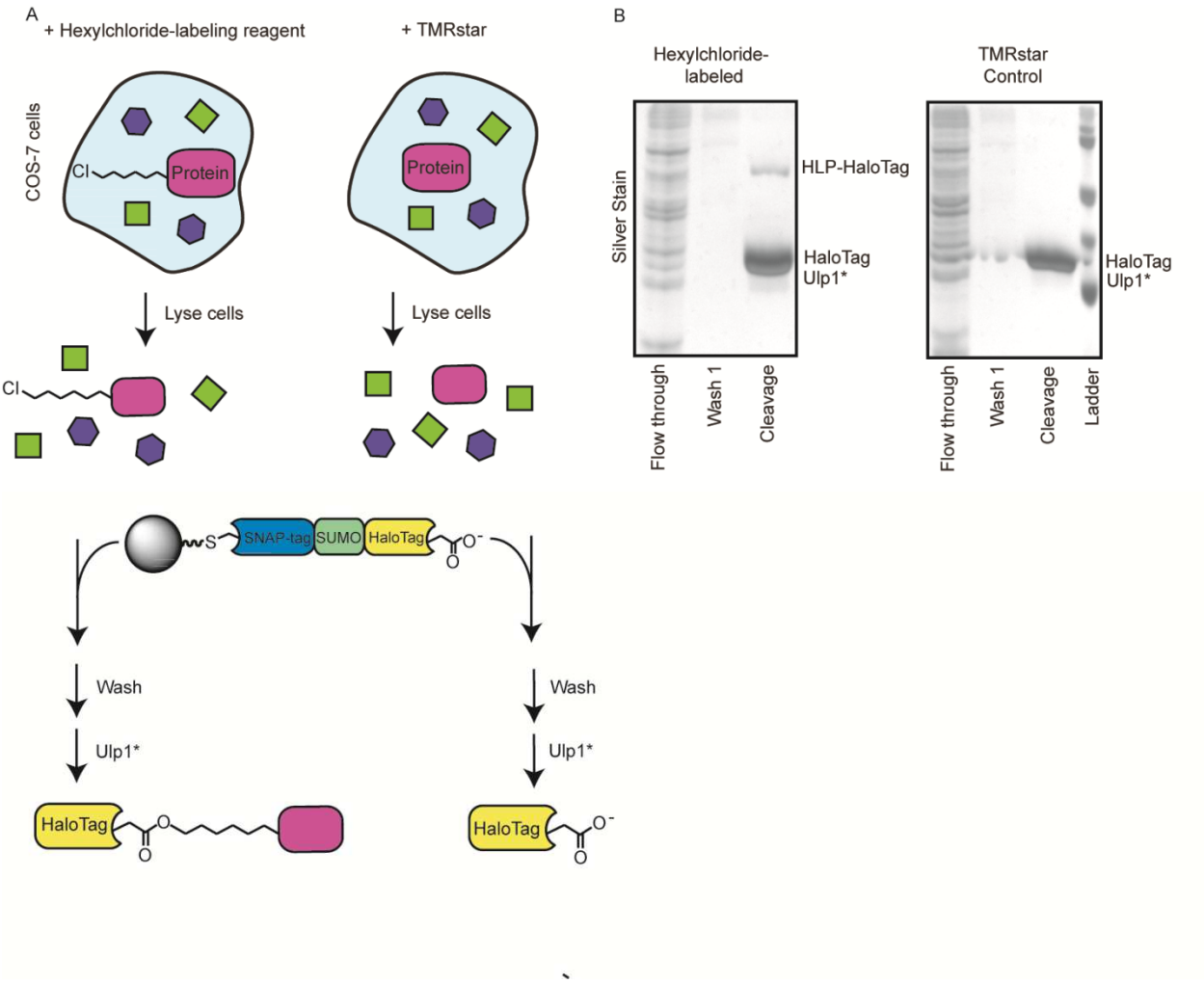


Figure 1- 7. Catch-and-release of an HLP from mammalian cells. (A) Schematic for the catch-and-release of an HLP from mammalian cells. COS-7 cells were transfected with a model protein capable of being singly-labeled with a hexylchloride tag. Transfected cells were either incubated with a hexylchloride-tagging reagent or a control compound. After washing, cells were lysed and the lysate was incubated with immobilized ASH*. Captured proteins were released by incubating the solid support with Ulp1*. (B) Silver stain analysis of the experiment described in (A).

C. Enrichment of hexylchloride-labeled protein kinases

A chemical proteomic application for our catch-and-release system was next investigated. We are particularly interested in chemical probes that are able to covalently label the ATP-binding sites of protein kinases. Affinity-based probes of this type can be used for inhibitor selectivity profiling and for the identification of aberrant kinase activities. Towards exploring

this application, a kinase-directed photo-crosslinker based on the ATP-competitive inhibitor **3** was generated (Figure 1-8). Inhibitor **3** is a potent inhibitor of the tyrosine kinase BTK and the SRC-family of protein kinases.[31] A crystal structure of the catalytic domain of BTK bound to **3** shows that this inhibitor occupies the ATP-binding site of this kinase, with the pyrazolopyrimidine scaffold forming the same hydrophobic and hydrogen-bonding interactions as the adenine ring of ATP (Figure 1-9). Interestingly, the ATP-binding site of BTK adopts an inactive conformation when bound to **3** in which a salt bridge between a conserved glutamate

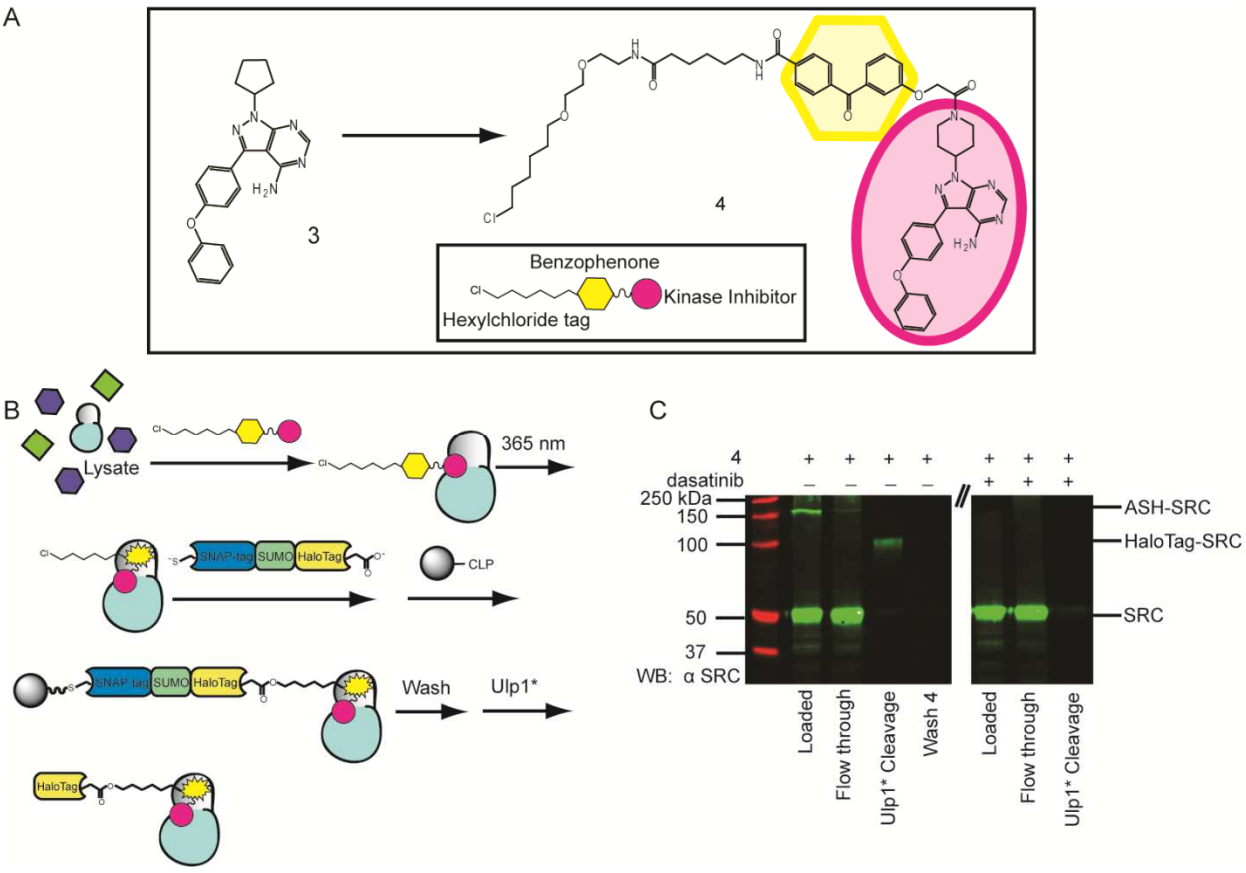


Figure 1- 8. Catch-and-release of proteins labeled by the kinase-directed photo-crosslinker **4**. (A) Chemical structure of Probe 4. Kinase inhibitor **3** was converted into a photo-crosslinker by attaching a photo-activatable benzophenone moiety and a hexylchloride tag. (B) Pulldown schematic for the catch-and-release of proteins photo-labeled by probe **4**. Cell lysate was incubated with probe **4**, followed by irradiation with UV light. The lysate was then successively incubated with ASH* and CLP-resin. After extensive washing, captured proteins were released from the resin by incubating with Ulp1*. A parallel experiment was performed in the presence of an active site competitor (dasatinib). (C) Western blot analysis (anti-SRC) of the catch-and-release of the tyrosine kinase SRC. Photo-labeled SRC is captured by ASH* and a SRC/HaloTag complex is released from the beads with Ulp1*.

and the catalytic lysine is disrupted by the movement of helix αC in the N-terminal lobe.[32,33] In order to convert inhibitor **3** into a kinase-directed probe that is able to covalently label kinase targets, it was derivatized with a benzophenone photo-crosslinker at a position that would not be expected to disrupt its interaction with the ATP-binding sites of protein kinases. A hexylchloride tag was also incorporated into the benzophenone moiety of probe **4** to allow enrichment of labeled kinase targets with our catch-and-release system (Figure 1-8A). Probe **4** was tested for its inhibitory activity in an *in vitro* assay with the tyrosine kinase SRC to confirm that modification of the kinase inhibitor scaffold does not adversely affect its ability to interact with protein kinases. Gratifyingly, probe **4** potently inhibits the catalytic activity of SRC (IC_{50} = 49 nM).

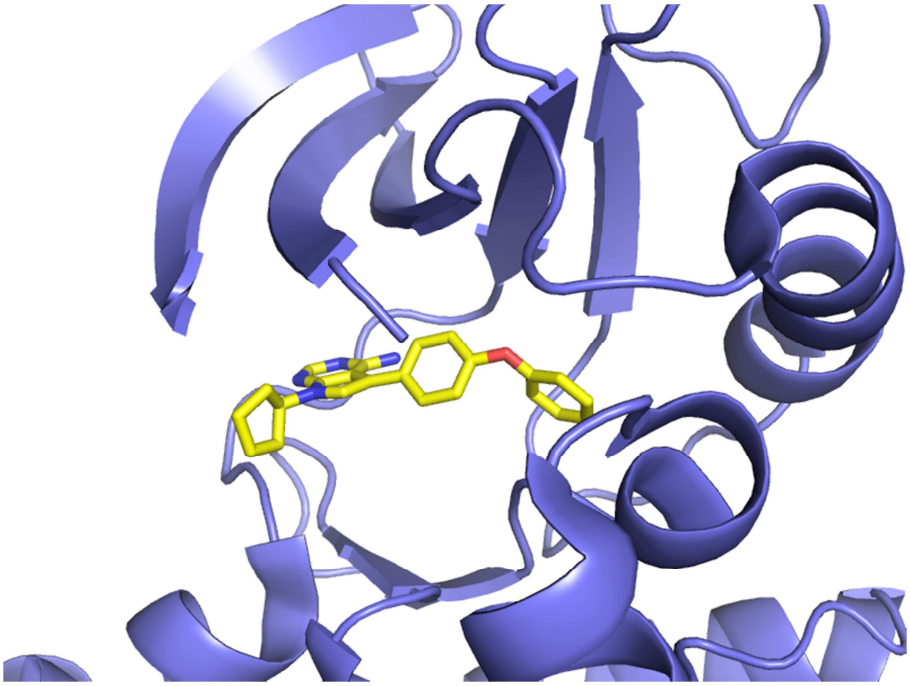


Figure 1- 9. Crystal structure of BTK bound to **3** (PDB: 3GEN). Inhibitor **3** binds to BTK in the SRC/Cdk-like inactive conformation. In this conformation, the movement of the helix-αC out of the ATP-binding site creates a hydrophobic pocket occupied by the biphenyl ether of inhibitor **3**.

The chemo-selectivity of the HaloTag labeling reaction allows for two different labeling and enrichment strategies with probe **4**. In the first strategy (pre-conjugation), probe **4** is

conjugated to the active site of HaloTag in the ASH* fusion protein. The ASH*-**4** conjugate is then incubated with cell lysate, and subsequently irradiated with UV light to covalently label any bound targets. The SNAP-tag portion of ASH* can be used to enrich any targets that are covalently linked to the ASH*-**4** conjugate. In the second strategy (post-conjugation), unconjugated probe **4** is incubated with cell lysate, and then irradiated with UV light. Labeled proteins can be captured using ASH*. With both strategies it is possible to determine the labeling efficiency of probe **4** for a specific protein target by subjecting samples to SDS-PAGE and immunoblotting with a target-specific antibody. Tethering a covalently modified target to ASH* results in a significant gel shift compared to the unlabeled protein and the percentage of crosslinked target protein can be determined with a gel-shift assay (Figure 1-10).

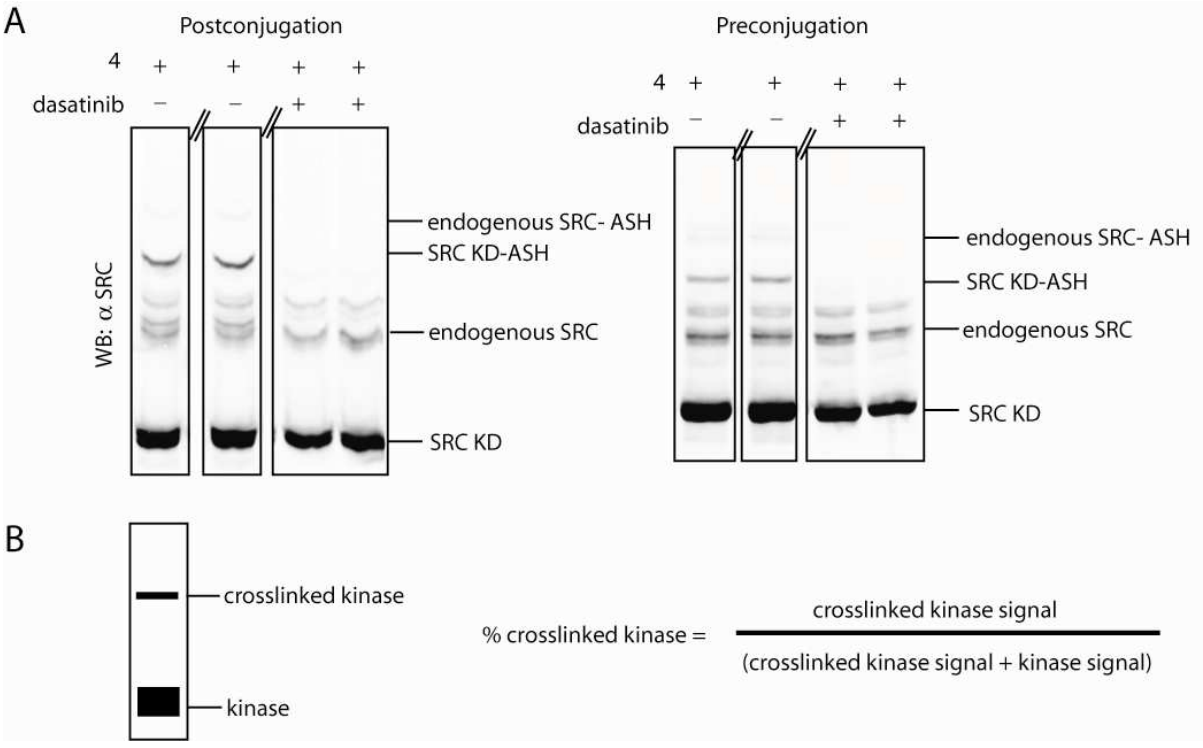
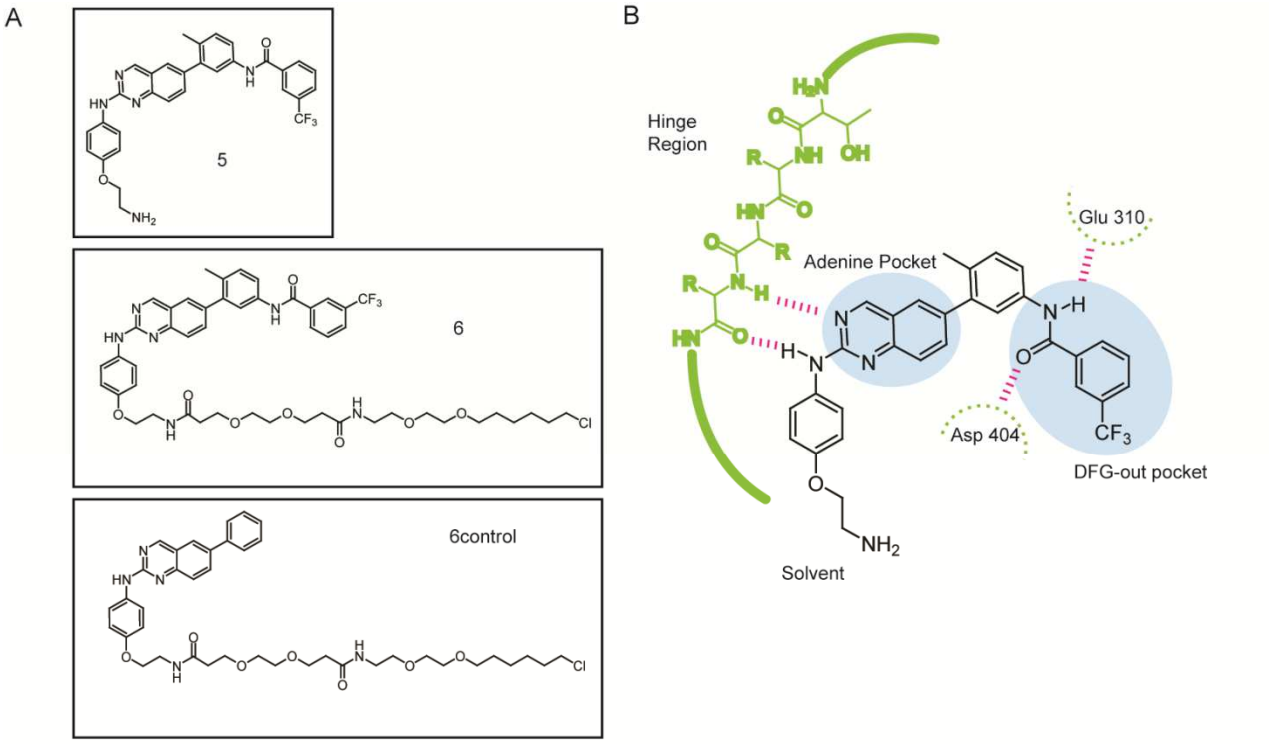


Figure 1- 10. Crosslinking efficiency of Probe **4**. (A) Western blot analysis (anti-SRC) of the pre- and post-conjugation crosslinking to SRC. The molecular weight shift of SRC shows that approximately 10% of total SRC is crosslinked in the postconjugation method while less than 3% of SRC is crosslinked in the preconjugation method. (B) Equation for calculating percentage of crosslinked kinase.

The crosslinking efficiency of both strategies was determined for probe **4** with the kinase domain of SRC (SRC KD) using a gel-shift assay. For the pre-conjugation method, the pre-assembled ASH*-**4** conjugate was added to mammalian cell lysate (1 mg mL^{-1}) supplemented with SRC KD and then irradiated with UV light for 20 minutes. Immunoblot analysis of the crosslinking reaction shows that 2.6% of SRC KD is labeled by ASH*-**4** (Figure 1-10). As expected, no gel-shifted SRC KD is observed when crosslinking is performed in the presence of a competitor (dasatinib, 10 μM), demonstrating that labeling is active site directed. When unconjugated probe **4** was first incubated with cell lysate (1 mg mL^{-1}) supplemented with SRC KD and then irradiated with UV light prior to conjugation to ASH* (post-conjugation method), a higher crosslinking efficiency was observed (10% of SRC KD). Like the pre-conjugation method, no photo-labeled SRC KD was observed in the presence of an active site competitor. Due to the higher crosslinking efficiency of the post-conjugation method, all subsequent experiments were performed with unconjugated probe **4**. Next, a catch-and-release experiment was performed with probe **4** in COS-7 cell lysate that had been supplemented with 50 nM SRC (Figure 1-8B). Lysate containing probe **4** was irradiated with UV light, and then incubated with ASH*. Western blot analysis showed that 11% of total SRC was labeled by probe **4** (Figure 1-8C). The irradiated lysate was then added to beads displaying CLP and the affinity matrix was then washed extensively prior to incubation with Ulp1* (Figure 1-8). Analysis of the flow-through of the capture step shows that nearly all of the ASH*-conjugated SRC was captured by the resin. Furthermore, almost all of the captured SRC was released from the beads as a SRC/HaloTag conjugate upon incubation with Ulp1*. Thus, the ASH* construct can be used to efficiently catch and release proteins that have been covalently labeled with a hexylchloride-derivatized probe.

D. Enrichment of Non-covalently Bound Kinases

Beyond the catch and release of covalently labeled proteins, we were interested in exploring whether the ASH* system can be used to enrich proteins that are non-covalently bound to a hexylchloride-tagged probe. To test this, catch-and-release experiments were performed with hexylchloride-derivatized ligands that stabilize a specific inactive form, called the DFG-out conformation, of the ATP-binding sites of protein kinases (Figure 1-11A). The DFG-out inactive conformation is characterized by an almost 180° rotation of the conserved Asp-Phe-Gly (DFG) motif, relative to the active conformation, which exposes a hydrophobic pocket that is accessible to small molecule inhibitors. A number of inhibitors, commonly referred to as type II inhibitors, have been identified that stabilize kinases in the DFG-out conformation (Figure 1-11B).[34-37] Type II ligands contain three conserved features: (1) A hetero-aromatic moiety that makes many of the same contacts as the adenine ring of ATP, (2) a hydrophobic group that occupies the pocket created by the movement of the DFG motif, and (3) a hydrogen bond donor/acceptor pair that interacts with the backbone of the DFG motif and a conserved glutamic acid in helix αC. In this study, we were interested in identifying the kinase targets of inhibitor **5**, which is an analogue of a series of potent type II inhibitors that target the SRC-family kinase LCK.[38] **5** potently inhibits the catalytic activities of protein kinases that have been structurally characterized in the DFG-out conformation (Figure 1-11C).



C

IC_{50}(nM)

	SRC	ABL	HCK	LCK	CSK	STK10	p38α	EPHA3	MAP3K5	PAK4	PAK5	IRAK4
5	<2.0	<2.0	<0.5	<2.0	8.4 ± 0.2	<20.0	<10.0	<10.0	>10,000	>10,000	>10,000	>10,000

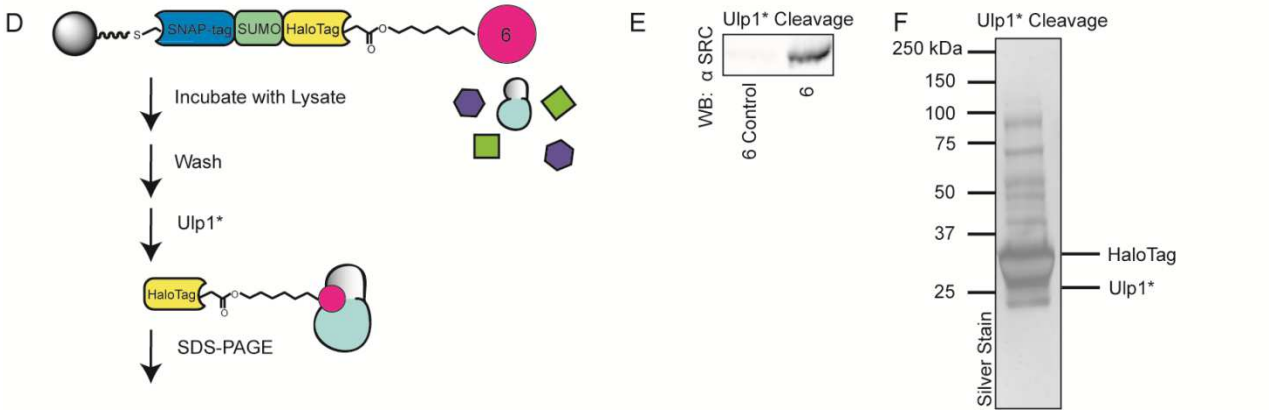


Figure 1- 11. Catch-and-release of protein kinases non-covalently bound to a probe that stabilizes an inactive ATP-binding site conformation. (A) Structures of inhibitors **5**, **6**, and **6control**. (B) A schematic of the binding interactions that the type II inhibitor **5** makes with the ATP-binding sites of protein kinases (residue numbering for the tyrosine kinase SRC is shown). (C) Inhibition data for **5** against a panel of protein kinases. (D) The overall strategy for the catch-and-release of lysate proteins bound to probes **6** and **6control**. HeLa lysate was added to pre-immobilized ASH* displaying **6** or **6control**. After extensive washing, bound proteins were released with Ulp1*. Eluted proteins were subjected to SDS-PAGE and then identified my tandem mass spectrometry (MS/MS). (E) Western blot analysis (anti-SRC) of a catch-and-release experiment performed with HeLa lysate that was enriched with the tyrosine kinase SRC. (F) Silver stain analysis of the catch-and-release experiment performed with **6**. Proteins that were selectively enriched with probe **6** are shown in Table 1-3.

Many type II inhibitors form high affinity interactions with their kinase targets and have slow association and dissociation kinetics. These features have made performing affinity chromatography with type II ligands challenging because it is difficult to elute inhibitor-bound kinases under non-denaturing conditions. For example, we previously developed a set of affinity reagents based on a general type II pyridinyl triazine inhibitor scaffold that are able to efficiently capture kinase targets, but we were unable to elute bound proteins with a soluble competitor.[39,40] Support-bound kinases could only be released in the presence of a detergent, which increases the number of non-specifically captured proteins that are eluted, and complicates mass spectrometric analysis. To determine whether this is also true for affinity reagents based on inhibitor **5**, we generated a support bound analogue of this ligand. Immobilized **5** was incubated with mammalian lysate supplemented with SRC kinase, washed extensively, and subjected to elution conditions of increasing stringency. Like the previous series of type II affinity reagents that we have characterized, none of the captured kinases could be eluted with a soluble competitor (100 μM of **5**). Elution with 0.5% and saturated SDS released a small percentage of SRC from the resin; only boiling the affinity matrix in SDS-loading buffer resulted in the release of a majority of the captured SRC kinase (Figure 1-12). Unfortunately, the latter conditions also led to the elution of a number of other proteins that were non-specifically retained by the affinity matrix. Clearly, conditions that allow selective release of probe-bound protein targets would greatly benefit the proteomic characterization of ligand **5**.

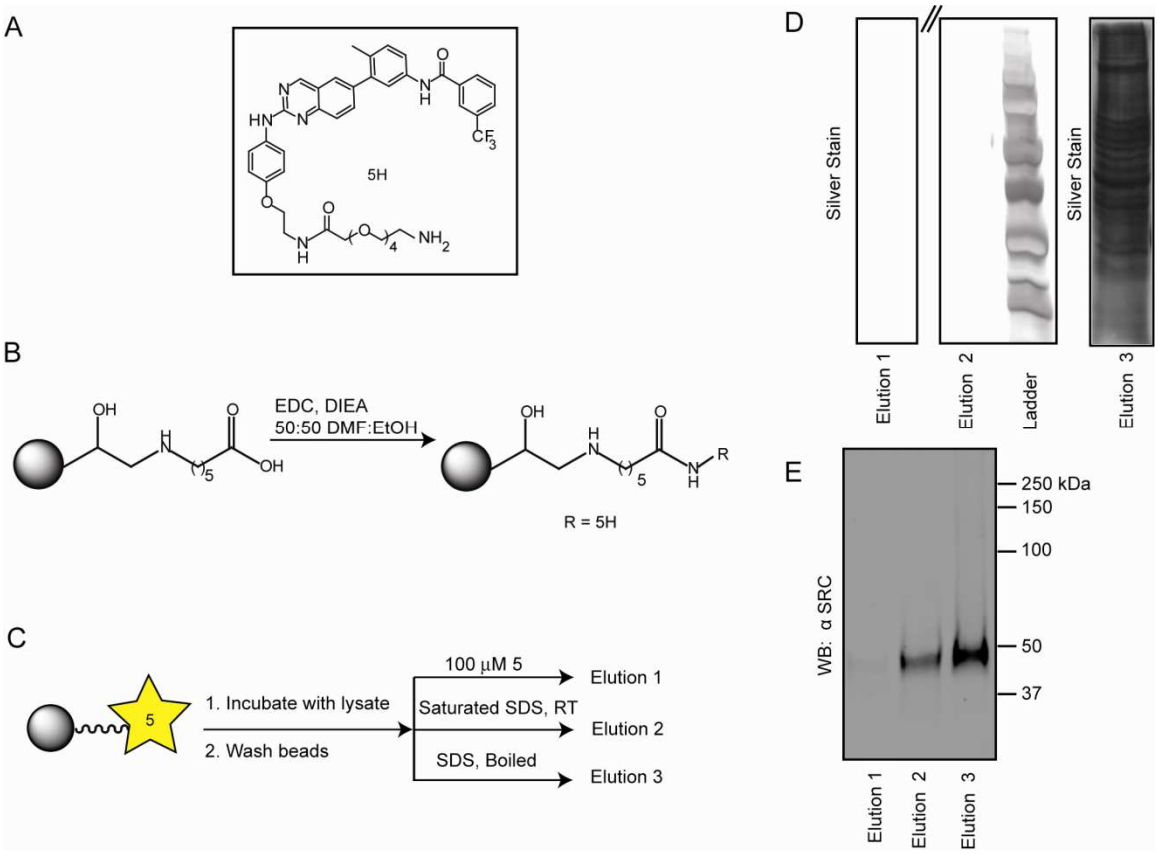


Figure 1- 12. Synthesis and characterization of affinity matrix **5H**. (A) Structure of inhibitor **5H**. (B) Procedure for generating affinity matrix with inhibitor **5H**. (C) Overall strategy for the catch-and-release of lysate proteins with affinity matrix **5H**. HeLa lysate was incubated with the affinity matrix and then thoroughly washed. Three different elution conditions were used to release bound proteins. (D) Silver stain analysis of the catch-and-release experiment described in (C). Elution conditions 1 and 2 resulted in retention of bound proteins on the resin. Elution condition 3 resulted in release of many proteins from the resin. (E) Western blot analysis of the catch and release experiment described in (C) (lysate was supplemented with 10 nM SRC). Elution condition 3 (boiling the resin in SDS loading buffer) resulted in the greatest release of SRC from the resin.

Modification of inhibitor **5** with a hexylchloride tag generated probe **6**. To aid in the identification of specific kinase targets during proteomic analysis, a control hexylchloride-tagged compound, **6control**, was also generated. **6control** contains the same core inhibitor scaffold as **6** but lacks the substituents that interact with the hydrophobic pocket created by movement of the DFG motif, which greatly reduces the affinity of most kinases for this inhibitor (Table 1-2). With the new hexylchloride-tagged compounds in hand, we optimized our catch-and-release

	IC_{50} (nM)	
	6	**6control**
SRC	**2.2**	**640**

Table 1- 2. IC_{50} values for compounds **6** and **6control** against SRC.

conditions with mammalian lysate that had been supplemented with SRC kinase. **6** and **6control** were each immobilized on resin displaying ASH* and then incubated with mammalian lysate. Both resins were washed to remove any non-specifically-bound proteins and then incubated with Ulp1* (Figure 1-11). Western blot analyses of the elutions from both catch-and-release experiments demonstrate that immobilized **6** is able to enrich SRC kinase, while immobilized **6control** cannot (Figure 1-11E).

Encouraged by these results, we tested the ability of ligand **6** to enrich endogenous kinases from HeLa lysate. Silver stain analysis of the eluted fractions from this pull-down experiment demonstrated that Ulp1* was able to release a number of proteins from ASH*-immobilized **6** (Figure 1-11F). The proteins eluted from catch-and-release experiments performed with **6** and **6control** were separated by SDS-PAGE, trypsinized, and then analyzed by mass spectrometry for identification. In total, there were 13 protein kinases that were enriched by ASH*-immobilized **6** but not by ASH*-immobilized **6control** (Table 1-3). The majority of the selectively enriched kinases are members of the Tyrosine Kinase group. Three of the tyrosine kinases, SRC, LCK, and CSK, have been structurally characterized in the DFG-out conformation.[38,41-43] In addition to SRC and LCK, three additional SRC-family kinase members, YES1, LYN, and FRK, were also enriched. While these three kinases have not been observed in the DFG-out conformation, the similar sensitivity of SRC-family members to type II inhibitors makes it highly likely that they are able to adopt this inactive form. All other enriched kinases

Kinase	# of peptides identified	Kinase Group
PTK2	3	Tyrosine
EPHA2	2	Tyrosine
FRK	1	Tyrosine
SRC	8	Tyrosine
YES1	2	Tyrosine
LYN	3	Tyrosine
CSK	18	Tyrosine
CRKRS	1	CMGC
CDK3	1	CMGC
ULK3	1	Other
EIF2AK2	1	Other
SMG1	1	Atypical
ZAK	1	Tyrosine Kinase-like

Table 1- 3. Kinases selectively enriched from HeLa lysate with ASH*-immobilized **6**.

have not been structurally characterized in the DFG-out conformation, although EPHA3 and EPHA7, which are highly homologous to EPHA2, have been observed in this inactive form.[44] To verify that newly identified kinases are sensitive to the type II inhibitor displayed from immobilized **6**, we performed activity assays with PTK2 and EIF2AK2 in the presence of **5** and **5control** (Table 1-4). Consistent with their selective enrichment by ASH*-immobilized **6**, **5** is a submicromolar inhibitor of both PTK2 and EIF2AK2. Furthermore, neither kinase is inhibited by the highest concentration of **5control** (10,000 nM) tested.

	IC_{50} (nM)	
	5	**5control**
PTK2	**450**	**>10000**
EIF2AK2	**250**	**>10000**

Table 1- 4. IC_{50} values for compounds **5** and **5control** against PTK2 and EIF2AK2.

III. CONCLUSION

We have developed a new catch-and-release strategy that allows the selective enrichment of hexylchloride-labeled probes and proteins. This new enrichment strategy leverages the rapid

and selective reaction between the self-labeling protein HaloTag and alkylchlorides. By using a HaloTag fusion protein that contains an immobilization domain (SNAP-tag) and a protease cleavage site, it is possible to selectively release captured proteins with a protease. We demonstrate that our strategy is able to efficiently capture and selectively release proteins that have been labeled with a hexylchloride tag. Furthermore, we show that kinases that are non-covalently bound to a hexylchloride-derivatized probe can be enriched with our catch-and-release system. The ability to selectively elute probe-bound proteins creates numerous opportunities for proteomic analysis, including the identification of the targets of small molecules that show interesting properties in phenotypic screens and the active site profiling of low abundance enzyme families.

IV. EXPERIMENTAL

A. SYNTHESIS

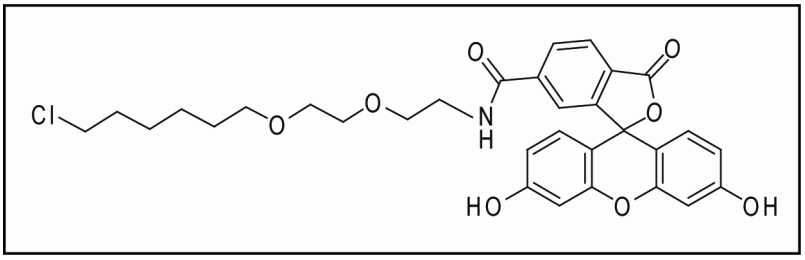
i. General information

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. ^{1}H-NMR spectra were obtained on a Bruker AV-300, AV-301, or Bruker 500 instrument at RT. Chemical shifts are reported in ppm, and coupling constants are reported in Hz. Mass spectrometry was performed on a Bruker Esquire Ion Trap MS instrument.

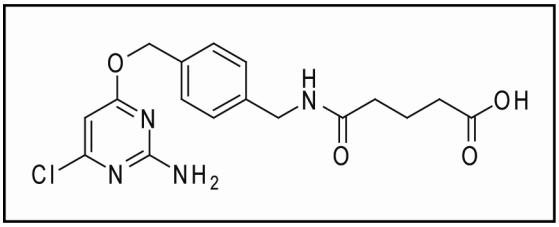
General HPLC purification conditions: Samples were injected on a preparatory reverse-phase C_{18} column (250 x 21 mm) run over 60 min. at 8 mL/min (Acetonitrile/Water–0.05% TFA gradient: 1:99 to 100:0). Purified products were detected by UV (254 nm).

General analytical HPLC conditions: General Analytical HPLC Conditions: C_{18} column (150 x 2.1 mm), Acetonitrile/Water-0.05% CF_3CO_2H gradient: 1:99 to 100:0 over 30 min. Methanol/Water–0.05% CF_3CO_2H gradient: 1:99 to 100:0 over 30 min. Flow rate = 1mL/min; 254 nM detection for 30 min.

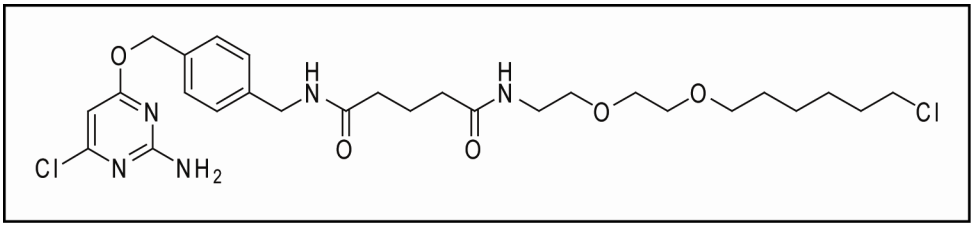
ii. Synthetic procedures



[1] Synthesized as previously reported .[23]



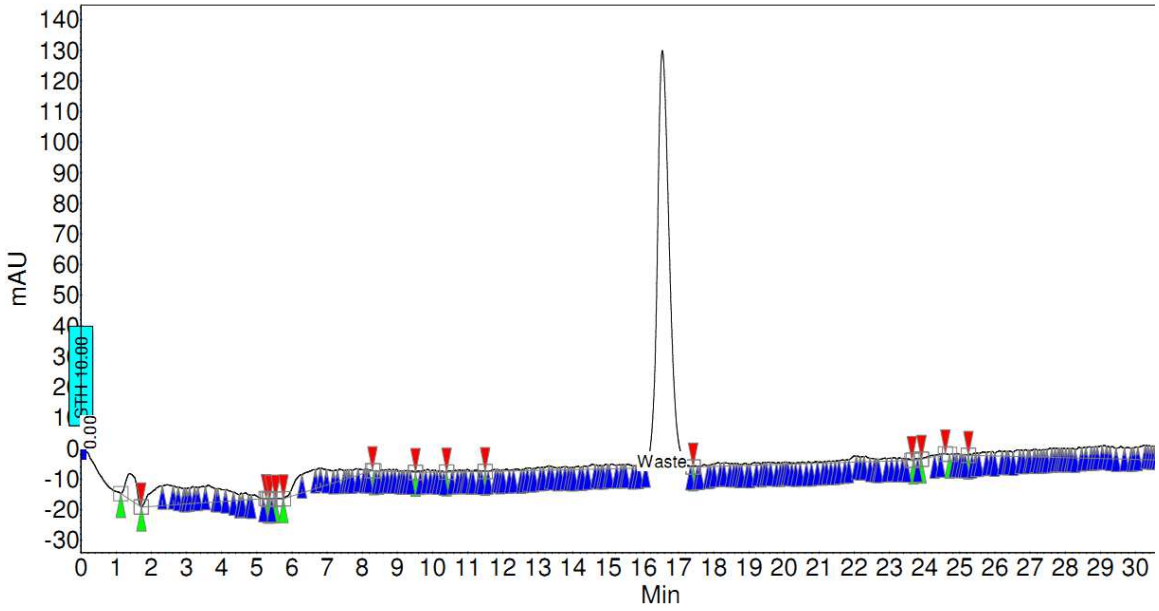
[2A] Synthesized as previously reported.[45]



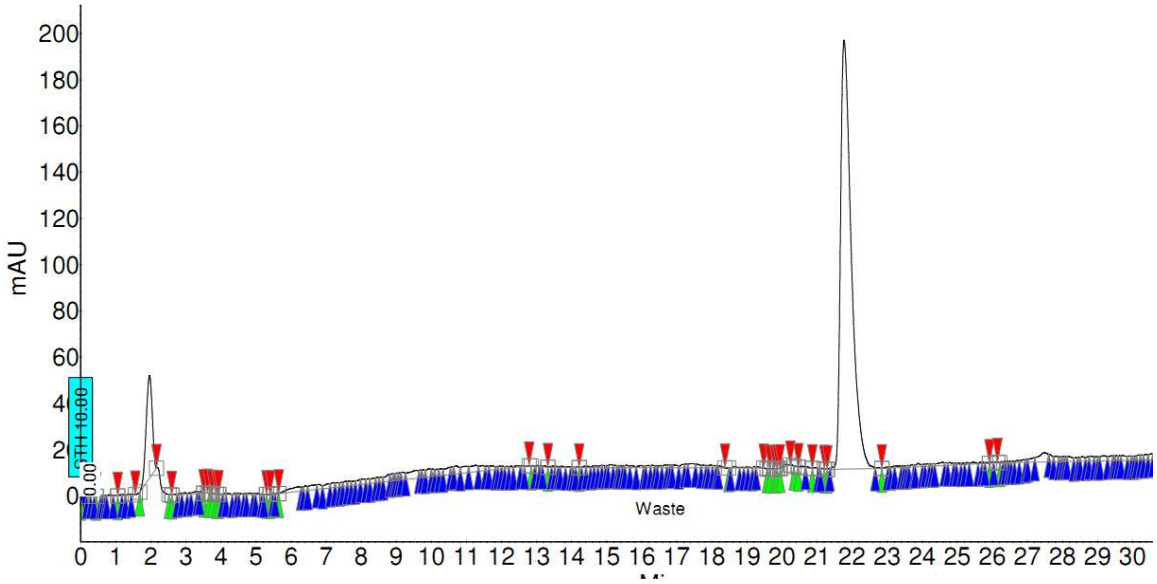
[2] 2A (5.0 mg, 0.013 mmol, 1 equiv.) and 2-(2-(6-chlorohexyloxy)ethoxy)ethanamine (8.9 mg, 0.026 mmol, 2 equiv.) were dissolved in DMF (80 µl) at 0 °C. HOBt (2.2 mg, 0.015 mmol, 1.1

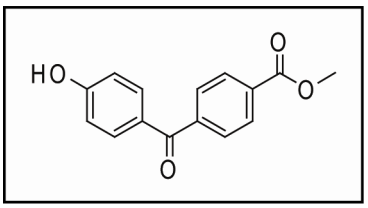
equiv.), DIEA (8 µl, 0.046 mmol, 3.5 equiv.) and EDCI (3.7 mg, 0.014 mmol, 1.3 equiv.) were added sequentially and the reaction was warmed to RT overnight. The reaction was diluted in acetonitrile/water and purified using General HPLC Purification conditions. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.46 – 7.29 (m, 4H), 6.19 (s, 1H), 5.33 (s, 2H), 4.46 (d, *J* = 6.1 Hz, 2H), 3.67 – 3.51 (m, 8H), 3.50 – 3.34 (m, 4H), 2.34 – 2.24 (m, 2H), 2.11 – 1.97 (m, 2H), 1.88 – 1.73 (m, 2H), 1.66 – 1.29 (m, 8H). Calculated for $C_{27}H_{39}Cl_2N_5O_5$ ($M+H^+$): 584.2; found 584.4.

Analytical HPLC trace of **2**: Acetonitrile/Water

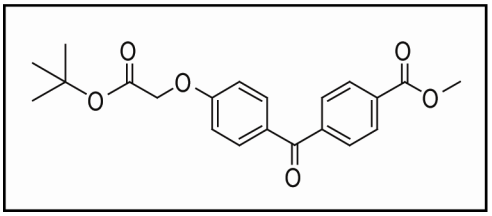


Analytical HPLC trace of **2**: Methanol/Water



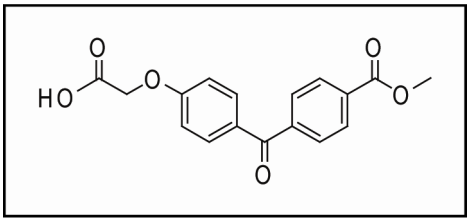


[4A] Synthesized as previously reported.[46]

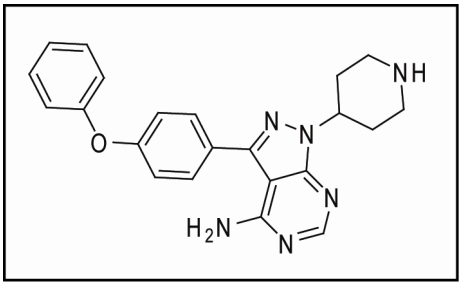


[4B] Under a nitrogen atmosphere, **4A** (284 mg, 1.11 mmol, 1 equiv.), tert-butyl bromoacetate (216 mg, 1.11 mmol, 1 equiv.), potassium carbonate (306 mg, 2.22 mmol, 2 equiv.), tetrabutylammonium iodide (36.9 mg, 0.100 mmol, 0.09 equiv.), and acetonitrile (2.3 mL) were combined, and the resulting mixture was heated to 70 °C. After 3 h, the reaction was cooled to

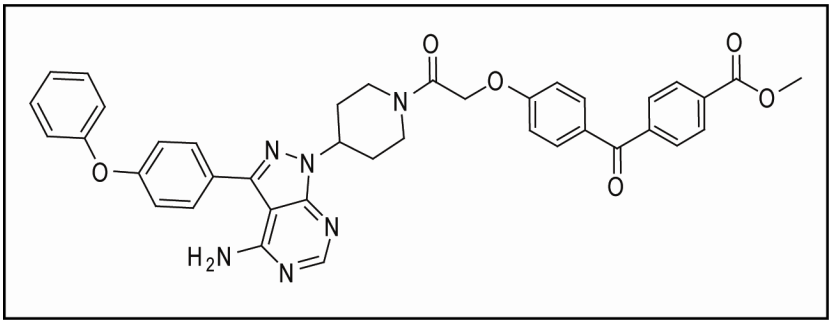
25 °C using an ice-water bath. The reaction was concentrated under reduced pressure. The concentrate was dissolved in dichloromethane, water was added, and the layers were separated. The organic layer was washed with water, dried over Na_2SO_4 and the solvent evaporated under reduced pressure to yield a pale yellow solid (91% yield) ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.08 (d, *J* = 9 Hz, 2H), 7.76 – 7.71 (m 4H), 6.90 (d, *J* = 9 Hz, 2H), 4.54 (s, 2H), 3.90 (s, 3H), 1.43 (s, 9H) Calculated for $C_{21}H_{22}O_6$ ($M+H^+$): 371.1; found 371.3.



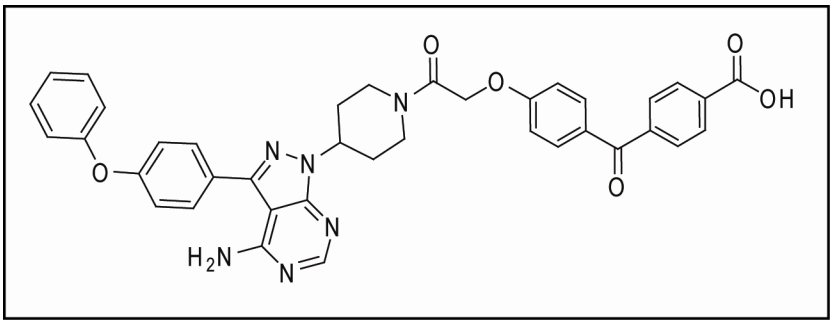
[4C] Under a nitrogen atmosphere, **4B** (373 mg, 1.01 mmol) and dichloromethane (1.9 mL) were combined at 0 °C. Trifluoroacetic acid (970 µl, 33% TFA) was added dropwise. The reaction was warmed to RT and stirred for 3 h. Toluene (1 mL) was added to the reaction and the resultant mixture was concentrated under vacuum. The product was used in the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.17 (d, *J* = 7.9 Hz, 2H), 7.88 – 7.81 (m, 4H), 7.04 (d, *J* = 9.1 Hz, 2H), 4.81 (s, 2H), 3.99 (s, 3H) Calculated for $C_{17}H_{14}O_6$ ($M+H^+$): 315.1; found 315.4.



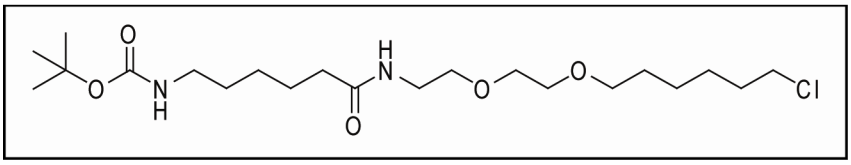
[4D] Synthesized as previously reported.[47]



[4E] 4D (25.6 mg, 0.050 mmol, 1 equiv.) and **4C** (20.3 mg, 0.064 mmol, 1.3 equiv.) were dissolved in DMF (243 µl) at 0 °C. HOAt (8.8 mg, 0.064 mmol, 1.3 equiv.), DIEA (26 µl, 0.15 mmol, 3 equiv.) and EDCI (12.3 mg, 0.064 mmol, 1.3 equiv.) were added sequentially and the reaction was warmed to RT overnight. The reaction was diluted with ethyl acetate and washed with K_2CO_3 and saturated NH_4Cl. The product was then purified on silica, eluting with dichloromethane/methanol to yield the product (62%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.39 (s, 1H), 8.15 (d, *J* = 8.5 Hz, 2H), 7.82 (dd, *J* = 12.0, 8.7 Hz, 4H), 7.64 (d, *J* = 8.7 Hz, 2H), 7.45 – 7.37 (m, 2H), 7.24 – 7.02 (m, 7H), 5.08 – 5.04 (m, 1H), 4.86 (s, 2H), 4.78 – 4.74 (m, 1H), 4.22 – 4.18 (m, 1H), 3.98 (s, 3H), 3.41 – 3.37 (m, 1H), 3.00-2.95 (m, 1H), 2.46 – 2.07 (m, 4H). Calculated for $C_{39}H_{34}N_6O_6$ ($M+H^+$): 683.7; found 683.4.

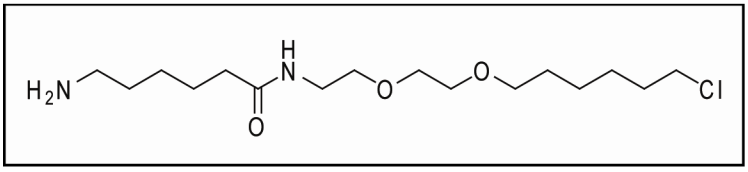


[**4F] 4E** (51.5 mg, 0.0756 mmol, 1 equiv.) was dissolved in dioxane. 2.2 equivalents of lithium hydroxide monohydrate (6.98 mg, 0.166 mmol) and 300 µl of water were added to the reaction mixture. The reaction was stirred at RT for 6 h. The reaction was taken up in water and washed with ethyl acetate. The aqueous layer was acidified with 3 M HCl and extracted with ethyl acetate. Extracted ethyl acetate layers were combined, dried with Na_2SO_4, and concentrated. The product was used in the next step without further purification. Calculated for $C_{38}H_{32}N_6O_6$ ($M+H^+$): 669.2; found 669.4.

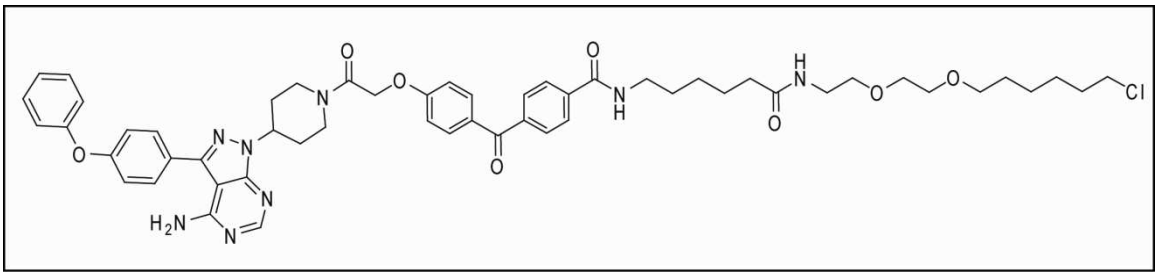


[4G] 2-(2-(6-chlorohexyloxy)ethoxy)ethanamine (63 mg, 0.186 mmol, 1 equiv.) and Boc-6-aminohexanoic acid (56 mg, 0.242 mmol, 1.3 equiv.) were dissolved in DMF (93 µl) at 0 °C. HOAt (33 mg, 0.242 mmol, 1.3 equiv.), DIEA (97 µl, 0.558 mmol, 3 equiv.) and EDCI (46 mg, 0.242 mmol, 1.3 equiv.) were added sequentially and the reaction was warmed to RT overnight. The reaction was diluted in ethyl acetate and washed with K_2CO_3 and saturated NH_4Cl. The product was then purified on silica, eluting with dichloromethane/methanol to yield the product. ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.69 – 3.56 (m, 6H), 3.56 – 3.44 (m, 4H), 3.19 – 3.07 (m, 2H),

2.24 – 2.16 (m, 2H), 2.00 – 1.89 (m, 2H), 1.85 – 1.75 (m, 2H), 1.74 – 1.63 (m, 6H), 1.46 (s, 9H), 1.42 – 1.32 (m, 6H). Calculated for $C_{21}H_{41}ClN_2O_5$ ($M+H^+$): 436.3; found ($M+Na^+$): 459.5.

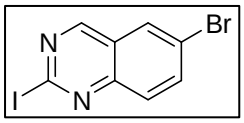


[4H] 4G (81 mg, 0.186 mmol) was dissolved in 30% TFA/DCM (9.3 mL) on ice. The reaction was stirred at RT for 3 h. Toluene (l mL) was added and the reaction was concentrated under reduced pressure. The reaction was diluted in acetonitrile/water and purified using General HPLC Purification conditions. ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.71 – 3.49 (m, 10H), 3.50 – 3.34 (m, 4H), 2.21 (t, *J* = 7.0 Hz, 2H), 1.86 – 1.73 (m, 2H), 1.72 – 1.53 (m, 6H), 1.51 – 1.29 (m, 6H). Calculated for $C_{16}H_{33}ClN_2O_3$ ($M+H^+$): 337.2; found 337.5.

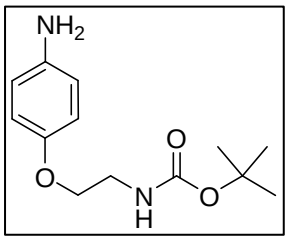


[4] 4F (7.0 mg, 0.010 mmol, 1 equiv.) and **4H** (9.4 mg, 0.021 mg, 2 equiv.) were dissolved in DMF (52 µl) at 0 °C. HOAt (1.84 mg, 0.014 mmol, 1.3 equiv.), DIEA (6.4 µl, 0.031 mmol, 3 equiv.) and EDCI (2.6 mg, 0.014 mmol, 1.3 equiv.) were added sequentially and the reaction was warmed to RT overnight. The reaction was diluted in acetonitrile/water and purified using General HPLC Purification conditions to obtain 3.3 mg of **4**. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.23 (s, 1H), 7.89 (d, *J* = 8.0 Hz, 2H), 7.82 (d, *J* = 8.6 Hz, 2H), 7.78 (d, *J* = 8.1 Hz, 2H), 7.55 (d, *J* = 8.4 Hz, 2H), 7.44 – 7.41 (m, 2H), 7.41 – 7.26 (m, 1H), 7.15 (d, *J* = 8.4 Hz, 2H), 7.09 (d, *J* = 7.9

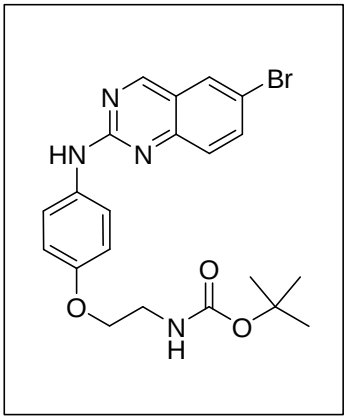
Hz, 2H), 7.05 (d, *J* = 8.6 Hz, 2H), 4.85 (s, 2H), 3.66 – 3.44 (m, 19H), 2.42 – 2.27 (m, 2H), 2.23 (t, *J* = 7.0 Hz, 2H), 2.18 – 2.06 (m, 2H), 1.83 – 1.53 (m, 8H), 1.50 – 1.25 (m, 6H). Calculated for $C_{54}H_{63}ClN_8O_8$ ($M+H^+$): 987.5; found 987.6.



[5A] Synthesized as previously described.[38]

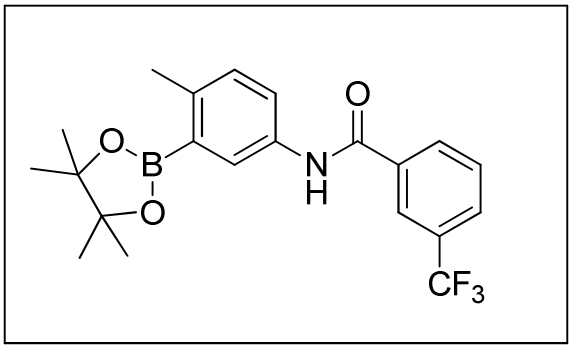


[5B] Synthesized as previously described.[39]

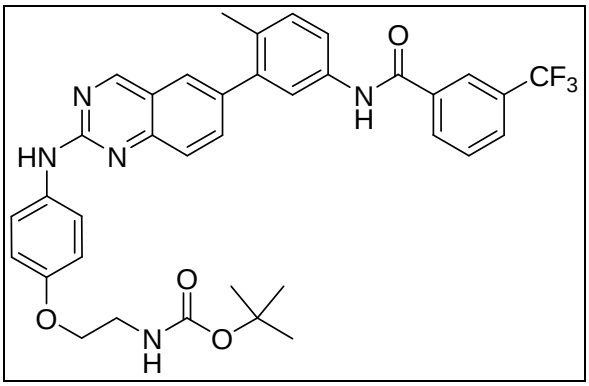


[5C] In a resealable Pyrex tube, **5A** (588.5 mg, 1.76 mmol) and **5B** (532 mg, 2.11 mmol) were dissolved in isopropanol (13.5 mL). TFA (262 µL, 3.52 mmol) was added and the tube was sealed. The suspension was stirred overnight at 70 °C. Triethylamine (1.5 mL) was added to

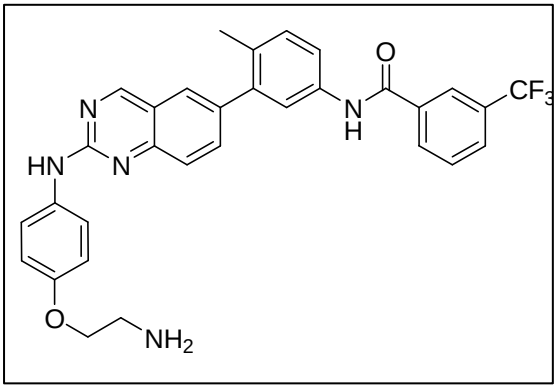
neutralize the mixture and the mixture was concentrated. The residue was purified by column chromatography (50% ethyl acetate in hexanes) to afford 378.5 mg of compound **5C** (47% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.17 (s, 1H), 8.01 – 7.94 (m, 2H), 7.75 (d, *J* = 9.0 Hz, 1H), 7.57 (d, *J* = 9.0 Hz, 2H), 6.98 – 6.92 (m, 2H), 5.00 (br s, 1H), 4.06 (t, *J* = 5.1 Hz, 2H), 3.59 – 3.53 (m, 2H), 1.48 (s, 9H). Calculated for $C_{21}H_{23}BrN_4O_3$ ($M+H^+$): 459.1; found 459.2.



[5D] 5-amino-2-methylphenylboronic acid pinacol ester (0.21 g, 0.86 mmol), 3-(trifluoromethyl)benzoic acid (0.21 g, 1.11 mmol), HOBt (0.17 g, 1.11 mmol), EDCI (0.21 g, 1.11 mmol) and DIEA (450 µL, 2.58 mmol) were dissolved in DMF (2.5 mL) and stirred overnight at RT. The crude mixture was diluted in ethyl acetate and washed with NH_4Cl and Na_2CO_3. The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* to afford 0.33 g of compound **5D** (96% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.14 (s, 1H), 8.07 (d, *J* = 6.0 Hz, 1H), 7.98 (d, *J* = 9.0 Hz, 1H), 7.84 – 7.81 (m, 2H), 7.68 – 7.62 (m, 2H), 7.23 (d, *J* = 9.0 Hz, 1H), 2.55 (s, 3H), 1.37 (s, 12H). Calculated for ($C_{21}H_{23}BF_3NO_3$) ($M+H^+$): 406.2; found 406.4.



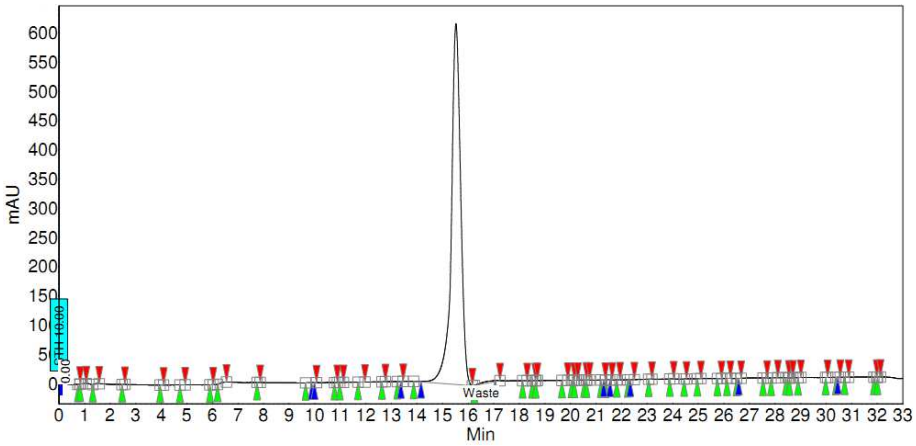
[5E] A mixture of compound **5C** (120 mg, 0.26 mmol), compound **5D** (127 mg, 0.31 mmol), tetrakis(triphenylphosphine)palladium (9.3 mg, 7.8 µmol) and sodium carbonate (60.9 mg, 0.57 mmol) was dissolved in a 3:1 mixture of DME/water (1 mL). The mixture was heated overnight at 85 °C. The crude mixture was cooled to RT, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo* and the resultant crude product was purified by column chromatography (5% methanol in dichloromethane) to afford 132 mg of compound **5E** (77 % yield). ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.03 (s, 1H), 8.23 – 8.14 (m, 2H), 7.85 – 7.80 (m, 1H), 7.71 – 7.62 (m, 7H), 7.26 – 7.20 (m, 1H), 6.89 – 6.86 (m, 2H), 6.71 – 6.67 (m, 1H), 3.98 – 3.85 (m, 2H), 3.42 – 3.37 (m, 2H), 2.23 – 2.15 (m, 3H), 1.42 (d, *J* = 3.0 Hz, 9H). Calculated for ($C_{36}H_{34}F_3N_5O_4$) ($M+H^+$): 658.3; found 658.4.



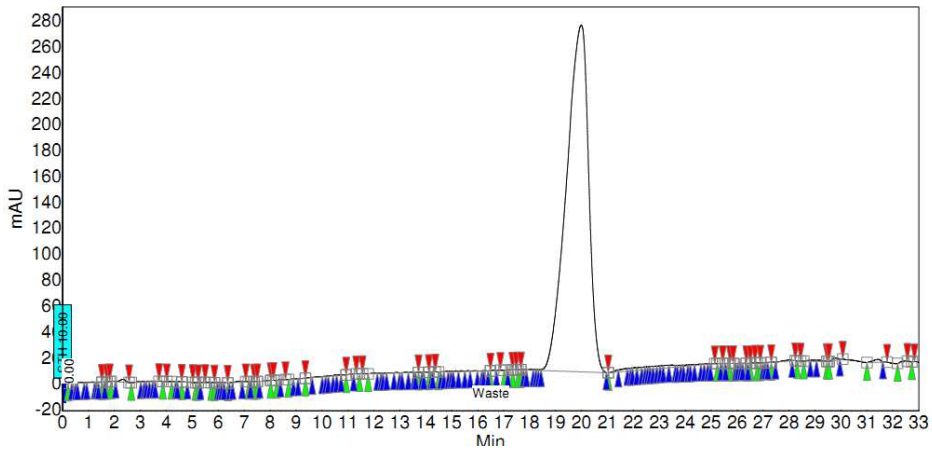
[5] Compound **5E** (40 mg, 0.062 mmol) was dissolved in CH_2Cl_2 (1 mL) and TFA (440 µL) and stirred for 3 h at RT. The reaction was concentrated and purified by reverse phase

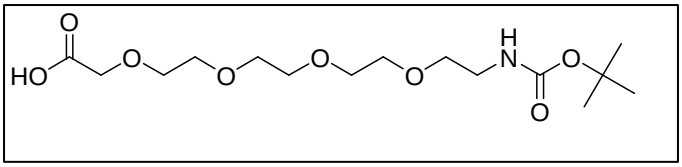
chromatography (HPLC) to obtain 20 mg (58 % yield) of the desired product **5**. ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.27 (s, 1H), 8.28 – 8.21 (m, 2H), 7.93 – 7.85 (m, 2H), 7.80 – 7.73 (m, 4H), 7.67 – 7.48 (m, 2H), 7.43 – 7.28 (m, 2H), 7.21 (d, *J* = 8.9 Hz, 1H), 7.08 (d, *J* = 9.1 Hz, 1H), 5.50 (s, 1H), 4.33 – 4.26 (m, 2H), 3.44 – 3.37 (m, 2H), 2.32 – 2.28 (m, 3H). Calculated for ($C_{31}H_{26}F_3N_5O_2$) ($M+H^+$): 558.2; found 558.4.

Analytical HPLC trace of **5**: Acetonitrile/Water

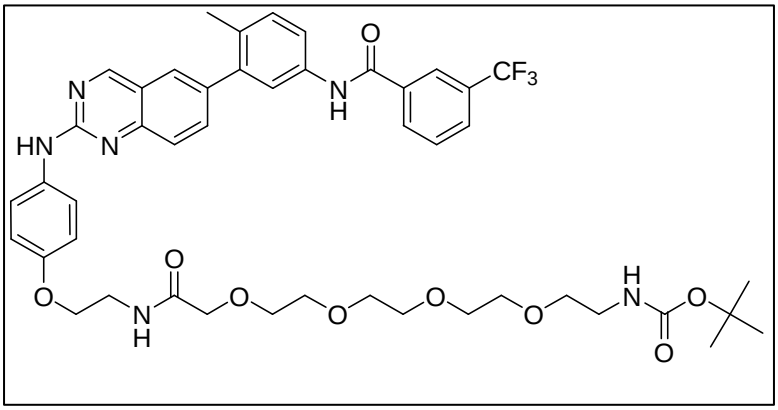


Analytical HPLC trace of **5**: Methanol/Water

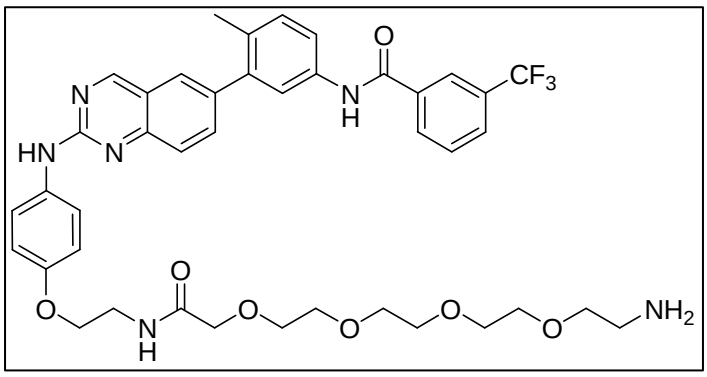




[5F] Synthesized as previously reported.[39]

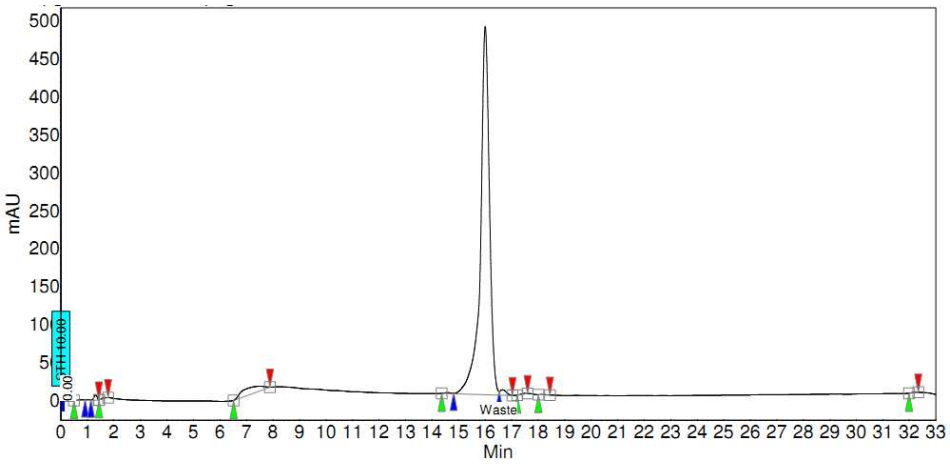


[5G] Compound **5** (62 mg, 0.112 mmol), compound **5F** (51 mg, 0.146 mmol), HOBt (23 mg, 0.146 mmol), EDCI (28.5 mg, 0.146 mmol) and DIEA (60 µL, 0.336 mmol) were dissolved in DMF (330 µL) and stirred overnight at RT. The crude mixture was diluted in ethyl acetate and washed with NH_4Cl and Na_2CO_3. The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* to obtain compound **5G**, which was used in the next step without further purification. Calculated for ($C_{46}H_{53}F_3N_6O_9$) ($M+H^+$): 891.4; found ($M+Na^+$): 913.4.

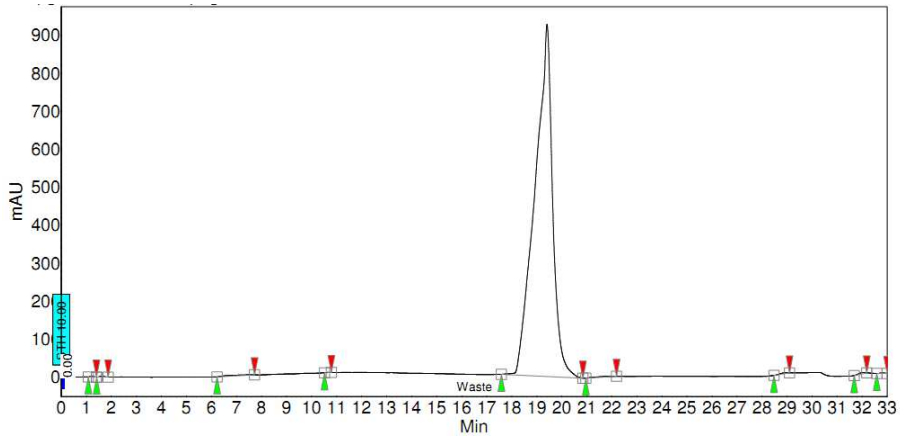


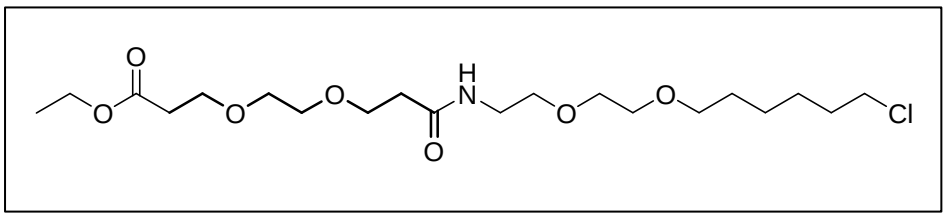
[5H] Compound **5G** (0.112 mmol) was dissolved in CH_2Cl_2 (1.9 mL) and TFA (0.8 mL) and stirred for 5 h at RT. The reaction was concentrated and purified by reverse phase chromatography (HPLC) to obtain 17.4 mg (20% yield) of the desired product **5H**. ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.26 (s, 1H), 8.28 – 8.21 (m, 2H), 7.92 – 7.84 (m, 2H), 7.77 – 7.43 (m, 3H), 7.40 – 7.26 (m, 3H), 7.25 – 7.09 (m, 2H), 7.14 (d, *J* = 8.4 Hz, 1H), 7.02 (d, *J* = 8.8 Hz, 1H), 5.80 (s, 1H), 4.20 – 4.01 (m, 4H), 3.77 – 3.60 (m, 16H), 3.17 – 3.10 (m, 2H), 2.32 – 2.28 (m, 3H). Calculated for ($C_{41}H_{45}F_3N_6O_7$) ($M+H^+$): 791.3; found 791.6.

Analytical HPLC trace of **5H**: Acetonitrile/Water

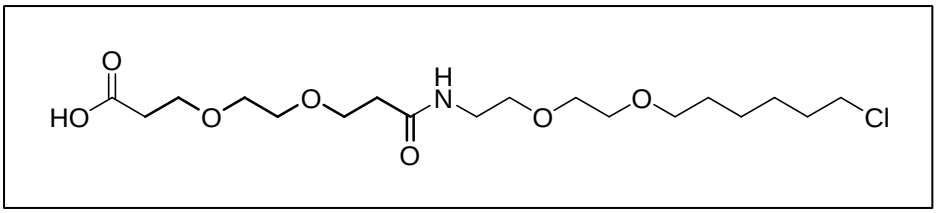


Analytical HPLC trace of **5H**: Methanol/Water

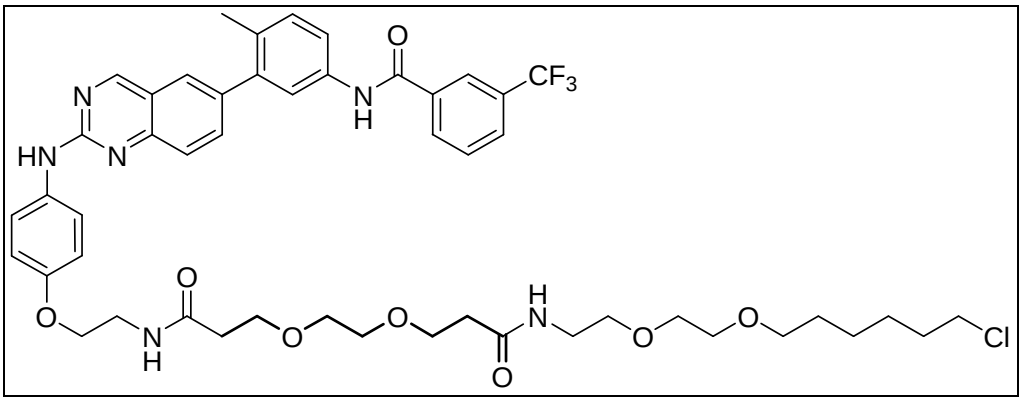




[6A] 2-(2-(6-chlorohexyloxy)ethoxy)ethanamine (74.6 mg, 0.33 mmol), 3-(2-(3-ethoxy-3-oxopropoxy)ethoxy)propanoic acid (101.6 mg, 0.43 mmol), HOBt (68.7 mg, 0.43 mmol), EDCI (84.9 mg, 0.43 mmol) and DIEA (180 µL, 1.0 mmol) were dissolved in DMF (980 µL) and stirred overnight at RT. The crude mixture was diluted in ethyl acetate and washed with NH_4Cl and Na_2CO_3. The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* to afford 106 mg of compound **6A** (73% yield). ^{1}H NMR (300 MHz, Methanol-d_4) δ 4.04 (q, *J* = 6.0 Hz, 2H), 3.70 – 3.58 (m, 4H), 3.56 – 3.40 (m, 12H), 3.38 – 3.29 (m, 4H), 2.51 – 2.46 (m, 2H), 2.38 – 2.34 (m, 2H), 1.70 – 1.63 (m, 2H), 1.53 – 1.48 (m, 2H), 1.39 – 1.24 (m, 4H) 1.17 (t, *J* = 6.0 Hz, 3H). Calculated for ($C_{20}H_{38}ClNO_7$) ($M+H^+$): 440.2; found 440.4.

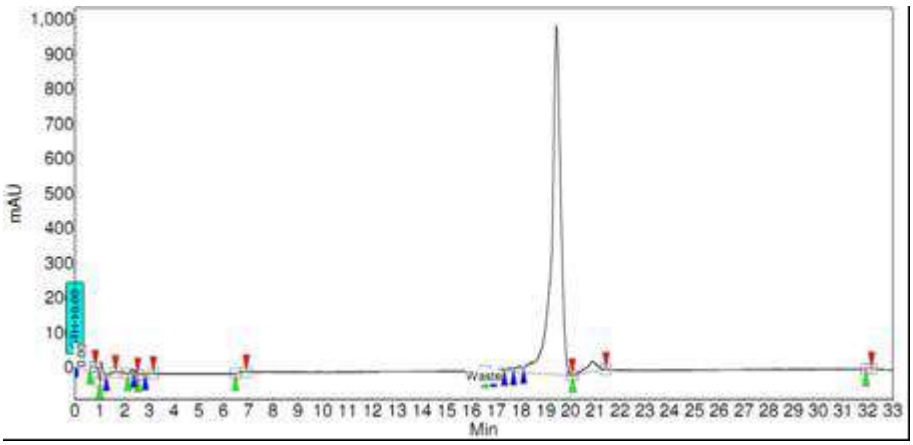


[6B] Compound **6A** (106 mg, 0.24 mmol) and $LiOH{\cdot}H_2O$ (30.5 mg, 0.73 mmol) was dissolved in a 1:1 mixture of THF/water (2.4 mL) and stirred for 2.5 h at RT. The reaction was concentrated and the pH was adjusted to 2 with 5 M HCl. The aqueous layer was extracted with ethyl acetate (4X), dried over Na_2SO_4 and concentrated *in vacuo* to afford 95.6 mg of compound **6B** (96% yield). Calculated for ($C_{18}H_{34}ClNO_7$) ($M+H^+$): 412.2; found 412.8.

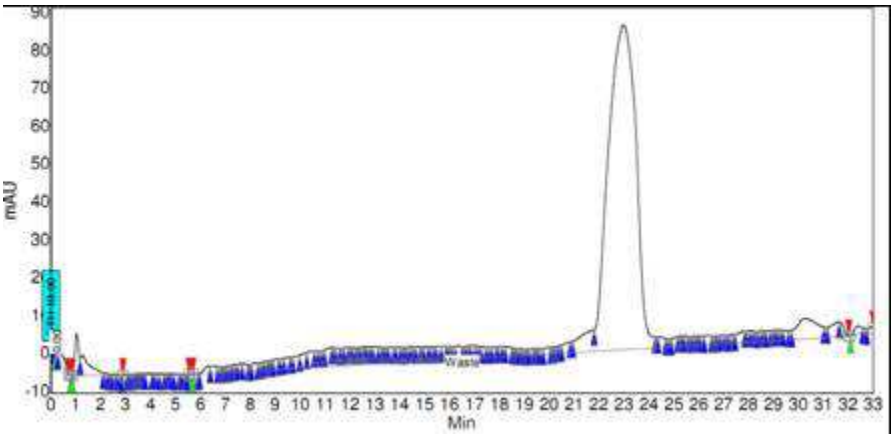


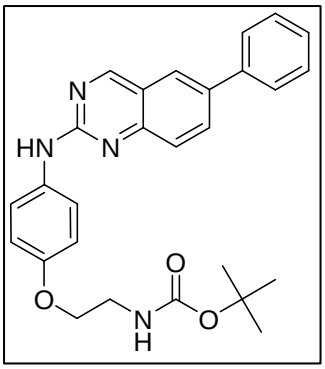
[6] Compound **5** (18.4 mg, 0.033 mmol), compound **6B** (17.6 mg, 0.043 mmol), HOBt (6.8 mg, 0.043 mmol), EDCI (8.4 mg, 0.043 mmol) and DIEA (17 µL, 0.099 mmol) were dissolved in DMF (100 µL) and stirred overnight at RT. The reaction was concentrated and purified by reverse phase chromatography (HPLC) to obtain 14.1 mg (45% yield) of the desired product **6**. ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.37 (s, 1H), 8.24 (s, 1H), 8.19 (d, *J* = 6.0 Hz, 1H), 7.98 – 7.87 (m, 2H), 7.78 – 7.68 (m, 2H), 7.63 – 7.45 (m, 4H), 7.35 – 7.27 (m, 3H), 7.13 – 7.03 (m, 2H), 4.12 – 4.08 (m, 2H), 3.74 – 3.66 (m, 3H), 3.62 – 3.48 (m, 14H), 3.46 – 3.42 (m, 2H), 3.36 – 3.30 (m, 2H), 2.49 – 2.39 (m, 4H), 2.29 – 2.25 (m, 3H), 1.85 – 1.67 (m, 2H), 1.64 – 1.51 (m, 2H), 1.49 – 1.28 (m, 4H). Calculated for ($C_{49}H_{58}ClF_3N_6O_8$) ($M+H^+$): 951.4; found 951.7.

Analytical HPLC trace of **6**: Acetonitrile/Water

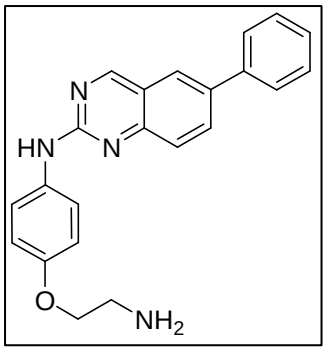


Analytical HPLC trace of **6**: Methanol/Water

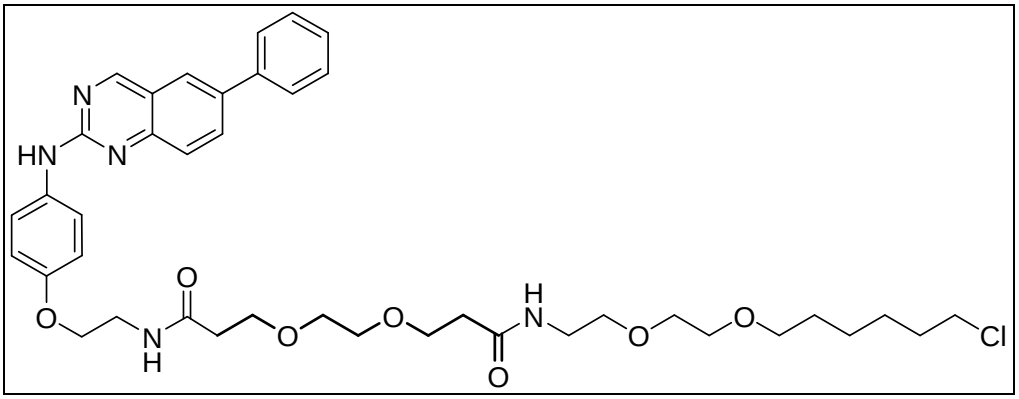




[5control A] A mixture of compound **5C** (45.1 mg, 0.098 mmol), phenylboronic acid pinacol ester (14.7 mg, 0..118 mmol), tetrakis(triphenylphosphine)palladium (3.5 mg, 3.0 µmol) and sodium carbonate (23 mg, 0.217 mmol) was dissolved in a 3:1 mixture of DME/water (390 µL). The mixture was heated overnight at 85 °C. The crude mixture was cooled to RT, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo* and the resultant crude product was purified by column chromatography (50% ethyl acetate in hexanes) to afford 44 mg of compound **5control A** (98 % yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.16 (s, 1H), 8.06 (dd, *J* = 6.0 Hz, *J* = 3.0 Hz, 1H), 7.94 (d, *J* = 3.0 Hz, 1H), 7.81 (d, *J* = 9.0 Hz, 1H), 7.74 – 7.69 (m, 4H), 7.52 (t, *J* = 9.0 Hz, 2H), 7.44 – 7.39 (m, 1H), 6.96 (d, *J* = 9.0 Hz, 2H), 4.07 (t, *J* = 6.0 Hz, 2H), 3.60 – 3.54 (m, 2H), 1.49 (s, 9H). Calculated for ($C_{27}H_{28}N_4O_3$) ($M+H^+$): 457.2; found 457.4.



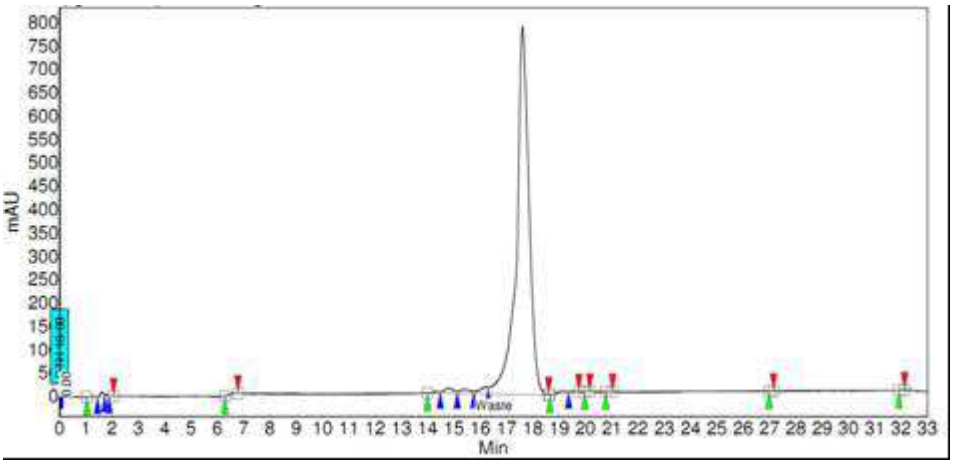
[5control B] Compound **5control A (**0.098 mmol) was dissolved in CH_2Cl_2 (1.6 mL) and TFA (0.7 mL) and stirred for 3 h at RT. The reaction was concentrated to obtain 34.1 mg (99% yield) of the desired product **5control B**. Crude product was azeotroped with toluene and used in the next step without further purification. Calculated for ($C_{22}H_{20}N_4O$) ($M+H^+$): 357.2; found 357.5.



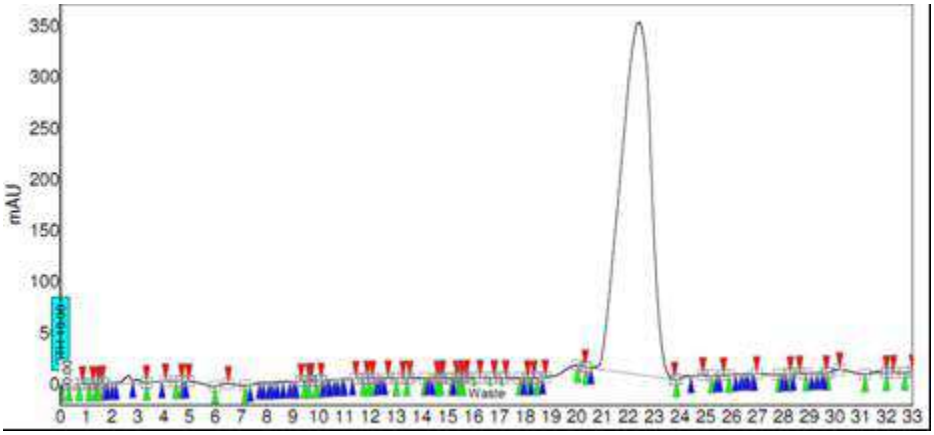
[6control] Compound **5control B** (26.3 mg, 0.074 mmol), **6B** (39.6 mg, 0.096 mmol), HOBt (15.2 mg, 0.096 mmol), EDCI (18.8 mg, 0.096 mmol) and DIEA (39 µL, 0.222 mmol) were dissolved in DMF (220 µL) and stirred overnight at RT. The reaction was concentrated and purified by reverse phase chromatography (HPLC) to obtain 26.0 mg (47% yield) of the desired product **6control**. ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.22 (s, 1H), 8.31 – 8.23 (m, 1H), 8.12 – 7.97 (m, 2H), 7.81 – 7.64 (m, 4H), 7.54 – 7.45 (m, 2H), 7.43 – 7.36 (m, 1H), 7.02 – 6.91 (m, 2H), 4.12 – 4.02 (m, 2H), 3.79 – 3.67 (m, 4H), 3.66 – 3.50 (m, 14H), 3.49 – 3.42 (m, 2H), 3.39 –

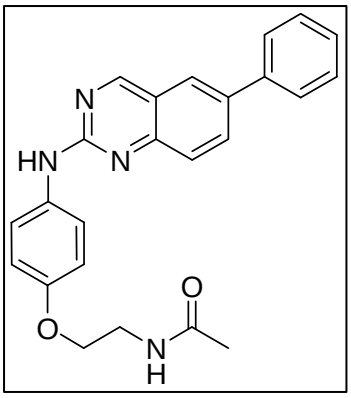
3.37 (m, 2H), 2.51 (t, *J* = 6.0 Hz, 2H), 2.44 (t, *J* = 6.0 Hz, 2H), 1.79 – 1.67 (m, 2H), 1.63 – 1.51 (m, 2H), 1.47 – 1.30 (m, 4H). Calculated for ($C_{40}H_{52}ClN_5O_7$) ($M+H^+$): 750.4; found 750.5.

Analytical HPLC trace of **6control**: Acetonitrile/Water

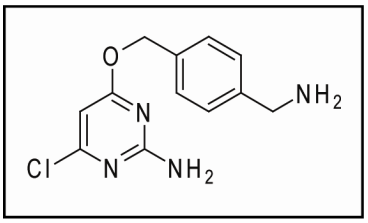


Analytical HPLC trace of **6control**: Methanol/Water





[**5control] 5control B** (6.3 mg, 0.0177 mmol) was dissolved in acetic anhydride (0.35 mL). Triethylamine (10 µL, 0.0708 mmol) was added and the reaction was stirred at RT for 2.5 h. The reaction was concentrated and purified by reverse phase chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d_4) δ 9.15 (s, 1H), 8.04 (dd, *J* = 6.0 Hz, *J* = 3.0 Hz, 1H), 7.94 (d, *J* = 3.0 Hz, 1H), 7.86 – 7.63 (m, 5H), 7.52 (t, *J* = 9.0 Hz, 2H), 7.44 – 7.39 (m, 1H), 6.97 (d, *J* = 9.0 Hz, 2H), 4.08 (t, *J* = 6.0 Hz, 2H), 3.75 – 3.65 (m, 2H), 2.04 (s, 3H). Calculated for ($C_{24}H_{22}N_4O_2$) ($M+H^+$): 399.2; found 399.6.



[7] Synthesized as previously described.[45]

iii. Generation of chloropyrimidine (CLP) resin

0.5 mL of NHS-activated Sepharose™ 4 Fast Flow (GE Healthcare) was washed with 10 bed volumes of 1:1 DMF/EtOH. 50 µL of 0.75 mM **7** was then added to the resin followed by 5 µL of DIEA. The reaction mixture was rotated overnight at RT, then drained and the resin washed with 2 bed volumes of DMF/EtOH. Unreacted NHS-activated esters were capped by incubating the resin with 0.1 M Tris-HCl, pH 8.5 overnight. The resin was washed with 10 bed volumes of

Wash Buffer A (0.1 M Tris-HCl, pH 8.5) then Wash Buffer B (0.1 M sodium acetate, pH 5, 0.5 M NaCl). This wash cycle was repeated 3 times followed by 10 bed volumes of 20% EtOH (1x). 0.5 mL of 20% EtOH was added to the resin resulting in a 50% slurry for storage at 4 °C.

B. PROTEIN EXPRESSION AND PURIFICATION

i. ASH-fusion protein design, expression, and purification

a. ASH-fusion protein design

The AGT gene from the SNAP source plasmid pss26b (Covalys) was amplified using primers that included a 3' end complementary to the 5' end of SUMO. The SUMO gene was amplified from a plasmid containing the SMT3/SUMO gene (Plasmid 16092: pT-35 Addgene) with primers that included a 5' end complementary to the 3' end of AGT and a 3' end complementary to the 5' end of HaloTag. The HaloTag gene from the pFC8K vector (Promega) was amplified using primers that included a 5' end complementary to the 3' end of SUMO. The two double-stranded DNA fragments from AGT and SUMO containing complementary overlapping regions were fused over 35 cycles using the LIC AGT Fwd and SUMO HaloTag Rev primers to give the AGT-SUMO construct. Likewise, the double-stranded AGT-SUMO fusion product and the HaloTag product were fused over 35 cycles using the LIC Fwd and LIC Rev primers to give the final construct AGT-SUMO-HaloTag (ASH). The complete construct was then incorporated into the vector pMCSG7 (Midwest Center for Structural Genomics) by ligation independent cloning.

b. Expression of ASH and ASH* fusion proteins

Single colonies of BL21(DE3) cells transformed with the ASH plasmid were grown overnight in 6 mL LB supplemented with ampicillin (100 μg/mL) at 37 °C. Overnight cultures were then

centrifuged at 3000 x g for 6 min. at 4 °C. The media was decanted and the cell pellet was resuspended in 2 mL of LB and transferred to 1 L LB (100 μg/mL ampicillin). The bacterial culture was grown at 37 °C to an OD_{600nm} of 0.8-1.0. The temperature was then decreased to 18 °C before expression of ASH was induced with 1 mM IPTG. The culture was grown for an additional 16 h and then was centrifuged at 5000 x g for 20 min. Cell pellets were stored at -80 °C until purification.

c. Purification of ASH and ASH* fusion proteins

Cell pellets from 250 mL cultures were re-suspended in His6 Wash Buffer (50 mM HEPES, 10 mM imidazole, pH 7.5) and PMSF (100 μg/mL) was added. Cells were lysed by sonication and centrifuged at 10,000 rpm for 20 min. at 4 °C to clear the lysate. The cleared lysate was then added to Promega HisLink™ Protein Purification resin and rotated at 4 °C for 30 min. The resin was washed with His6 Wash Buffer (3X) and then the protein was eluted with His6 Elution Buffer (100 mM HEPES, 600 mM imidazole, pH 7.5). The collected fractions were then dialyzed into storage buffer (50 mM HEPES, 1 mM DTT, pH 7.5) and concentrated.

ii. Ulp1 expression and purification

a. Ulp1 and Ulp1* expression

The gene encoding for Ulp1 (GenScript) was cloned into the vector pMCSG7 (Midwest Center for Structural Genomics). Single colonies of BL21(DE3) cells transformed with the Ulp1 plasmid were grown overnight in 6 mL of LB supplemented with ampicillin (100 μg/mL) at 37^{o} C. Overnight cultures were then centrifuged at 3000 x g for 6 min. at 4 °C. The media was decanted and the cell pellet was resuspended in 2 mL of LB and transferred to 1 L LB (100

μg/mL ampicillin). The bacterial culture was grown at 37 °C to an OD_{600nm} of 0.6 – 0.8. The temperature was then decreased to 25 °C before expression of Ulp1 was induced with 1 mM IPTG. The culture was grown for an additional 4 h and then was centrifuged at 5000 x g for 20 min. Cell pellets were stored at -80 °C until purification.

b. Ulp1 and Ulp1* purification

Cell pellets from 250 mL cultures were re-suspended in His6 Wash Buffer (50 mM HEPES, 10 mM imidazole, pH 7.5) and PMSF (100 μg/mL) was added. Cells were lysed by sonication and centrifuged at 10,000 rpm for 20 min at 4 °C to clear the lysate. The cleared lysate was then added to Promega HisLink™ Protein Purification resin and rotated at 4 °C for 30 min. The resin was washed with His6 Wash Buffer (3X) and then the protein was eluted with His6 Elution Buffer (100 mM HEPES, 600 mM imidazole, pH 7.5). The collected fractions were then dialyzed into storage buffer (75 mM Tris (pH 8.0), 1 mM DTT, 2 mM EDTA) and concentrated.

iii. ASH* primer design

SUMO mutations (R64T and R71E) were generated using site-directed mutagenesis (Stratagene).

Fwd R64T: 5’ – GAA ATG GAC TCC TTA ACC TTC TTG TAC GAC GGT – 3’

Rev R64T: 5’ – ACC GTC GTA CAA GAA GGT TAA GGA GTC CAT TTC – 3’

Fwd R71E: 5’ – TTG TAC GAC GGT ATT GAA ATT CAA GCT GAT CAG – 3’

Rev R71E: 5’ – CTG ATC AGC TTG AAT TTC AAT ACC GTC GTA CAA – 3’

iv. Ulp1* primer design

Ulp1 mutations (D451S, T452G, and E455S) were generated using two sequential site-directed mutagenesis reactions (Stratagene).

Ulp1 QC Fwd1: 5’ – CCG CGT CGC TGG CTG AAT AGT GGC ATC ATC GAA TTT TTC ATG – 3’

Ulp1 QC Rev1: 5’ – CAT GAA AAA TTC GAT GAT GCC ACT ATT CAG CCA GCG ACG CGG – 3’

Ulp1 QC Fwd2: 5’ – CTG AAT AGT GGC ATC ATC AGT TTT TTC ATG AAA TAC ATC – 3’

Ulp1 QC Rev2: 5’ – GAT GTA TTT CAT GAA AAA ACT GAT GAT GCC ACT ATT CAG – 3’

v. SRC expression and purification

The catalytic domain of chicken c-SRC (SRC KD, residues 251-533) and c-SRC that contains the SH1, SH2 and SH3 domains (SRC) were expressed and purified using a previously published procedure.[48] This method generates unphosphorylated SRC.

C. MAMMALIAN CELL CULTURE

i. Cell line maintenance

HeLa cells were cultured in DMEM-Low Glucose media supplemented with 10% fetal bovine serum. COS-7 cells were cultured in DMEM-High Glucose media supplemented with 10% fetal bovine serum. Cells were incubated at 37 °C in a humidified atmosphere containing 5% CO_2.

ii. Transfection protocol

Cells were transfected with mammalian-optimized genes in pDEST26 (Invitrogen) mammalian expression vector using FuGene HD transfection reagent (Promega). 2 x 10^5 cells were plated in each well of a 12 well plate. Cells were grown overnight to 50-70% confluency then washed with serum free media and incubated with 1 mL of Opti-MEM Media (Gibco). Each well of cells received 2 µg of DNA and 6 µL of FuGene reagent in 100 µL of Opti-MEM Media. Cells were incubated for 24 h to allow for protein expression.

iii. Preparation of mammalian cell lysate

COS-7 cells or HeLa cells were grown to confluency in 15 cm plates and pelleted. Five pellets were then combined and resuspended in 1.5 mL of cell lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM PMSF). The cells were lysed using a Dounce homogenizer and cleared via centrifugation for 20 min at 16,000 rpm. The lysate concentration was determined using a Bradford assay.

D. *IN VITRO* PULLDOWN METHODS

i. Fluorescence assay for determination of catch-and-release efficiency

Purified ASH or ASH* (30 μM) was labeled with **1** (45 μM) in 50 mM HEPES buffer, pH=7.5, 100 mM NaCl, and 1 mM DTT for 1 h at RT. Fluorescently-labeled protein was then rotated with CLP resin for 1.5 h at RT. After incubation, the resin was washed and then incubated with Ulp1 or Ulp1* protease. Eluted samples were separated by SDS-PAGE and scanned with a GE Typhoon FLA 9000 fluorescent scanner. The intensities of fluorescently-labeled protein bands were quantified with ImageQuant software.

ii. Generation of a singly-labeled hexylchloride protein

Purified SNAP-tag (15 μM) was incubated with **2** (22.5 μM) in labeling buffer (50 mM Tris buffer, pH = 7.5, 100 mM NaCl, 0.1% Tween 20 and 1 mM DTT) for 1.5 h at RT. The singly-labeled SNAP-tag was then separated from the unconjugated labeling reagent **2** by running the reaction mixture through a Bio-Rad Micro Bio-Spin© column equilibrated with labeling buffer.

iii. Catch-and-release of singly-labeled hexylchloride proteins from cell lysate

ASH protein (2 μM) was added to a 200 μL mixture of singly-labeled protein (SNAP-tag) (500 nM), 50 μg mammalian cell lysate, 1X protease inhibitor cocktail (Roche complete), and 1 mM DTT in 50 mM HEPES, pH=7.5. The mixture was incubated at RT for 1 h and then incubated for 1.5 h with 60 μL (50% slurry) CLP resin. The beads were then washed twice with Cleavage Buffer (50 mM Tris, pH=8.0, 150 mM NaCl, 0.2% NP40, and 1 mM DTT) and then incubated with Ulp1 (1:80 mass ratio) in Cleavage Buffer for 2 h at 30 °C. Samples were then subjected to

SDS-PAGE and immunoblot analysis (anti-His, abcam) to determine capture and release efficiency. Blots were quantified with Li-cor Odyssey software.

iv. Catch-and-release of proteins labeled in cells by a single hexylchloride tag

HeLa cells were grown in a 12-well plate, transfected with SNAP-tag (New England BioLabs), and cultured for 24 h. Cells were then treated with **2** (10 μM) in 1 mL of serum-free media for 1 h at 37 °C. Control cells were incubated with TMRstar (10 μM) (New England Biolabs). Cells were then washed with media (3X, 10 min) and PBS (2X). Cells were then transferred to a 1.5 mL microcentrifuge tube and sonicated to lyse the cells. The lysate was then incubated with resin displaying preimmobilized ASH*. After a series of washes (50 mM Tris, pH=8.0, 300 mM NaCl, 0.1% Tween), immobilized proteins were eluted from the resin with Ulp1* (1:20 mass ratio Ulp1*:ASH*). Samples were separated on 10% SDS-PAGE gels and either silver stained (Invitrogen SilverXpress Staining Kit) or immunoblotted (anti-His, abcam). Blots were quantified with Li-cor Odyssey software.

v. Crosslinking procedures

a. Preconjugation

All crosslinking experiments were performed in a 96-well, U-bottom plate. Purified SRC was diluted in PBS to 25 μL at a concentration of 100 nM with a concentration of 2 mg/mL mammalian cell lysate. ASH*, covalently labeled with **4**, was added at a final concentration of 500 nM, such that the final volume per well was 50 μL ($[kinase]_f$ = 50 nM). Dasatinib was added to control wells only at a final concentration of 10 μM. Samples were irradiated at 365 nm on ice for 10 min by placing a Spectroline ENF-260C UV lamp directly on top of the plate**.** Samples

were run on 10% SDS-PAGE gels and immunoblotted (Src (36D10) antibody, Cell Signaling). The scanned blots were quantified with Li-cor Odyssey software to determine crosslinking efficiency.

b. Postconjugation

Purified SRC was diluted in PBS to 50 μL at a concentration of 50 nM with a concentration of 1 mg/mL mammalian cell lysate. Compound **4**, was added at a final concentration of 500 nM. Dasatinib was added to control wells only at a final concentration of 10 μM. Samples were irradiated at 365 nm on ice for 10 min by placing a Spectroline ENF-260C UV lamp directly on top of the plate**.** After irradiation, 1 μM ASH* was added to the crosslinked sample and incubated at RT for 1 h. Samples were run on 10% SDS-PAGE gels and immunoblotted (SRC (36D10) antibody, Cell Signaling). The scanned blots were quantified with Li-cor Odyssey software to determine crosslinking efficiency.

vi. Catch-and-release of hexylchloride-labeled protein kinases

After crosslinking using the post-conjugation technique, 4x50 μL wells were combined and DTT was added to a final concentration of 2 mM. The sample was then incubated with CLP resin for 90 min. After a series of washes (50 mM Tris, pH=8.0, 300 mM NaCl, 0.1% Tween), immobilized proteins were eluted from the resin with Ulp1* (1:20 mass ratio Ulp1*:ASH*). Samples were separated on 10% SDS-PAGE gels and immunoblotted (SRC (36D10) antibody, Cell Signaling). Blots were quantified with Li-cor Odyssey software.

E. AFFINITY RESIN METHODS

i. Generation of affinity resin

Affinity matrix **5H** was generated from ECH Sepharose resin using a previously published procedure [39] (Figure S7).

ii. General method for affinity resin catch-and-release experiments

HeLa lysate (500 μg of protein) was added to 25 μL of affinity resin and incubated for 2 h at 4 °C. The resin was then washed 4X by incubating with 10 bed volumes of Wash Buffer (50 mM HEPES, pH 7.5, 0.5% Triton X-100, 1 mM EDTA, 1 mM EGTA, and 100 mM NaCl) for 30 min at 4 °C. After the final wash, one of three elution conditions was used: (1.) Incubation with 200 μL of free inhibitor, **5**, at a final concentration of 100 μM for 30 min at RT; (2.) 100 μL of saturated SDS in binding buffer was added to the resin and incubated for 10 min at RT. The eluted protein was collected and this process was repeated once; (3.) 200 μL 1X SDS Loading Buffer was added to the resin and boiled for 20 min. Samples were run on 10% SDS-PAGE gels and either silver stained (Invitrogen SilverXpress Staining Kit) or immunoblotted (Src (36D10) antibody, Cell Signaling). The scanned blots were quantified with Li-cor Odyssey software to determine elution efficiency.

iii. Enrichment of endogenous kinases from HeLa lysate

ASH* (0.2 mg) was immobilized on CLP resin and then incubated with 200 μL of **6** (final concentration = 8.5 μM) or **6control** (final concentration = 8.5 μM) in Buffer A (50 mM Tris, pH=7.5, 100 mM NaCl) for 1 h at RT. Excess small molecule was removed through a series of

washes with Buffer A. HeLa lysate was then incubated with the resin for 2 h at RT. After a series of washes with Buffer B (50 mM Tris, pH=8.0, 300 mM NaCl, 0.1% Tween), the beads were incubated with Ulp1* (1:20 mass ratio Ulp1*:ASH*). Eluted samples were separated on 10% SDS-PAGE gels. Samples were then subjected to the Mass Spectrometry protocol described below.

F. ACTIVITY ASSAYS

In vitro activity assays for ABL, CSK, HCK, IRAK4, LCK, STK10, MAP3K5, p38α, PAK4, PAK5, EPHA3 and SRC were performed using previously published protocols.[40]

i. PTK2

Inhibitors (initial concentration = 10 µM, 3-fold serial dilutions down to 0.2 nM) were assayed in triplicate against PTK2 (Invitrogen) (final concentration = 1 ng/µL) in assay buffer containing 50 mM HEPES, pH = 7.5, 60 mM $MgCl_2$, 1 mM EGTA, 2 mM Na_3VO_4, 100 µg/mL BSA, $\gamma^{32}P$ ATP (0.2 µCi/well) and poly(Glu, Tyr) (Sigma) as substrate (final concentration = 200 µg/mL). The final volume of each assay well was 30 µL. The enzymatic reaction was run at RT for 2 h and then terminated by spotting 4.6 µL of the reaction mixture onto a nitrocellulose membrane. Membranes were washed with 0.5% phosphoric acid (4 x 5 min each wash), dried and the radioactivity was determined by phosphorimaging with a GE Typhoon FLA9000 scanner. The scanned membranes were quantified with ImageQuant and converted to percent inhibition. Data was analyzed using Prism Graphpad software and IC_{50} values were determined using non-linear regression analysis.

ii. EIF2AK2

Inhibitors (initial concentration = 10 µM, 3-fold serial dilutions down to 0.2 nM) were assayed in triplicate against EIF2AK2 (Invitrogen) (final concentration = 165 pg/µL) in assay buffer containing 50 mM HEPES, pH = 7.5, 60 mM $MgCl_2$, 1 mM EGTA, 2 mM Na_3VO_4, 100 µg/mL BSA, γ32P ATP (0.2 µCi/well) and myelin basic protein as substrate (final concentration = 200 µg/mL). The final volume of each assay well was 30 µL. The enzymatic reaction was run at RT for 100 min and then terminated by spotting 4.6 µL of the reaction mixture onto a phosphocellulose membrane. Membranes were washed with 0.5% phosphoric acid (4 x 5 min each wash), dried and the radioactivity was determined by phosphorimaging with a GE Typhoon FLA9000 scanner. The scanned membranes were quantified with ImageQuant and converted to percent inhibition. Data was analyzed using Prism Graphpad software and IC_{50} values were determined using non-linear regression analysis.

G. MASS SPECTROMETRY

i. In-gel trypsin digest of enriched DFG-out proteins

Each lane of a 10% SDS-PAGE gel above 30 kDa was divided into three sections. Each section was excised from the gel and placed in 1.5 mL eppendorf tubes. Each gel slice was washed three times: first with 500 µL of 100 mM ammonium bicarbonate by rotating for 15 min at RT then with 500 µL acetonitrile (15 min, RT). Gel slices were then speedvaced to dryness and rehydrated on ice for 45 min with 50 µL of a trypsin solution (20 µg dissolved in 1 mL 50 mM ammonium bicarbonate) (Sigma, Proteomics Grade, BioReagent, Dimethylated). 50 mM

ammonium bicarbonate was then added to cover the expanded gel slice and then the gel slice was incubated overnight at 37 °C.

ii. Mass spectrometry analysis

Samples were submitted to Mass Spectrometry Center at the University of Washington, School of Pharmacy. Data were generated with an LTQ Velos (Thermo Scientific).

V. REFERENCES

1 Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* **25**, 1035-1044 (2007).

2 Daub, H. *et al.* Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol Cell* **31**, 438-448 (2008).

3 Dulla, K., Daub, H., Hornberger, R., Nigg, E. A. & Korner, R. Quantitative site-specific phosphorylation dynamics of human protein kinases during mitotic progression. *Mol Cell Proteomics* **9**, 1167-1181 (2010).

4 Oppermann, F. S. *et al.* Large-scale proteomics analysis of the human kinome. *Mol Cell Proteomics* **8**, 1751-1764 (2009).

5 Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408-411 (1996).

6 Ong, S. E. *et al.* Identifying the proteins to which small-molecule probes and drugs bind in cells. *Proc Natl Acad Sci U S A* **106**, 4617-4622 (2009).

7 Rix, U. & Superti-Furga, G. Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* **5**, 616-624 (2009).

8 Evans, M. J. & Cravatt, B. F. Mechanism-Based Profiling of Enzyme Families. *Chem Rev* **106**, 3279-3301 (2006).

9 Taylor, M. T., Blackman, M. L., Dmitrenko, O. & Fox, J. M. Design and synthesis of highly reactive dienophiles for the tetrazine-trans-cyclooctene ligation. *J Am Chem Soc* **133**, 9646-9649 (2011).

10 Blackman, M. L., Royzen, M. & Fox, J. M. Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J Am Chem Soc* **130**, 13518-13519 (2008).

11 Mahal, L. K., Yarema, K. J. & Bertozzi, C. R. Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis. *Science* **276**, 1125-1128 (1997).

12 Sampathkumar, S.-G., Li, A. V., Jones, M. B., Sun, Z. & Yarema, K. J. Metabolic installation of thiols into sialic acid modulates adhesion and stem cell biology. *Nat Chem Biol* **2**, 149-152 (2006).

13 Dondoni, A. The Emergence of Thiol–Ene Coupling as a Click Process for Materials and Bioorganic Chemistry. *Angew Chem Int Edit* **47**, 8995-8997 (2008).

14 Saxon, E. & Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **287**, 2007-2010, doi:10.1126/science.287.5460.2007 (2000).

15 Best, M. D. Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. *Biochemistry* **48**, 6571-6584 (2009).

16 Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383-414, doi:10.1146/annurev.biochem.75.101304.124125 (2008).

17 Salisbury, C. M. & Cravatt, B. F. Activity-based probes for proteomic profiling of histone deacetylase complexes. *Proc Natl Acad Sci U S A* **104**, 1171-1176, doi:10.1073/pnas.0608659104 (2007).

18 Saghatelian, A., Jessani, N., Joseph, A., Humphrey, M. & Cravatt, B. F. Activity-based probes for the proteomic profiling of metalloproteases. *Proc Natl Acad Sci U S A* **101**, 10000-10005 (2004).

19 Besanceney-Webler, C. *et al.* Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study. *Angew Chem Int Edit* **50**, 8051-8056, doi:10.1002/anie.201101817 (2011).
20 Speers, A. E., Adam, G. C. & Cravatt, B. F. Activity-Based Protein Profiling in Vivo Using a Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J Am Chem Soc* **125**, 4686-4687 (2003).
21 Praul, C. A., Brubaker, K. D., Leach, R. M. & Gay, C. V. Detection of Endogenous Biotin-Containing Proteins in Bone and Cartilage Cells with Streptavidin Systems. *Biochem Bioph Res Co* **247**, 312-314 (1998).
22 Yang, Y.-Y., Grammel, M., Raghavan, A. S., Charron, G. & Hang, H. C. Comparative Analysis of Cleavable Azobenzene-Based Affinity Tags for Bioorthogonal Chemical Proteomics. *Chem Biol* **17**, 1212-1222 (2010).
23 Los, G. V. *et al.* HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem Biol* **3**, 373-382 (2008).
24 Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C. & Johnsson, K. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng Des Sel* **19**, 309-316, doi:10.1093/protein/gzl014 (2006).
25 Keppler, A. *et al.* Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods* **32**, 437-444, doi:10.1016/j.ymeth.2003.10.007 (2004).
26 Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol* **21**, 86-89, doi:10.1038/nbt765 (2003).
27 Chidley, C., Mosiewicz, K. & Johnsson, K. A Designed Protein for the Specific and Covalent Heteroconjugation of Biomolecules. *Bioconjugate Chem* **19**, 1753-1756 (2008).
28 Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032 (1999).
29 Butt, T. R., Edavettal, S. C., Hall, J. P. & Mattern, M. R. SUMO fusion technology for difficult-to-express proteins. *Protein Expres Purif* **43**, 1-9 (2005).
30 Liu, L., Spurrier, J., Butt, T. R. & Strickler, J. E. Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expres Purif* **62**, 21-28 (2008).
31 Marcotte, D. J. *et al.* Structures of human Bruton's tyrosine kinase in active and inactive conformations suggest a mechanism of activation for TEC family kinases. *Protein Sci* **19**, 429-439 (2010).
32 Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602-609 (1997).
33 Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595-602 (1997).
34 Pargellis, C. *et al.* Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Biol* **9**, 268-272 (2002).
35 Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat Chem Biol* **2**, 358-364 (2006).
36 Wan, P. T. *et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855-867 (2004).
37 Schindler, T. *et al.* Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* **289**, 1938-1942 (2000).

38 DiMauro, E. F. *et al.* Discovery of Aminoquinazolines as Potent, Orally Bioavailable Inhibitors of Lck: Synthesis, SAR, and in Vivo Anti-Inflammatory Activity. *J Med Chem* **49**, 5671-5686 (2006).
39 Ranjitkar, P., Brock, A. M. & Maly, D. J. Affinity Reagents that Target a Specific Inactive Form of Protein Kinases. *Chem Biol* **17**, 195-206 (2010).
40 Ranjitkar, P. *et al.* Affinity-Based Probes Based on Type II Kinase Inhibitors. *J Am Chem Soc* **134**, 19017-19025 (2012).
41 Seeliger, M. A. *et al.* Equally Potent Inhibition of c-Src and Abl by Compounds that Recognize Inactive Kinase Conformations. *Cancer Res* **69**, 2384-2392, doi:10.1158/0008-5472.can-08-3953 (2009).
42 Simard, J. R. *et al.* A new screening assay for allosteric inhibitors of cSrc. *Nat Chem Biol* **5**, 394-396 (2009).
43 Dar, A. C., Lopez, M. S. & Shokat, K. M. Small molecule recognition of c-Src via the Imatinib-binding conformation. *Chem Biol* **15**, 1015-1022, doi:10.1016/j.chembiol.2008.09.007 (2008).
44 Choi, Y. *et al.* Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg Med Chem Lett* **19**, 4467-4470 (2009).
45 Hill, Z. B., Perera, B. G. K., Andrews, S. S. & Maly, D. J. Targeting Diverse Signaling Interaction Sites Allows the Rapid Generation of Bivalent Kinase Inhibitors. *ACS Chem Biol* **7**, 487-495 (2012).
46 Krishnamurty, R. *et al.* Active site profiling reveals coupling between domains in SRC-family kinases. *Nat Chem Biol* **9**, 43-50, doi:10.1038/nchembio.1118 (2013).
47 Johnson, S. M. *et al.* Development of Toxoplasma gondii Calcium-Dependent Protein Kinase 1 (TgCDPK1) Inhibitors with Potent Anti-Toxoplasma Activity. *J Med Chem* **55**, 2416-2426 (2012).
48 Seeliger, M. A. *et al.* High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci* **14**, 3135-3139 (2005).

Chapter 2: Promising new kinase targets for the treatment of Human African Trypanosomiasis

I. INTRODUCTION

Trypanosoma brucei is a unicellular protozoan parasite which causes Human African Trypanosomiasis (HAT), commonly known as African Sleeping Sickness. HAT occurs only in sub-Saharan Africa and is endemic to 36 countries.[1] The disease affects mostly poor rural populations and when left untreated is usually fatal.[2] In 2009, 9878 new cases were reported, although the World Health Organization estimates the actual number of new infections to be three-fold higher.[3]

The mode of transmission, through the bite of an infected tsetse fly, was discovered shortly after the initial descriptions of the parasite in the early twentieth century.[4] The complex life cycle of *T. brucei*, which undergoes at least 13 distinct developmental stages in tsetse flies and mammalian hosts, has been the topic of numerous and continuing studies as researchers focus on treatment and elimination of HAT.[4-8] _ENREF_5

The parasite life cycle in the tsetse fly begins when the insect takes a blood meal from an infected mammalian host and ingests bloodstream trypomastigotes.[5] Bloodstream trypomastigotes (stumpy forms of the parasite) migrate through the fly into the midgut, where three discrete developmental stages have been identified. The parasite multiplies through binary fission, then leaves the fly midgut as an epimastigote, migrating to the salivary gland. There, *T. brucei* continues to multiply through binary fission and undergoes at least another four stages of development. Upon maturation, the parasite detaches as a metastatic trypomastigote and is ready to be transferred to a new mammalian host.

Humans, domestic livestock, and wild game animals become infected with *T. brucei* when a tsetse fly injects metastatic trypomastigotes into their skin. In the mammalian host the parasite undergoes further development and multiplies through binary fission in the bloodstream, lymph system, and spinal fluids.[5]

Two different subspecies of trypanosomes cause HAT infection in humans: *T. b. gambiense* leads to slow development of the disease (often asymptomatic for months or years), and *T. b. rhodesiense* causes an acute form, which is fatal within months or even weeks when left untreated.[9] In humans, the disease progresses from a mild hemolymphatic stage, where symptoms include headache, enlarged lymph nodes, fever, itching, and in some cases a canker sore at the location of the tsetse fly bite, to the encephalitic stage. At this stage, the parasite crosses the blood-brain barrier and invades the central nervous system, causing severe headaches, disturbance of the sleep cycle, abnormal behavior, loss of consciousness, and coma.[5]

Current treatments for HAT are complicated and often difficult to administer. Subspecies of parasite, along with the stage of the infection, determine which of four main drugs should be administered. Pentamidine is used to treat the hemolymphatic stage of *T. b. gambiense* infection but is not absorbed by the gut and, therefore, must be administered parenterally.[10] While the exact mechanism of action of pentamidine has not been fully characterized, this drug is known to bind the minor groove of kinetoplast DNA (kDNA) in the mitochondrion of the parasite, inducing changes in DNA topology and inhibiting topoisomerases.[11,12] In addition, pentamidine inhibits S-adenosylmethionine decarboxylase in *T. brucei brucei*,[13] which in turn inhibits polyamine synthesis and results in a lack of the redox regulator trypanothione.

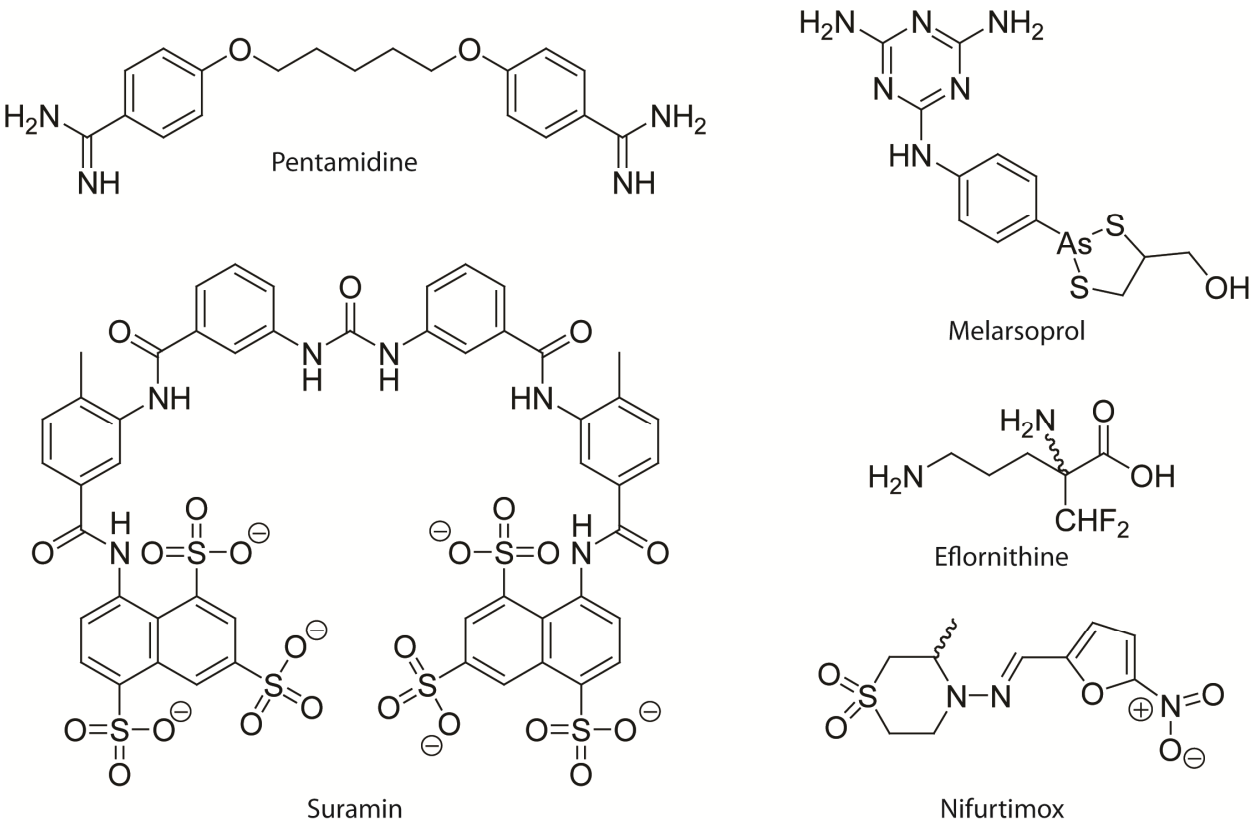


Figure 2- 1. Structures of current T. brucei chemotherapeutics. Structures of early-stage therapeutics (pentamidine and suramin) and advanced-stage therapeutics (melarsoprol and eflornithine and the Chagas disease therapeutic nifurtimox).

Intravenous injection of the drug suramin is used for treatment of early stage *T. b. rhodesiense*.[14] Suramin has been shown to inhibit both glycolytic enzymes and receptor-mediated LDL uptake most likely limiting the parasite's supply of cholesterol and phospholipids.[15] While both of these early disease stage drugs are toxic, they are less so than advanced stage drugs and are also easier to administer.[9]

Melarsoprol is used for the treatment of both subspecies of parasite in the advanced encephalitic stage, but it is so toxic that it is fatal for 3-10% of patients.[9] Treatment with melarsoprol involves ten daily injections, but unlike pentamidine and suramin, this drug effectively crosses the blood-brain barrier.[16] Melarsoprol is degraded in the parasite to melarsen

oxide, a highly toxic compound that causes rapid lysis of the parasites.[17,18] The drug also interacts with trypanothione, forming a stable complex which is an inhibitor of trypanothione reductase.[19,20] The decrease in available trypanothione and inhibition of trypanothione reductase makes the parasite especially vulnerable to oxidative stress.[21]

A newer advanced-stage drug, eflornithine, also crosses the blood-brain barrier and is less toxic than melarsoprol, but it is effective in treating only one sub-species (*T. b. gambiense).*[10] Eflornithine is an irreversible inhibitor of the enzyme ornithine decarboxylase (ODC), which is involved in polyamine biosynthesis.[22] Inhibiting ODC results in decreased levels of trypanothione.[23] Eflornithine, although less toxic than melarsoprol, is considerably more difficult to administer and also more expensive. Each eflornithine kit (developed by the WHO to treat late stage *T. b. gambiense*) contains only two treatments, weighs 40 kg, and costs $1420 U.S.[1] In addition to the cost and logistical problems of having treatment kits available in rural settings, eflornithine must be diluted and given for 14 days with 4 intravenous perfusions per day.[16] In 2009, a less cumbersome treatment option where eflornithine is used in combination with nifurtimox (an inhibitor of trypanothione reductase that has been shown to generate toxic free radicals) was approved. Nifurtimox is taken orally and is used in the treatment of Chagas disease.[10] Its efficacy as a mono-treatment in *T. b. gambiense* is limited, but clinical trials show that in combination with eflornithine it is just as effective as mono-treatment with eflornithine.

Although the chemotherapeutics described above can be effective treatments for HAT, drug resistant strains of *T. b. rhodesiense* and *T. b. gambiense* threaten successful control of the disease. Livestock and wild animals serve as infection reservoirs and contribute to the challenge of controlling HAT; civil wars, social upheavals, and difficult to reach rural areas further impair the eradication of the parasite.[1] Except for those currently used in chemotherapy, no other drugs

are available, meaning that drug resistance could become a considerable problem affecting the people of sub-Saharan Africa.

There are multiple mechanisms of drug resistance, including (1) altered metabolic pathways, which can change drug target dependence or prevent a drug from reaching its target, (2) alteration of the target's binding site such that it can no longer bind the drug, and (3) reduced drug accumulation through changes in drug import, export, or permeability.[24] Current reduced sensitivity and resistance in *T. b. rhodesiense* and *T. b. gambiense* are frequently mediated by reduced drug import via loss of purine transporter P2.[24,25] Both pentamidine and melarsoprol are transported by P2 into the parasite, and loss of P2 results in a 2-3 fold increase in resistance.[26] The overexpression of the efflux pump TbMRPA (*T. brucei* multidrug resistance-associated protein A) in the parasite also leads to a 10-fold increase in resistance to melarsoprol, most likely conferring resistance by reducing the accumulation of the drug in the parasite.[27]

Other resistance mechanisms unrelated to drug accumulation have been observed for eflornithine.[28] An *in vitro* eflornithine-resistant *T. b. brucei* strain demonstrates an increased uptake of ornithine, the substrate for ODC. Ornithine competes with eflornithine for the active site of ODC, leading to sufficient synthesis of polyamine and trypanothione.[29,30]

Current efforts for the development of new anti-trypanosomal drugs include natural product and small-molecule library screening against parasites *in vitro*, derivatization of current therapeutics, and target-specific structure-based drug design.[24,31-33] Despite current efforts in drug discovery, no new drugs have been approved in twenty years. Clearly there is a strong need for research into novel drugs that are safe, effective, and overcome drug resistance.

Our approach to developing new inhibitors was to target the kinome, or the kinase complement of the proteome. The genome of *T. brucei* encodes 182 kinases, most of whose

functions are poorly understood.[34] However, in other systems kinases have been shown to play central roles in critical signaling pathways and have proven to be druggable targets; in fact, at least ten small-molecule kinase inhibitors have been approved for use as effective therapies for disease.[35] Perhaps the best known example is Gleevec (imatinib), a highly selective type II inhibitor that is used to treat patients with chronic myelogenous leukemia and other cancers.[36]

Compared to how much is known about human protein kinases, astonishingly little is known about those in *T. brucei*. Very few have close human orthologs, and their individual roles in larger signaling networks are almost completely unknown. While these and other challenges have greatly hindered *T. brucei* kinase drug development, some progress has been made; notably, SCYX-5070 has recently been shown to target MAPK-like and CDK-like kinases to effectively treat *T. brucei* in mouse model systems.[32]

Here we describe a novel small molecule drug that potently kills the bloodstream form of *T. brucei.* Using advanced proteomic strategies we found that the targets of this drug are protein kinases, some of which are critical for parasite survival. These studies establish that protein kinases are valid targets for HAT drug development and demonstrate the utility of new proteomic tools and analyses for target identification.

II. RESULTS AND DISCUSSION

A. Compound 1 potently inhibits the growth of *T. brucei*

Recently, we identified a compound (**1**, Figure 2.2A) that effectively inhibits the growth of bloodstream form *T. brucei in vitro* with an EC_{50} of 50 nM**. 1** is an analogue of a series of type II compounds that have been shown to potently inhibit mammalian protein kinases.[37,38] Type II ligands stabilize a specific inactive conformation of kinases known as the DFG-out conformation

and contain, in addition to a moiety that occupies the adenosine-binding pocket, a functional group that sits in a hydrophobic pocket that is revealed by the translocation of the kinase's activation loop (Figure 2-2B). The hallmark outward rotation of the conserved Asp-Phe-Gly motif at the base of the activation loop is the reason behind the "DFG-out" moniker.

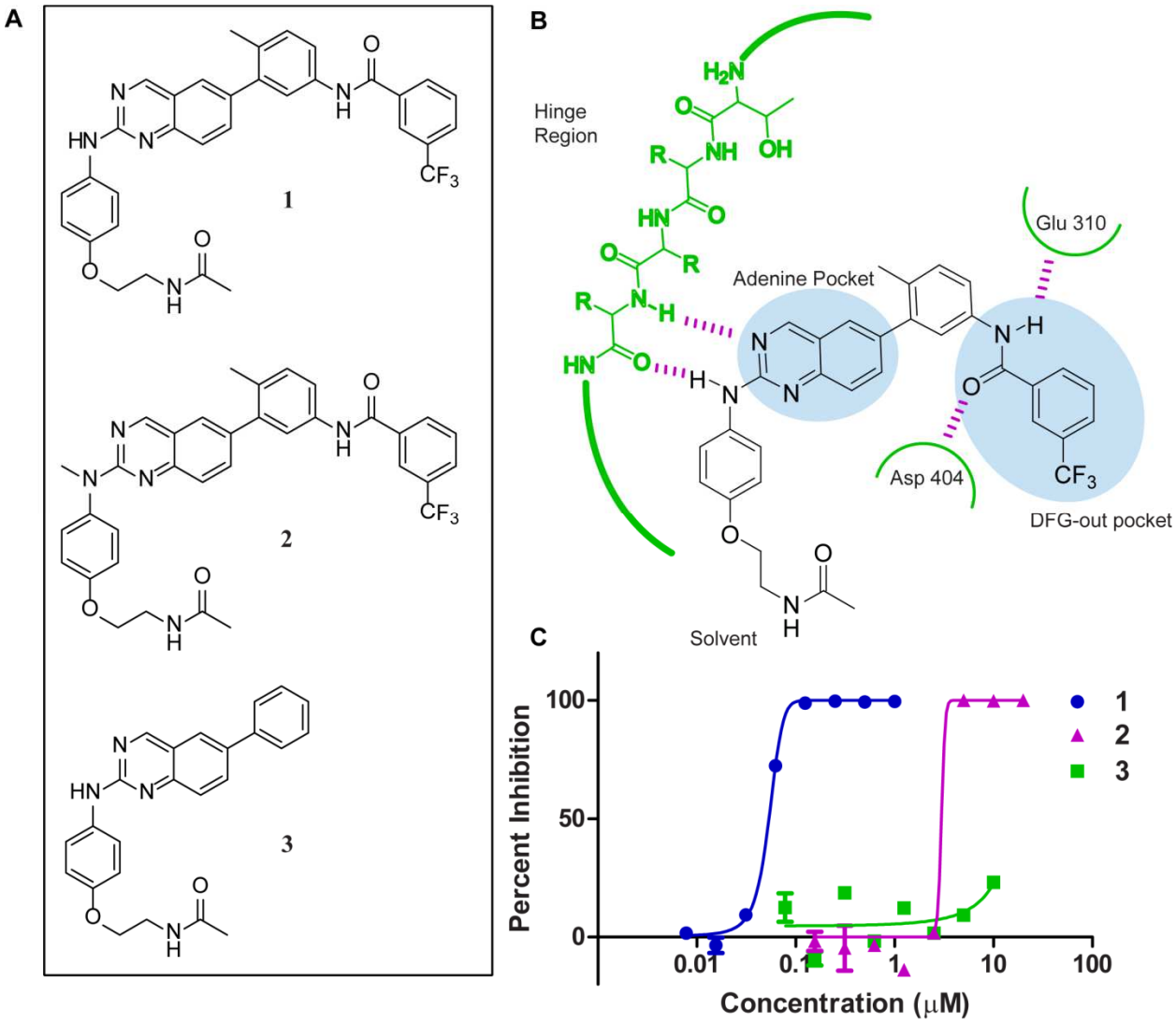


Figure 2- 2. Compound **1** inhibits protein kinases that adopt the DFG-out inactive conformation. (A) Structures of compounds **1**, **2**, and **3**. (B) A schematic of the binding interactions that the type II inhibitor **1** makes with the ATP-binding sites of protein kinases (residue numbering for the mammalian tyrosine kinase SRC is shown). (C) Bloodstream form *T. brucei* was treated with compounds **1**, **2**, or **3** and their inhibitory effect on growth was measured in triplicate.

First, in order to determine whether the observed growth inhibition phenotype results from compound **1** inhibiting a protein kinase, a control analogue was designed and generated that disrupts one of its key interactions with the hinge region of protein kinase active sites. Control compound **2** (Figure 2.2B) contains a methyl group at the secondary amine off the quinazoline ring, thereby precluding one of the hydrogen bonds that is characteristic of type II inhibitors. Therefore, **2** should not bind to protein kinases as tightly as compound **1** (Figure 2-2A). Indeed, an activity assay against the mammalian kinase SRC Y527F shows that compound **1** has an IC_{50} of 0.17 nM while **2** exhibits a significant decrease in potency with an IC_{50} of 81 nM. Likewise, we observed that compound **2** has an EC_{50} of 3400 nM, a greater than 60-fold knockback in EC_{50} compared to compound **1**. This expected decrease in potency of compound **2** supports that the target of **1** is a protein kinase.

We next set about generating a control compound that is not able to exploit interactions with the hydrophobic pocket created by movement of the DFG motif (Figure 2-2A). To do this, control compound **3**, which lacks a terminal benzamide group, was synthesized. **3** has a considerably lower affinity for mammalian kinases that adopt the DFG-out inactive conformation than compound **1**. ENREF 38[37] Therefore, we also expected to see a decrease in the potency of compound **3** compared to **1** in our *T. brucei* growth inhibition assay. Gratifyingly, compound **3** inhibits the growth of the parasite only at concentrations greater than 10,000 nM. Together, the structure/activity relationships of compounds **1**-**3** are compelling evidence that the phenotypic target of **1** is a protein kinase that adopts a specific inactive conformation.

B. ASH* enrichment and identification of compound 1 target proteins

We next set out to identify the targets of compound **1**. We decided to use the ASH* (SNAP-tag(AGT)/SUMO/HaloTag) system that was recently developed in our lab for the catch and release of target proteins. Previously, we demonstrated that we could use the ASH* system to enrich and identify mammalian kinases that adopt the DFG-out inactive conformation.[37] The ASH* system utilizes the self-labeling properties of SNAP-tag for immobilization onto resin displaying a suitable substrate (chloropyrimidine) and the self-labeling reaction of HaloTag to display small molecules that have been derivatized with an alkylchloride substrate (Figure 2-3A). SUMO* (small ubiquitin-like modifier), a recognition site for the protease Ulp1*, is located between the two self-labeling domains in the fusion protein. This protease cleavage site allows captured proteins to be selectively released under mild conditions.

In order to use the ASH* catch-and-release system developed in our lab, compound **1** was modified with a hexylchloride tag at the solvent-exposed amine, generating **1 Probe** (Figure 2-3B). Modification of control compound **3** with a hexylchloride tag resulted in **3 Probe**. Equipped with these probes, we set about using the ASH* system to identify the protein targets of the *T. brucei* inhibitor **1**.

1 Probe and **3 Probe** were each immobilized on resin displaying ASH* and then incubated with bloodstream form *T. brucei* lysate. Both resins were washed to remove non-specifically bound proteins and then incubated with the protease Ulp1* in order to cleave HaloTag and any proteins bound to the inhibitor displayed from HaloTag. The silver stain of the elution from **1 Probe** shows two bands corresponding to the molecular weights of HaloTag and Ulp1* along with several higher molecular weight bands (Figure 2-3C). Notably, most of the higher molecular weight bands are not present in the silver stain elution from the control **3**

Probe. In order to identify the proteins in the higher molecular weight bands, the elutions from the catch-and-release experiments performed with **1 Probe** and **3 Probe** were separated by SDS-PAGE, trypsinized, and then analyzed by mass spectrometry. The pulldown experiments were performed five times with increasing stringency of washes. Altogether, we identified 26 protein kinases from 6 different kinase groups that were selectively enriched by **1 Probe** over **3 Probe** (Table 2-1). Eleven of these kinases were enriched multiple times, indicating true binding to Compound **1** rather than detection of false positives due to non-specific binding.

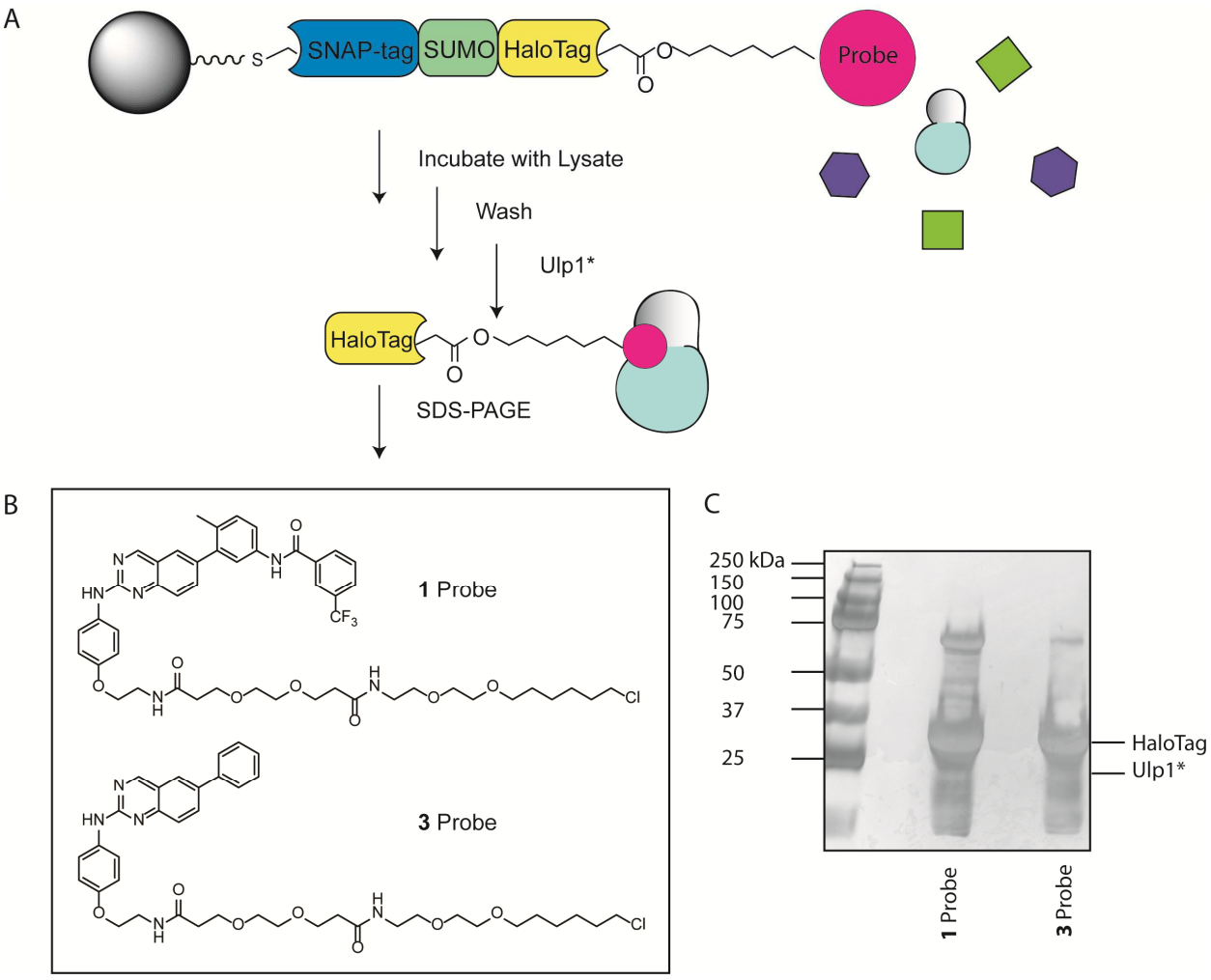


Figure 2- 3. Catch-and-release of protein kinases non-covalently bound to **1 Probe**. (A) The overall strategy for the catch and release of lysate proteins bound to probes **1 Probe** and **3 Probe**. *T. brucei* lysate was added to pre-immobilized ASH* displaying **1 Probe** or **3 Probe**. After extensive washing, bound proteins were released with Ulp1*. Eluted proteins were separated by SDS-PAGE and then identified by tandem mass spectrometry (MS/MS). (B) Structures of **1 Probe** and **3 Probe**. (C) Silver stain analysis of the catch-and-release experiment performed with **1 Probe** and **3 Probe**. Proteins that were selectively enriched with **1 Probe** are shown in Table 2-1.

Kinase	Kinase Group	# of times enriched by 1 Probe	# of times enriched by 3 Probe
Tb427.08.6810	STE	4	
Tb427.10.10350	STE	2	
Tb427.08.5730	STE	2	
Tb427.10.14780	STE	2	1
Tb427tmp.211.3820	STE	2	1
Tb427.10.2040	STE	1	
Tb427.08.3770	CMGC	5	2
Tb427.06.1780	CMGC	3	1
Tb427.10.12040	CMGC	2	
Tb427.08.3550	CMGC	2	1
Tb427.10.3230	CMGC	1	
Tb427.10.16030	CMGC	1	
Tb427.03.4670	CMGC	1	
Tb427.05.1650	CMGC	1	
Tb427.10.1070	CMGC	1	
Tb427.04.5310	NEK	5	
Tb427.05.2820	NEK	1	
Tb427.10.1460	NEK	1	
Tb09.160.4770	CAMK	2	1
Tb427.05.4430	CAMK	1	
Tb427.07.6220	CAMK	1	
Tb427.02.1820	CAMK	1	
Tb11.01.2290	CAMK	1	
Tb427.10.15300	Other	1	
Tb427.10.15880	Other	1	
Tb11.02.3640	AGC	1	

Table 2- 1. Kinases enriched from bloodstream form *T. brucei* lysate with ASH*-immobilized **1 Probe** and **3 Probe** (control).

C. Validation of enriched targets using quantitative mass spectrometry

As a complementary approach to both validate targets found by the ASH* system and identify additional kinase targets, a chemical proteomics method employing a kinase affinity matrix and quantitative mass spectrometry was used. This proprietary solid support displays a general ligand that captures a substantial portion of the mammalian kinome and has been used

for kinase inhibitor profiling in cell lysates.[39,40] In these experiments, affinity matrix is added to lysate that is has been pre-incubated with a kinase inhibitor of interest; target kinases bind to the inhibitor and, therefore, should be unable to bind the affinity matrix. In contrast, non-target proteins show no difference in their abilities to bind to the kinase affinity matrix compared to a DMSO control. The bound proteins are then nonspecifically eluted from the matrix, trypsinized, and analyzed by LC/MS. By performing this assay at varying concentrations of free inhibitor and isobarically tagging each concentration, half maximal residual binding (RB_{50}) curves can be generated for each protein to quantitate their respective affinities (Figure 2-4). Recently, this strategy was employed to profile bloodstream form *T. brucei* kinases: Urbaniak *et al.* used four different "kinobead" platforms displaying various ATP-competitive small molecules to identify 57 kinases from *T brucei* cell lysates.[41]

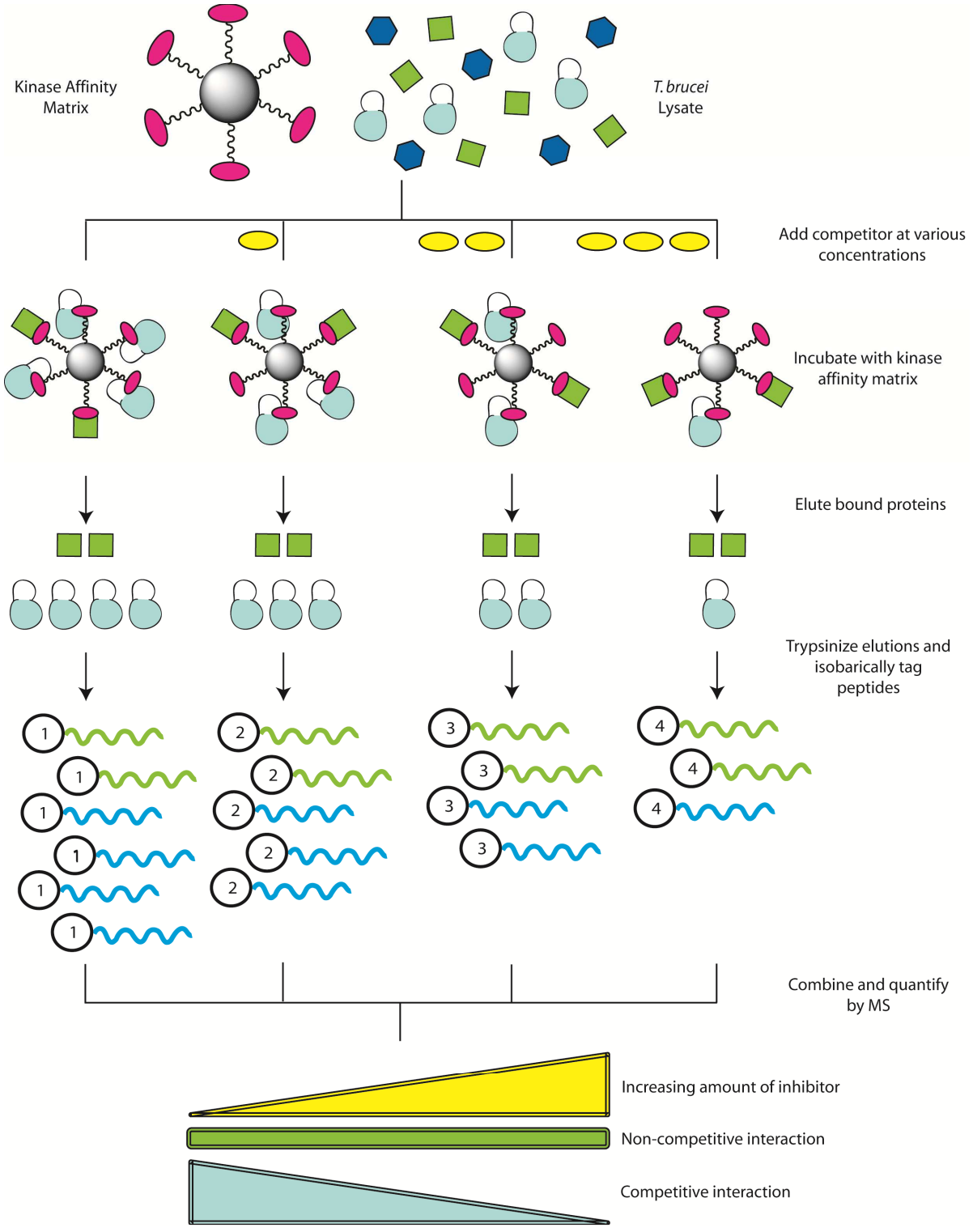


Figure 2- 4. Chemical proteomics strategy for profiling the *T. brucei* targets of kinase inhibitors. *T. brucei* lysate is incubated with varying concentrations of inhibitor before being added to a kinase affinity matrix used for the enrichment of kinases. The presence of inhibitor prevents the target of the inhibitor from binding to the kinase affinity matrix. Bound proteins are eluted, trypsinized, isobarically labeled, and quantified by mass spectrometry. Half maximal residual binding (RB_{50}) curves can be generated for each kinase observed.

We used a comparable kinase affinity matrix to profile bloodstream form *T. brucei* lysate and were able to observe 120 protein kinases across all nine kinase groups. This accounts for 102 of the 124 protein kinases that have been previously observed in the bloodstream form of the parasite, plus 18 newly observed protein kinases.[41] We hypothesize that our enhanced method of peptide separation accounts for the increase in the number of observed kinases compared to the kinobead experiments.

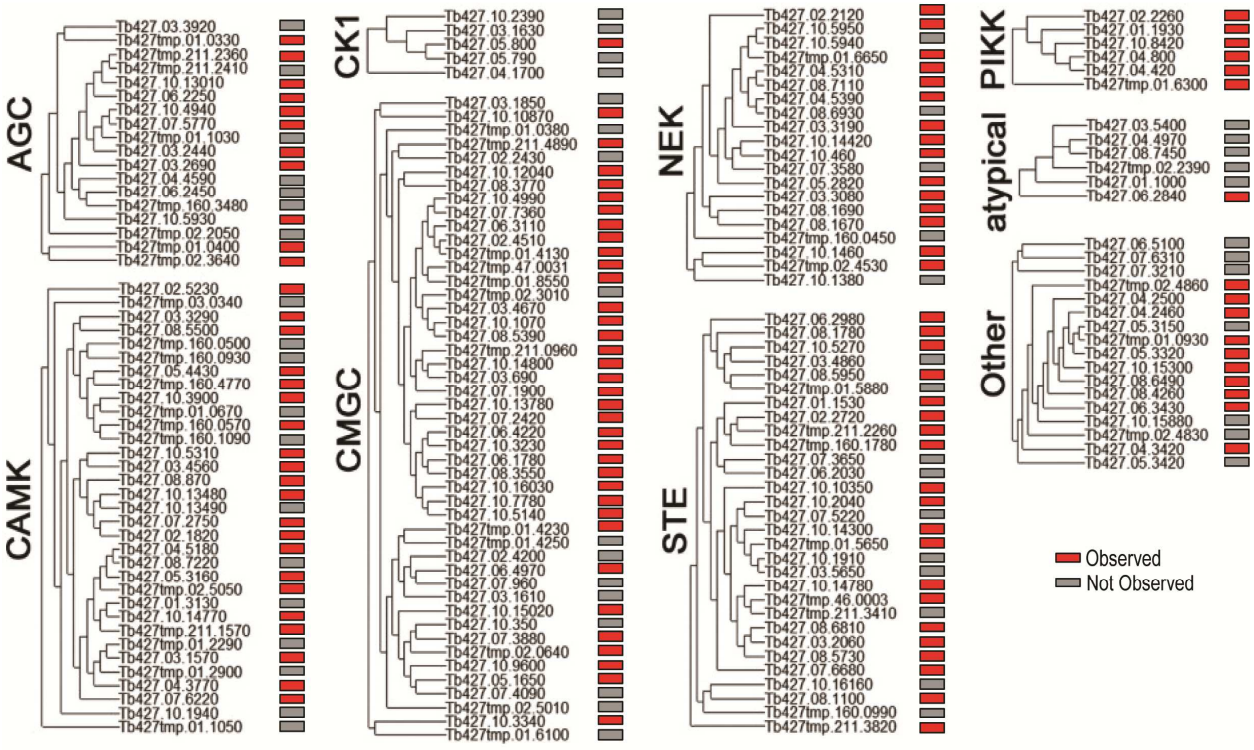


Figure 2- 5. Kinases identified from bloodstream form *T. brucei* using a kinase affinity matrix displaying a general inhibitor.

Satisfied by the large number of kinases enriched with the affinity matrix, we quantified their affinities to compounds **1** and **3**. During this experiment, approximately 2000 proteins were identified, 107 being kinases. Compound **1** binds to 15 *T. brucei* protein kinases (Table 2-2); two of these kinases also bind to control compound **3**, thus leaving 13 remaining kinases that are

potentially responsible for the observed growth inhibition of the bloodstream form of the parasite. Furthermore, only eight non-kinase proteins were competed away by **1**, indicating that protein kinases are the primary target for the inhibitor. Figure 2-6 compares the competition profiles of ligands **1** and **3**, both at 100 µM. The upper left and bottom right quadrants contain kinases that are effectively competed away by compounds **1** and **3**, respectively. The upper right contains proteins that show no competition with either ligand, while the bottom left quadrant shows the kinases that bind to both compounds. The kinases of highest interest are those in the top left quadrant; they are successfully competed away from the affinity matrix by ligand **1** but not **3**.

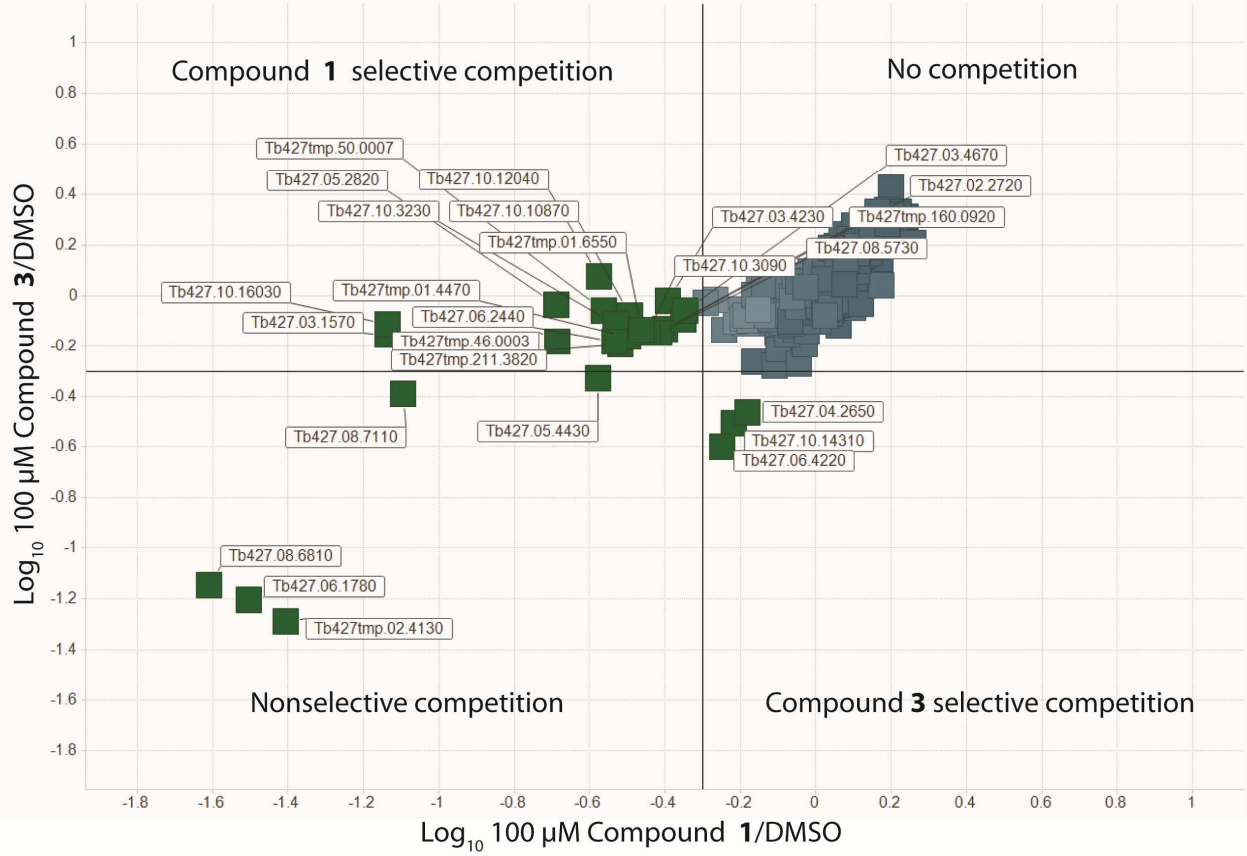


Figure 2- 6. Identification of compound 1-selective kinases. Comparison plot of kinase selectivities of compounds **1** and **3**.

Comparison of the two proteomics methods showed good agreement in kinases that were sensitive to compound **1** (Figure 2-7A). The majority of the kinases identified by both screens were members of the STE, CMGC, NEK, or CAMK kinase groups. STE kinases that were strongly competed (RB_{50} < 60 μM) by compound **1** in the kinase affinity matrix screen (Tb427tmp.211.3820, Tb427.08.5730, and Tb427.08.6810) were also identified more than once in the ASH* catch-and-release method (Table 2-2). Likewise, of the six CMGC group kinases identified via the affinity matrix screen, five were identified in the original ASH* studies. We did expect, though, that some kinases found using the kinase affinity matrix method would not be detected by the ASH* system, primarily because we have shown that the addition of a linker, even at solvent exposed functional groups, can lead to a decrease in the affinity of the kinase for the probe.[42,43] Thus, kinases that bind to compound **1** may demonstrate reduced affinity for **1 Probe** compared to the free inhibitor. We hypothesize that the weakest binders of **1** experienced a large enough decrease in affinity for **1 Probe** that they were not enriched by the ASH* catch-and-release system.

Technically, only one of the two NEK group kinases identified in the kinase affinity matrix screen (Tb427.05.2820) was also identified in the ASH* proteomic study. However, the other kinase (Tb427.08.7110) is highly similar (85% sequence identity) to a NEK group kinase that *was* selectively enriched by **1 Probe** in all five sets of ASH* catch-and-release experiments (Tb427.4.5310). It is possible that Tb427.08.7110 was also enriched by the ASH* system but could not be identified in the mass spectrometry analysis because no unique peptides were identified, and the peptides common to both NEK kinases were assigned to Tb427.4.5310.

Lastly, the CAMK group kinase 427.05.4430 was mutually identified; however, another kinase in this group (Tb427.03.1570) was only identified in the kinase affinity matrix screen.

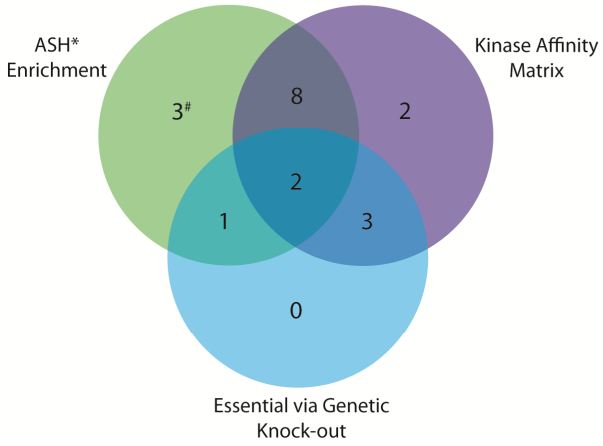
This kinase is strongly competed by compound **1** (RB_{50} = 2 µM) and is quite disparate in sequence compared to any other kinase in *T. brucei*, making it more difficult to explain why it was not observed in both proteomic methods. One possibility is that the abundance of Tb427.03.1570 in the parasite is very low, and hence the trypsinized peptides were not detected by the LC/MS in the ASH* elutions. Another complication is that during the elution of ASH* catch-and-release samples, HaloTag is cleaved from the resin along with the proteins of interest. The peptides from HaloTag are in high abundance compared to the peptides from enriched proteins and may decrease the sensitivity of the MS analysis.

Kinase	Kinase Group	RB_{50} 1	RB_{50} 3
Tb427tmp.211.3820	STE	25	>100
Tb427.08.5730	STE	50	>100
Tb427.02.2720	STE	60	>100
Tb427tmp.46.0003	STE	100	>100
Tb427.08.6810	STE	<1	<10
Tb427.10.16030	CMGC	10	>100
Tb427.10.12040	CMGC	30	>100
Tb427.10.10870	CMGC	40	>100
Tb427.10.3230	CMGC	40	>100
Tb427.03.4670	CMGC	60	>100
Tb427.06.1780	CMGC	<1	<10
Tb427.08.7110	NEK	10	60
Tb427.05.2820	NEK	20	>100
Tb427.03.1570	CAMK	2	>100
Tb427.05.4430	CAMK	20	80

Table 2- 2. Protein kinases selectively competed by **1** in the kinase affinity matrix screen. RB_{50} values are listed.

Of the kinases identified only by the ASH* system (16), only three were selectively enriched by **1 Probe** more than once, a majority of which came from experiments that used very mild washing conditions. This suggests that many of these kinases were non-specifically binding to the resin and are not true targets of compound **1**. Therefore, only kinases that were identified selectively by the ASH* system more than once were considered potential targets.

A



B

Kinase	Indications of Essesntiality
Tb427tmp.211.3820	Conditional null strong growth defect
Tb427.08.5730	Conditional null strong growth defect
Tb427.10.10870	Conditional null growth defect
Tb427tmp.46.0003	null essential; RNAi growth defect
Tb427.02.2720	null essential
Tb09.160.4770	null growth defect
Tb427.08.7110	currently underway
Tb427.03.1570	currently underway
Tb427.10.12040	currently underway
Tb427.05.4430	currently underway
Tb427.4.5310	currently underway
Tb427.8.3770	currently underway
Tb427.10.10350	currently underway
Tb427.10.16030	null viable
Tb427.03.4670	null viable
Tb427.08.6810	null viable
Tb427.05.2820	null viable
Tb427.10.3230	null viable
Tb427.06.1780	null viable

Figure 2- 7. Compound **1**-sensitive kinases essential for cell proliferation in *T. brucei*. (A) Venn diagram summarizing the mutually identified targets of compound **1**. 14 proteins (green) were identified as targets of compound 1 via the ASH* enrichment method, and 15 kinases (purple) were identified via the kinase affinity matrix screen. Using genetic knockouts, 6 kinases (blue) have been identified as important for parasite growth. #12 additional kinases were identified using the ASH* enrichment method but only under the mildest wash conditions (B) Table of essential kinases and genetic indications of essentiality.

D. Genetic analysis of compound 1-sensitive kinases

Very little is known about the 19 kinases identified in our proteomic screens. Their roles in cell signaling pathways in the bloodstream form of the parasite have not been studied, and the necessity of each kinase for normal cell growth is not known. In order to elucidate the functions

of these kinases in the parasite and assess the potential of each as a druggable target, we performed a series of genetic studies.

We first attempted to create gene knockouts for each of the kinases identified in our proteomic studies. In order to make knockouts, cells are transfected by electroporation with a linear DNA construct that contains an antibiotic resistance gene and flanking sequences that target recombination with the gene of interest.[44] The result is that one copy of the endogenous gene is replaced by the gene for antiobiotic resistance. Treating the cells with antibiotic selects for cells that have undergone recombination. The second allele can be replaced in a second transfection step in which the DNA construct has a different antibiotic resistance gene. Cells with copies of both alleles knocked out can be selected for by treating with the two selection antibiotics. The successful creation of a null cell line in which both alleles have been knocked out is strong evidence that the gene is non-essential.

Cell lines that have both alleles knocked out for Tb427.08.6810 and Tb427.06.1780 show no growth inhibition and thus seem to be non-essential. This result corroborates our proteomic data: Tb427.08.6810 and Tb427.06.1780 bound tightly to **1** but also the control compound **3** in the kinase affinity matrix screen. Because of the divergent phenotypes displayed by the parasite when treated with either inhibitor **1** or **3**, non-selectivity indicates that the kinase must not be essential for normal *T. brucei* cell growth. In addition to the two kinases that we expected to be non-essential, null cell lines for Tb427.10.16030, Tb427.03.4670, Tb427.05.2820 and Tb427.10.3230 were also created; that they exhibited normal growth phentotypes indicates that these compound **1**-selective kinases are not essential for normal *T. brucei* cell proliferation (Figure 2-7B). Therefore, these six non-essential kinases were rejected as druggable targets for the treatment of HAT.

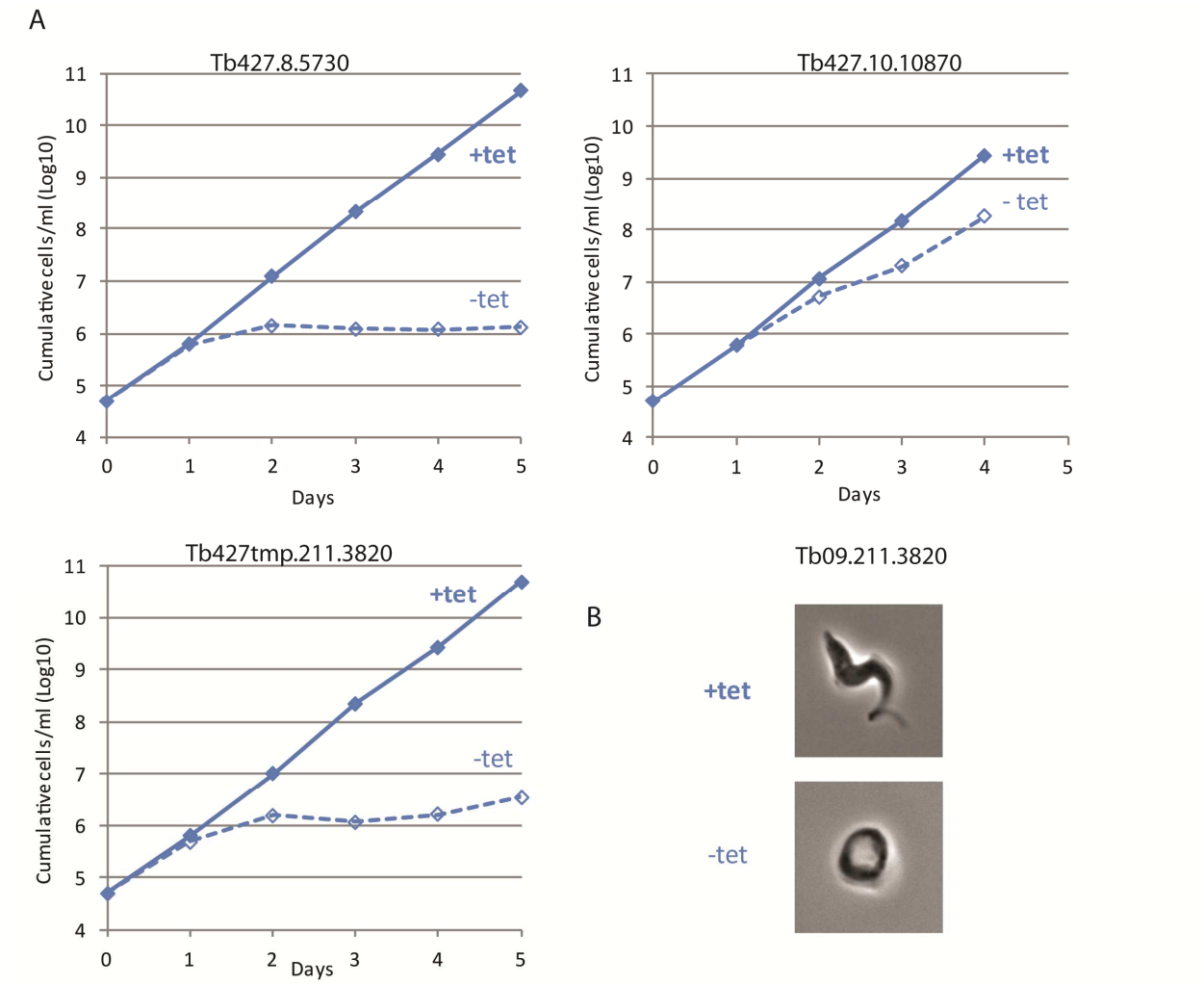


Figure 2- 8. Conditional null cell lines show growth defects upon withdrawal of tetracycline. (A) Growth curves of conditional null cell lines (Tb427.8.5730, Tb427.10.10870, Tb427tmp.211.3820) show growth rate inhibition upon withdrawal of tetracycline. (B) Tb427tmp.211.3820 conditional null cell lines show a distinct morphology in the absence of tetracycline.

Despite ruling out these kinases, we were encouraged by the results of the genetic experiments indicating that others may be essential for *T. brucei* growth. Stable null cell lines cannot be created for essential genes without biological redundancy; attempts to knock out the second allele of Tb427tmp.211.3820, Tb427.08.5730, Tb427.10.10870, Tb427tmp.46.0003, and Tb427.02.2720 resulted in no viable stable cell lines after multiple attempts (data not shown), indicating that they may be essential for growth (Figure 2-7B). In addition to the five kinases so

far that appear to be essential to the bloodstream form of the parasite, null cell lines of Tb09.160.4770 show a growth defect that suggests the importance of this kinase for normal cellular function.

To further validate the essentiality of these kinases, conditional null cell lines were prepared. Conditional null cell lines are generated by integrating a tetracycline-regulatable gene copy prior to the knockout of endogenous alleles.[44] The second endogenous allele is knocked out in the presence of tetracycline so that the regulatable copy of the kinase is expressed. After both endogenous copies of the gene are knocked out, cells can be assayed in the presence and absence of tetracycline and monitored for changes in growth, cell viability, and kinase expression level. Currently, conditional null cell lines for three kinases (Tb427tmp.211.3820, Tb427.08.5730, Tb427.10.10870[2]) have been prepared; each shows a growth defect, thus providing strong evidence that these kinases are essential for normal *T. brucei* cell proliferation (Figure 2-8). Work on creating genetic knockouts and conditional null cell lines for the other seven kinases of interest is underway.

E. Structure/activity relationship studies

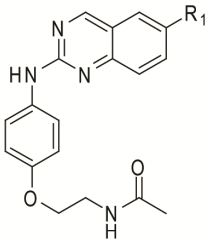
Concurrent with the genetic screens, a second strategy using structure/activity relationships in combination with proteomics was designed to determine which of the identified kinase targets are suitable for drug development. First we designed, synthesized, and evaluated a panel of compounds based on the quinazoline inhibitor scaffold of compound **1** for their ability to inhibit the growth of the bloodstream form of *T. brucei*. We aimed to create two molecules that are very structurally similar to compound **1** but that lose their activity against the parasite,

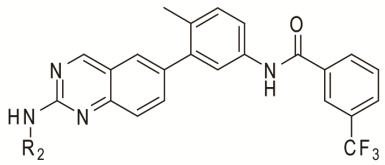
[2]Conditional null cell lines of Tb427.10.10870 show a slightly weaker growth defect compared to the conditional null cell lines for Tb427tmp.211.3820 and Tb427.08.5730.

and one analogue of **1** that maintains its potency against the parasite despite having a moderately different moiety occupying the DFG-out pocket. We then envisioned profiling the three compounds in the kinase affinity matrix and comparing the identified kinases, allowing rapid discernment of the essential targets of interest. Kinases displaced in the original screen by compound **1** and by either of the two compounds that show no growth phenotype in the parasite can be eliminated as druggable targets, just as Tb427.08.6810 and Tb427.06.1780 were eliminated based on their non-selectivity between compounds **1** and **3**. Kinases that are displaced by both **1** and the potent analogue may be essential for normal parasite activity.

We first studied the effect of the R_1 substituent (the moiety that occupies the hydrophobic pocket created by the displacement of the DFG-motif in the inactive conformation of the kinase) on inhibitor potency against the parasite (Table 2-3). The relatively small structural changes of shifting the position of the trifluoromethyl group (**4**) and extending the trifluoromethyl phenyl group deeper into the hydrophobic pocket (**5**, **6**) resulted in correspondingly small changes in EC_{50}. Replacement of the trifluoromethyl phenyl group with a cyclopropyl group resulted in compound **7**, which retains reasonable potency against the parasite (EC_{50} = 155 nM). Because this compound contains a much smaller group at the R_1 position, yet retains its potency against the parasite, it was chosen as one of the compounds for profiling in the kinase affinity matrix screen. Interestingly, when the cyclopropyl group is replaced by a methyl group (compound **8**) potency against the parasite is completely lost ((EC_{50} = >20,000 nM). Compound **8** was chosen for profiling because the R_1 group is still similar to that of **1** (it maintains the amide group that can form hydrogen bonds to the Asp residue of the DFG-motif and the conserved Glu residue in helix αC), yet it shows a complete loss of inhibition in the cellular assay.

For the second non-potent analogue we examined the effect of changing the R_2 group on potency against the parasite. As a general trend, as we decreased the size of the R_2 group, a reduction in EC_{50} was observed. The notable exception to this trend was compound **17** (EC_{50} = 487 nM). Compound **17** contains a solvent-exposed free amine; in compound **1** this amine has been acetylated and thus is slightly larger yet it is nearly ten times more potent than **17**. In mammalian activity assays, compounds **1** and **17** (data not shown) have very similar IC_{50}s. However, compound **18**, which also contains a basic amine has a similar EC_{50} as **17**. The reason for the large discrepancy between **17** and **18** is not clear. From the group of R_2 analogues, we chose to proceed with compound **12** as our inactive analog that is structurally similar to **1**. **12** maintains all of the characteristic hydrogen bonds to the hinge region of kinases, but the R_2 position has been truncated to a methyl group. This truncation results in a 40-fold decrease in cellular potency compared to **1** (EC_{50} = 2,022 nM) but maintains all of the characteristics of a type II inhibitor.





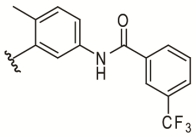
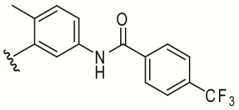
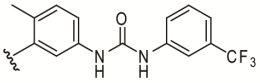
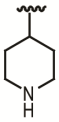
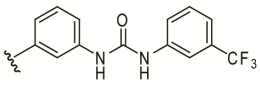
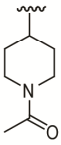
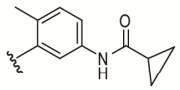
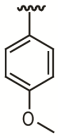
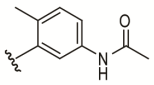
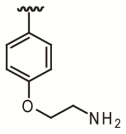
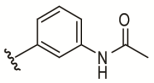
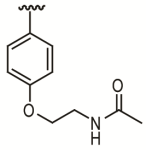
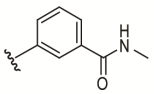
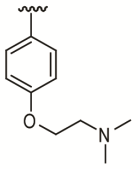
R_1	Cmpd	EC_{50} (nM)	R_2	Cmpd	EC_{50} (nM)
	**1**	50		**11**	2466
				**12**	2022
	**4**	128		**13**	900
	**5**	95		**14**	<156
	**6**	184		**15**	210
	**7**	155		**16**	150
	**8**	>20,000		**17**	487
	**9**	>10,000		**1**	50
	**10**	>20,000		**18**	42
	**3**	>10,000			

Table 2- 3. *T. brucei* proliferation (EC_{50}) results for compounds with variable R_1 and R_2 substructures.

F. Biochemical characterization of Tb427tmp.46.0003

Next we wanted to further characterize the *T. brucei* kinases identified as essential in our genetic screens. However, traditional activity assays are challenging because the substrates of most *T. brucei* kinases are unknown, and many *T. brucei* kinases that have been expressed in *E. coli* have very low catalytic activity. Nevertheless, we screened several purified kinases for activity using a number of commonly used substrates including myelin basic protein (MBP), histone H3, casein, and poly(Glu,Tyr). Unfortunately, no activity was observed for any of the conditions used in our screen. Instead, we used a fluorescent binding assay that was developed in our lab to screen for the DFG-out conformational accessibility of mammalian kinases. This binding assay allows us to directly measure the ligand affinities for the kinases of interest independent of catalytic activity. Ranjitkar *et al.* conjugated the fluorophore BODIPY to a solvent-exposed position on the general type II inhibitor DSA7 in order to generate the probe **DSA-BODIPY**. Conjugation of a fluorophore to this position of DSA7 did not significantly affect its potency against protein kinases in activity assays (Figure 2-9A).[43] Importantly, **DSA-BODIPY** shows a concentration-dependent increase in fluorescence when bound to the ATP-binding sites of protein kinases. This fluorescence assay was used to measure the binding kinetics of several mammalian kinases and later to study various active site mutants of mitogen-activated protein kinases (MAPKs).[45]

DSA7 is similar in structure to compound **1**: both compounds make the same type II-characteristic hydrogen bond contacts to the kinase, and both possess the same moiety that occupies the hydrophobic pocket created by the rotation of the DFG-motif (Figure 2-9B). Because of the structural similarity between **DSA7** and **1**, and the well-established generality of the **DSA7** probe, we hypothesized that we could use this fluorescent assay to further characterize

purified *T. brucei* kinases. The *T. brucei* kinase Tb427tmp.46.0003 was initially characterized with this assay because chemical proteomic data shows that this kinase is a specific target of inhibitor **1** and it appears to be essential for parasite viability. Furthermore, a soluble recombinant form of this kinase can be expressed in *E. coli* and purified to near homogeneity using IMAC. **DSA-BODIPY** (20 nM) was incubated with increasing concentrations of Tb427tmp.46.0003 (2-2000 nM) and then scanned for fluorescence. A concentration-dependent increase in fluorescence was observed (K_d = 55 nM), indicating that **DSA-BODIPY** binds to Tb427tmp.46.0003 (Figure 2-9C). Next, the experiment was repeated in the presence of a competitor, compound **1** (10 µM), and as expected, no increase in fluorescence was observed.

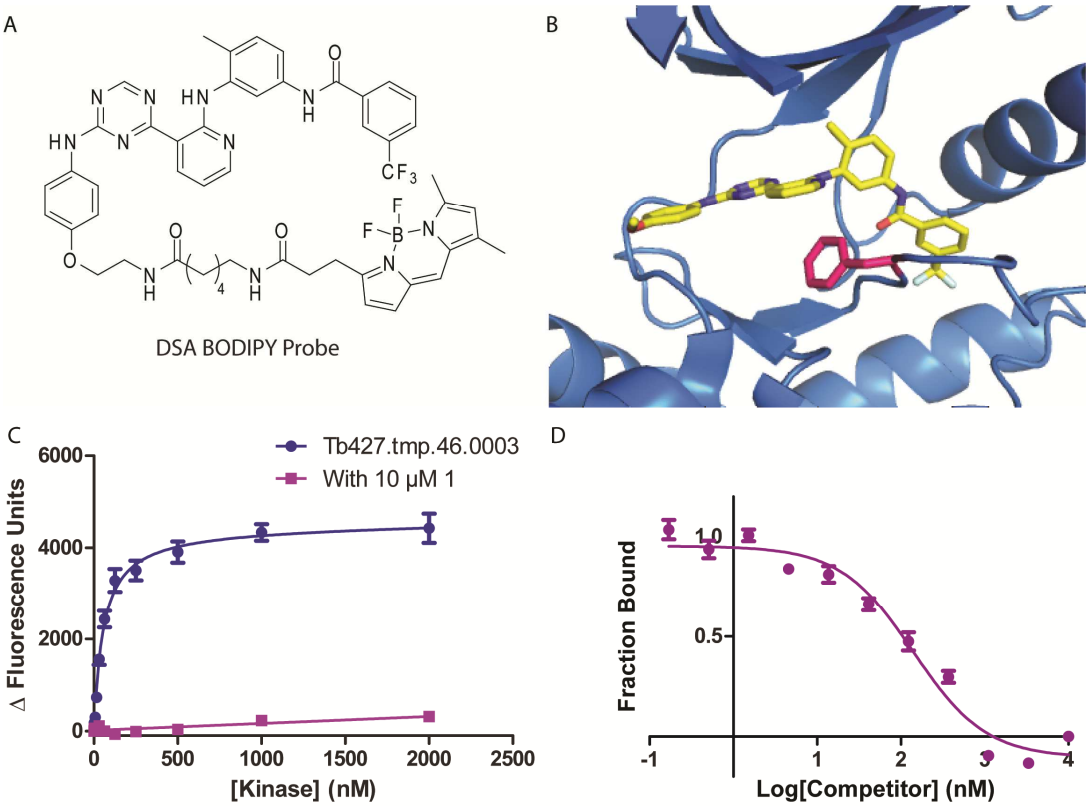


Figure 2- 9. Biophysical characterization of *T. brucei* kinase Tb427tmp46.0003. (A) Structure of the BODIPY-labeled DFG-out selective compound **DSA-BODIPY**. (B) Crystal structure of the DFG-out inhibitor DSA7 (inhibitor scaffold of **DSA-BODIPY**) bound to catalytic domain of STK-10 (LOK) (PDB ID: 4EQU). The Phe of the DFG-out motif (maroon) points out of the pocket. Part of the N-lobe is hidden for clarity. (C) Change in fluorescence observed with increasing concentrations of kinase (Tb427tmp.46.0003) in the presence of **DSA BODIPY** (20 nM) and in the presence of competitor 1 (10 µM). (D) Change in fluorescence observed with increasing concentrations of competitor 1 incubated with 50 nM kinase (Tb427tmp.46.0003) and 20 nM **DSA BODIPY.**

Encouraged by the results of this binding assay, the fluorescent probe was also used in competition assays with several of the inhibitors generated during the SAR studies. In this experiment the concentration of **DSA-BODIPY** (20 nM) and kinase (50 or 100 nM) was held constant. The competitor was titrated in a three-fold dilution series from (100 µM to 0.2 nM). An IC_{50} for each compound was determined by fitting the data using non-linear regression analysis (Figure 2-9D). The results of the biophysical characterization assay showed that compounds **1** and **7** bind tightly to the inhibitor, exhibiting IC_{50}s less than 100 nM (Table 2-4). This enzyme inhibition is in close agreement with our cell proliferation assay data. Likewise, the control compound **3** and inhibitor **12** only weakly inhibit Tb427tmp.46.0003, corresponding to their weak inhibitory effect on *T. brucei* cell proliferation. However, compound **8**, tightly binds to the kinase (IC_{50} = 172 nM) in direct contrast to the results of growth inhibition assay (EC_{50} = >20,000). One explanation for this contrast may be that compound **8** has reduced drug accumulation in the parasite because of low cell permeability or differences in drug import or export compared to compounds **1** and **7**.

Cmpd	IC_{50} (nM)	EC_{50} (nM)
1	94	50
3	4095	>10,000
7	<100	155
8	172	>20,000
12	2189	2022

Table 2- 4. IC_{50} values against Tbtmp.46.0003 and *T. brucei* proliferation (EC_{50}) results for compounds with variable R_1 and R_2 substructures. IC_{50} values were determined via competition assays with **DSA-BODIPY**

III. CONCLUSION

The high molecular weight (violating Lipinski's rules) and low parasite kinase selectivity of compound **1** make it an unattractive drug candidate for the treatment of HAT. However,

compound **1** and its analogues have provided valuable insight into new protein kinases that can be targeted for the treatment of HAT. Through our proteomic and genetic screens we were able to identify several kinases, primarily from the STE kinase group, that seem to be good targets for small-molecule therapies. Importantly, this study demonstrates the druggability of the *T. brucei* kinome for treatments of HAT and should spur research in the area. Further studies should enable the development of inhibitors that bind selectively to *T. brucei* kinases, serving as valuable starting points for the development of HAT therapies that are not toxic to humans.

IV. EXPERIMENTAL

A. SYNTHESIS

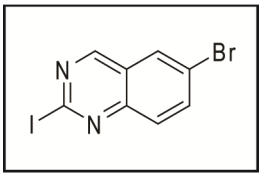
i. General information

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. ^{1}H-NMR spectra were obtained on a Bruker AV-300, AV-301, or Bruker 500 instrument at room tempaerature. Chemical shifts are reported in ppm, and coupling constants are reported in Hz. Mass spectrometry was performed on a Bruker Esquire Ion Trap MS instrument.

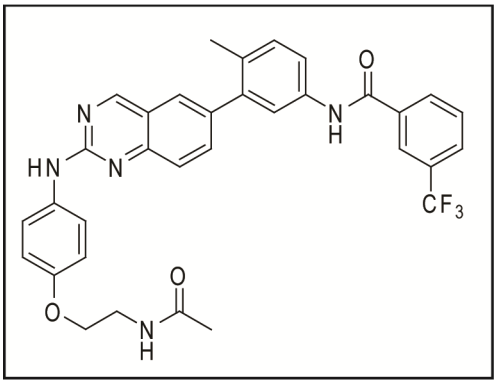
General HPLC purification conditions: Samples were injected onto a preparatory reverse-phase C_{18} column (250 x 21 mm) run over 60 minutes at 8 mL/min (Acetonitrile/Water–0.05% TFA gradient: 1:99 to 100:0). Purified products were detected by UV (254 nm).

General analytical HPLC conditions: General Analytical HPLC Conditions: C_{18} column (150 x 2.1 mm), Acetonitrile/Water-0.05% CF_3CO_2H gradient: 1:99 to 100:0 over 30 min. Methanol/Water–0.05% CF_3CO_2H gradient: 1:99 to 100:0 over 30 min. Flow rate = 1mL/min; 254 nM detection for 30 min.

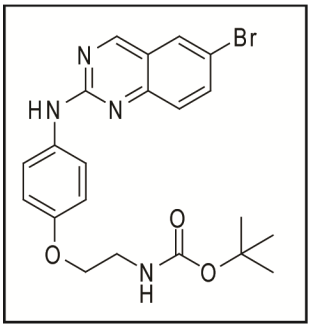
ii. Synthetic procedures



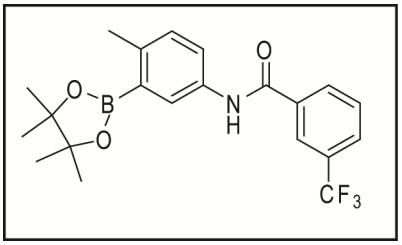
[6-Bromo-2-iodoquinazoline] Synthesized as previously described.[38]



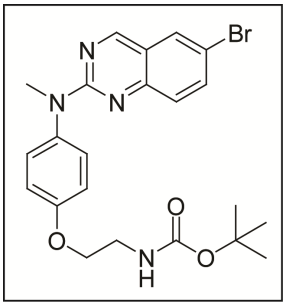
[1] Synthesized as previously described.[45]



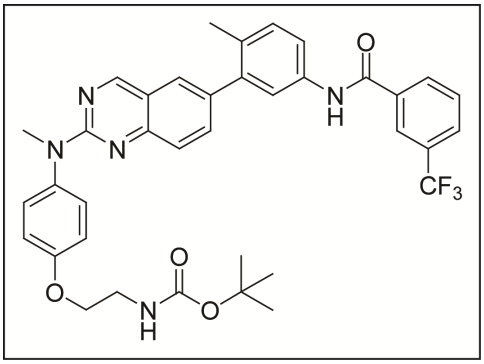
[2A] Synthesized as previously described.[37]



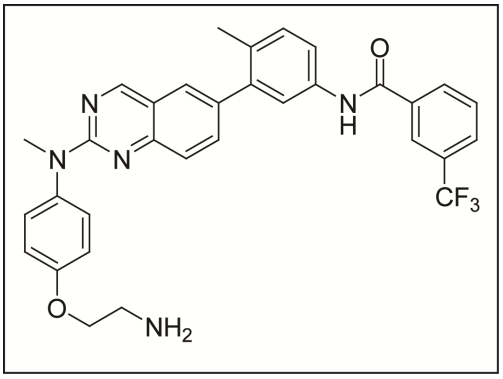
[2B] Synthesized as previously described.[37]



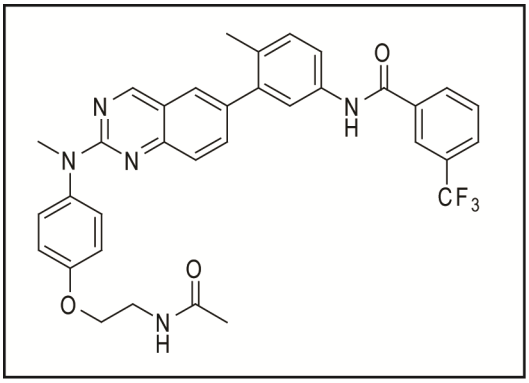
[2C] Compound **2A** (32.8 mg, 0.0716 mmol) was dissolved in DMF (240 µL) and cooled to 0 °C on ice. Sodium hydride (60%) (4.3 mg, 0.107 mmol) was added and the reaction was stirred at 0 °C for 5 min. Methyl iodide (4.9 µL, 0.0787 mmol) was added dropwise and the reaction was warmed to room temperature. The reaction was monitored by TLC and stirred until the starting material was no longer detected. The reaction was diluted in ethyl acetate, washed with water, dried over Na_2SO_4, and concentrated *in vacuo*. The product was used in the next reaction without further purification. Calculated for ($C_{22}H_{25}BrN_4O_3$) ($M+H^+$): 473.1; found 473.3.



[2D] A mixture of compound **2C** (34 mg, 0.0716 mmol), **2B** (34.8 mg, 0.0859 mmol), tetrakis(triphenylphosphine)palladium (2.6 mg, 2.25 µmol), and sodium carbonate (16.7 mg, 0.158 mmol) was dissolved in a 3:1 mixture of DME/water (280 µL). The mixture was heated overnight at 85 °C. The reaction was diluted in ethyl acetate, washed with water, dried over Na_2SO_4, and concentrated *in vacuo*. The product was used in the next reaction without further purification. Calculated for ($C_{37}H_{36}F_3N_5O_4$) ($M+H^+$): 672.3; found 672.4.



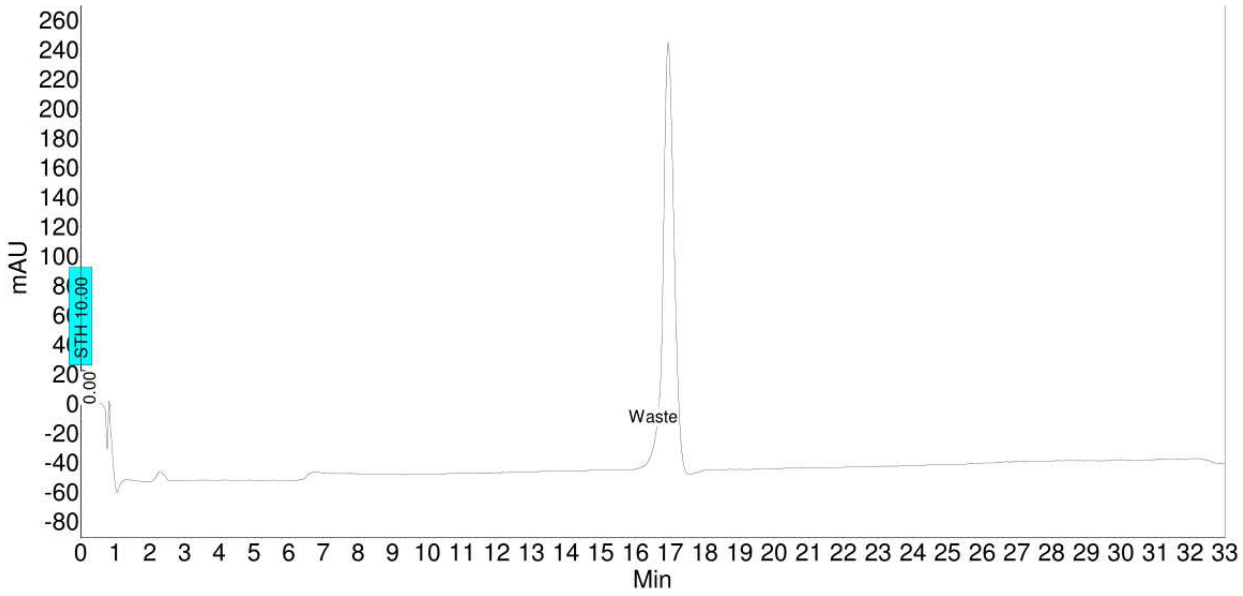
[2E] Compound **2D** (48 mg, 0.0716 mmol) was dissolved in CH_2Cl_2 (1.2 mL) and TFA (0.51 mL) and stirred for 3 h at room temperature. The reaction mixture was concentrated *in vacuo* and azeotroped with toluene. The resulting product **2E** was used without further purification. Calculated for ($C_{32}H_{28}F_3N_5O_2$): 572.2; found 572.3.



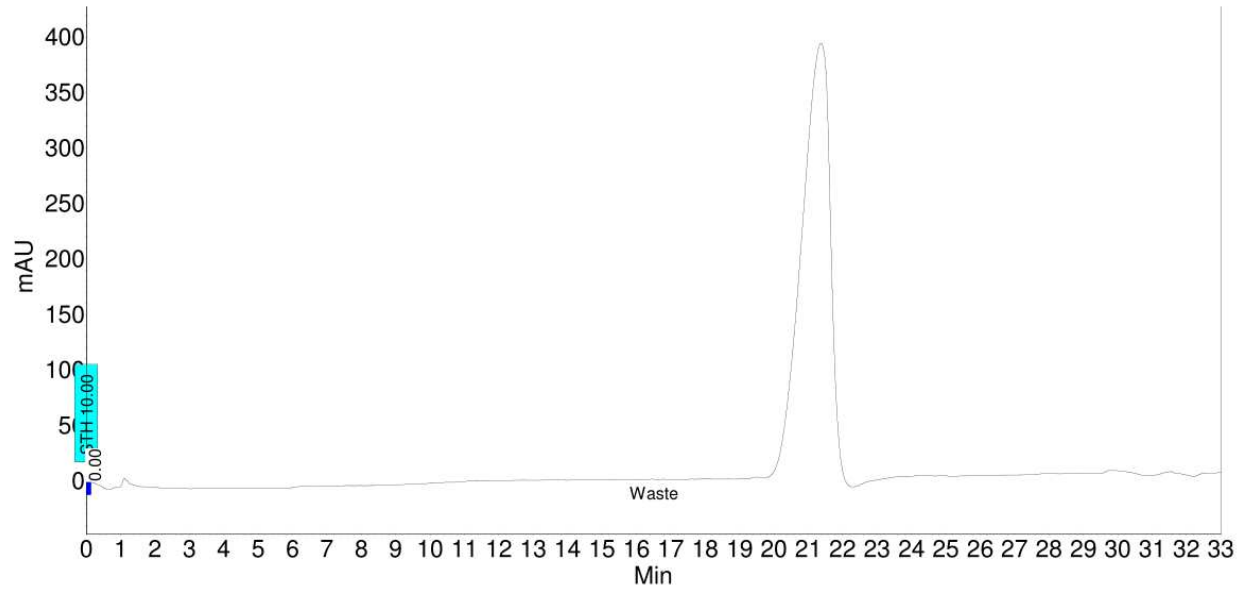
[2] Compound **2E** (41 mg, 0.0716 mmol) was added to acetic acid anhydride (1.4 mL) on ice. Triethylamine (40 µL, 0.287 mmol) was added and the reaction was warmed to room temperature and stirred for 1 h. The resulting crude product was concentrated *in vacuo* and purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.48 (s, 1H), 8.31 – 8.18 (m, 2H), 8.14 – 7.97 (m, 1H), 7.94 – 7.67 (m, 3H), 7.64 – 7.49 (m, 2H), 7.48 – 7.28 (m, 4 H), 7.22 – 7.12 (m, 2H), 5.73 (s, 1H), 4.20 – 4.10 (m, 2H), 3.77 (s, 1H), 3.68 –

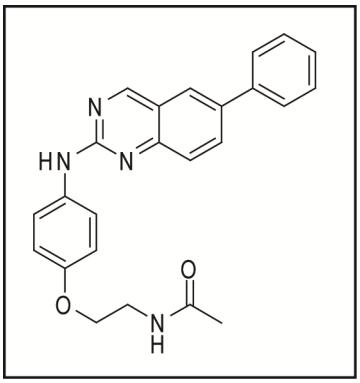
3.56 (m, 5H), 2.29 (m, 3H), 2.00 (m, 3H). Calculated for ($C_{34}H_{30}F_3N_5O_3$) ($M+H^+$): 614.2; found 614.6.

Analytical HPLC trace of **2**: Acetonitrile/Water

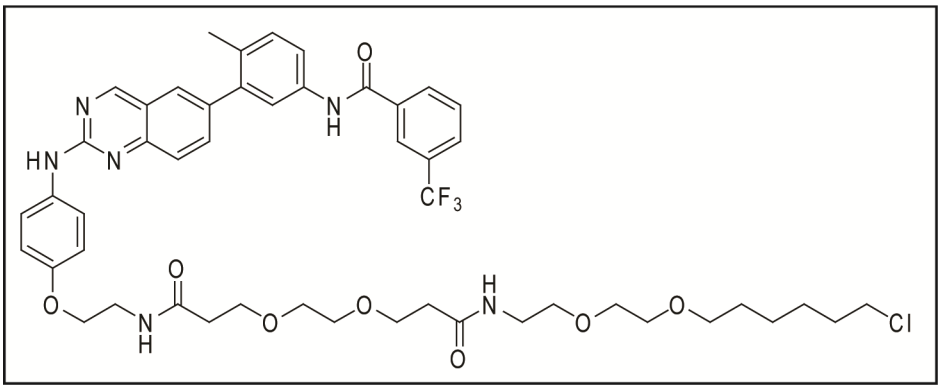


Analytical HPLC trace of **2**: Methanol/Water\

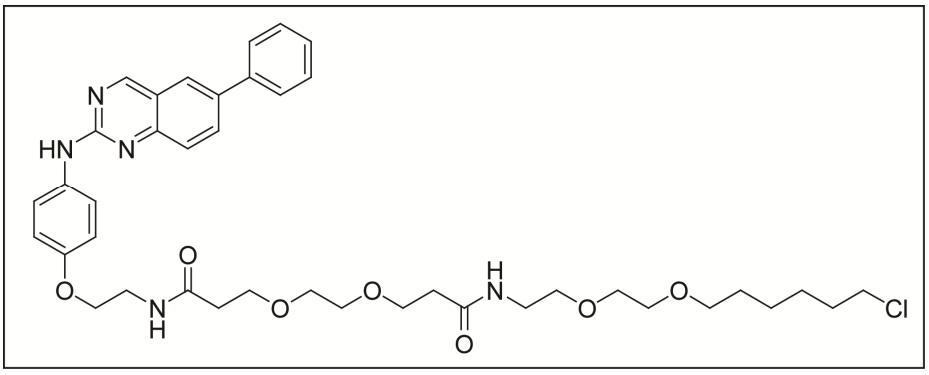




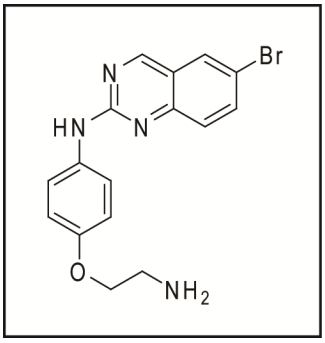
[3] Synthesized as previously described.[37]



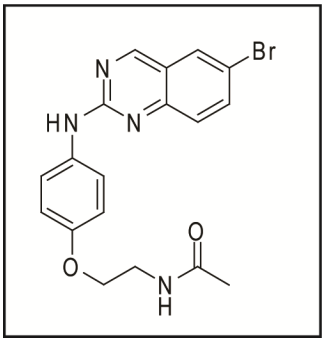
[1 Probe] Synthesized as previously described.[37]



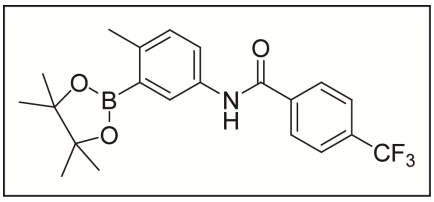
[3 Probe] Synthesized as previously described.[37]



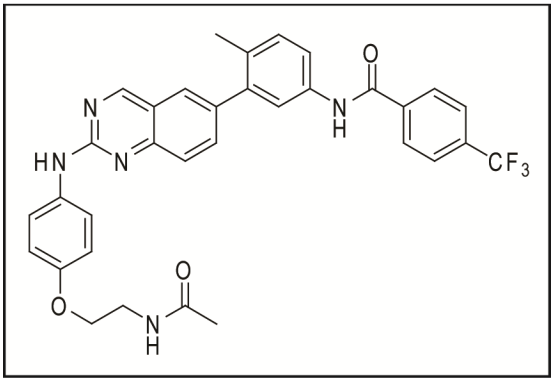
[4A] Compound **2A** (107 mg, 0.234 mmol) was dissolved in CH_2Cl_2 (4.0 mL) and TFA (1.9 mL) and stirred for 3 h at room temperature. The reaction mixture was concentrated *in vacuo* and azeotroped with toluene. The resulting product **4A** was used without further purification. Calculated for ($C_{16}H_{15}BrN_4O$): 359.0; found 359.1.



[4B] Compound **4A** (229 mg, 0.640 mmol) and DIEA (262 µL, 1.92 mmol) were dissolved in dichloromethane (2.52 mL) and cooled to 0 °C on ice. Acetic anhydride (73 µL, 0.768 mmol) was added and the reaction was allowed to warm to room temperature and stirred overnight at room temperature. The reaction was concentrated *in vacuo* and used without further purification. ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.08 (s, 1H), 7.98 – 7.84 (m, 2H), 7.70 – 7.53 (m, 3H), 6.99 – 6.88 (m, 2H), 4.06 (t, J = 5.1 Hz, 2H), 3.72 – 3.66 (m, 2H), 2.03 (s, 3H). Calculated for ($C_{18}H_{17}BrN_4O_2$) ($M+H^+$): 401.1; found 401.5.



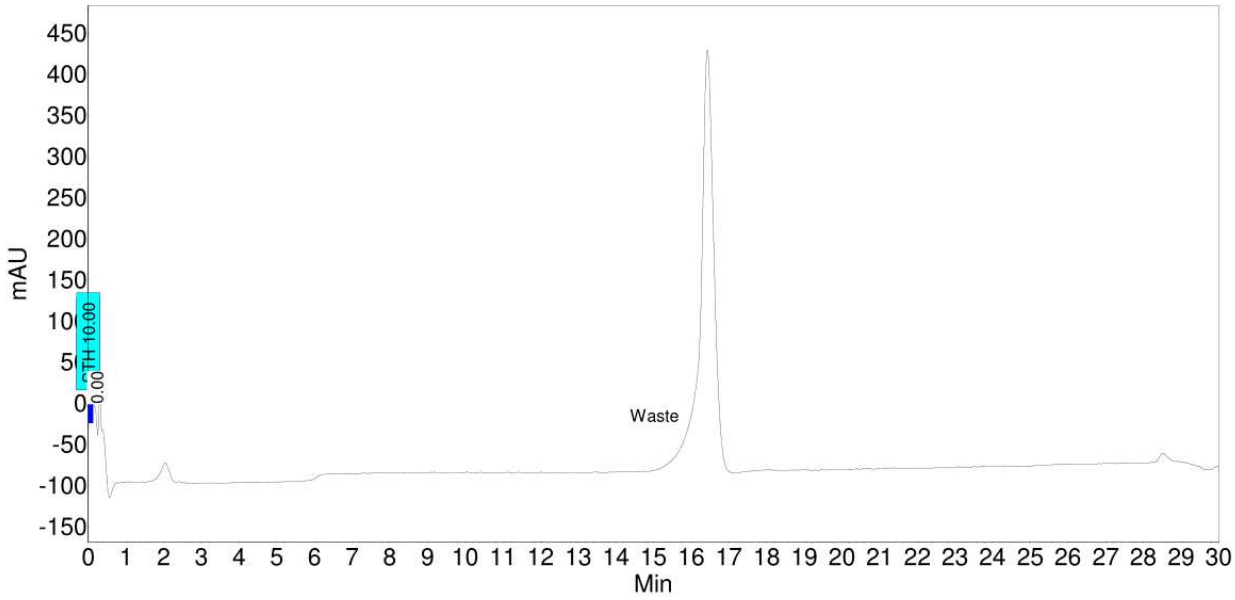
[4C] 4-(trifluoromethyl)benzoic acid (100 mg, 0.526), 5-amino-2-methylphenylboronic acid pinacol ester (159 mg, 0.683 mmol), DIEA (275 µL, 1.58 mmol), and HOAt (93 mg, 0.684 mmol) were dissolved in DMF (2.56 mL) and cooled to 0 °C on ice. EDCI (131 mg, 0.684 mmol) was added and the reaction was warmed to room temperature and stirred overnight. The crude mixture was diluted in ethyl acetate and washed with brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by column chromatography. ^{1}H NMR (300 MHz, Methanol-d4) δ 8.08 (d, J = 8.1 Hz, 2H), 7.94 (d, J = 2.4 Hz, 1H), 7.81 (d, J = 8.2 Hz, 2H), 7.71 (dd, J = 8.2, 2.4 Hz, 1H), 7.18 (d, J = 8.3 Hz, 1H), 2.49 (s, 3H), 1.35 (s, 12H). Calculated for ($C_{21}H_{23}BF_3NO_3$) ($M+H^+$): 406.2; found 406.3.



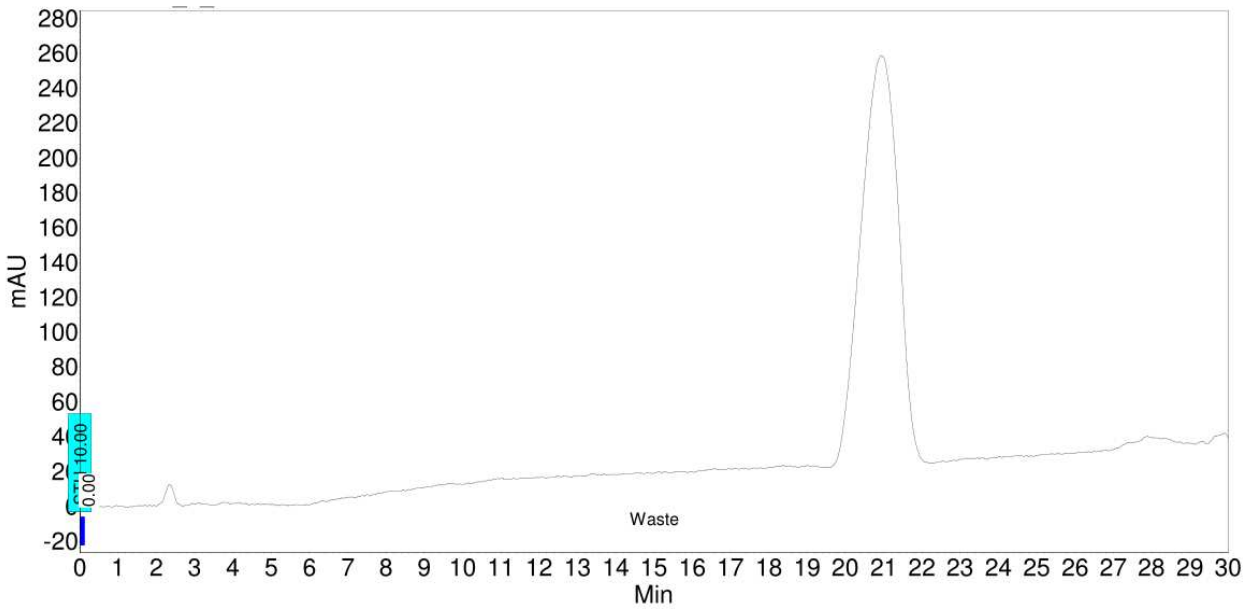
[4] A mixture of compound **4B** (8.32 mg, 0.0207 mmol), **4C** (10.08 mg, 0.0249 mmol), tetrakis(triphenylphosphine)palladium (0.72 mg, 0.623 µmol), and sodium carbonate (4.83 mg, 0.0456 mmol) was dissolved in a 3:1 mixture of DME/water (300 µL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was purified by reverse phase chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.23 (s, 1H), 8.16 – 8.08 (m, 2H),

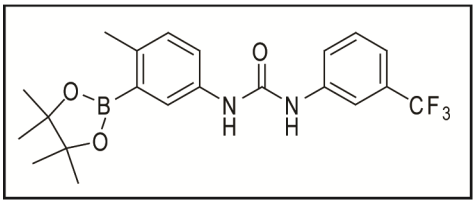
7.88 – 7.80 (m, 3H), 7.78 – 7.62 (m, 4H), 7.60 – 7.47 (m, 1H), 7.40 – 7.28 (m, 2H), 7.18 – 6.97 (m, 2H), 4.16 – 4.05 (m, 2H), 3.64 – 3.55 (m, 2H), 2.32 (s, 3H), 2.01 – 1.97 (m, 3H). Calculated for ($C_{33}H_{28}F_3N_5O_3$) ($M+H^+$): 600.6; found 600.9.

Analytical HPLC trace of **4**: Acetonitrile/Water

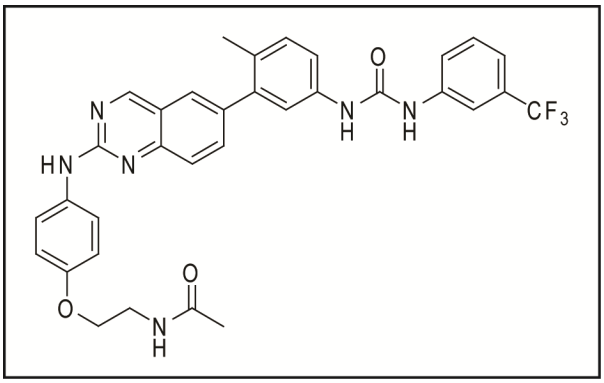


Analytical HPLC trace of **4**: Methanol/Water



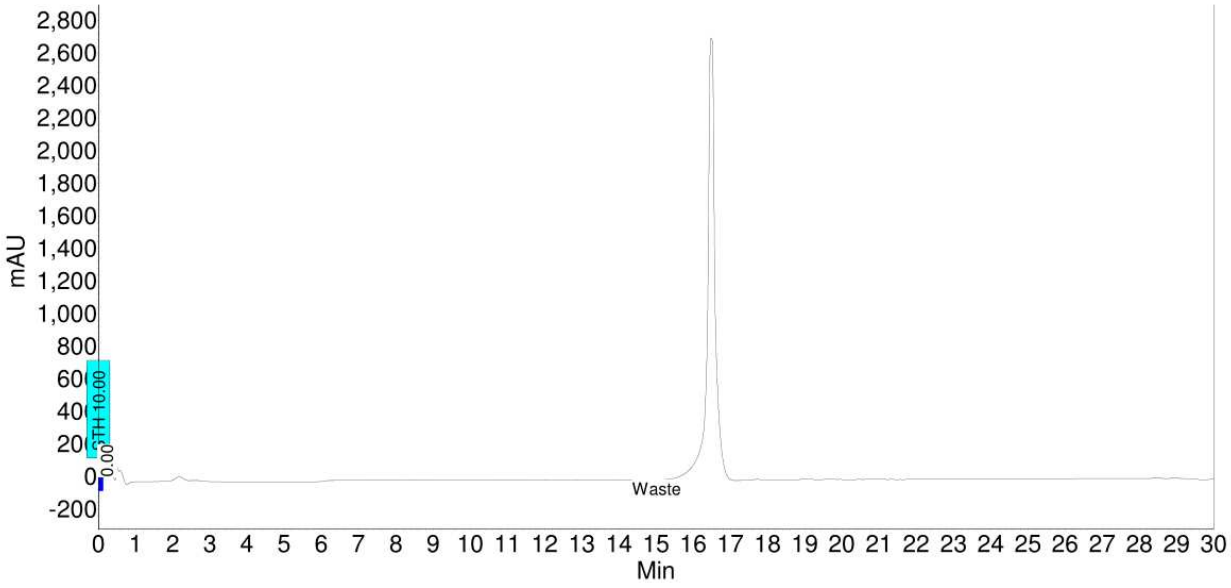


[5A] 5-amino-2-methylphenyl boronic acid pinacol ester (50 mg, 0.214 mmol) and 3-(trifluoromethyl)phenyl isocyanide (38 µL, 0.278 mmol) were dissolved in THF (1.79 mL) and stirred overnight at room temperature. The crude mixture was diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by column chromatography. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.51 (s, 1H), 7.35 – 7.21 (m, 3H), 7.19 – 7.08 (m, 1H), 7.03 – 6.93 (m, 1H), 6.84 (d, J = 8.3 Hz, 1H), 2.19 (s, 3H), 1.06 (s, 12H). Calculated for ($C_{21}H_{24}BF_3N_2O_3$) ($M+H^+$): 421.2; found 421.5.

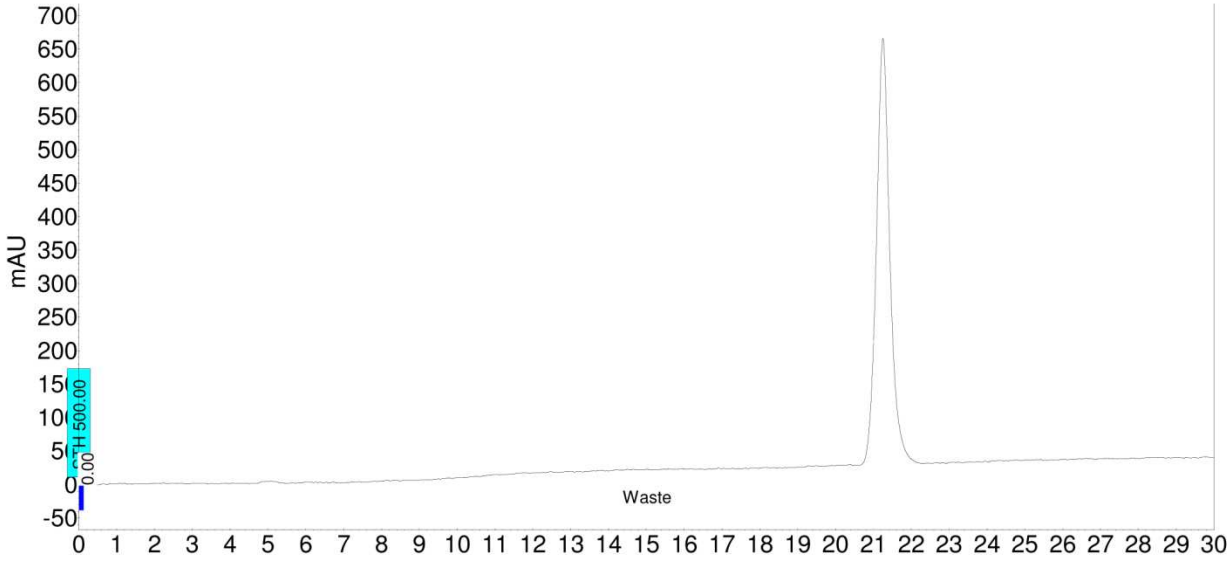


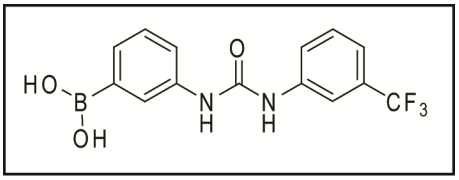
[5] A mixture of compound **4B** (10.0 mg, 0.0249 mmol), **5A** (12.6 mg, 0.0299 mmol), tetrakis(triphenylphosphine)palladium (0.86 mg, 0.744 µmol), and sodium carbonate (5.81 mg, 0.0548 mmol) was dissolved in a 3:1 mixture of DME/water (800 µL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was cooled to room temperature, concentrated *in vacuo*, and purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.44 (s, 1H), 8.05 – 7.89 (m, 1H), 7.79 (d, J = 8.7 Hz, 1H), 7.66 – 7.43 (m, 5H), 7.42 – 7.22 (m, 5H), 7.19 – 7.07 (m, 2H), 5.80 (s, 1H), 4.13 (t, J = 5.5 Hz, 2H), 3.67 – 3.57 (m, 2H), 2.31 – 2.21 (m, 3H), 2.04 – 1.96 (m, 3H). Calculated for ($C_{33}H_{29}F_3N_6O_3$) ($M+H^+$): 615.2; found 615.5.

Analytical HPLC trace of **5**: Acetonitrile/Water

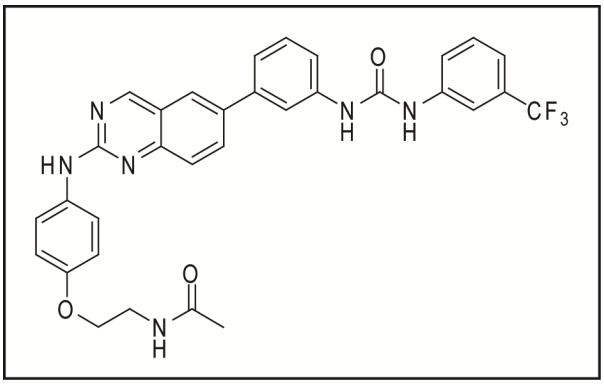


Analytical HPLC trace of **5**: Methanol/Water





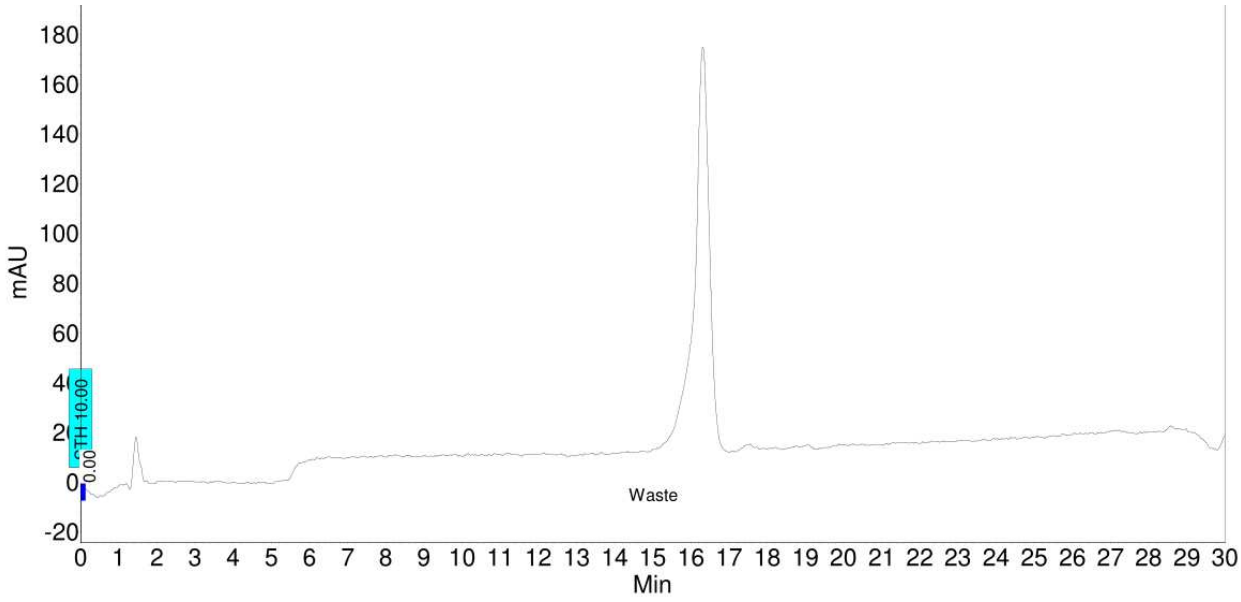
[6A] 5-aminophenyl boronic acid (50 mg, 0.365 mmol) and 3-(trifluoromethyl)phenyl isocyanide (65 µL, 0.478 mmol) were dissolved in THF (3.6 mL) and stirred overnight at room temperature. The crude mixture was diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by column chromatography. ^{1}H NMR (300 MHz, Methanol-d4) δ 7.93 – 7.87 (m, 1H), 7.66 – 7.56 (m, 2H), 7.55 – 7.48 (m, 1H), 7.47 – 2.42 (m, 1H), 7.35 – 7.25 (m, 3H). Calculated for ($C_{14}H_{12}BF_3N_2O_3$) ($M+H^+$): 325.1; found 325.3.



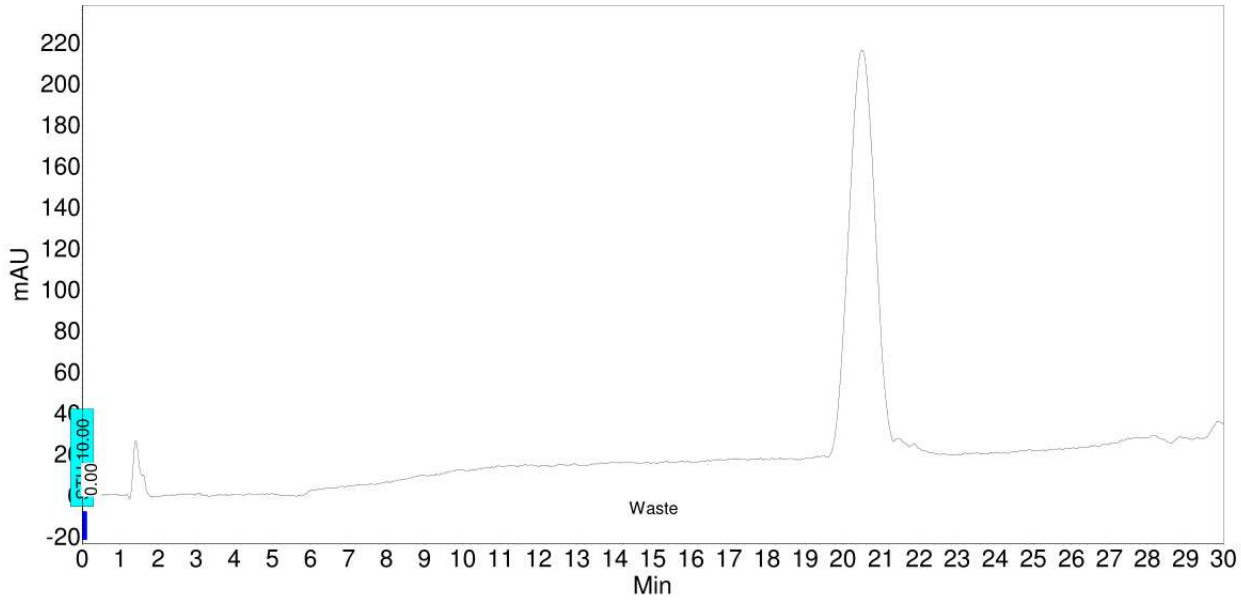
[6] A mixture of compound **4B** (8.32 mg, 0.0207 mmol), **6A** (8.06 mg, 0.0249 mmol), tetrakis(triphenylphosphine)palladium (0.72 mg, 0.623 µmol), and sodium carbonate (4.83 mg, 0.456 mmol) was dissolved in a 3:1 mixture of DME/water (300 µL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.35 (s, 1H), 8.23 – 8.14 (m, 1H), 8.02 – 7.91 (m, 2H), 7.81 – 7.71

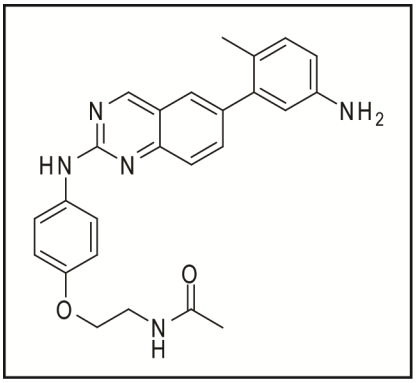
(m, 1H), 7.61- 7.56 (m, 2H), 7.52 – 7.24 (m, 7H), 7.16 – 6.99 (m, 2H), 5.81 (s, 1H), 4.14 – 4.06 (m, 2H), 3.58 (t, J = 5.6 Hz, 2H), 1.99 – 1.96 (m, 3H). Calculated for ($C_{32}H_{27}F_3N_6O_3$) ($M+H^+$): 601.2; found 601.6.

Analytical HPLC trace of **6**: Acetonitrile/Water



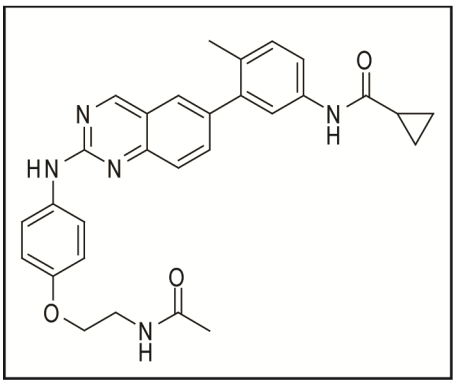
Analytical HPLC trace of **6**: Methanol/Water





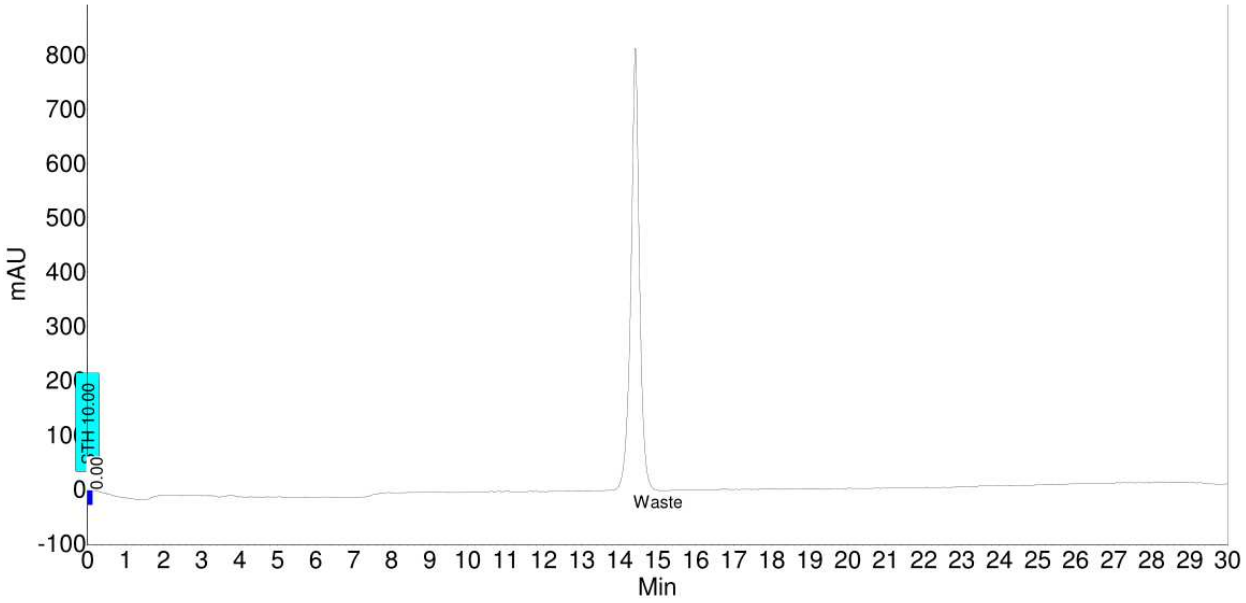
[7A] A mixture of compound **4B** (23 mg, 0.0573 mmol), 5-amino-2-methylphenylboronic acid pinacol ester (16 mg, 0.0686 mmol), tetrakis(triphenylphosphine)palladium (2 mg, 1.73 µmol), and sodium carbonate (13.4 mg, 0.126 mmol) was dissolved in a 3:1 mixture of DME/water (1.2 mL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic

layer was dried over Na_2SO_4, concentrated *in vacuo*, and used in the next step without further purification. Calculated for ($C_{25}H_{25}N_5O_2$) ($M+H^+$): 428.2; found 428.5.

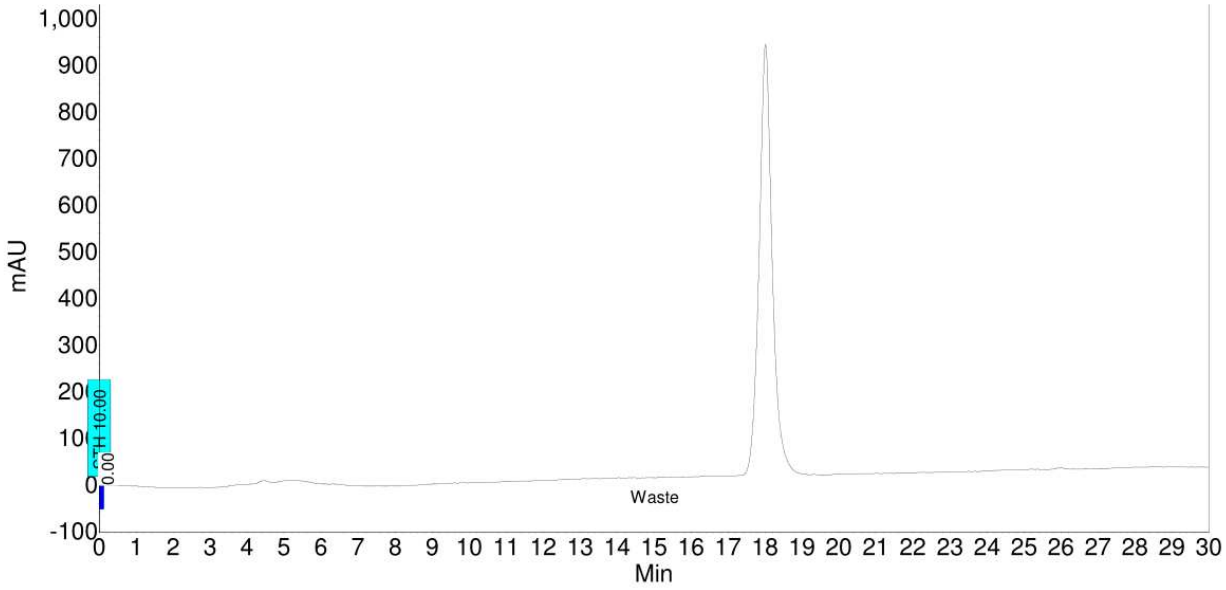


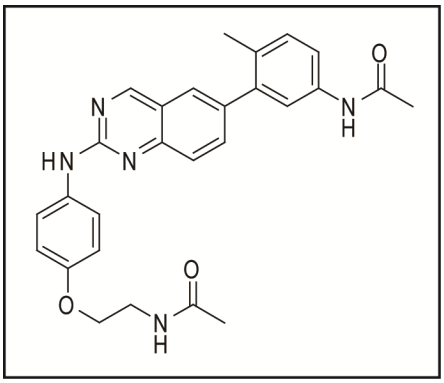
[7] Compound **7A** (10.7 mg, 0.0250 mmol) was dissolved in dichloromethane (600 µL) and cooled to 0 °C on ice. Triethylamine (2.5 µL, 0.0275 mmol) was added. Cyclopropanecarbonyl chloride was dissolved in dichloromethane, cooled to 0 °C, and added dropwise. The reaction was warmed to room temperature, stirred for 1 h, then concentrated *in vacuo*. The resulting product was purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.27 (d, *J* =10.6 Hz, 1H), 7.99 – 7.83 (m, 3H), 7.74 – 7.38 (m, 4H) 7.13 – 7.02 (m, 1H) 6.99 – 6.85 (m, 2H), 5.30 (s, 1H), 4.08 (t, *J* = 5.1 Hz, 2H), 3.74 – 3.66 (m, 2H), 2.23 (s, 3H), 2.08 – 2.02 (m, 3H), 1.12 – 1.05 (m, 2H), 0.93 – 0.84 (m, 3H). Calculated for ($C_{29}H_{29}N_5O_3$) ($M+H^+$): 496.2; found 496.2.

Analytical HPLC trace of **7**: Acetonitrile/Water



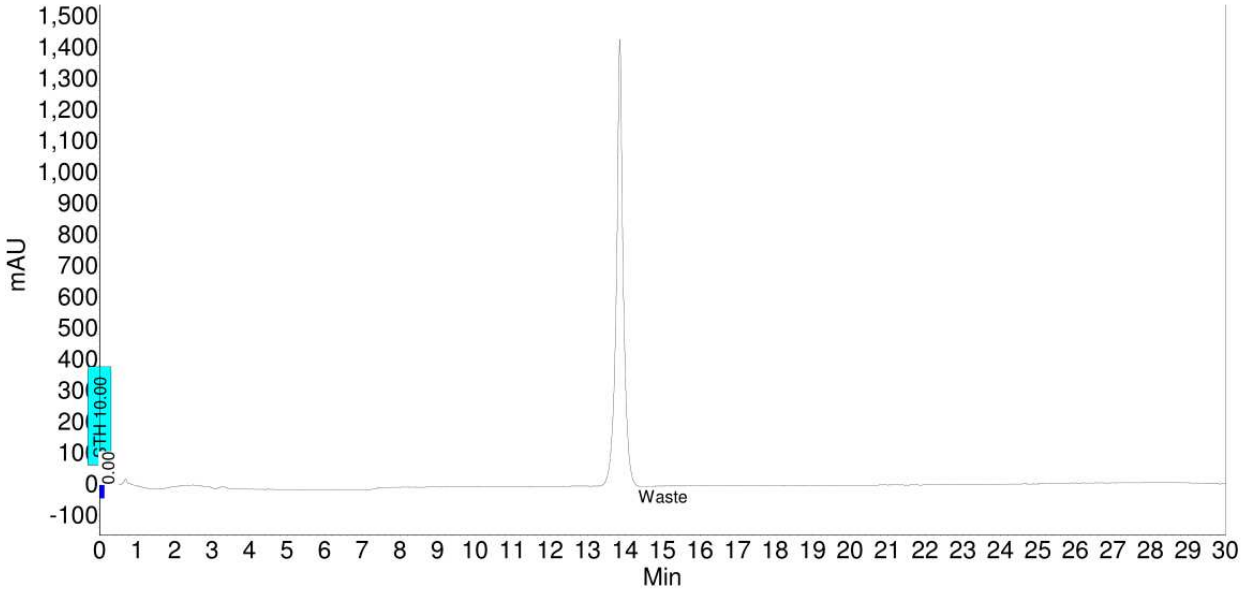
Analytical HPLC trace of **7**: Methanol/Water



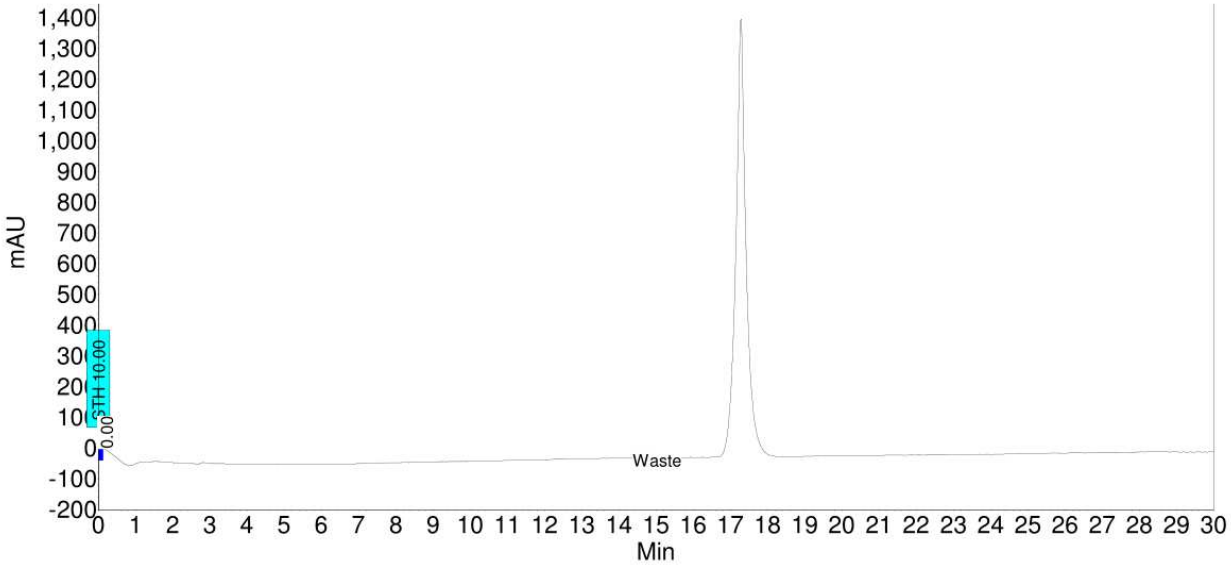


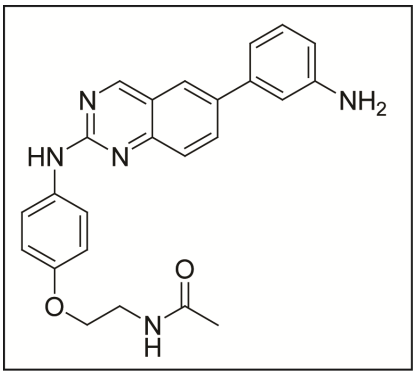
[8] Compound **7A** (8.2 mg, 0.0192 mmol) and DIEA (10 µL, 0.0575 mmol) were dissolved in dichloromethane (100 µL) and cooled to 0 °C on ice. Acetic anhydride (5 µL, 0.0526 mmol) was added and the reaction was allowed to warm to room temperature and stirred overnight at room temperature. The reaction was concentrated *in vacuo* and used without further purification. ^{1}H NMR (300 MHz, Chloroform-d) δ 9.31 – 9.25 (m, 1H), 8.01 – 7.82 (m, 3H), 7.66 – 7.37 (m, 3H), 7.42 – 7.33 (m, 1H), 7.11 – 7.04 (m, 1H), 7.02 – 6.84 (m, 2H), 5.30 (s, 1H), 4.08 (t, J = 4.9 Hz, 2H), 3.75 – 3.68 (m, 2H), 2.23 (s, 3H), 2.20 (s, 3H), 2.06 (s, 3H). Calculated for ($C_{27}H_{27}N_5O_3$) ($M+H^+$): 470.2; found 470.5.

Analytical HPLC trace of **8**: Acetonitrile/Water

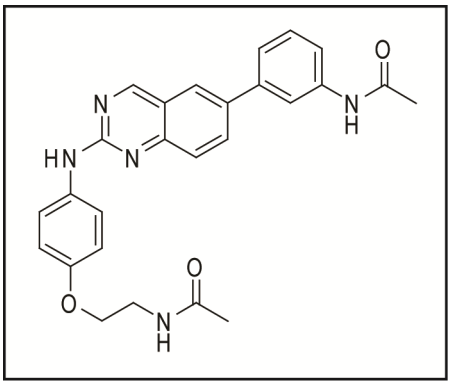


Analytical HPLC trace of **8**: Methanol /Water





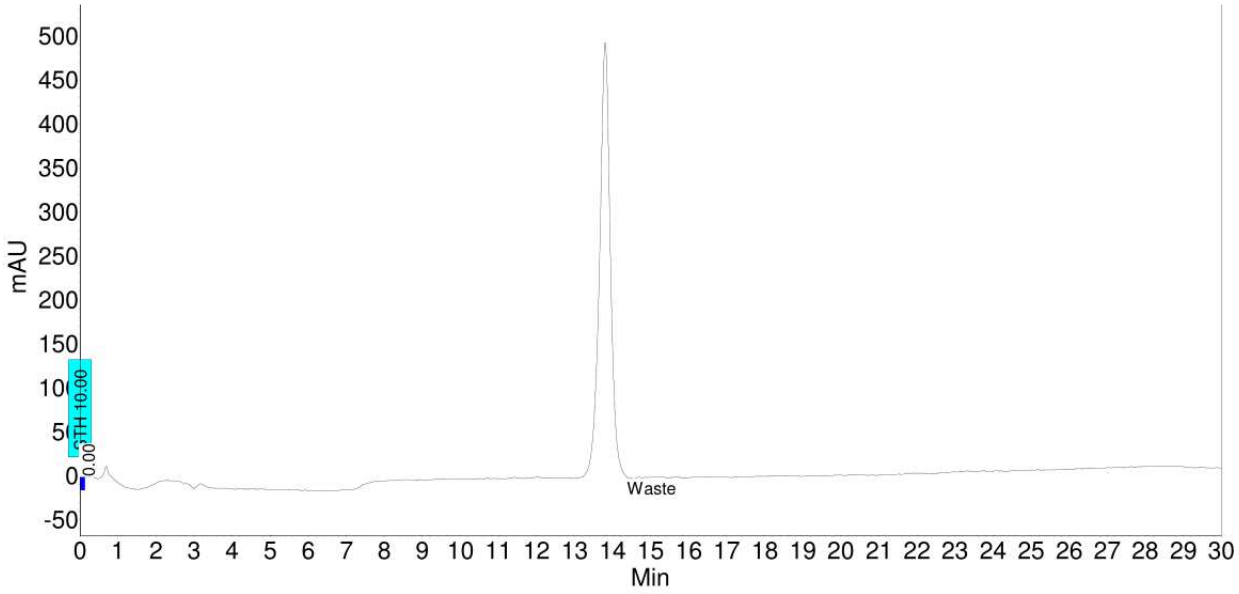
[9A] A mixture of compound **4B** (19.6 mg, 0.0489 mmol), 3-carboxyphenylboronic acid pinacol ester (8.03 mg, 0.0586 mmol), tetrakis(triphenylphosphine)palladium (1.7 mg, 1.47 µmol), and sodium carbonate (11.4 mg, 0.108 mmol) was dissolved in a 3:1 mixture of DME/water (1.2 mL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by reverse phase chromatography (HPLC). Calculated for ($C_{24}H_{23}N_5O_2$) ($M+H^+$): 414.2; found 414.5.



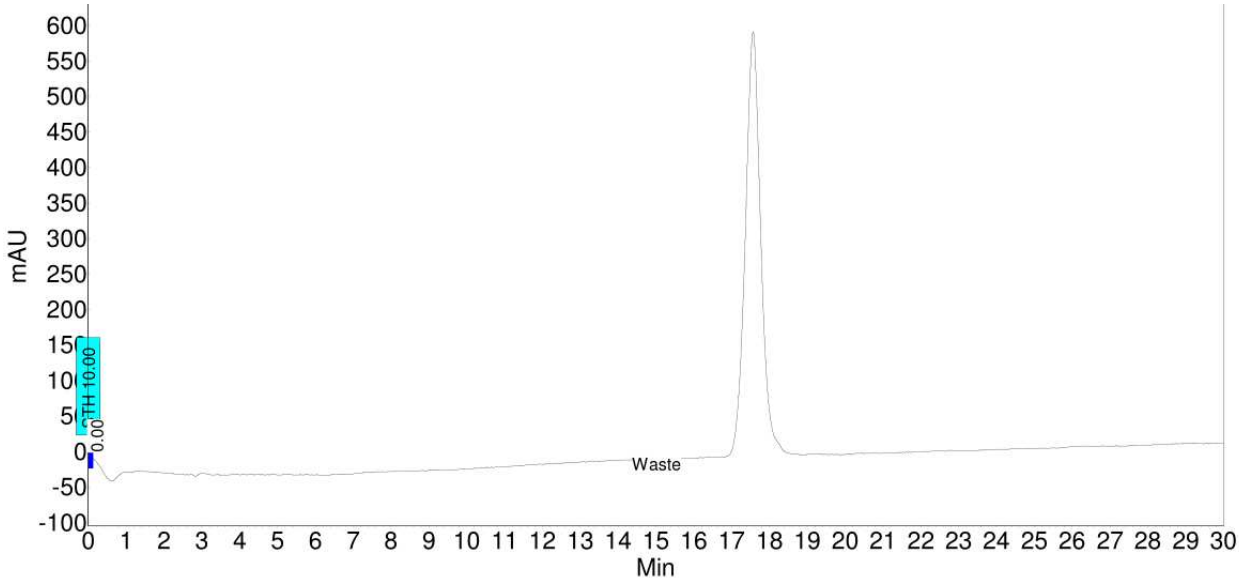
[9] Compound **9A** (6 mg, 0.0145 mmol) was dissolved in dichloromethane (100 µL) and cooled to 0 °C on ice. DIEA (7.6 µL, 0.0435 mmol) and acetic acid anhydride (1.7 µL, 0.0174 mmol) were added. The reaction was warmed to room temperature and stirred overnight. The reaction was purified by reverse phase column chromatography. ^{1}H NMR (300 MHz, Methanol-d4) δ

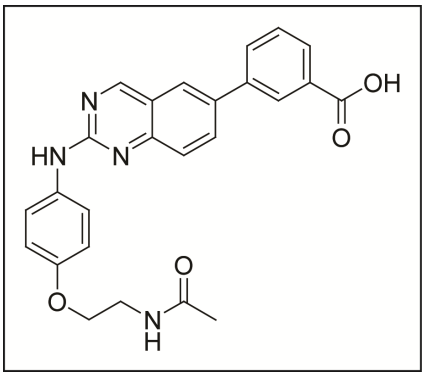
9.35 (s, 1H), 8.21 – 8.11 (m, 1H), 8.06 – 7.96 (m, 1H), 7.81 – 7.62 (m, 2H), 7.57 – 7.27 (m, 5H), 7.19 – 7.11 (m, 1H), 7.10 – 7.01 (m, 1H), 5.82 (s, 1H), 4.18 – 4.07 (m, 2H), 3.64 – 3.57 (m, 2H), 2.21 – 2.15 (m, 3H), 2.05 – 1.96 (m, 3H). Calculated for ($C_{26}H_{25}N_5O_3$) ($M+H^+$): 456.2; found 456.5

Analytical HPLC trace of **9**: Acteonitrile/Water

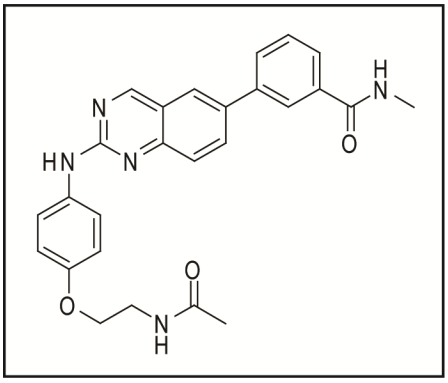


Analytical HPLC trace of **9**: Methanol/Water





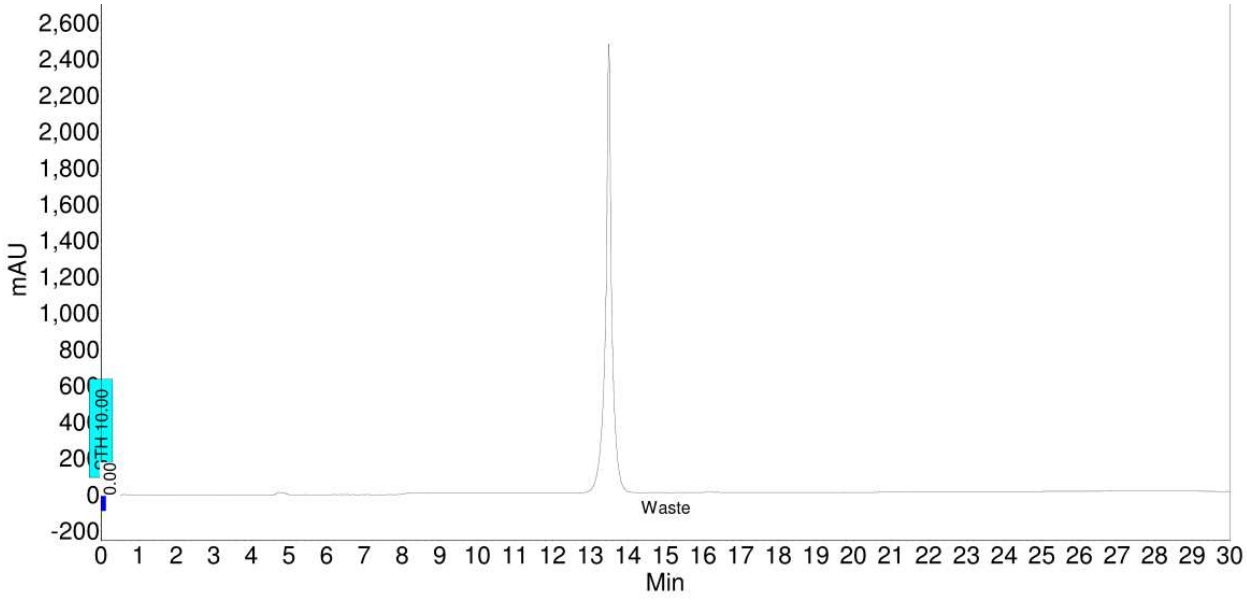
[10A] A mixture of compound **4B** (19.6 mg, 0.0489 mmol), 3-carboxyphenylboronic acid pinacol ester (14.5 mg, 0.0584 mmol), tetrakis(triphenylphosphine)palladium (1.7 mg, 1.47 µmol), and sodium carbonate (11.4 mg, 0.108 mmol) was dissolved in a 3:1 mixture of DME/water (1.2 mL). The mixture was heated for 3 h at 100 °C in the microwave. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by reverse phase column chromatography (HPLC). Calculated for ($C_{25}H_{22}N_4O_4$) ($M+H^+$): 443.2; found 443.5.



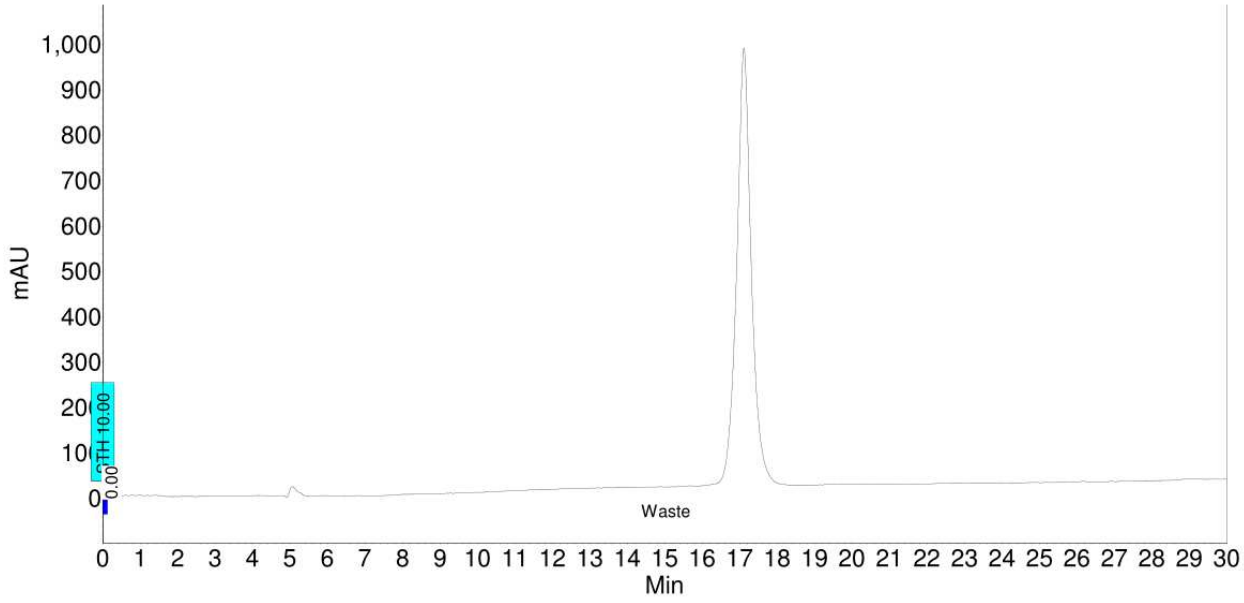
[10] Compound **10A** (8.4 mg, 0.0190 mmol), methylamine hydrochloride (1.54 mg, 0.0228 mmol), DIEA (9.9 µL, 0.570 mmol), and HOAt (3.1 mg, 0.0228 mmol) were dissolved in DMF (200 µL) and cooled to 0 °C on ice. EDCI (4.4 mg, 0.0229 mmol) was added and the reaction

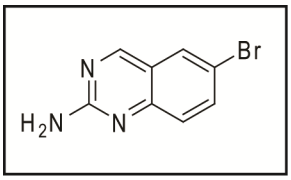
was warmed to room temperature and stirred overnight. The reaction was purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.49 (s, 1H), 8.04 – 8.26 (m, 1H), 8.22 – 8.04 (m, 1H), 8.01 – 7.71 (m, 4H), 7.66 – 7.47 (m, 1H), 7.37 – 7.26 (m, 2H), 7.14 – 7.06 (m, 2H), 5.80 (s, 1H), 4.09 (t, J = 5.5 Hz, 2H), 3.62 – 3.54 (m, 2H), 3.00 – 2.82 (m, 3H), 1.96 (s, 3H). Calculated for ($C_{26}H_{25}N_5O_3$) ($M+H^+$): 456.2; found 456.6

Analytical HPLC trace of **10**: Acetonitrile/Water

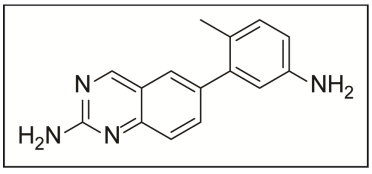


Analytical HPLC trace of **10**: Methanol/Water



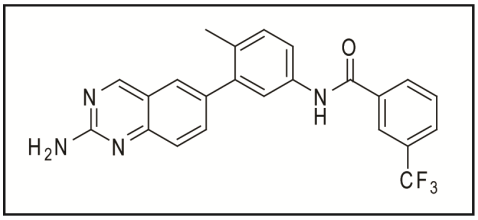


[11A] Synthesized as previously described.[38]



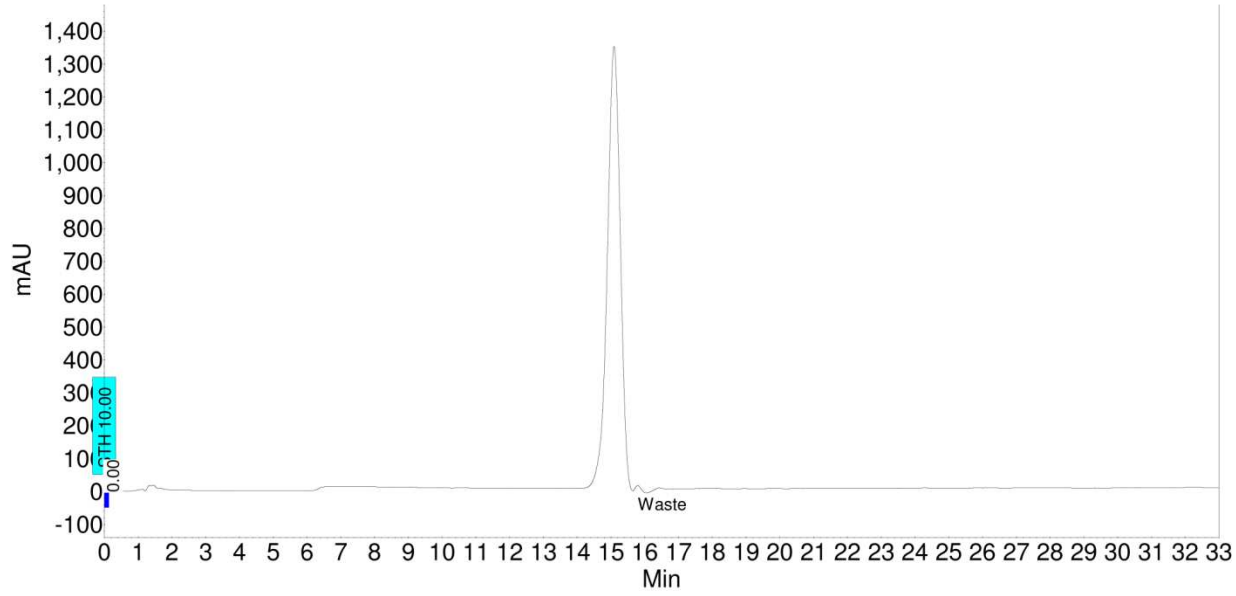
[11B] A mixture of compound **11A** (100 mg, 0.45 mmol), 5-amino-2-methylphenylboronic acid pinacol ester (118 mg, 0.49 mmol), tetrakis(triphenylphosphine)palladium (15.8 mg, 14 µmol), and sodium carbonate (104 mg, 0.98 mmol) was dissolved in a 3:1 mixture of DME/water (1.75 mL). The mixture was heated overnight at 85 °C. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was

dried over Na_2SO_4, concentrated *in vacuo* and the resultant crude product was purified by column chromatography to afford 60 mg of compound **11B** (54% yield). ^{1}H NMR (300 MHz, $CDCl_3$) 9.07 (s, 1H), 7.68 – 7.65 (m, 2H), 7.53 – 7.50 (m, 2H), 7.03 – 7.00 (m, 1 H), 6.70 – 6.65 (m, 1H), 2.18 (s, 3H). Calculated for ($C_{15}H_{14}N_4$) ($M+H^+$): 251.1; found 251.2.

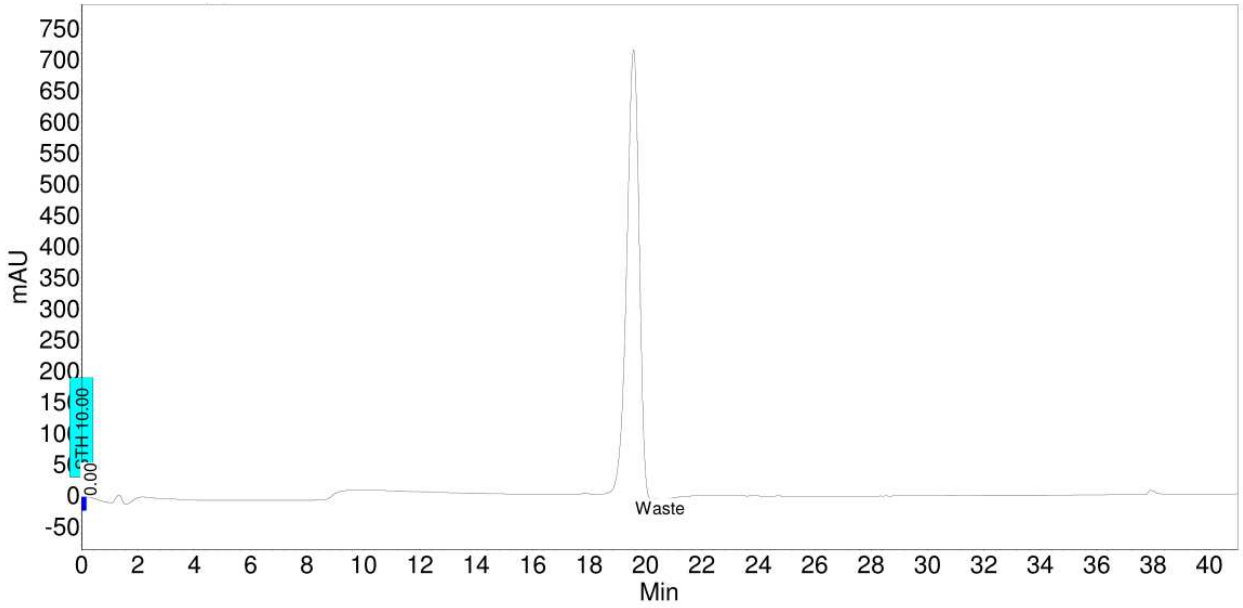


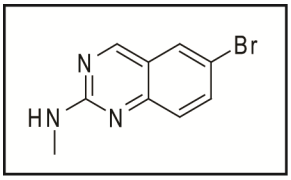
[11] Compound **11B** (59.9 mg, 0.239 mmol), 3-(trifluoromethyl)benzoic acid (60.4 mg, 0.311 mmol), HOBt (49.2 mg, 0.311 mmol), EDCI (60.9 mg, 0.311 mmol) and DIEA (125 µL, 0.718 mmol) were dissolved in DMF (600 µL) and stirred overnight at room temperature. The reaction was purified by reverse phase chromatography (HPLC) to obtain product **11**. ^{1}H NMR (300 MHz, $CDCl_3$) 9.61 – 9.45 (m, 1H) 8.33 – 8.11 (m, 2H), 7.97 – 7.52 (m, 4H) 7.42 – 7.39 (m, 2H), 7.37 – 7.15 (m, 2H), 5.83 (s, 1H), 2.24 (s, 3H). Calculated for ($C_{23}H_{17}F_3N_4O$) ($M+H^+$): 423.1; found 423.3.

Analytical HPLC trace of **11**: Acetonitrile/Water

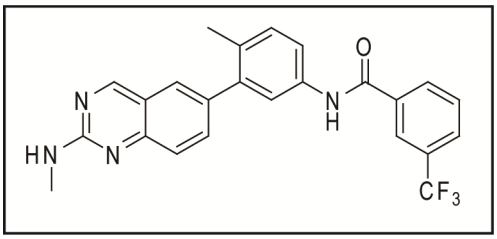


Analytical HPLC trace of **11**: Methanol/Water



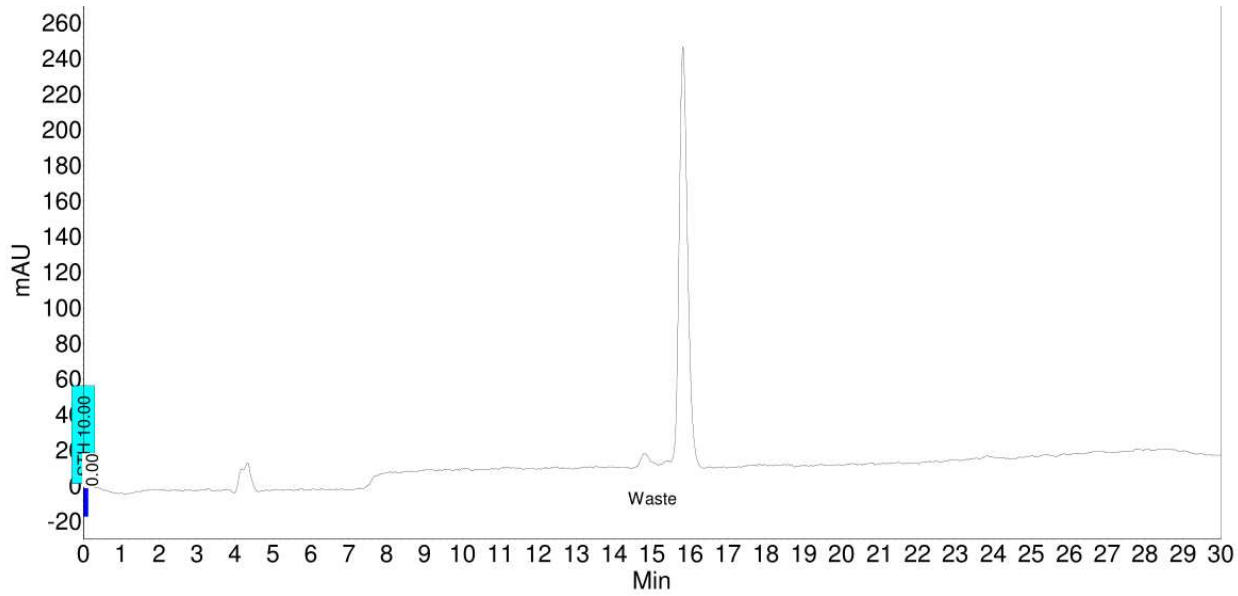


[12A] 6-Bromo-2-iodoquinazoline (100 mg, 0.299 mmol) was dissolved in isopropyl alcohol (1.36 mL). Methanol amine (40% in water) (0.23 mL, 2.99 mmol) was added and the microwave tube was sealed. The mixture was stirred for 6 h at 120^{o} C in the microwave. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo* and the resultant crude product was purified by column chromatography. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.92 (s, 1H), 7.88 – 7.71 (m, 2H), 7.57 – 7.26 (m, 1H), 3.25 – 3.08 (m, 3H). Calculated for ($C_9H_8BrN_3$) ($M+H^+$): 238.0; found 238.3.

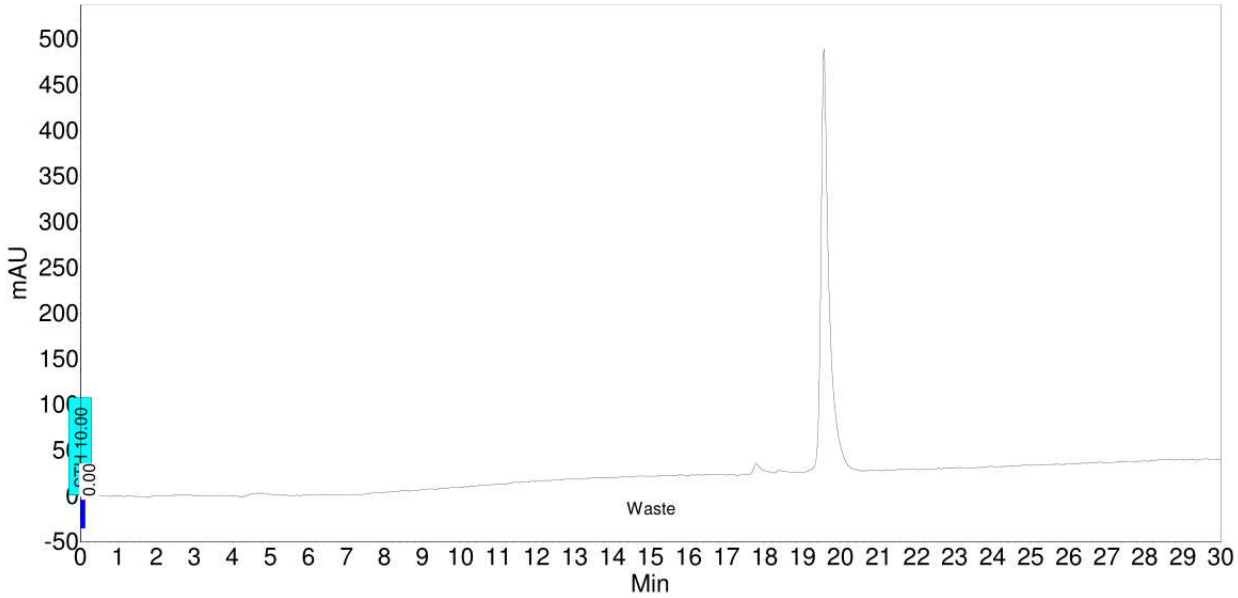


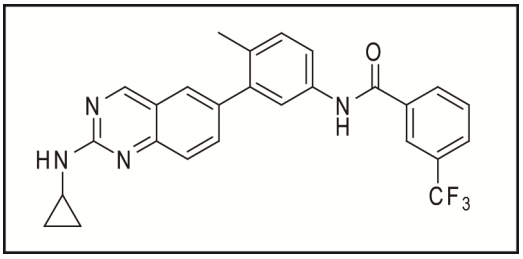
[12] A mixture of compound **12A** (15.0 mg, 0.063 mmol), **2B** (30.8 mg, 0.076 mmol), tetrakis(triphenylphosphine)palladium (2.2 mg, 1.9 µmol), and sodium carbonate (14.8 mg, 0.140 mmol) was dissolved in a 3:1 mixture of DME/water (0.7 mL). The mixture was stirred for 3 h at 100° C in the microwave. The crude product was purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.45 (s, 1H), 8.29 – 8.17 (m, 2H), 8.16 – 8.01 (m, 1H), 7.92 – 7.85 (m, 1H), 7.80 – 7.72 (m, 1H), 7.70 – 7.60 (m, 1H), 7.60 – 7.53 (m, 2H), 7.50 – 7.43 (m, 1H), 7.40 – 7.27 (m, 1H), 3.06 (s, 3H), 2.27 (s, 3H). Calculated for ($C_{24}H_{19}F_3N_4O$) ($M+H^+$): 437.2; found 437.5.

Analytical HPLC trace of **12**: Acetonitrile/Water

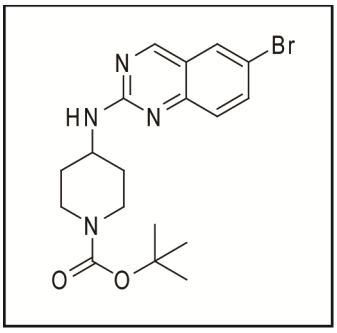


Analytical HPLC trace of **12**: Methanol/Water

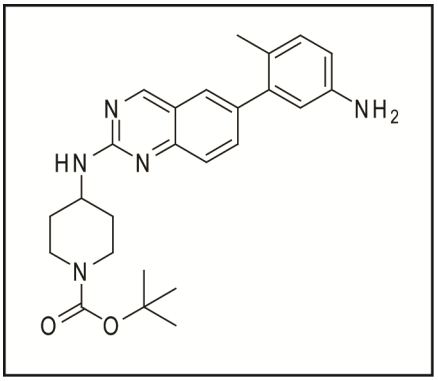




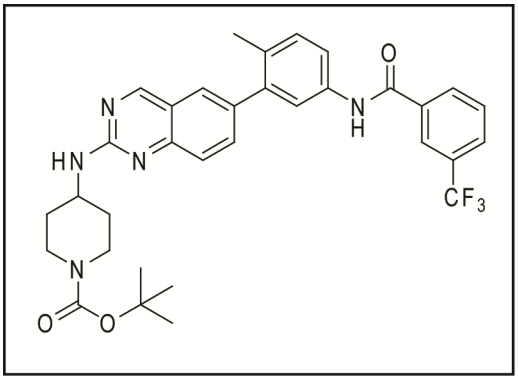
[13] Synthesized as previously described.[45]



[14A] In a resealable Pyrex tube, 6-Bromo-2-iodoquinazoline (50.0 mg, 0.15 mmol) and tertbutyl 4-aminopiperidine-1-carboxylate (153 mg, 0.75 mmol) were taken up in isopropyl alcohol (1.1 mL). DIEA (40 uL, 0.22 mmol) was added and the tube was sealed. The suspension was stirred for 2 h at 80 °C. The mixture was concentrated *in vacuo* and purified by column chromatography (50% ethyl acetate in hexanes) to afford 37.1 mg of compound 7 (61% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.90 (s, 1H), 7.81 (d, J = 2.2 Hz, 1H), 7.78 – 7.69 (m, 1H), 7.46 (d, J = 8.9 Hz, 1H), 5.29 (d, J = 7.7 Hz, 1H), 4.17 - 4.07 (m, 3H), 3.01 (t, J = 12.5 Hz, 2H), 2.10 (t, J = 12.0 Hz, 2H), 1.50- 1.43 (m, 12H). Calculated for ($C_{18}H_{23}BrN_4O_2$) ($M+H^+$): 407.1; found 407.3.

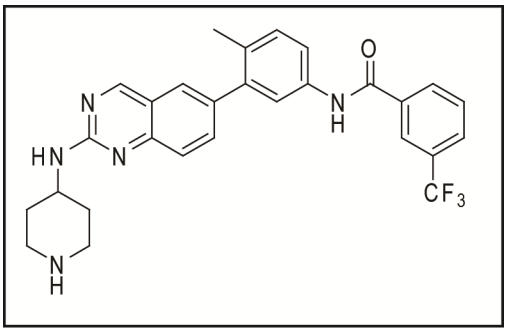


[14B] A mixture of compound **14A** (37.1 mg, 0.09 mmol), 5-Amino-2-methylphenylboronic acid pinacol ester (24.1 mg, 0.10 mmol), Tetrakis(triphenylphosphine)palladium (3.2 mg, 2.7 µmol) and sodium carbonate (21.2 mg, 0.20 mmol) was dissolved in a 3:1 mixture of DME/water (0.36 mL). The mixture was heated overnight at 85 °C. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo* to obtain **14B** (79% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.06 (s, 1H), 7.71 – 7.54 (m, 3H), 7.05 (d, J= 8.8 Hz, 1H), 6.88 – 6.63 (m, 2H), 3.07 – 2.98 (m, 2H), 3.39 (s, 1H), 2.16 (s, 3H), 2.07 – 2.04 (m, 4H), 1.93 (s, 1H), 1.51 – 1.45 (m, 12H). Calculated for ($C_{25}H_{31}N_5O_2$) ($M+H^+$): 434.3; found 434.4.



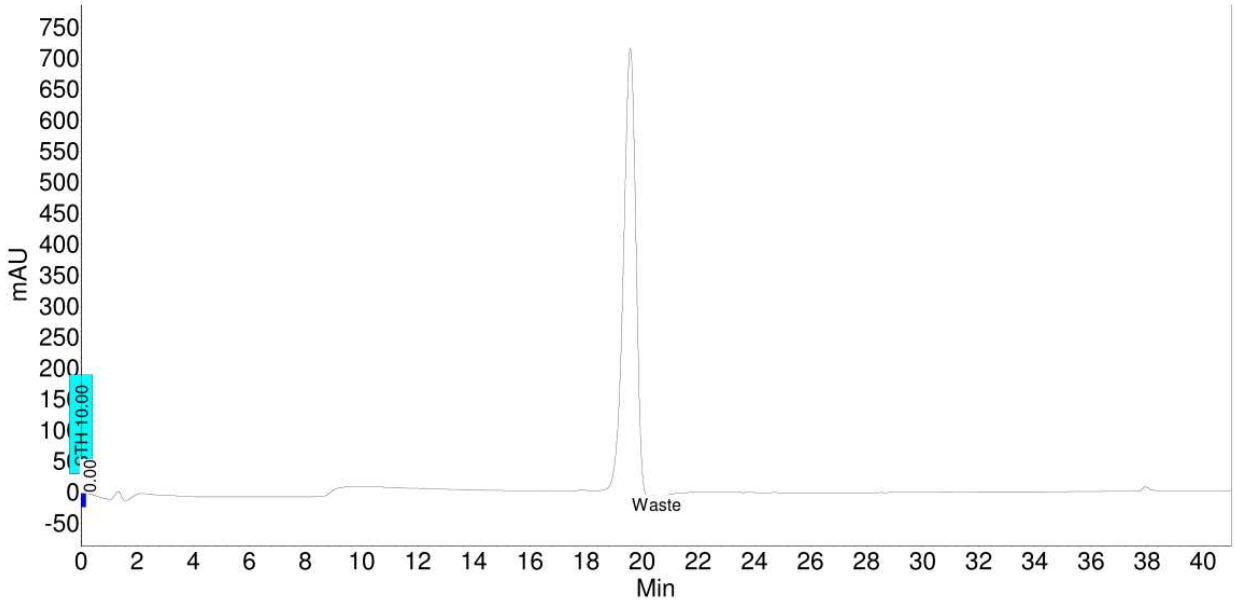
[14C] Compound **14B** (53.1 mg, 0.12 mmol), 3-(trifluoromethyl)benzoic acid (30.9 mg, 0.16 mmol), HOBt (25.2 mg, 0.16 mmol), EDCI (31.1 mg, 0.16 mmol) and DIEA (60 µL, 0.37 mmol)

were dissolved in DMF (360 µL) and stirred overnight at room temperature. The crude mixture was diluted in ethyl acetate and washed with NH_4Cl and Na_2CO_3. The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* to yield compound **14C**, which was used in the next step without further purification. Calculated for ($C_{33}H_{34}F_3N_5O_3$) ($M+H^+$): 606.3; found 606.3

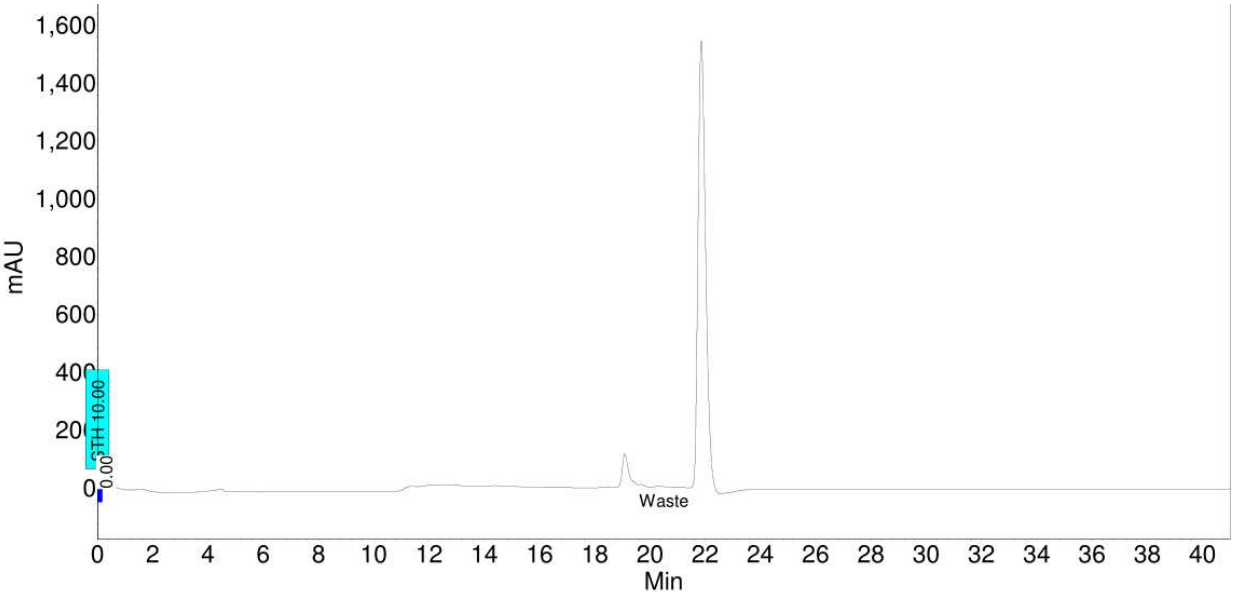


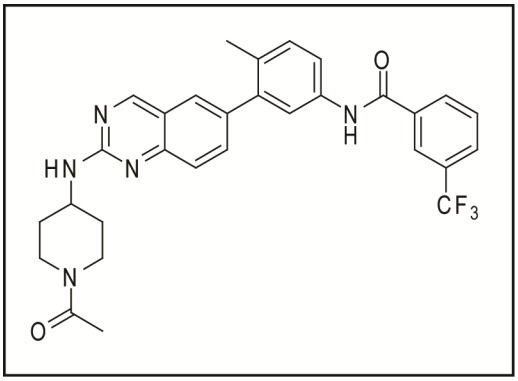
[14] Compound **14C** (73 mg, 0.12 mmol) was dissolved in CH_2Cl_2 (2.1 mL) and TFA (0.9 mL) and stirred for 3 h at room temperature. The reaction mixture was concentrated *in vacuo*, diluted in acetonitrile/water mixture and purified by reverse phase chromatography (HPLC) to obtain 25.0 mg of the desired product **14** (40% yield after 2 steps). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.32 (s, 1H), 8.28 (s, 1H), 8.22 (d, J = 9.0 Hz, 1H), 7.90 – 7.93 (m, 2H), 7.81 – 7.68 (m, 3H), 7.6- - 7.44 (m, 2H), 7.37 (d, J = 8.5 Hz, 1H), 4.45 - 4.36 (m, 1H), 3.59 – 3.52 (m, 2H), 3.29 – 3.17 (m, 2H), 2.41 – 2.34 (m, 2H), 2.31 – 2.28 (m, 3H), 2.00 – 1.86 (m, 2H). Calculated for ($C_{28}H_{26}F_3N_5O$) ($M+H^+$): 506.2; found 506.3.

Analytical HPLC trace of **14**: Acetonitrile/Water



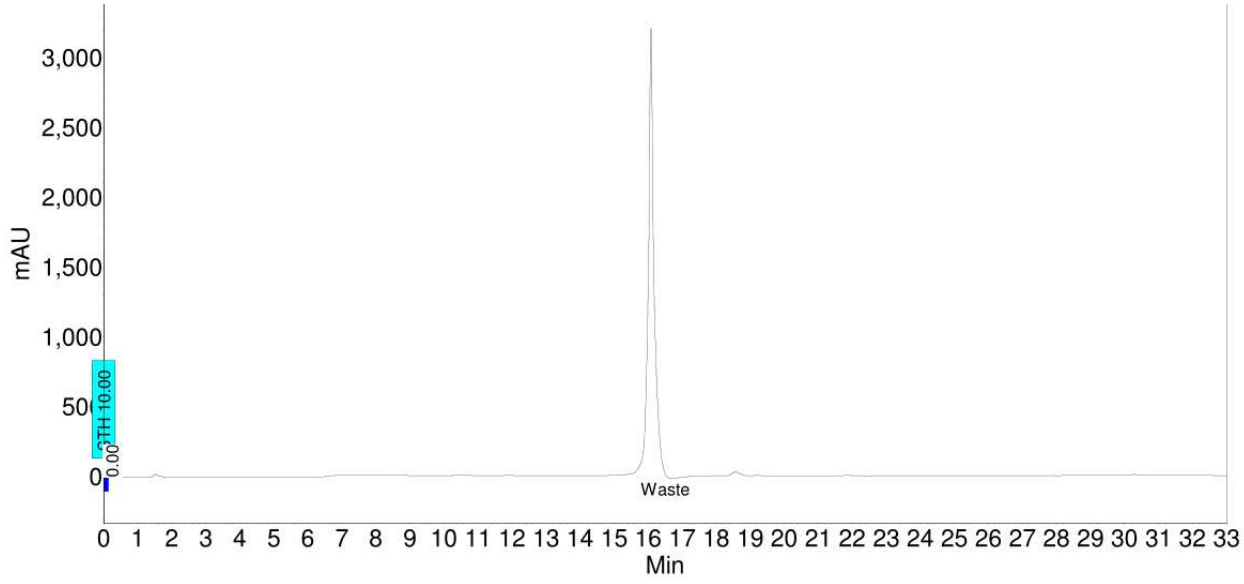
Analytical HPLC trace of **14**: Methanol/Water



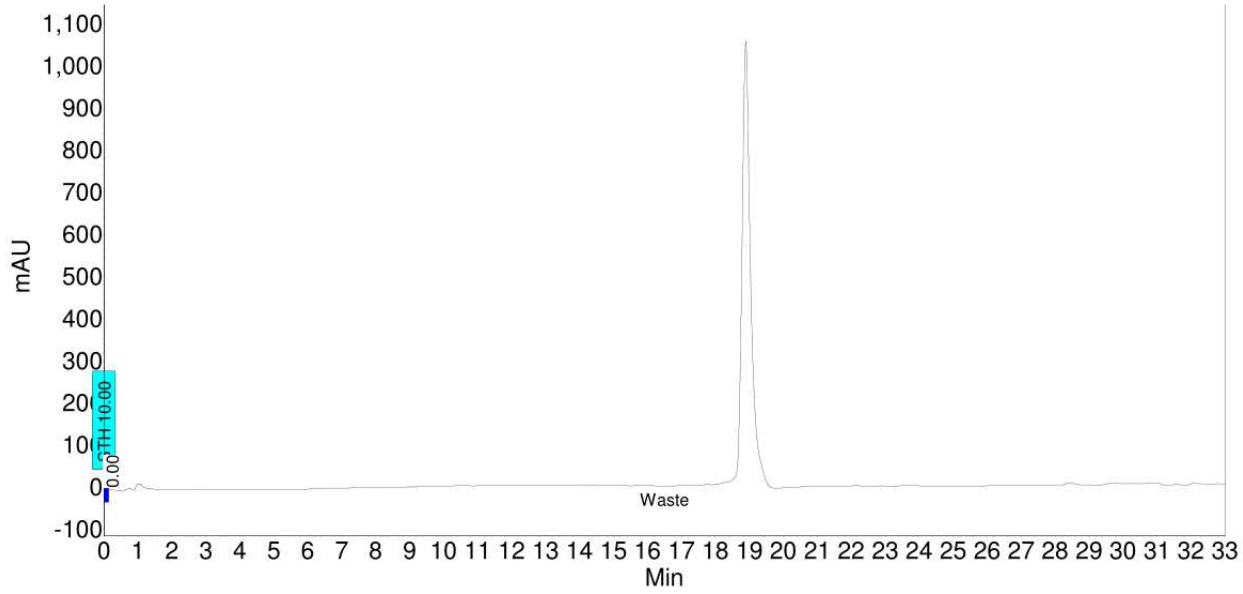


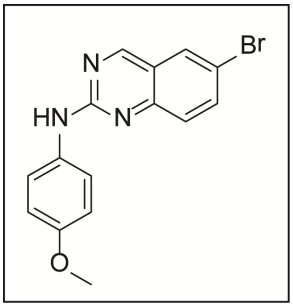
[15] Compound **14** (47.7 mg, 0.0945 mmol), acetic anhydride (1.89 mL), and triethylamine (50 µL, 0.358 mmol) were combined and stirred at room temperature for 1.5 h. Reaction was concentrated and purified by reverse phase column chromatography (HPLC) to yield compound **15**. ^{1}H-NMR (300 MHz, Methanol-d4) δ 9.50 (s, 1H), 8.30 – 8.19 (m, 2H), 8.14 – 8.04 (m, 1H), 7.91 (d, J = 8.1 Hz, 1H), 7.81 – 7.63 (m, 3H), 7.61 – 7.45 (m, 1H), 7.40 – 7.31 (m, 2H), 5.86 (s, 1H), 4.65 – 3.89 (m, 2H), 3.47 – 3.34 (m, 1H) 2.99 – 2.80 (m, 1H), 2.34 – 2.26 (m, 3H), 2.20 – 2.12 (m, 3H), 2.12 – 2.03 (m, 2H), 1.84 – 1.46 (m, 2H).Calculated for ($C_{30}H_{28}F_3N_5O_2$) ($M+H^+$): 548.2; found 548.2.

Analytical HPLC trace of **15**: Acetonitrile/Water

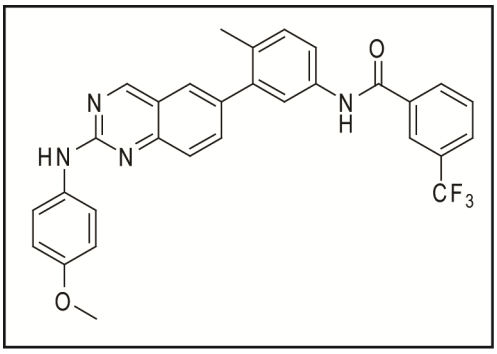


Analytical HPLC trace of **15**: Methanol/Water



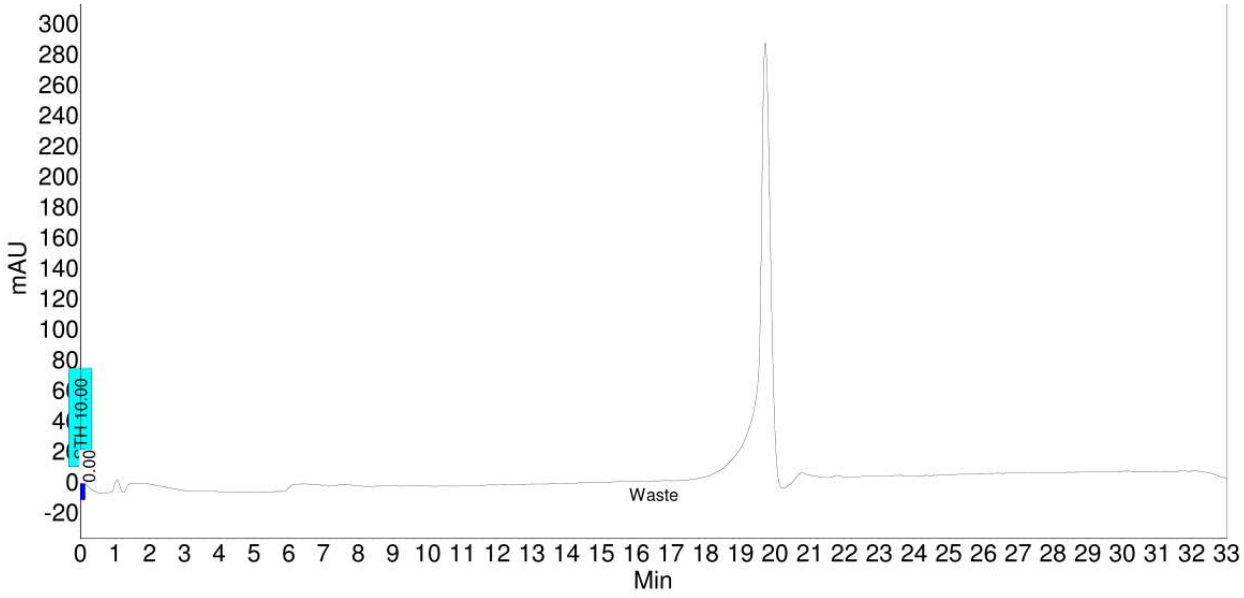


[16A] In a resealable Pyrex tube, 6-Bromo-2-iodoquinazoline (100 mg, 0.299 mmol) and *p*-Anisidine (45 mg, 0.365 mmol) were dissolved in isopropyl alcohol (2.3 mL). TFA (44.5 µL, 0.597 mmol) was added and the reaction was stirred overnight at 70 °C. Triethylamine was added to neutralize the reaction and the reaction was concentrated *in vacuo*. The resultant crude product was purified by column chromatography to yield **16A** (51% yield). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.09 (s, 1H), 8.04 – 8.00 (m, 1H), 7.77 – 7.70 (m, 1H), 7.77 – 7.70 (m, 2H), 7.57 – 7.51 (m, 1H), 6.99 – 6.89 (m, 2H), 3.82 (s, 3H). Calculated for ($C_{15}H_{12}BrN_3O$) ($M+H^+$): 330.0; found 330.2.

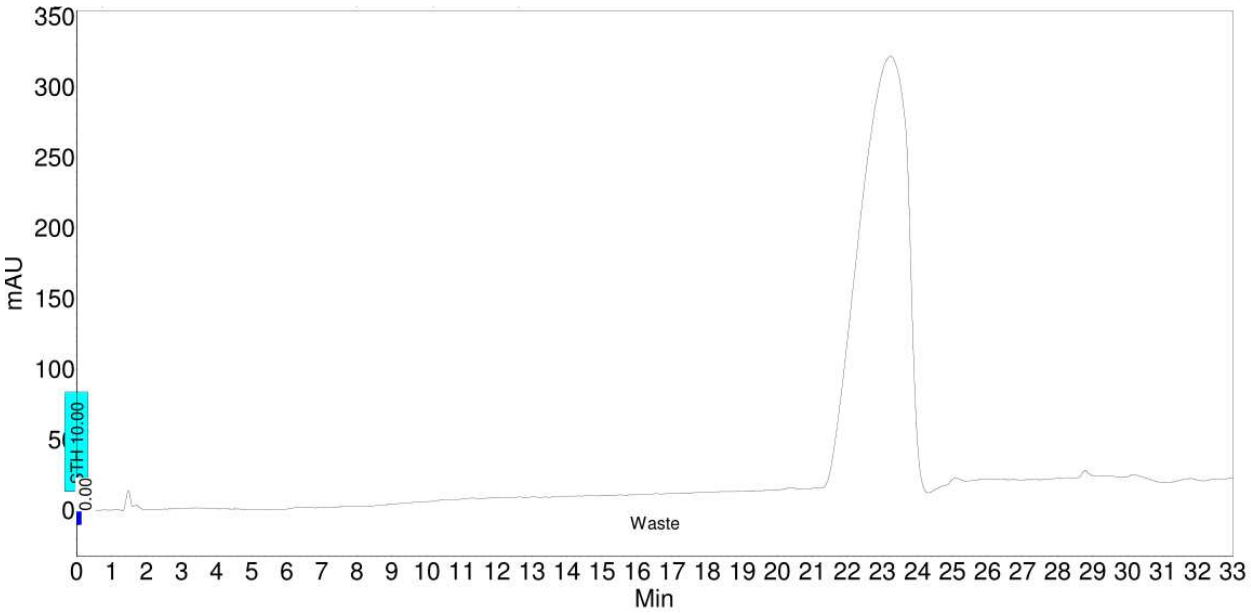


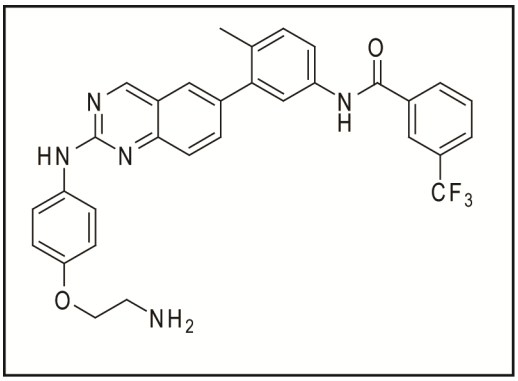
[16] A mixture of compound **16A** (27.5 mg, 0.083 mmol), **2B** (40.5 mg, 0.100 mmol), tetrakis(triphenylphosphine)palladium (3 mg, 2.6 µmol), and sodium carbonate (19.4 mg, 0.183 mmol) was dissolved in a 3:1 mixture of DME/water (0.33 mL). The reaction was stirred overnight at 85 °C. The crude product was purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.32 (s, 1H), 8.29 – 8.12 (m, 2H), 7.95 – 7.83 (m, 2H), 7.78 – 7.51 (m, 4H), 7.49 – 7.39 (m, 1H), 7.37 – 7.23 (m, 3H), 7.14 – 6.95 (m, 2H), 5.84 – 5.68 (s, 1H), 3.91 – 3.75 (s, 3H), 2.25 (s, 3H). Calculated for ($C_{30}H_{23}F_3N_4O_2$) ($M+H^+$): 529.2; found 529.5.

Analytical HPLC trace of **16**: Acetonitrile/Water

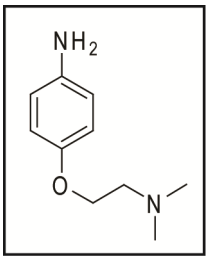


Analytical HPLC trace of **16**: Methanol/Water

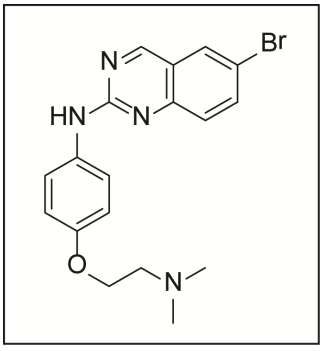




[17] Synthesized as previously described.[37]

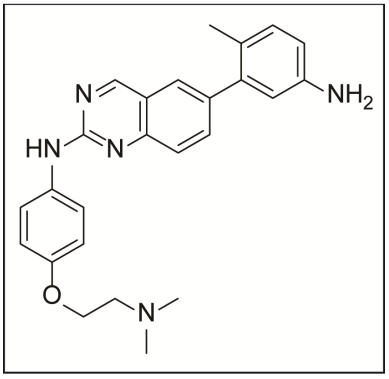


[18A] A mixture of 4-aminophenol (1.02 g, 9.35 mmol), 2-chloro-N,N-dimethylethylamine hydrochloride (1.63 g, 11.3 mmol), NaOH (0.945 g, 23.6 mmol) in DMF (16.1 mL) was stirred for 2 h at 75 °C under a nitrogen atmosphere. The mixture was then cooled to room temperature and filtered through a glass fritted funnel. The filtrate was diluted with dichloromethane and washed with brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and then purified by column chromatography to yield **18A**. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 6.81 – 6.70 (m, 2H), 6.69 – 6.55 (m, 2H), 3.98 (t, J = 5.8 Hz, 2H), 2.67 (t, J = 5.8 Hz, 2H), 2.31 (s, 6H). Calculated for ($C_{10}H_{16}N_2O$) ($M+H^+$): 181.1; found 181.1.

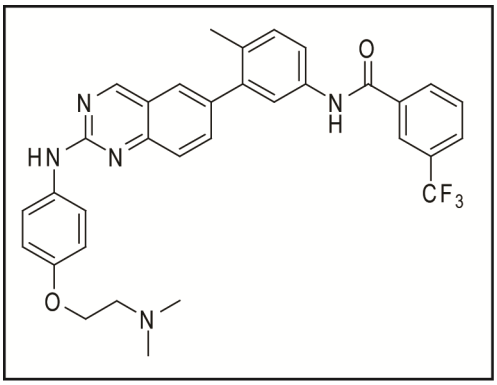


[18B] In a resealable Pyrex tube, 6-Bromo-2-iodoquinazoline (100 mg, 0.299 mmol) and **18A** (65 mg, 0.361 mmol) were dissolved in isopropyl alcohol (2.3 mL). TFA (50 µL, 0.673 mmol) was added and the reaction was stirred overnight at 70 °C. Triethylamine was added to neutralize the reaction and the reaction was concentrated *in vacuo*. The resultant crude product was purified by column chromatography to yield **18B**. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.82 (s, 1H), 7.69 (s,

1H), 7.59 (d, J = 7.7 Hz, 1H), 7.42 – 7.21 (m, 1H), 6.94 – 6.51 (m, 4H), 4.13 – 3.91 (m, 2H), 3.58 – 3.35 (m, 2H), 2.39 (s, 6H). Calculated for ($C_{18}H_{19}BrN_4O$) ($M+H^+$): 387.1; found 387.2.

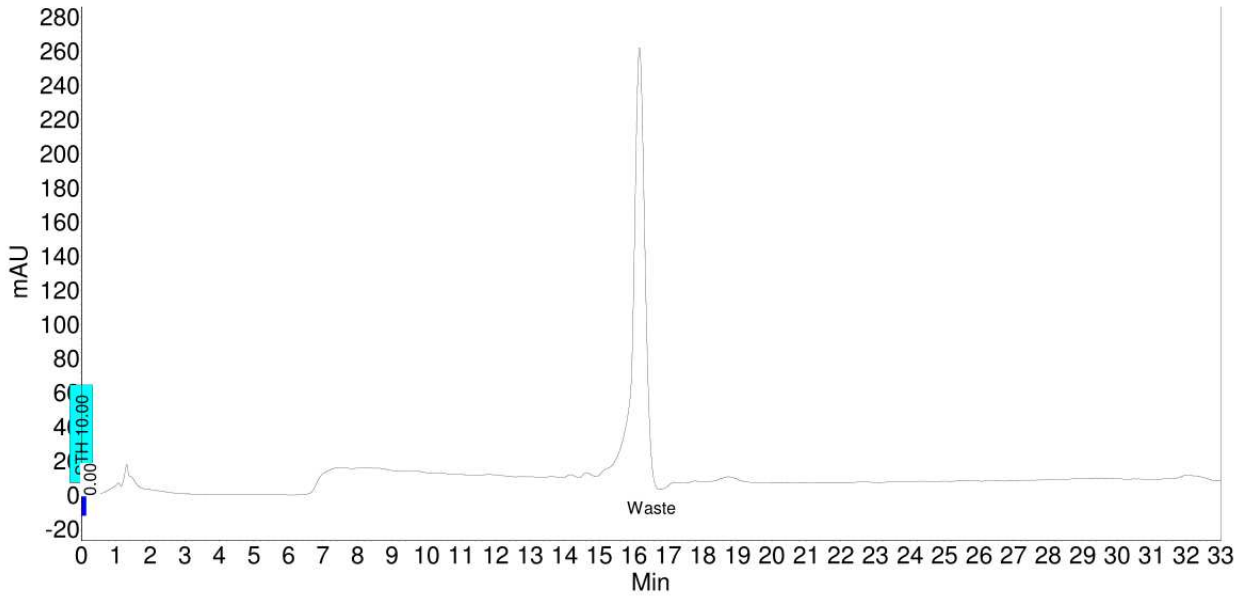


[18C] A mixture of compound **18B** (32 mg, 0.0826 mmol), 5-amino-2-methylphenylboronic acid pinacol ester (22 mg, 0.0944 mmol), tetrakis(triphenylphosphine)palladium (3 mg, 2.60 µmol), and sodium carbonate (19.5 mg, 0.184 mmol) was dissolved in a 3:1 mixture of DME/water (0.33 mL). The mixture was heated overnight at 85 °C. The crude mixture was cooled to room temperature, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and used in the next step without further purification. Calculated for ($C_{25}H_{27}N_5O$) ($M+H^+$): 414.2; found 414.5.

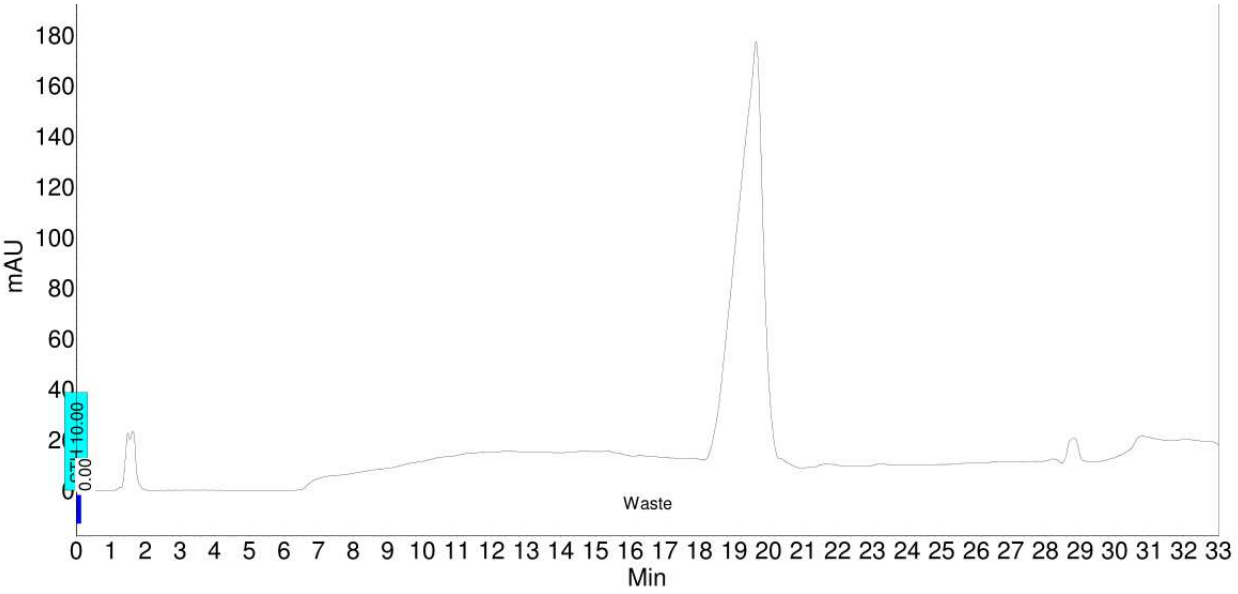


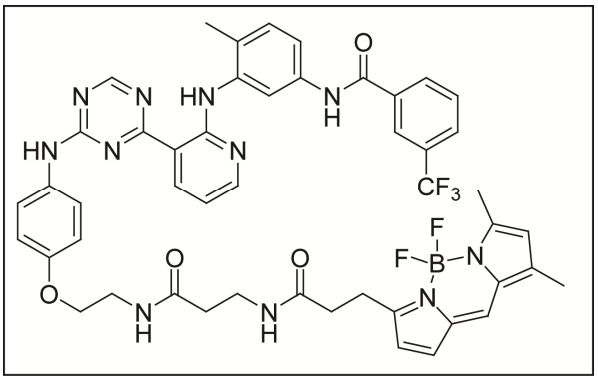
[18] Compound **18C** (55.1 mg, 0.133 mmol), 3-(trifluoromethyl)benzoic acid (21.1 mg, 0.111 mmol), HOBt (17.2 mg, 0.111 mmol), EDCI (21.3 mg, 0.111 mmol) and DIEA (40 µL, 0.230 mmol) were dissolved in DMF (240 µL) and stirred overnight at room temperature. The crude mixture was diluted in ethyl acetate and washed with NH_4Cl and Na_2CO_3. The organic layer was dried over Na_2SO_4, concentrated *in vacuo*, and purified by reverse phase column chromatography (HPLC). ^{1}H NMR (300 MHz, Methanol-d4) δ 9.22 (s, 1H), 8.28 – 8.15 (m, 2H), 7.93 – 7.78 (m, 4H), 7.75 – 7.67 (m, 2H), 7.63 – 7.43 (m, 2H), 7.41 – 7.23 (m, 2H), 7.19 (d, J = 8.9 Hz, 1H), 7.05 (d, J = 9.1 Hz, 1H), 5.77 (s, 1H), 4.45 – 4.31 (m, 2H), 3.67 – 3.53 (m, 2H), 2.99 (s, 6H), 2.33 – 2.22 (m, 3H). Calculated for ($C_{33}H_{30}F_3N_5O_2$) ($M+H^+$): 586.2; found 586.3.

Analytical HPLC trace of **18**: Acetonitrile/Water



Analytical HPLC trace of **18**: Methanol/Water





[DSA-BODIPY] Synthesized as previously described.[46]

B. *T. brucei* cell proliferation assay

T. brucei (bloodstream form strain 427 from K. Stuart, Seattle Biomedical Research Institute, Seattle, WA) was cultured in HMI-9 medium containing 10% fetal bovine serum, penicillin, and streptomycin at 37°C with 5% CO_2. Drug sensitivity of the T. brucei strain was determined in 96-well microtiter plates in triplicate with an initial inoculum of 1 x 10^4 trypomastigotes per well. Compound stock solutions were prepared in DMSO at 20mM and added in 7 serial dilutions for a final volume of 200 µL/well. Parasite growth was quantified at 48 h by the addition of Alamar Blue (Alamar Biosciences).[47]

C. *T. brucei* lysate preparation

0.8 liters of *T.brucei* strain 427 were collected mid-log phase ($2.5*10^6$ cells/mL ± $1*10^6$ cells/mL). Cells were pelleted and washed twice in Iscove's Modified Dulbecco's Medium (IMDM) media. The pellet was then resuspended in 1 mL of lysate buffer with protease inhibitor (20 mM HEPES [pH 7.5], 50 mM NaCl, 0.25% [v/v] NP-40, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 10% glycerol, and Complete protease inhibitor cocktail [Roche]), and sonicated. Sample

was centrifuged at 13k RPM at 4 °C in a tabletop centrifuge. The supernatant was removed and flash frozen in EtOH/CO_2(s) and stored at -80 °C.

D. ASH pulldown methods

i. ASH* and Ulp1* expression and purification

ASH* was expressed using a previously published protocol.[37]

ii. Enrichment of kinases from *T. brucei* lysate

ASH* (0.2 mg) was immobilized on CLP resin and then incubated with 200 µL of **1 Probe** (final concentration = 8.5 µM) or **3 Probe** (final concentration = 8.5 µM) in Buffer A (50 mM Tris [pH 7.5], 100 mM NaCl) for 1 h at room temperature. Excess small molecule was removed through a series of washes with Buffer A. *T. brucei* lysate was then incubated with the resin for 2 h at room temperature. After a series of washes with Buffer B (50 mM Tris [pH 8.0], 300 mM NaCl, 0.1% [v/v] Tween), the beads were incubated with Ulp1* (1:20 mass ratio Ulp1*:ASH*). Eluted samples were separated on 10% SDS-PAGE gels. Samples were then subjected to the mass spectrometry protocol described below.

iii. Mass spectrometry

a. In-gel trypsin digest of enriched proteins

Each lane of a 10% SDS-PAGE gel above 30 kDa was divided into three sections. Each section was excised from the gel and placed in 1.5 mL eppendorf tubes. Each gel slice was washed three times: first with 500 µL of 100 mM ammonium bicarbonate by rotating for 15 min at room temperature then with 500 µL acetonitrile (15 min, room temperature). Gel slices were then

speedvaced to dryness and rehydrated on ice for 45 min with 50 μL of a trypsin solution (20 μg dissolved in 1 mL 50 mM ammonium bicarbonate) (Sigma, Proteomics Grade, BioReagent, Dimethylated). 50 mM ammonium bicarbonate was then added to cover the expanded gel slice and then the gel slice was incubated overnight at 37 °C.

b. Mass spectrometry analysis

Samples were submitted to Mass Spectrometry Center at the University of Washington, School of Pharmacy. Data were generated with an LTQ Velos (Thermo Scientific).

E. Kinase affinity resin competition experiments

Kinase affinity resin competition experiments were performed by the Target Identifcation Proteomics (TIP) team at Novartis Insititute for Biomedical Research.

F. Genetic experiments

Kinase knockouts and conditional cell lines were generated using previously published protocols.[44]

G. Kinase expression and purification

i. Src Y527F

Src Y527F was expressed and purified from *E. coli* using a previously published protocol.[48]

ii. Tbtmp.46.0003

Tbtmp.46.0003 was amplified from *T. brucei* genomic DNA and cloned into the ligation independent cloning (LIC) site of expression vector AVA0421. Tbtmp.46.0003 was transformed into Rosetta2(DE3) *E. coli* cells (Novagen) and expressed using Studier auto-induction protocols

at 20 °C.[49] The kinase was purified by IMAC on a Ni^{2+}–NTA (Qiagen, Valencia, CA) column as previously described.[50] The binding buffer is composed of 20 mM HEPES (pH 7.25), 500 mM NaCl, 5% glycerol, 30 mM imidazole, 0.5% CHAPS, and 1 mM TCEP. Purified proteins were eluted with the same buffer supplemented with 250 mM imidazole.

Tbtmp.46.0003 Sequence:

MAHHHHHHMGTLEAQTQGPGSMKLRADSRGGGEQDIRASFGHSTPRDTGESPSIQPVE
GPIPLNFQYDPKVTSLEGALLFVGRTSSANVNSVNQKHLHGNSGVAIPNSPELVGGNQPL
ASARSSVADNSDQSSGKKKSAVGTGSIGGRHHDPNAPAAKIIVNETQGLNKPKSAAADI
VNETPHAEGTGVKSAFWAVAKNVSDRTPLSGQSKPPPRKNSNDGSPTPDHAGDEPIDVR
ACNFKFKIIRELGGGGFGKVYQAILGDGRFVAVKKMKITSHDKVIDREVRVLSTLPPHK
HCVRYLGSKRSSNHYYIFMEYVSGGSIRYIRKSAGVFEEPVMRRCVKMVLEGLQHIHRH
DIVHRDIKGENVLLDEKGCVKIVDFGACKVLNSGHNTVGSVGTPYWMSPEVCRGEAAT
EKSDVWGVGCLCLEMTNESGIPWEFHSTANNTQAVLYSIAAAKNPPKIPQHLSPAARDF
IACTLRVDPKDRPTVDKLLQHPFFSQ

H. Activity assays

In vitro activity assays for Src Y527F were performed using previously published protocols.[46]

I. DSA-BODIPY assays

i. Affinity measurements

Purified kinase (initial concentration = 2000 nM, 2-fold serial dilution down to 2 nM) was incubated with 20 nM **DSA-BODIPY** in 120 µl of buffer containing 20 mM Tris (pH 7.5), 2% DMSO, 100 mM KCl, 10% glycerol, 1 mM DTT, 0.05% Pluronic F-68, and 1 mM $MgCl_2$. Fluorescence was read after 1 h of incubation at room temperature (ex./em. 485/535 nm). The K_d was determined by fitting the data to non-linear regression analysis (one site total binding) with Prism GraphPad software.

ii. Competition assays

Competitor (initial concentration = 100 µM, 3-fold serial dilution down to 0.2 nM) was incubated with 20 nM **DSA-BODIPY** and either 50 nM or 100 nM purified kinase in 120 µl of buffer containing 20 mM Tris (pH 7.5), 2% DMSO, 100 mM KCl, 10% glycerol, 1 mM DTT, 0.05% Pluronic F-68, and 1 mM $MgCl_2$. Fluorescence was read after 1 h. of incubation at room temperature (excitation = 485 nm; emission = 535 nm). The IC_{50} was determined by fitting the data to non-linear regression analysis (one site – fit log IC50) with Prism GraphPad software.

V. REFERENCES

1 Simarro, P. P., Diarra, A., Ruiz Postigo, J. A., Franco, J. R. & Jannin, J. G. The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000-2009: the way forward. *PLoS Negl Trop Dis* **5**, e1007, doi:10.1371/journal.pntd.0001007 (2011).

2 *Human African trypanosomiasis: number of new cases drops to historically low level in 50 years*, <http://www.who.int/neglected_diseases/integrated_media/integrated_media_hat_june_2010/en/index.html> (2010).

3 Jannin, J. S., P.P.; Franco, J.R. in *The Causes and Impacts of Neglected Tropical and Zoonotic Diseases: Opportunities for Integrated Intervention Strategies* (National Academies Press http://www.ncbi.nlm.nih.gov/books/NBK62507/, 2011).

4 Sharma, R. *et al.* The heart of darkness: growth and form of Trypanosoma brucei in the tsetse fly. *Trends Parasitol* **25**, 517-524 (2009).

5 *Trypanosomiasis, African*, <http://www.dpd.cdc.gov/dpdx/HTML/TrypanosomiasisAfrican.htm> (2009).

6 Peacock, L. *et al.* Identification of the meiotic life cycle stage of Trypanosoma brucei in the tsetse fly. *Proc Natl Acad Sci U S A* **108**, 3671-3676, doi:10.1073/pnas.1019423108 (2011).

7 Matthews, K. R. The developmental cell biology of Trypanosoma brucei. *J Cell Sci* **118**, 283-290, doi:10.1242/jcs.01649 (2005).

8 Hamm, B., Schindler, A., Mecke, D. & Duszenko, M. Differentiation of Trypanosoma brucei bloodstream trypomastigotes from long slender to short stumpy-like forms in axenic culture. *Mol Biochem Parasitol* **40**, 13-22 (1990).

9 *Trypanosomiasis, Human African (sleeping sickness) Fact sheet N°259*, <http://www.who.int/mediacentre/factsheets/fs259/en/> (2013).

10 *WHO Human African Trypanosomiasis: Drugs*, <http://www.who.int/trypanosomiasis_african/drugs/en/index.html> (

11 Fox, K. R., Sansom, C. E. & Stevens, M. F. Footprinting studies on the sequence-selective binding of pentamidine to DNA. *FEBS Lett* **266**, 150-154 (1990).

12 Edwards, K. J., Jenkins, T. C. & Neidle, S. Crystal structure of a pentamidine-oligonucleotide complex: implications for DNA-binding properties. *Biochemistry* **31**, 7104-7109 (1992).

13 Bitonti, A. J., Dumont, J. A. & McCann, P. P. Characterization of Trypanosoma brucei brucei S-adenosyl-L-methionine decarboxylase and its inhibition by Berenil, pentamidine and methylglyoxal bis(guanylhydrazone). *Biochem J* **237**, 685-689 (1986).

14 Bacchi, C. J. Chemotherapy of Human African Trypanosomiasis. *Interdiscip Perspect Infect Dis* **2009**, doi:10.1155/2009/195040 (2009).

15 Wang, C. C. Molecular mechanisms and therapeutic approaches to the treatment of African trypanosomiasis. *Annu Rev Pharmacol Toxicol* **35**, 93-127, doi:10.1146/annurev.pa.35.040195.000521 (1995).

16 *WHO Improving access to the best treatment for second stage T. b. gambiense*, <http://www.who.int/neglected_diseases/disease_management/hat_access_treatment/en/index.html> (2009).

17 Keiser, J., Ericsson, O. & Burri, C. Investigations of the metabolites of the trypanocidal drug melarsoprol. *Clin Pharmacol Ther* **67**, 478-488, doi:10.1067/mcp.2000.105990 (2000).
18 Barrett, M. P., Boykin, D. W., Brun, R. & Tidwell, R. R. Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. *Br J Pharmacol* **152**, 1155-1171, doi:10.1038/sj.bjp.0707354 (2007).
19 Fairlamb, A. H., Henderson, G. B. & Cerami, A. Trypanothione is the primary target for arsenical drugs against African trypanosomes. *Proc Natl Acad Sci U S A* **86**, 2607-2611 (1989).
20 Cunningham, M. L., Zvelebil, M. J. & Fairlamb, A. H. Mechanism of inhibition of trypanothione reductase and glutathione reductase by trivalent organic arsenicals. *Eur J Biochem* **221**, 285-295 (1994).
21 Fairlamb, A. H. & Cerami, A. Metabolism and functions of trypanothione in the Kinetoplastida. *Annu Rev Microbiol* **46**, 695-729, doi:10.1146/annurev.mi.46.100192.003403 (1992).
22 Poulin, R., Lu, L., Ackermann, B., Bey, P. & Pegg, A. E. Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by alpha-difluoromethylornithine. Characterization of sequences at the inhibitor and coenzyme binding sites. *J Biol Chem* **267**, 150-158 (1992).
23 Delespaux, V. & de Koning, H. P. Drugs and drug resistance in African trypanosomiasis. *Drug Resist Updat* **10**, 30-50, doi:10.1016/j.drup.2007.02.004 (2007).
24 Gehrig, S. & Efferth, T. Development of drug resistance in Trypanosoma brucei rhodesiense and Trypanosoma brucei gambiense. Treatment of human African trypanosomiasis with natural products (Review). *Int J Mol Med* **22**, 411-419 (2008).
25 Maser, P., Luscher, A. & Kaminsky, R. Drug transport and drug resistance in African trypanosomes. *Drug Resist Updat* **6**, 281-290 (2003).
26 Matovu, E. *et al.* Mechanisms of arsenical and diamidine uptake and resistance in Trypanosoma brucei. *Eukaryot Cell* **2**, 1003-1008 (2003).
27 Shahi, S. K., Krauth-Siegel, R. L. & Clayton, C. E. Overexpression of the putative thiol conjugate transporter TbMRPA causes melarsoprol resistance in Trypanosoma brucei. *Mol Microbiol* **43**, 1129-1138 (2002).
28 Ouellette, M. Biochemical and molecular mechanisms of drug resistance in parasites. *Trop Med Int Health* **6**, 874-882 (2001).
29 Bellofatto, V., Fairlamb, A. H., Henderson, G. B. & Cross, G. A. Biochemical changes associated with alpha-difluoromethylornithine uptake and resistance in Trypanosoma brucei. *Mol Biochem Parasitol* **25**, 227-238 (1987).
30 Phillips, M. A. & Wang, C. C. A Trypanosoma brucei mutant resistant to alpha-difluoromethylornithine. *Mol Biochem Parasitol* **22**, 9-17 (1987).
31 Frearson, J. A. *et al.* N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature* **464**, 728-732, doi:10.1038/nature08893 (2010).
32 Jacobs, R. T. *et al.* SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. *PLoS Negl Trop Dis* **5**, e1151, doi:10.1371/journal.pntd.0001151 (2011).
33 Oduor, R. O. *et al.* Trypanosoma brucei glycogen synthase kinase-3, a target for anti-trypanosomal drug development: a public-private partnership to identify novel leads. *PLoS Negl Trop Dis* **5**, e1017, doi:10.1371/journal.pntd.0001017 (2011).

34 Nett, I. R. *et al.* The phosphoproteome of bloodstream form Trypanosoma brucei, causative agent of African sleeping sickness. *Mol Cell Proteomics* **8**, 1527-1538, doi:10.1074/mcp.M800556-MCP200 (2009).
35 Knight, Z. A., Lin, H. & Shokat, K. M. Targeting the cancer kinome through polypharmacology. *Nat Rev Cancer* **10**, 130-137, doi:10.1038/nrc2787 (2010).
36 Zimmermann, J., Buchdunger, E., Mett, H., Meyer, T. & Lydon, N. B. Potent and selective inhibitors of the Abl-kinase: phenylamino-pyrimidine (PAP) derivatives. *Bioorg Med Chem Lett* **7**, 187-192 (1997).
37 Brigham, J. L., Perera, B. G. K. & Maly, D. J. A Hexylchloride-Based Catch-and-Release System for Chemical Proteomic Applications. *ACS Chem Biol* **8**, 691-699, doi:10.1021/cb300623a (2013).
38 DiMauro, E. F. *et al.* Discovery of Aminoquinazolines as Potent, Orally Bioavailable Inhibitors of Lck: Synthesis, SAR, and in Vivo Anti-Inflammatory Activity. *J Med Chem* **49**, 5671-5686 (2006).
39 Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* **25**, 1035-1044 (2007).
40 Bantscheff, M. & Drewes, G. Chemoproteomic approaches to drug target identification and drug profiling. *Bioorg Med Chem* **20**, 1973-1978, doi:10.1016/j.bmc.2011.11.003 (2012).
41 Urbaniak, M. D. *et al.* Chemical proteomic analysis reveals the drugability of the kinome of Trypanosoma brucei. *ACS Chem Biol* **7**, 1858-1865, doi:10.1021/cb300326z (2012).
42 Hill, Z. B., Perera, B. G. & Maly, D. J. Bivalent inhibitors of the tyrosine kinases ABL and SRC: determinants of potency and selectivity. *Mol Biosyst* **7**, 447-456, doi:10.1039/c0mb00108b (2011).
43 Ranjitkar, P., Brock, A. M. & Maly, D. J. Affinity Reagents that Target a Specific Inactive Form of Protein Kinases. *Chem Biol* **17**, 195-206 (2010).
44 Merritt, C. & Stuart, K. Identification of essential and non-essential protein kinases by a fusion PCR method for efficient production of transgenic Trypanosoma brucei. *Mol Biochem Parasitol* **190**, 44-49, doi:10.1016/j.molbiopara.2013.05.002 (2013).
45 Hari, Sanjay B., Merritt, Ethan A. & Maly, Dustin J. Sequence Determinants of a Specific Inactive Protein Kinase Conformation. *Chem Biol* **20**, 806-815 (2013).
46 Ranjitkar, P. *et al.* Affinity-Based Probes Based on Type II Kinase Inhibitors. *J Am Chem Soc* **134**, 19017-19025 (2012).
47 Shibata, S. *et al.* Selective inhibitors of methionyl-tRNA synthetase have potent activity against Trypanosoma brucei Infection in Mice. *Antimicrob Agents Chemother* **55**, 1982-1989, doi:10.1128/AAC.01796-10 (2011).
48 Seeliger, M. A. *et al.* High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci* **14**, 3135-3139 (2005).
49 Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).
50 Ojo, K. K. *et al.* Structure determination of glycogen synthase kinase-3 from Leishmania major and comparative inhibitor structure-activity relationships with Trypanosoma brucei GSK-3. *Mol Biochem Parasitol* **176**, 98-108, doi:10.1016/j.molbiopara.2010.12.009 (2011).