UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors

Permalink

<https://escholarship.org/uc/item/75p4s454>

Author

Miller, Rand Micah

Publication Date

2014

Peer reviewed|Thesis/dissertation

Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors

by

Rand M. Miller

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

Copyright (2014)
by
Rand M. Miller

Acknowledgements

I must begin by thanking my thesis advisor, Jack Taunton, for providing endless advice, motivation and inspiration for this project. I could not have asked for a better balance between rigor, intensity, exciting results, hard work and ultimately reward. It is a commonly held belief is that the relationship between graduate student and advisor must be strained and contentious. I believe what some experience as “strained and contentious” is often better described as passion and investment, and there is no doubt that Jack is extremely passionate and invested in every student he advises. My interactions with Jack were never anything but productive, even in disagreement, and I leave the lab a far better scientist as a result. His commitment to improving my technique and to securing me a future career in science will be forever appreciated.

Scientifically, I must thank Ville Paavilainen, Iana Serafimova, Shyam Krishnan, Jeff Henise, Nir London, and Jesse McFarland for providing reagents and advice directly related to my project. I must also thank several eras of Taunton Lab members for their insightful conversation and ideas: Nathan Gushwa, Sarah Maifeld, Sumit Prakash, Julia Rumpf, Jonathan Choy, Geoff Smith, Jordan Carelli and Rebecca Maglathlin. I also thank my thesis committee members Kevan Shokat and Jim Wells for their direction and their time. Without input from each of these people, none of the work presented in this thesis would have been possible.

My family has been motivational and supportive in every way possible: emotionally, financially, and intellectually. My mom and dad have endured nearly

three decades of my nonsense, and through it all have been a never-ending flood of kindness, caring, and support. I cannot thank them enough, nor will I ever be able to repay them for the sacrifices they have made for my sisters and I. I must also thank them, Aubrey and Lyndelle, for being there for me and for providing me with a fun, creative, intellectually stimulating and relatively painless upbringing.

I owe special thanks to Justin and Jessica Fraga, who have been my closest friends since the day I started graduate school. The three of us have fought to succeed in each of our lives, and it has been incredibly motivational to watch the two of them do so even as I floundered through the hardest periods of my Ph. D. Without their guidance by example, and by direct advice at times, I would be nothing. Finally, I am continually humbled by the breadth and depth of the talent exuded by the rest of the people I call my friends. Thanks goes to anyone who helped me focus on science, distract me from science, and those who simply spent time with me when I needed a friend.

To Mom, Dad, Justin and Jessica.

Abstract

Protein kinases are an important class of enzymes that are ubiquitously involved in cellular signal transduction. The misregulation of protein kinases is implicated in many diseases, motivating the development of selective inhibitors for use as drugs or chemical probes. However, a large fraction of the roughly 500 human kinases remains untargeted by small molecule inhibitors. Fragment-based ligand design and covalent targeting of noncatalytic cysteines have been successfully employed to develop potent and selective kinase inhibitors. Here, we combine these approaches, starting with a panel of low-molecular weight, heteroaryl-substituted cyanoacrylamides, which we have previously shown to form reversible covalent bonds with cysteine thiols at physiological pH. Using this strategy, we identified electrophilic fragments with sufficient ligand efficiency and selectivity to serve as starting points for the first reported inhibitors of the MSK1 C-terminal kinase domain.

While these first generation cyanoacrylamides served as novel chemotypes for inhibiting a previously untargeted protein, they suffered from thermal and metabolic instability. Exploration of chemical space around the amide electron withdrawing substituent enabled the thermal and metabolic stabilization of these electrophiles. The most successful electrophiles retained a nitrile as one electron withdrawing group, while an electron-deficient triazole played the role of the second electron-withdrawing substituent. Further optimization of both the electrophile and the noncovalent scaffold enabled the

development of RMM-64 (**29**), a low-nM inhibitor of RSK and MSK kinases with prolonged target occupancy, high aqueous solubility, and oral bioavailability.

Finally, we adapted the computational digital screening program DOCK to enable the large scale, rapid virtual screening and ranking of cyanoacrylamide fragments against RSK and MSK kinases. With the help of collaborators, we constructed a library of more than 12,000 cyanoacrylamide-containing fragments based on commercially available aldehydes. Computational screening of this library against RSK2 and the RSK2 T493M mutant, with the constraint of a covalent bond to C436, enabled the identification of novel, potent and cell-active inhibitors of RSK2 and MSK1. We anticipate that the cyanoacrylamide fragment screening method described in this thesis, both experimental and computational, will be broadly applicable to any protein target containing a cysteine near a suitable binding site.

Table of Contents

List of Figures	x
List of Tables	xiv
Chapter 1. Introduction	1
1.1 MSK1 and MSK2 Kinases	2
1.2 References	10
Chapter 2. Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors	14
2.1 Abstract	15
2.2 Introduction	16
2.3 Results and Discussion	19
2.4 Methods	34
2.5 References	53
Chapter 3. Tuning and Stabilization of Doubly-Activated Olefins by Variation of the Non-Nitrile Electron Withdrawing Substituent	58
3.1 Abstract	59
3.2 Introduction	60
3.3 Results and Discussion	63
3.4 Methods	79
3.5 References	97
Chapter 4. Computational, Fragment-Based Design of Reversible Covalent, Cell-Active Cyanoacrylamide Inhibitors of RSK2 and MSK1	99
4.1 Abstract	100

4.2 Introduction	101
4.3 Results and Discussion	104
4.4 Methods	117
4.5 References	128
Chapter 5. Targeting Protein Kinases with Electrophilic Inhibitors	136
5.1 Abstract	137
5.2 Introduction	138
5.3 Irreversible, Covalent Inhibitors of RSK1, 2 and 4	141
C-terminal Kinase Domains	
5.4 Reversible Covalent Inhibitors	151
5.5 Using Promiscuous Covalent Inhibitors as Chemical Probes	155
5.6 Conclusion	160
5.7 References	163
Chapter 6. Conclusions and Future Directions	169
Appendix A: Reproduced NMR spectra of Synthesized Compounds	173

List of Figures

Figure 1-1 RSK and MSK kinases signal downstream of mitogen and stress activated kinase signaling cascades.	3
Figure 1-2 The model for activation of MSKs downstream of ERK or p38.	4
Figure 1-3 Chemical structures of the few frequently used inhibitors of MSK1, each targeting the AGC-family N-terminal kinase domain.	6
Figure 1-4 A rare, nonconserved cysteine residue (shown in red) is found in only 11 human kinases.	8
Figure 2-1 Reversible Covalent Targeting of Nonconserved Cysteines.	17
Figure 2-2 Cyanoacrylamide-based electrophilic fragments.	19
Figure 2-3 Intact cyanoacrylamide electrophile is required for potent inhibition of RSK.	22
Figure 2-4 Crystal Structures of fragments bound to RSK2 CTD inform inhibitor design.	23
Figure 2-5 Omit map confirms binding mode of compound 8 .	23
Figure 2-6 Glycine-rich loop dynamics may influence selectivity of cysteine-targeted inhibitors.	24
Figure 2-7 Compound 12 is selective for RSK2 and RSK2T493M over PLK1 and NEK2.	26
Figure 2-8 Crystal Structures of compound 12 bound to RSK2 confirm expected binding mode.	28

Figure 2-9 Binding of 12 to RSK2 T493M is rapidly reversible upon unfolding of the protein with guanidinium chloride.	29
Figure 2-10 Inhibition of MSK1 autophosphorylation in PMA stimulated COS7 cells.	30
Figure 2-11 MSK2, RSK2 and RSK3 are also inhibited by 12 in cellular assays.	31
Figure 2-12 Compound 12 inhibits signaling downstream of RSK and MSK in cells.	32
Figure 2-13 MSK1/2 CTD activity is required for CREB phosphorylation upon PMA stimulation.	32
Figure 3-1 Cyanoacrylamides 1 and 2 are unstable when incubated at 37 °C in PBS at pH 7.6.	63
Figure 3-2 Increasing steric bulk and substitution of the amide nitrogen of the cyanoacrylamide electrophile improves their thermal stability dramatically in buffered solution.	64
Figure 3-3 Facile substitution at the amide nitrogen with a variety of substituents enables the rapid generation of an SAR series on the same scaffold.	65
Figure 3-4 Varied amide nitrogen substitution fails to stabilize inhibitors to the cellular environment.	66
Figure 3-5 Synthesis of triazole-activated indazole 26 .	67
Figure 3-6 Compounds 12 and 26 potently inhibit the RSK2 CTD <i>in vitro</i> , but are much less potent against NEK2 and PLK1 despite the presence of a homologous cysteine.	68

Figure 3-7	69
Triazole 26 competes with BODIPY-FMK after 24 hours in cells, while cyanoacrylamide 12 is inactive after the same duration.	
Figure 3-8	70
Triazole 26 dissociates from RSK2 CTD upon denaturation with a $t_{1/2}$ of about 60 minutes.	
Figure 3-9	70
A co-crystal structure of triazole 23 bound to RSK2 CTD reveals the covalent and noncovalent interactions required for potent inhibition.	
Figure 3-10	72
Selectivity may arise because 26 is accommodated by L546 of RSK2 (blue), but is incompatible with a phenylalanine at this position in NEK2 and PLK1 (yellow) as well as R136 in PLK1 (red).	
Figure 3-11	74
Initial, linear synthesis of triazole-activated indazole 29 .	
Figure 3-12	75
Reorganized synthetic route to solubilized triazole-activated indazole 29 , with the installation of the highly reactive electrophile moved to the beginning of the synthesis.	
Figure 3-13	77
Compound 29 is equipotent to the precursor triazole 26 in spite of the bulky N-Me-piperazyl substituent.	
Figure 4-1	104
Experimental outline for covalent docking.	
Figure 4-2	106
Blind docking predictions of two cyanoacrylamide fragments from Chapter 2 covalently bound to RSK2.	
Figure 4-3	107
DOCKcovalent predictions accurately recapitulate the crystallographic poses of optimized cyanoacrylamide inhibitors of RSK2.	
Figure 4-4	108
Chemical structures of cyanoacrylamide fragments selected for synthesis and testing.	

Figure 4-5 Docking rank and <i>in vitro</i> IC ₅₀ values for cyanoacrylamides 32 – 39 against RSK2 WT and T493M mutant C-terminal kinase domain.	110
Figure 4-6 Docking predictions for cyanoacrylamide hits 34 and 37 .	111
Figure 4-7 Compounds 37 and 34 inhibit the activating autophosphorylation of RSK2 and MSK1 in stimulated mammalian cells.	112
Figure 5-1 Chemical structures of initial irreversible probes of RSK	143
Figure 5-2 Solubilized compound fmk-mea is selective for RSK CTDs..	145
Figure 5-3 Occupancy of RSK1 and RSK2 by fmk-mea assessed in the heart tissue of living animals using BODIPY-fmk .	146
Figure 5-4 Cyanoacrylamides and cyanoacrylates form reversible covalent bonds with thiols at physiological pH.	151
Figure 5-5 Cyanoacrylamide fragment screening enables the identification of selective, ligand efficient compounds.	154
Figure 5-6 The natural product hypothemycin may be adapted as a chemical probe.	156
Figure 5-7 An adenosine-like scaffold can target two different nucleophiles to interrogate inhibition of Src family kinases in cells.	158

List of Tables

Table 2-1 In vitro kinase assay IC ₅₀ values (μM) for compounds 1 – 10	20
Table 2-2 Kinase assay data for compounds 1 – 10 against C436V and T493M RSK2 CTD	21
Table 2-3 In vitro kinase assay IC ₅₀ values (nM) for 11 and 12	25
Table 2-4 In vitro kinase profiling data of 12	27
Table 2-5 Crystallographic Data Collection and Refinement Statistics for Compounds 1 , 8 and 12 bound to RSK CTD kinases	52
Table 3-1 Crystallographic Data Collection and Refinement Statistics for compound 26 to RSK2 T493M CTD kinase.	80
Table 4-1 Crystallographic Data Collection and Refinement Statistics for compound 37 bound to RSK2 T493M.	121

Chapter 1. Introduction

1.1 The MSK1 and MSK2 Kinases

Protein kinases are enzymes that catalyze the transfer of the γ -phosphate from ATP to the hydroxyl of serine, threonine or tyrosine residues on their protein or peptide substrates. While relatively conservative, this post-translational modification is capable of drastically altering protein function and is a ubiquitous mediator of intracellular signaling¹. Protein phosphorylation is often associated with cell growth, proliferation, and differentiation, and thus the misregulation of kinase signaling is implicated in many diseases. For this reason, protein kinases have remained a highly sought target class for small-molecule inhibition. Unfortunately, by virtue of their architectural similarity and sequence homology, targeting protein kinases selectively remains a major challenge.

The RSK and MSK kinases comprise a set of six closely related, dual kinase domain proteins. Each member of these families contains an N-terminal AGC-like kinase domain (NTD) and a C-terminal CAMK-like kinase domain (CTD), connected by a short, regulatory hydrophobic linker peptide.² While the RSK kinases (RSK1-4) are activated exclusively downstream of the Ras-Raf-MEK-ERK signaling sequence, the two MSK kinases (MSK1 and MSK2) can be activated by both p38 MAPK and ERK, thereby integrating both mitogen- and stress-activated signaling (Figure 1-1).³

Figure 1-1: RSK and MSK kinases signal downstream of mitogen and stress activated kinase signaling cascades. RSKs 1-4 are activated exclusively by ERK, whereas MSK1 and MSK2 are activated by both ERK and p38.

Propagation of signaling through each of these kinases follows a complex activation scheme, culminating in the full activation of the NTD and consequent phosphorylation of its downstream targets (Figure 1-2).⁴ Upon stimulation, ERK (in the case of RSK) or either ERK or p38 (in the case of MSK) binds to the C-terminal tail of the kinase and phosphorylate the C-terminal kinase domain on its activation loop (pT581 in human MSK1). This activation causes the CTD to autophosphorylates several residues on the linker peptide (pS381 and pS376) that ultimately lead to signaling through the NTD.⁵ Full RSK activation requires docking of PDK1 to the phosphorylated hydrophobic linker, followed by PDK1-mediated phosphorylation of the N-terminal kinase domain activation loop.⁶

Figure 1-2: The model for activation of MSKs downstream of ERK or p38. (1) Erk binds and activates the CTD by phosphorylating S581. (2) The CTD then phosphorylates two residues on the linker peptide (S376 and S381) as well as the activation loop of the NTD (S212). Note that PDK is required to bind and phosphorylate the NTD in RSK. (3) The NTD is allosterically activated and capable of phosphorylating downstream targets like CREB. Figure adapted from Iana M. Serafimova's Ph.D. thesis.

Unlike RSKs, MSK1 does not require any additional kinases like PDK1 for full activation of the NTD. Instead, the MSK CTD is capable of phosphorylating the activation loop of the NTD at S212 in addition to the two sites on the regulatory linker.⁵ The MSK CTD is the only known kinase capable of fully phosphorylating its hydrophobic linker and NTD, and is therefore required for full signaling capacity of the MSKs.

Unlike the RSK family, which has been studied extensively, the role of the closely related MSKs remains poorly understood. MSK1 and MSK2 have been shown to phosphorylate CREB and ATF1 in response to either mitogen- or stress-activation of ERK or p38.⁷ The activation of these CREB-family transcription factors results in the expression of immediate early genes like *fos*, *jun*, and *nur77*.⁸ MSK1 and 2 are localized to the nucleus, and have been shown to phosphorylate histones at H3S10 and H3S28.⁹ These histone marks are typically associated with an open chromatin state and are associated with transcriptionally active genes. Whether there exist other important substrates for the MSKs, and how the context of the known phosphorylation events effects the cellular output, remains to be seen.

There is limited evidence to support a role for MSK1 in the progression of several cancers. A single report has shown that MSK1 activity was upregulated in FLT3 ITD-driven AML cell lines, and that knockdown of MSK1 activity was sufficient to inhibit growth of these cell lines relative to non-FLT3 mutant cells.¹⁰ Additionally, the expression of many genes known to be regulated by MSK1 are increased in Ras-transformed murine fibroblasts.¹¹ Cells devoid of MSK1 activity were unable to exhibit anchorage-independent growth, a hallmark of the malignant tumor phenotype of these transformed cells. Finally, MSK1 activity is known to mediate the expression of miR-132,¹² a microRNA that has been shown to regulate VEGF-dependent tumor angiogenesis in mouse models.¹³ Thus, MSK1 has become an emerging target for cancer chemotherapy.

One of the best studied roles for MSK1 and MSK2 is the regulation of the innate immune system. MSK1/2 knockout mice, and the cell lines derived thereof, have been used extensively to establish a role for these kinases in regulating inflammation.¹⁴ Notably, these double knockout mice are healthy, develop normally, and exhibit no overt phenotype. Macrophages derived from MSK1/2 knockout mice demonstrate increased TNF, IL-6 and IL-12 levels and decreased IL-10 production relative to wild-type cells, indicating that MSK1 and 2 play an anti-inflammatory role in Toll-like Receptor (TLR) response.¹⁵ To date, a clear role for the MSKs in the regulation of the adaptive immune system has not emerged, though there is evidence to support a role in IL-17 production by T-cells.¹⁶ All the studies performed thus far have been carried out in MSK1/2 knockout cells or animals, and this genetic approach might confound the more nuanced roles of these kinases.

Selective small-molecule inhibition of the MSKs is a potential strategy for further investigation of the role of these kinases in cellular function and disease. To date, several inhibitors of the NTD of MSK1 have been reported in the literature (Figure 1-3).

Figure 1-3: Chemical structures of the few frequently used inhibitors of MSK1, each targeting the AGC-family N-terminal kinase domain.

The most widely employed inhibitor is H89, an isoquinoline sulfonamide originally reported as a PKA inhibitor.¹⁷ While this compound has been used extensively in the literature to block MSK1 activity, it inhibits a variety of AGC kinases like ROCK and S6K, and thus caution should be exercised when interpreting the results.¹⁸ Another frequently used PKC inhibitor, Ro-31-8220, exhibited even more promiscuous behavior than H89 when screened against a large fraction of the kinome¹⁹. More recently, GlaxoSmithKline has reported SB-747651A, a (1H-imidazo[4,5-c]pyridine-2-yl)-1,2,5-oxadiazolylamine that was originally reported as a selective MSK1 inhibitor.¹⁹ While this compound recapitulates many of the signaling phenotypes of MSK1/2 knockout cells, it still inhibits at least 13 kinases at 1 μ M, several of which are AGC-family kinases. To adequately study the intricate role of MSK1 and MSK2 in cells, a more selective inhibitor that does not affect upstream AGC kinases is necessary. To date, no inhibitors of the CAMK-like MSK1/2 C-terminal kinase domain have been reported, in spite of the potential for cleaner profile among AGC off-targets.

Covalent targeting of nucleophilic cysteine residues with electrophiles is a powerful strategy for designing selective inhibitors.^{20,21} A sequence alignment of all human kinases indicates the presence of a rare cysteine residue at the top of the ATP binding site, found in the CTD of each of the RSKs and MSKs (C458 in MSK1, Figure 1-4).²² This cysteine residue is present in only five other human protein kinases: MEKK1, NEK2, PLK1, PLK2, and PLK3.

Figure 1-4: A rare, nonconserved cysteine residue (shown in red) is found in only 11 human kinases. This residue could be exploited with electrophilic inhibitors for the design of MSK-selective molecules, but there are no obvious residues in the vicinity that would enable differentiation amongst those 11 kinases.

Our lab has successfully inhibited the CTDs of RSK1, 2, and 4 with compounds targeting this nonconserved cysteine and a small threonine gatekeeper.^{23,24} By virtue of these two “selectivity filters,” these pyrrolopyrimidine compounds inhibit only those 3 out of the 518 protein kinases in the human proteome, among the most selective inhibitors reported.

Covalently targeting the MSK1, MSK2 and RSK3 CTD kinases without inhibiting the other kinases that contain a homologous cysteine -- or any of the other 500 kinases for that matter -- presents a significant challenge. It is not immediately apparent based on crystal structures how one might design small molecules that can differentiate between these kinases in the absence of a small gatekeeper. In this thesis, we report a novel method to identify low molecular weight, covalent “fragments” capable of differentiating between highly related

kinases that contain a cysteine at the same position. We employ this method for the development of the first reported inhibitors of the CTD of MSK1 and MSK2, and demonstrate nearly 1000-fold selectivity for the MSKs and RSKs over other kinases with a cysteine at the same position. Using these molecules, we are able to prove that small molecule inhibition of the MSK1 CTD is sufficient to block disease-relevant signaling downstream of the NTD.

1.2 References

1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S.
Science, **2002**, 298, 1912 -- 1934.
2. Hauge, C.; Frodin, M. RSK and MSK in MAPK signalling. *Journal of Cell Science*, **2006**, 119, 3021 – 3023.
3. Deak, M.; Clifton, A. D.; Lucocq, J. M.; Alessi, D. R. *EMBO J.* **1998**, 17, 4426–4441.
4. Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J. & Cohen, P. *J Biol Chem* **1998**, 273, 1496-1505.
5. McCoy, C. E.; Campbell, D. G.; Deak, M.; Bloomberg, G. B.; Arthur, J. S. C. *Biochem. J.* **2005**, 387, 507–517.
6. Frodin, M., Jensen, C. J., Merienne, K. & Gammeltoft, S. *Embo J* **2000**, 19, 2924-2934
7. Wiggin, G. R.; Soloaga, A.; Foster, J. M.; Murray-Tait, V.; Cohen, P.; Arthur, J. S. C. *Mol. Cell. Biol.* **2002**, 22, 2871–2881.
8. Darragh, J.; Soloaga, A.; Beardmore, V.A.; Wingate, A. D.; Wiggin, G. R.; Pegg, M.; et al. *Biochem J* **2005**, 390, 749 - 759.

9. Dyson, M.H.; Thomson, S.; Inagaki, M.; Goto, H.; Arthur, S. J.; Nightingale, K.; Iborra, F. J.; Mahadevan, L. C. *J. Cell Sci.*, 2005, 118, 2247–2259.

10. Odgerel, T.; Kikuchi, J.; Wada, T.; Shimizu, R.; Kano, Y.; Furukawa, Y. *Leukemia* **2010**, 24, 1087–1090.

11. Pérez-Cadahía, B.; Drobic, B.; Espino, P. S.; He, S.; Mandal, S.; Healy, S.; Davie, J. R. *J. Biol. Chem.* **2010**, 286, 42–49.

12. Remeney, J.; Hunter, C. J.; Cole, C.; Ando, H.; Impey, S.; Monk, C. E.; Martin, K. J.; Barton, G. J.; Hutvagner, G.; Arthur, J. S. *Biochem J.* **2010**, 428, 281–291

13. Anand, S; Majeti, B. K.; Acevedo, L. M.; Murphy, E. A.; Mukthavaram, R.; Scheppke, L.; Huang, M.; Shields, D. J.; Lindquist, J. N.; Lapinski, P. E.; King, P. D.; Weis, S. M.; Cheresch, D. A. *Nat Med.* **2010**, 16, 909–914.

14. Wiggin, G.R. et al. *Mol. Cell. Biol.* **2002**, 22, 2871–2881

15. Ananieva O, Darragh J, Johansen C, Carr JM, McIlrath J, Park JM, Wingate A, Monk CE, Toth R, Santos SG, Iversen L, Arthur JS. *Nat Immunol.* **2008**, 9, 1028-1036.

16. Commodaro AG, Bombardieri CR, Peron JP, Saito KC, Guedes PM, Hamassaki D. E. et al. *Invest Ophthalmol Vis Sci* **2010**; *51*, 3567 - 3574
17. Chijiwa, T.; Mishima, A.; Hagiwara, M.; Sano, M.; Hayashi, K.; Inoue, T.; Naito, K.; Toshioka T.; Hidaka, H. *J. Biol. Chem.* **1990**, *265*, 5267-5272.
18. Lochner, A.; Moolman, J. A.; *Cardiovascular Drug Reviews* **2006**, *24*, 261-274.
19. Naqvi, S.; Macdonald, A.; McCoy, C. E.; Darragh, J.; Reith, A. D.; and Arthur, J. S. C. *Biochem. J.* **2012**, *441*, 347–357.
20. Potashman, M. H.; Duggan, M. E. *J. Med. Chem.* **2009**, *52*, 1231–1246.
21. Hagel, M.; Niu, D.; St Martin, T.; Sheets, M. P.; Qiao, L.; Bernard, H.; Karp, R. M.; Zhu, Z.; Labenski, M. T.; Chaturvedi, P.; Nacht, M.; Westlin, W. F.; Petter, R. C.; Singh, J. *Nat. Chem. Biol.* **2011**, *7*, 22–24.
22. Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28–39.

23. Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J.. *Science* **2005**, 308, 1318-1321.

24. Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. *Nat. Chem. Biol.* **2012**, 8, 471–476.

Chapter 2. Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors

2.1 Abstract

Fragment-based ligand design and covalent targeting of noncatalytic cysteines have been employed to develop potent and selective kinase inhibitors. Here, we combine these approaches, starting with a panel of low-molecular weight, heteroaryl-substituted cyanoacrylamides, which we have previously shown to form reversible covalent bonds with cysteine thiols. Using this strategy, we identified electrophilic fragments with sufficient ligand efficiency and selectivity to serve as starting points for the first reported inhibitors of the MSK1 C-terminal kinase domain. Guided by x-ray co-crystal structures, indazole fragment **1** was elaborated to afford **12** (RMM-46), a reversible covalent inhibitor that exhibits high ligand efficiency and selectivity for MSK/RSK-family kinases. At nanomolar concentrations, **12** blocked activation of cellular MSK and RSK, as well as downstream phosphorylation of the critical transcription factor, CREB.

Note: This chapter was co-written with Prof. Jack Taunton, and has been published. Miller, R.M.; Paavilainen, V. O.; Krishnan, S.; Serafimova, I. M.; Taunton, J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J Am Chem Soc*, **2013**, 135, 5298 - 5301.

2.2 Introduction

Fragment-based design is a powerful approach for developing ligands that modulate protein function, including protein kinase inhibitors.¹⁻³ This strategy relies on the identification of low-molecular weight compounds (100–300 Da) that form specific interactions with the target protein, often of low affinity ($K_d \sim 0.1$ –1 mM). Despite their low binding affinity, fragments often have higher ligand efficiency than typical high-throughput screening hits, which tend to be of higher molecular weight (>300 Da) and therefore have a higher probability of exhibiting steric clashes and other unfavorable interactions with the ligand binding site.⁴ Elaboration of the initial fragment hits guided by NMR or co-crystal structures affords a ligand that is ideally more potent and selective than the initial fragment, while retaining druglike physical properties.

Covalent bond formation between an electrophilic ligand and a poorly conserved, noncatalytic cysteine is another powerful strategy in drug discovery that has been exploited to enhance potency and selectivity.⁵⁻⁷ This is especially true for protein kinases, which are difficult to target selectively due to the high sequence and structural conservation of the active site. A large fraction of the 518 human protein kinases contain a solvent-exposed cysteine within or near the ATP binding site.^{8,9} We recently reported a series of compounds that inhibit the RSK2 C-terminal kinase domain (CTD) by forming a reversible covalent bond with a cysteine (C436) found in only eleven of the 518 human protein kinases (Figure 1-1A).¹⁰ This reversible covalent interaction is made possible by an α -cyanoacrylamide functionality, which forms β -thioether adducts that eliminate more rapidly than acrylamide-derived adducts due to the decreased pK_a of the α -

proton (Figure 1-1B). We sought to exploit this reversible cysteine-targeting chemistry in the context of fragment-based ligand design.

Figure 2-1: Reversible Covalent Targeting of Nonconserved Cysteines (A) Sequence alignment of 11 human kinases containing a cysteine at the same position as C436 of RSK2 (yellow). The gatekeeper position is highlighted in blue (threonine) or red (larger hydrophobic residues). (B) Michael adducts of thiols with acrylamides are kinetically stable, whereas cyanoacrylamides form rapidly reversible adducts with thiols.

MSK1 is a close relative of RSK2, possessing two kinase domains and a structurally homologous cysteine in its CTD. Despite this similarity, MSK1 is insensitive to our previously developed RSK inhibitors,¹⁰ most likely because it has a large methionine in the gatekeeper position (Figure 2-1A). Previous studies have suggested that the MSK1 CTD is essential for intramolecular phosphorylation and activation of the N-terminal kinase domain (NTD),¹¹ which subsequently phosphorylates transcription factors and histone H3.^{12,13} Although MSK1 has been implicated in various cancers,¹⁴⁻¹⁶ the few known inhibitors bind the NTD and show little discrimination among several other AGC-family kinases,

including S6K1, AKT1, PRK2, and ROCK2.¹⁷ No inhibitors of the MSK1 CTD have been reported to date.

In this communication, we describe an electrophilic fragment-based approach to ligand discovery. We have used this approach to develop the first reported inhibitor of the MSK1 CTD. This cyanoacrylamide-based inhibitor is active against closely related MSK/RSK-family kinases, but it is highly selective over NEK2 and PLK1, despite the presence of a homologous cysteine in these kinases.

2.3 Results and Discussion

Whereas our previous study started with a known RSK inhibitor,¹⁰ there was no obvious starting point for the current project. Moreover, it was not at all clear whether low-molecular weight electrophiles would be able to discriminate among different noncatalytic cysteines. To test the feasibility of an electrophilic fragment-based approach, we assembled a panel of ten aldehyde fragments (MW 96–250 Da), all with nitrogen-containing heterocycles that are commonly found in kinase inhibitors. Condensation with cyanoacetamide yielded the corresponding cyanoacrylamides **1-10** (Figure 2-2). We screened **1-10** against three human kinases, all of which contain a cysteine at the same position: RSK2 (C436), NEK2 (C22), and PLK1 (C67) (Table 2-1). To mimic intracellular redox conditions, glutathione (GSH, 10 mM) was included in all kinase assays, in addition to ATP (0.1 mM) and a peptide or protein substrate. At a molar excess of one million-fold over the kinase, glutathione also increased the stringency of our screen by acting as a competing nucleophile.

Figure 2-2: Cyanoacrylamide-based fragments used in this study.

	RSK2	NEK2	PLK1
1	0.36 ± 0.01	1.4 ± 0.2	0.57 ± 0.07
2	0.23 ± 0.01	3.4 ± 0.7	2.5 ± 0.4
3	38 ± 1	14 ± 0.7	37 ± 2
4	1.4 ± 0.1	96 ± 6	25 ± 2
5	21 ± 2	>150	> 300
6	6.0 ± 0.2	6.1 ± 0.5	45 ± 7
7	4.4 ± 0.6	1.1 ± 0.2	27 ± 0.4
8	0.12 ± 0.02	2.1 ± 0.4	> 10
9	8.2 ± 0.3	26 ± 0.2	< 0.62
10	68 ± 3	>100	53 ± 2

Table 2 -1: In vitro kinase assay IC₅₀ values (μM) for compounds **1** – **10**. Values reported as the mean ± range for duplicate measurements.

We observed significant potency differences among the ten cyanoacrylamides, with distinct patterns of inhibition noted for each kinase. Isomeric indazoles **1** and **2** and azaindole **8** inhibited RSK2 at submicromolar concentrations and high ligand efficiencies (LE = 0.54, 0.56, and 0.39, respectively; see ref. 18 for ligand efficiency calculation). By contrast, **2** and **8** were significantly less potent against NEK2 and PLK1. Azaindole **9** was selective for PLK1 over the other kinases in this screen, and purine **7** exhibited modest selectivity for NEK2. Mutation of the key cysteine in RSK2 (C436V) conferred >100-fold resistance to all of the cyanoacrylamides with submicromolar potency,

suggesting that covalent bond formation is required for potent inhibition (Table 2-2).

Entry	RSK2 C436V	RSK2 T493M
1	> 150	0.23
2	67	0.19
3	> 150	> 10
4	> 150	0.92
5	> 150	> 10
6	> 150	0.76
7	> 150	N/D
8	46	0.16
9	21	8.9
10	> 150	8.5

Table 2-2: Kinase assay data for compounds **1–10** against C436V and T493M RSK2 CTD. Kinase assays were run as described in the Supplemental Methods, and the values are the average of duplicate measurements.

Unlike cyanoacrylamide **1**, indazole itself failed to inhibit RSK2 at concentrations below 500 mM, and indazole-5-carboxaldehyde was weakly active against both wild-type and C436V RSK2 (IC₅₀ ~350 mM, Figure 2-3).

Figure 2-3: Intact cyanoacrylamide electrophile is required for potent inhibition of RSK. Neither indazole nor indazole-5-carboxaldehyde potently inhibits RSK2 WT or C436V.

Collectively, the structure-activity relationships in Table 2-1 suggest that the free energy of covalent bond formation is highly sensitive to structural differences in the fragments and the cysteine-containing kinase.

To gain structural insight into the extraordinary ligand efficiency of indazole **1**, we determined its co-crystal structure bound to RSK2 (PDB code 4JG6). The structure reveals the same kinase hinge-binding mode observed with other indazole inhibitors,^{19,20} in which the N1 proton donates a hydrogen bond to the carbonyl of L495 while N2 accepts a hydrogen bond from the amide of M496 (Figure 2-4A). Contiguous electron density was observed between C436 and the b-carbon of the cyanoacrylamide, consistent with covalent bond formation. Notably, indazole **1** does not extend into the hydrophobic pocket behind the gatekeeper T493 (Figure 2-4A). This observation suggested that, unlike our previously developed RSK2 inhibitors,¹⁰ **1** would be sterically compatible with a bulkier side chain in the gatekeeper position. Consistent with this hypothesis, indazole **1** inhibited the T493M gatekeeper mutant of RSK2 with an IC₅₀ of 230 nM.

Figure 2-4: Crystal Structures of fragments bound to RSK2 CTD inform inhibitor design. (A) Co-crystal structure of **1** bound to the RSK2 CTD. $2F_o - F_c$ map is contoured to 1σ . (B) Overlay of the structures of **1** (blue) and **8** (yellow) bound to RSK2, indicating the potential for aryl substitution at the 3-position of the indazole scaffold.

A co-crystal structure of azaindole **8** (PDB code: 4JG7) revealed that it binds RSK2 in a completely different manner from **1** (Figure 2-4B). The electron density for **8** was clear and unambiguous, allowing its accurate placement within the RSK2 active site (Figure 2-5).

Figure 2-5: Omit map confirms binding mode of compound **8**.

Unlike structurally related kinase inhibitors, the azaindole moiety of **8** does not engage the hinge. Rather, the ketone accepts a hydrogen bond from hinge residue M496, while the azaindole projects toward the entrance of the ATP pocket, sandwiched between the side chains of I428, M496 and L546. This interaction is accompanied by a conformational change in the glycine-rich loop, which contains I428 (Figure 2-6).

Figure 2-6: Glycine-rich loop dynamics may influence selectivity of cysteine-targeted inhibitors. The glycine-rich loop of RSK2 bound to **8** is shifted downward relative to apo-RSK2 (PDB code: 2QR8).

In both co-crystal structures, the nitrile and amide groups of the covalently bound inhibitors are solvent exposed and poorly resolved, precluding assignment of the newly formed α -stereocenter. Together, the structures suggest that potent RSK2 inhibition by cyanoacrylamide fragments **1** and **8** requires the simultaneous satisfaction of two geometric restraints: at least one hydrogen bond to the hinge and an unstrained covalent bond to C436.

An overlay of the two co-crystal structures suggested that aromatic substituents appended to the 3-position of indazole **1** would fill the hydrophobic cleft occupied by the azaindole moiety of **8** (Figure 3B). We therefore designed and synthesized the trimethoxyphenyl-substituted indazole **11** (Table 2-3). This provided a 20-fold increase in potency toward RSK2, but its selectivity over NEK2 and PLK1 was poor. Addition of a 1,1-dimethyl-2-hydroxyethyl substituent to the primary amide of **11** afforded our optimized inhibitor **12**, which retained potency against RSK2 and showed increased selectivity over NEK2 and PLK1 (Table 2-3, Figure 2-7). The IC₅₀ of **12** against T493M RSK2 was below 2.5 nM, the lower limit of our assay (LE >0.35).

	RSK2	RSK2 T493M	NEK2	PLK1
11	15 ± 2 ^b	3 ± 1	62 ± 8	37 ± 3
12	12 ± 3	< 2.5 ^c	530 ± 16	2200 ± 200

Table 2-3: In vitro kinase assay IC₅₀ values (nM) for **11** and **12**. Values reported as the mean ± SD (n ≥ 3) unless otherwise noted. All assays contain 10 mM GSH and 100 μM ATP. ^bValue is the mean ± range for duplicate measurements. ^cCalculated IC₅₀ is less than half the concentration of kinase present (5 nM).

Figure 2-7: Compound **12** is selective for RSK2 and RSK2T493M over PLK1 and NEK2. *In vitro* kinase assay dose response curves demonstrating the selectivity of **12** relative to NEK2 and PLK1 in spite of the presence of a homologous active site cysteine.

The selectivity of **12** was further defined by screening a commercial panel of 26 kinases (Table 2-4), twelve of which have a cysteine at various positions throughout the ATP-binding site (e.g., BTK, EGFR, JAK3, KDR). At a screening concentration of 1 μ M, NEK2 and PLK1 were the only kinases that were significantly inhibited (77% and 84% inhibition, respectively; [ATP] = 200 and 15 mM). We note that wild-type and T493M RSK2 were >40-fold more sensitive than NEK2 and PLK1 when tested under identical conditions in our laboratory (Table 2-3 and Figure 2-7, 0.1 mM ATP, 10 mM GSH). Kinases showing negligible inhibition by **12** include among others, MEK1, ERK1, and p38 MAPK, all of which are upstream regulators of RSK/MSK-family kinases (Table 2-4).

Kinase Tested	% Inhibition	Kinase Vendor	Kinase Vendor Part#	Enzyme Conc.	[ATP] Tested (Km app)	Incubation Time
	Mean			(nM)	(μ M)	(HR)
AKT3	-1	INVITROGEN	PV3185	0.78	100	3
AURORA-A	-9	AMPHORA	AMP-AurA	4.5	10	1
AURORA-B	10	AMPHORA	AMP-AUR-B	5	10	3
BTK	0	BPS	40405	2.5	16	3
CDK2_CYCLINA	0	UPSTATE	14-448	0.16	50	3
CRAF	3	BPS	40008	100	15	6
EGFR	5	BPS	40187	1	3	3
JAK1	-4	INVITROGEN	PV4774	5	70	17
JAK3	5	INVITROGEN	PV3855	2	2	3
KDR	-13	INVITROGEN	PV3660	0.5	80	1
LYNA	4	INVITROGEN	P2906	0.8	35	3
MAPK1	1	UPSTATE	14-173	1	1	3
MAPK14-P38-ALPHA	-2	AMPHORA	AMP-p38a	5	130	3
MEK1	3	UPSTATE	14-173	1	1	3
NEK2	77	UPSTATE	14-429	3.1	35	3
PLK1	84	AMPHORA	AMP-PLK	5	15	3
SRC	8	INVITROGEN	P3044	3	25	1.5
SYK	-9	BPS	40485	2	30	3
TEC	4	Millipore	14-801	4.6	12	3
TTK	64	INVITROGEN	PV3792	2.5	15	3
AKT1	1	UPSTATE	14-276	0.6	200	3
CAMK2A	1	INVITROGEN	PV3142	0.04	50	3
MSK1	0	UPSTATE	14-548M	0.08	30	3
PKA	1	SIGMA	C8482	0.03	5	3
PKC-THETA	0	INVITROGEN	P2996	0.04	130	3
RSK2	27	INVITROGEN	PV3323	0.08	20	3

Table 2-4: In vitro kinase profiling data of **12**. Assays were performed by Nanosyn (<http://www.nanosyn.com/>) at a single compound concentration of 1 μ M. The ATP concentration was set to the $K_{m,ATP}$ for each kinase. Note that this assay format does not support RSK or MSK CTDs.

A co-crystal structure of **12** bound to T493M RSK2 (PDB code: 4JG8) confirmed the binding mode anticipated by our structural analysis of **1** and **8** (Figure 2-8A,B).

Figure 2-8: Crystal Structures of compound **12** bound to RSK2 confirm expected binding mode. (A) and (B) Co-crystal structure of **12** bound to T493M RSK2.

Overlaying the co-crystal structures of **12** and **1** revealed that the indazole cores of both inhibitors were perfectly aligned, with the indazole of **12** packed against the methionine gatekeeper (M493). Similar to the azaindole of **8**, the trimethoxyphenyl substituent of **12** was nestled between the side chains of I428, M496 and L546. The 1,1-dimethyl-2-hydroxyethyl amide group of **12** extended beneath the glycine-rich loop (Figure 2-8A,B). Given the low sequence conservation and conformational heterogeneity of this loop,^{21,22} this interaction may account for the improved selectivity of **12** compared to **11** (Table 2-3). Finally, a covalent bond was observed between C436 and the electrophilic b-carbon of the cyanoacrylamide. Unfolding the **12**/RSK2 complex with guanidinium HCl resulted in the quantitative release of **12**, confirming that the

covalent bond is reversible and that its stability requires the intact folded structure of the kinase domain (Figure 2-9).

A

B

Figure 2-9: Binding of **12** to RSK2 T493M is rapidly reversible upon unfolding of the protein with guanidinium chloride. (A) HPLC traces of **12** treated with 3 M guanidinium chloride (top) and **12** bound to RSK2 T493M, followed by treatment with 3 M guanidinium chloride (bottom), indicating quantitative recovery of **12**. (B) Mass spectra obtained by integrating under the peaks in part A, confirming that **12** is recovered intact upon denaturation of the protein.

The ability of **12** to inhibit T493M RSK2 with more than 1,000-fold higher potency than our previous cyanoacrylamide inhibitors¹⁰ prompted us to test the related kinase, MSK1, which also has a methionine in the gatekeeper position (Figure 2-1). To date, it has not been possible to perform kinase assays with the isolated MSK1 CTD. We therefore asked whether **12** could prevent CTD-mediated autophosphorylation of full-length MSK1 in stimulated cells. Cells expressing HA-tagged MSK1 were treated with increasing concentrations of **12**, followed by stimulation with phorbol myristate acetate (PMA). As revealed by a phospho-specific antibody that recognizes pS376, **12** inhibited MSK1 CTD activity in the context of the full-length protein, with an IC₅₀ of ~100 nM (Figure 2-10).

Figure 2-10: Inhibition of MSK1 autophosphorylation in PMA stimulated COS7 cells.

In the same cell-based assay format, **12** was equipotent against the closely related kinases, MSK2, RSK2, and RSK3 (Figure 2-11).

Figure 2-11: MSK2, RSK2 and RSK3 are also inhibited by **12** in cellular assays. Each experiment was performed identically to that shown in Figure 2-10.

MSK1/2 are among several kinases that have been shown to regulate the cAMP-response element binding transcription factor, CREB, via phosphorylation of S133. Studies with knockout mice have demonstrated a role for MSK1/2 in CREB S133 phosphorylation in response to mitogens (e.g., PMA) and cellular stress (e.g, UV-C).¹³ However, until now, it has not been possible to test whether the CTD of endogenous MSK1/2 is required for CREB phosphorylation.

Importantly, **12** was found to be inactive toward several kinases previously shown to phosphorylate CREB, including AKT1, PKA, CAMK2, and PKCq (Table S2). Treatment of HeLa cells with **12** reduced CREB phosphorylation (IC₅₀ ~300 nM) under both PMA and UV-C stimulation conditions (Figure 2-12).

Figure 2-12: Inhibition of CREB phosphorylation in HeLa cells after stimulation with PMA or UV-C by compound **12**.

By contrast, a potent RSK inhibitor,¹⁰ which is inactive against MSK1/2, failed to block CREB phosphorylation under these conditions (Figure 2-13). These results indicate that the MSK1/2 CTD plays a critical role in promoting CREB phosphorylation by the NTD.

Figure 2-13: MSK1/2 CTD activity is required for CREB phosphorylation upon PMA stimulation. Compound **12** blocks CREB phosphorylation at S133, whereas the potent RSK1/2 inhibitor **CN-NHiPr** does not.

In this study, we have developed a new approach to fragment-based ligand design that exploits the reversible conjugate addition of thiols to cyanoacrylamides. This chemistry, when deployed against a poorly conserved

noncatalytic cysteine, provides enhanced selectivity and potency, while avoiding the formation of irreversible covalent adducts. Indeed, it is not clear that a fragment-based approach would be feasible with electrophiles that form irreversible covalent bonds. Inhibition/binding would occur under kinetic control and would be sensitive to the intrinsic nucleophilicity of a given cysteine.²³ Although intriguing results with acrylamide fragments have been reported,²⁴ it remains to be seen whether useful levels of kinetic discrimination can be achieved.

The structure-activity relationships observed with our collection of 10 cyanoacrylamides suggest that potency is not solely driven by the free energy of covalent bond formation. Rather, specific noncovalent interactions (hinge hydrogen bonds, van der Waals contacts, and steric complementarity) are required in concert with the covalent bond to cooperatively stabilize the complex. Given the geometric restraints imposed by an unstrained thioether bond, it may be possible to computationally predict the binding poses of fragment hits, potentially extending the utility of this approach to cases where experimental structures cannot be obtained. Relative to disulfide-based fragments,²⁵ the cyanoacrylamide electrophile is more compatible with the thiol-rich cellular environment and may therefore be retained during inhibitor optimization for use in biological studies. Finally, we anticipate that cyanoacrylamide fragments can be applied to cysteine-containing targets beyond kinases, enabling the discovery of new chemical probes.

2.4 Methods

LC-MS assay of compound reversibility upon protein denaturation

300 μ M RSK2 T493M was incubated with 250 μ M **12** in PBS in a total volume of 50 μ L. Subsequently, 50 μ L of 6M guanidinium chloride was added to denature the protein. After 1 min, 100 μ L of MeCN was added. The mixture was centrifuged and then sequentially passed through a 0.22 μ m filter and a 10 kD cutoff filter. An identical sample was prepared that lacked the kinase as a control for total compound recovery. Filtrates from both conditions were analyzed by LC-MS using a Waters UPLC-MS.

Protein Expression and Purification

WT, C436V, and T493M RSK2 CTD (human RSK2, 399-740) were expressed in *E. coli* and purified as described previously^{10,26}. An active mutant of full-length NEK2 (T125A) was expressed and purified as previously described^{27,28}.

In vitro Kinase Assays

WT, C436V, and T493M RSK2 CTD kinase assays were performed as previously reported¹⁰. Briefly, activated RSK kinase (5 nM) was incubated with varying concentrations of each inhibitor in buffered solution at pH 8.0 for 30 min in the presence of 100 μ M ATP and 10 mM GSH. Each reaction was initiated by the addition of 167 mM substrate peptide (RRQLFRGFSFVAK) and 0.3 μ Ci/ μ L γ -³²P-ATP in a final volume of 25 μ L for an additional 30 min. Reactions were

spotted on phosphocellulose sheets, washed once with 10% AcOH, twice with 0.1% H₃PO₄, and once with MeOH prior to drying.

NEK2 kinase assays were performed as previously reported²⁸. Briefly, NEK2 kinase (60 nM) was incubated with varying concentrations of each inhibitor in buffered solution at pH 7.6 for 30 minutes in the presence of 100 μM ATP and 10 mM GSH. Each reaction was initiated by the addition of 2.37 mg/mL β-casein (Sigma) and 0.4 μCi/μL γ-³²P-ATP in a final volume of 25 μL for an additional 30 minutes. Reactions were spotted on nitrocellulose sheets, washed once with 10% AcOH and twice with 0.1% H₃PO₄ prior to drying.

Active human PLK1 kinase (Millipore, catalog number 14-777M) (7.2 nM) was assayed as described above, except the substrate was 0.5 mg/mL dephosphorylated α-casein (Sigma).

Dried blots were exposed to a storage phosphor screen (Amersham Biosciences) for 30 min and scanned with a Typhoon imager (GE Life Sciences). Data were quantified using the SPOT program²⁹ and plotted as a fraction of DMSO controls using GraphPad Prism 4.0 software.

Protein Crystallography

His6-tagged RSK2 constructs (399 – 740, K591E) were expressed and purified as described previously¹⁰, then concentrated to 20 mg/mL (WT) and 27 mg/mL (T493M) respectively. Compounds **1**, **8**, and **13** were allowed to bind to each protein by mixing 19 μL of either WT or mutant protein with 1 μL of a 10 mM compound stock in DMSO. After incubation on ice for 30 min, these solutions

were filtered through 0.22 μm spin filters. Crystals were grown in hanging drops by mixing 1 μL of freshly filtered protein/compound solution with 1 μL of precipitant (0.1 M HEPES pH 7.0, 7.5% PEG3350, 50 mM ammonium sulfate). Crystals typically grew to their maximum dimensions within 5 days, at which point they were looped, transferred to cryoprotectant (LV Cryo Oil, Mitegen LLC, Ithaca, NY), and immediately frozen in a stream of liquid nitrogen. Data for RSK2 WT and RSK2 T493M in complex with compounds **8** and **12**, respectively, were collected in the UCSF Macromolecular Structure Group Crystallography Facility using a Rigaku RU-200 rotating anode system, fitted with Varimax-HR optics and an Raxis-IV image plate detector using Structure Studio control software. Data for compound **1** in complex with RSK2 WT were collected on the HHMI beamline 8.2.1 at the Advanced Light Source at Lawrence Berkeley National Laboratories. Diffraction data were processed using XDS³⁰ and the structures solved by molecular replacement using the Phaser module in Phenix³¹. Models were iteratively built and refined using Coot³² and Phenix.refine, respectively. Due to high resolution incompleteness observed as a consequence of an extremely long unit cell dimension, refinement for each dataset was limited to a lower resolution than collected. The data collection and refinement statistics are summarized in Table 2-4.

Cell Culture

COS-7 and HeLa cells (ATCC) were maintained in DMEM (Gibco) supplemented with 10% FBS (Axenia Biologix), 100 units/mL penicillin, and 100

μg/mL streptomycin (Gibco). COS-7 cells were maintained at 37 °C in a 10% CO₂ atmosphere, while HeLa cells were maintained at 37 °C in a 5% CO₂ atmosphere.

Stimulation of RSK/MSK autophosphorylation in transfected COS-7 cells

COS-7 cells were grown to confluence in a T150 flask and transfected in 15 mL antibiotic-free media overnight by addition of the following: 3 mL OptiMEM containing 60 μL Lipofectamine 2000 (Invitrogen) and 24 μg plamid encoding HA-tagged MSK1, MSK2, RSK2 or RSK3. Transfected cells were seeded into six-well plates in 2 mL of antibiotic-free media and grown overnight prior to drug treatment. Cells were serum starved for two hours. Each well was treated with 2 μL of 1000x inhibitor dilutions in DMSO for 2 h and then stimulated with 100ng/mL PMA for 30 min. Cells were washed with PBS and frozen at -80 °C prior to lysis.

Hela cell culture and stimulation

Cells were seeded into six-well plates at an initial density of 200,000/well in 2 mL of complete media and allowed to grow overnight. Cells were deprived of serum for 18 h, treated with 2 μL of 1000x inhibitor dilutions in DMSO for 2 h, and then stimulated with either PMA (400ng/mL for 15 minutes) or UV-C (tissue culture hood UV lamp, 5 min irradiation, followed by 30 min of incubation in serum free media at 37 °C). Cells were washed with PBS and frozen at -80 °C prior to lysis.

Western blotting

Cells were lysed in 60 μ L of lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% Triton X100 supplemented with phosphatase and protease inhibitor cocktails (Roche)). Lysates were clarified by centrifugation at 14,000 rpm, normalized by Bradford assay, denatured in SDS and separated by 7.5% acrylamide SDS-PAGE. The gels were transferred to nitrocellulose membranes, stained with Ponceau, and blocked for 1 h in Odyssey Blocking Buffer. Blots were probed overnight at 4 °C with 1:1000 dilutions of rabbit primary antibodies (Cell Signaling): α -pS376 MSK1, α -pS360 MSK2, α -pS380 RSK2, α -pS385 RSK3, or α -pS133 CREB. Note that the α -pS133 CREB antibody also recognizes an identical epitope surrounding pS63 of ATF1, a related transcription factor. Blots were subsequently probed with 1:1000 mouse α -HA (12CA5, Roche) or 1:1000 mouse α -CREB (Cell Signaling) for 1 h at room temperature. Blots were subsequently washed with TBST and probed with 1:10,000 fluorescent secondary antibodies (Odyssey) for 1 hour at room temperature. Blots were washed thoroughly with TBST and imaged with a LiCOR imaging system (Odyssey).

Kinase Screen

In vitro kinase assay screening against 26 protein kinases was performed by Nanosyn, Inc.³³ by treating each kinase with 1 μ M **12** for the amount of time indicated in Table 2-4.

Supplemental Chart: Structures of intermediates not shown in main text of Chapter 2.

Supplemental Scheme: Synthesis of optimized inhibitors **11** and **12**^a

^aReagents and conditions: (a) NaOH, I₂, Dioxane. (b) 3,4,5-trimethoxyphenyl boronic acid, K₂CO₃, Pd(PPh₃)₄, 5:1 Dioxane:H₂O. (c) DBU, cyanoacetamide, THF, rt. (d) DBU, **15**, THF, rt.

Chemical Synthesis:

All purchased chemicals were used as received without further purification. Solvents were dried by passage through columns (either alumina or activated molecular sieves) on a Glass Contour solvent system. NMR spectra were obtained on a Varian Inova 400 MHz spectrometer and referenced to the residual solvent peak. LC-MS analysis was performed on a Waters Acquity LCT UPLC equipped with a TUV detector (monitored at 254 nm) and a Waters Acquity UPLC 1.7 μ m C-18 column, eluting at 0.6 mL/min with a 5 minute water:MeCN (with 0.1% formic acid) gradient method: 0 – 0.1 min, 5% MeCN;

0.1 – 3.8 min, 5 – 95% MeCN; 3.8 – 4.5 min, 95% MeCN; 4.5 – 4.7 min, 95 – 5% MeCN; 4.7 – 5.0 min, 5% MeCN. High resolution mass spectrometry (HRMS) was performed at the QB3 Mass Spectrometry Facility at UC Berkeley.

2-cyano-3-(1H-indazol-5-yl)acrylamide (1). To a solution of 5-formyl-1H-indazole (102.5 mg, 0.7 mmol) in THF (2 mL) was added 2-cyanoacetamide (75 mg, 0.9 mmol) and piperidine (20 μ L, 0.2 mmol). The reaction was stirred for 21 hours at room temperature. The yellowish precipitate was filtered and the solids were washed with THF and dried *in vacuo* to afford 65 mg (44%) of 2-cyano-3-(1H-indazol-5-yl)acrylamide as a yellow solid. ^1H NMR (400 MHz, DMSO- d_6): δ 7.50 - 7.90 (overlaps peak at 7.72, br d, 2H), 7.72 (d, 1H, J = 8.8 Hz), 8.02 (dd, 1H, J = 8.8, 1.6 Hz), 8.29 (s, 1H), 8.30 (s, 1H), 8.42 (s, 1H). ^{13}C NMR (100 MHz, DMSO- d_6): δ 103.51, 111.14, 117.17, 123.10, 124.48, 125.65, 126.76, 135.43, 140.95, 151.48, 163.12. HPLC purity > 99%, R_t = 1.1 min. HRMS (ESI $^-$) Calc. exact mass: 211.0625, M/z found: 211.0624 (M-H) $^-$.

2-cyano-3-(1*H*-indazol-6-yl)acrylamide (2). To a solution of 6-formyl-1*H*-indazole (102.5 mg, 0.7 mmol) in THF (2 mL) was added 2-cyanoacetamide (78.2 mg, 0.7 mmol) and piperidine (50 μ L, 0.5 mmol). The reaction was stirred for 18 hours at room temperature with production of a white precipitate. The reaction was filtered and the solids were washed with THF and dried *in vacuo* to afford 69 mg (46%) of 2-cyano-3-(1*H*-indazol-6-yl)acrylamide as a white solid. ^1H NMR (DMSO- d_6): δ 7.67 (dd, 1H, J = 8.8, 1.6 Hz), 7.70-7.95 (overlaps peak at 7.92, br d, 2H) 7.92 (d, 1H, J = 8.8 Hz), 8.18 (s, 1H), 8.19 (s, 1H), 8.32 (s, 1H). ^{13}C NMR: δ 106.17, 112.87, 116.77, 121.37, 121.42, 124.60, 129.46, 133.86, 139.59, 151.29, 162.83. HPLC purity > 99%, R_t = 1.25 min. HRMS (ESI $^-$) Calc.: 211.0625, M/z found: 211.0632 ($M-H$) $^-$.

2-cyano-3-(quinolin-3-yl)acrylamide (3). To a solution of 3-formylquinoline (104 mg, 0.662 mmol) in 2:1 THF:iPrOH (6 mL) was added 2-cyanoacetamide (65 mg, 0.773 mmol) and piperidine (0.1 mL, 1.01 mmol). The reaction was stirred overnight at room temperature with evolution of a white precipitate. The reaction was filtered and the solids were washed with THF and dried *in vacuo* to afford 110 mg (69%) as a white solid. ^1H NMR (400 MHz, DMSO- d_6): δ 7.73 (t, 1H, J = 7.6 Hz), 7.85 – 8.10 (overlaps peak at 7.92, br d, 2H) 7.92 (t, 1H, J = 7.6 Hz), 8.11 (t, 2H, 7.6 Hz), 8.41 (s, 1H), 8.94 (d, 1H, J = 2.0 Hz). ^{13}C NMR (100 MHz,

DMSO- d_6): δ 108.74, 116.28, 125.49, 126.70, 127.85, 128.87, 129.27, 132.03, 137.50, 147.76, 150.47, 162.24. HPLC purity > 99%, R_t = 1.35 min. HRMS (ESI⁺) Calc.: 224.0818, M/z found: 224.0827 (M+H)⁺.

2-cyano-3-(quinolin-4-yl)acrylamide (4). To a solution of 4-formylquinoline (102 mg, 0.662 mmol) in 2:1 THF:iPrOH (6 mL) was added 2-cyanoacetamide (67 mg, 0.796 mmol) and piperidine (0.1 mL, 1.01 mmol). The reaction was stirred for 48 hours at room temperature with production of a white precipitate. The reaction was filtered and the solids were washed with THF and dried *in vacuo* to afford 106 mg (66%) of 2-cyano-3-(quinolin-4-yl)acrylamide as a white solid. ¹H NMR (400 MHz, DMSO- d_6): δ 7.38 (ddd, 1H, J = 8.4, 6.8, 1.2 Hz), 7.85 – 7.90 (m, 2H), 7.95 – 8.35 (br d, 2H), 8.14 (ddd, 1H, J = 8.4, 1.2, 0.8 Hz), 8.18 (ddd, 1H, J = 8.4, 1.2, 0.4 Hz), 8.82 (dd, 1H, J = 1.2, 0.8 Hz), 9.07 (d, 1H, J = 4.4 Hz). ¹³C NMR (100 MHz, DMSO- d_6): δ 114.76, 115.17, 120.22, 124.31, 124.92, 127.75, 129.72, 130.31, 137.92, 146.62, 147.77, 150.38, 161.94. HPLC purity > 99%, R_t = 1.05 min. HRMS (ESI⁺) Calc.: 224.0818, M/z found: 224.0827 (M+H)⁺.

3-(2-aminothiazol-5-yl)-2-cyanoacrylamide (5). To a solution of 2-aminothiazole-5-carbaldehyde (97 mg, 0.757 mmol) in 2:1 THF:iPrOH (6 mL) was added 2-cyanoacetamide (82 mg, 0.975 mmol) and piperidine (0.14 mL, 1.14 mmol). The reaction was stirred for 48 hours at room temperature with production of a bright yellow precipitate. The reaction was filtered and the solids were washed with THF and dried *in vacuo* to afford 68 mg (45 %) of 3-(2-aminothiazol-5-yl)-2-cyanoacrylamide as a yellow solid. ^1H NMR (400 MHz, DMSO- d_6): δ 7.42 (br s, 2H), 7.85 (s, 1H), 8.18 (s, 1H), 8.37 (br s, 2H). ^{13}C NMR (100 MHz, DMSO- d_6): δ 94.69, 117.79, 119.77, 143.05, 155.54, 163.42, 175.05. HPLC purity > 99%, R_t = 2 isomers, 0.65 and 0.75 min. HRMS (ESI $^-$) Calc.: 193.0190, M/z found: 193.0193 (M-H) $^-$.

2-cyano-3-(4-(pyridin-4-yl)phenyl)acrylamide (6). To a solution of 4-(pyridin-4-yl)benzaldehyde (100 mg, 0.545 mmol) in THF:iPrOH (2:1) (6 mL) was added 2-cyanoacetamide (57 mg, 0.660 mmol) and piperidine (0.1 mL, 1.015 mmol). The reaction was stirred for 24 hours at room temperature with production of a white precipitate. Filtration of the reaction afforded 68 mg (50%) of 2-cyano-3-(4-(pyridin-4-yl)phenyl)acrylamide as a white solid. ^1H NMR (400 MHz, DMSO- d_6): δ

7.81 (d, 2H, $J = 6$ Hz), 7.85 – 8.00 (overlaps peak at 7.81, br d, 2H), 8.03 (d, 2H, $J = 8.4$ Hz), 8.09 (d, 2H, $J = 8.4$ Hz), 8.25 (s, 1H), 8.69 (d, 2H, $J = 6$ Hz). ^{13}C NMR (100 MHz, DMSO- d_6): δ 107.19, 116.40, 121.26, 127.56, 130.75, 132.59, 140.43, 145.54, 149.69, 150.39, 162.56. HPLC purity > 99%, $R_t = 0.8$ min. HRMS (ESI $^+$) Calc.: 250.0975, M/z found: 250.0984 ($M+H$) $^+$.

3-(4-(9H-purin-6-yl)phenyl)-2-cyanoacrylamide (7). 4-(9-(tetrahydro-2H-pyran-2-yl)-9H-purin-6-yl)benzaldehyde was synthesized as reported previously³⁴, and spectral data corresponded to reported values. ^1H NMR (400 MHz, CDCl_3): δ 1.65 – 1.93 (m, 3H), 2.02 - 2.28 (m, 3H), 3.78 – 3.88 (m, 1H), 4.18 - 4.27 (m, 1H), 5.88 (dd, 1H, $J = 10, 2.4$ Hz), 8.07 (d, 2H, $J = 8.4$ Hz), 8.39 (s, 1H), 8.98 (d, 2H, $J = 8.4$ Hz), 9.08 (s, 1H), 10.13 (s, 1H). To a solution of this 4-(9-(tetrahydro-2H-pyran-2-yl)-9H-purin-6-yl)benzaldehyde (170 mg, 0.55 mmol) in 6 mL THF was added 20 drops of aqueous 1N HCl. The reaction was stirred at room temperature overnight. After 18 hours a white precipitate had formed and all starting material had been consumed as determined by TLC. The reaction was diluted with EtOAc, washed with saturated NaHCO_3 , and the organic layer was dried with Na_2SO_4 . Rotary evaporation afforded 100 mg (81%) of the deprotected

product as a light yellow solid that was carried on without further purification. A slurry of this deprotected 6-(*p*-formylphenyl)-purine (10.9 mg, 0.049 mmol), 2-cyanoacetamide (5.4 mg, 0.064 mmol), and PPh₃ (13 mg, 0.049 mmol) in THF (0.5 mL) was heated to 80°C in a sealed vial. The mixture was stirred at 80°C for 12 hours and concentrated *in vacuo*. The residue was chromatographed on silica with 9:1 CH₂Cl₂:MeOH, affording 9.6 mg (87%) of 3-(4-(9*H*-purin-6-yl)phenyl)-2-cyanoacrylamide as a light yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ 7.80 – 8.05 (br d, 2H), 8.15 (d, 2H, *J* = 8 Hz), 8.27 (s, 1H), 8.71 (s, 1H), 8.97 (br d, 2H, *J* = 8 Hz), 9.01 (s, 1H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ 107.8, 116.3, 129.7, 130.2, 133.8, 138.8, 145.7, 149.7, 150.4, 151.9, 154.0, 162.6. HPLC purity: 95%, *R*_t = 1.26 min. HRMS (ESI[−]) Calc.: 289.0843, *M/z* found: 289.0836 (*M*-H)[−].

3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)benzaldehyde (**13**). 7-azaindole (1.01 g, 8.5 mmol) and AlCl₃ (5.4 g, 40 mmol) were combined in dry CH₂Cl₂ (40 mL) in a flame-dried round bottom flask under argon. The insoluble brownish mixture was allowed to stir at room temperature for one hour, at which time freshly prepared 3-formyl benzoyl chloride (1.68 g, 10 mmol) was added dropwise via a syringe under positive pressure of argon. After two additional hours of stirring at room temperature, the reaction was diluted with EtOAc and water was added cautiously to quench excess acid chloride. The organic layer was dried (Na₂SO₄)

and concentrated *in vacuo* to afford a tan solid, which was triturated with 20 mL CH₂Cl₂ to afford 525 mg (26%) of 3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)benzaldehyde as a tan solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ 7.30 (dd, 1H, *J* = 8.0, 4.8 Hz), 7.75 (t, 1H, *J* = 7.6 Hz) 8.10 (d, 1H, *J* = 1.6 Hz), 8.12 (d, 1H, *J* = 1.6 Hz), 8.18 (s, 1H), 8.30 (t, 1H, *J* = 1.6 Hz), 8.37 (dd, 1H, *J* = 4.8, 1.6 Hz), 8.53 (dd, 1H, *J* = 8.0, 1.6 Hz), 10.10 (s, 1H). ¹³C NMR (100MHz, DMSO-*d*₆): δ 113.4, 118.8, 118.6, 129.6, 129.8, 130.0, 131.3, 134.1, 136.2, 136.4, 140.2, 144.7, 149.1, 188.7, 193.0. HPLC purity: 92%, *R*_t = 1.65 min. HRMS (ESI⁻) Calc.: 249.0670, *M/z* found: 249.0677 (*M*-H)⁻.

3-(3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)phenyl)-2-cyanoacrylamide (**8**). To a solution of 3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)benzaldehyde (75 mg, 0.3 mmol) in THF (1 mL) was added 2-cyanoacetamide (23 mg, 0.3 mmol) and DBU (45 μL, 0.3 mmol) and the slurry was solubilized with a few drops of MeOH. The clear solution was stirred at room temperature overnight with evolution of a white precipitate. The precipitate was collected by filtration, dissolved in a EtOAc, washed with 1M HCl, saturated NaHCO₃, and brine. The organics were dried and concentrated in vacuo to afford 21 mg (22%) of 3-(3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)phenyl)-2-cyanoacrylamide as a white solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ 7.32 (dd, 1H, *J* = 8.0, 4.8 Hz), 7.75 (t, 1H, *J* = 8.0 Hz), 7.80 – 8.00 (br d,

2H), 8.00 (d, 1H, $J = 7.6$ Hz), 8.15 (d, 1H, $J = 8.0$ Hz), 8.21 (s, 1H), 8.32 (s, 1H), 8.36 (s, 1H), 8.39 (dd, 1H, $J = 4.8, 1.6$ Hz), 8.57 (dd, 1H, $J = 7.6, 1.6$ Hz). ^{13}C NMR (100MHz, DMSO- d_6): δ 107.75, 113.306, 116.36, 118.38, 118.63, 129.63, 129.72, 129.84, 131.95, 132.20, 132.51, 136.53, 140.28, 144.70, 149.19, 150.13, 162.55, 188.70. HPLC purity: 98%, $R_t = 1.60$ min. HRMS (ESI $^-$) Calc.: 315.0887, M/z found: 315.0895 (M-H) $^-$.

2-cyano-3-(1H-pyrrolo[2,3-b]pyridin-3-yl)acrylamide (9). To a solution of *N*-Boc-3-formyl-7-azaindole (107 mg, 0.435 mmol) in THF (6 mL) was added 2-cyanoacetamide (45 mg, 0.535 mmol) and piperidine (0.1 mL, 0.74 mmol). The reaction was stirred for hours at room temperature with production of a white precipitate. The reaction was filtered and the solids were washed with THF and dried *in vacuo* to afford 45 mg (49 %) of the *N*-deprotected 2-cyano-3-(1H-pyrrolo[2,3-b]pyridin-3-yl)acrylamide as a white solid. ^1H NMR (400 MHz, DMSO- d_6): δ 7.30 (ddd, 1H, $J = 8.0, 4.8, 1.2$ Hz), 7.50 – 7.85 (br d, 2H), 8.35 – 8.45 (m, 2H), 8.45 (s, 1H), 8.49 (s, 1H). ^{13}C NMR (100 MHz, DMSO- d_6): δ 99.29, 108.29, 117.64, 118.49, 119.23, 127.63, 130.70, 142.14, 144.75, 148.52, 163.66. HPLC purity = 99%, $R_t = 1.05$ min. HRMS (ESI $^-$) Calc.: 211.0625, M/z found: 211.0630 (M-H) $^-$.

2-cyano-3-(1H-pyrazol-4-yl)acrylamide (10). To a slurry of 1H-pyrazole-4-carbaldehyde (96 mg, 1 mmol) in dilute NaHCO₃ solution (2 mL of a tenfold dilution of a saturated solution) was added 2-cyanoacetamide (100 mg, 1.2 mmol). The reaction became homogeneous after several hours of stirring at room temperature, followed by evolution of precipitate. After continued stirring overnight, the solids were collected by filtration and dried *in vacuo* to afford 69 mg (43%) of 2-cyano-3-(1H-pyrazol-4-yl)acrylamide as a white solid. ¹H NMR (400 MHz, DMSO-*d*₆ + 1% TFA): δ 7.46 – 7.76 (br d, 2H), 8.15 (d, 1H, *J* = 0.4 Hz), 8.28 (d, 2H, *J* = 0.4 Hz). ¹³C NMR (100 MHz, DMSO-*d*₆ + 1% TFA): δ 101.58, 115.10, 117.65, 136.35, 143.32, 163.06. HPLC purity = 94%, *R*_t = 0.85 min. HRMS (ESI⁺) Calc.: 161.0469, *M/z* found: 161.0472 (*M*-H)⁺.

3-(3,4,5-trimethoxyphenyl)-1H-indazole-5-carbaldehyde (14). 3-iodo-5-formylindazole² (645 mg, 2.37 mmol), trimethoxyphenyl boronic acid (608 mg, 2.86 mmol) and K₂CO₃ (987 mg, 7.14 mmol) were combined in 5:1 Dioxane:H₂O (4 mL) and degassed with bubbling argon for 20 minutes. Pd(PPh₃)₄ (270 mg,

0.233 mmol) was added, and the reaction vessel purged with argon. The reaction was microwaved twice at 150°C for five minutes each. The crude reaction mixture was diluted with EtOAc (40 mL), washed with 1M HCl (20 mL), H₂O (20 mL) and brine (20 mL), dried with Na₂SO₄, and concentrated. The crude residue was purified on silica, eluting with 1:1 Hex:EtOAc to afford 409 mg (55 %) of 3-(3,4,5-trimethoxyphenyl)-1*H*-indazole-5-carbaldehyde as a white solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ 3.75 (s, 3H), 3.92 (s, 6H), 7.24 (s, 2H), 7.72 (d, 1H, *J* = 8.8 Hz), 7.88 (d, 1H, *J* = 8.8, 1.6 Hz), 8.70 (s, 1H), 10.10 (s, 1H), ¹³C NMR (100MHz, DMSO-*d*₆): δ 56.06, 60.13, 104.58, 111.53, 119.89, 124.036, 128.31, 128.52, 130.58, 137.90, 143.88, 145.77, 153.37, 192.62. HPLC purity = 99%, *R*_t = 1.95 min. HRMS (ESI⁺) Calc.: 313.1183, *M/z* found: 313.1194 (*M*+H)⁺.

2-cyano-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide (11). To a solution of 3-(3,4,5-trimethoxyphenyl)-1*H*-indazole-5-carbaldehyde **14** (52.0 mg, 0.17 mmol) and 2-cyanoacetamide (15.0 mg, 0.18 mmol) in THF (1 mL) was added DBU (16 μL, 0.16 mmol). The colorless solution became bright yellow upon addition of DBU, and slowly turned bright red over the course of 15 minutes. After stirring overnight (~17 hours), the reaction was concentrated under a stream of air, resuspended in EtOAc (~0.5 mL). The resulting yellow precipitate

was collected by filtration and washed with EtOAc, THF, and EtOAc (~0.5 mL each) to afford 35 mg (58%) of 2-cyano-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide as a bright yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ 3.74 (s, 3H), 3.92 (s, 6H), 7.22 (s, 2H), 7.65 – 7.89 (overlaps peak at 7.75, br d, 2H), 7.75 (d, 1H, *J* = 8.8 Hz), 7.96 (d, 1H, *J* = 8.8, 1.6 Hz), 8.39 (s, 1H), 8.84 (s, 1H). ¹³C NMR (100MHz, DMSO-*d*₆): δ 56.02, 60.12, 103.52, 104.29, 111.52, 117.37, 120.32, 124.49, 125.07, 128.30, 128.35, 137.78, 142.62, 145.12, 151.55, 153.37, 163.04. HPLC purity: 98%, *R*_t = 1.75 min. HRMS (ESI⁺) Calc.: 377.1267, *M/z* found: 377.1255 (M-H)⁺.

2-cyano-N-(1-hydroxy-2-methylpropan-2-yl)acetamide. Ethyl cyanoacetate (2.13 g, 18.8 mmol, 1.0 equiv.) was heated with 2-amino-2-methylpropanol (1.68 g, 1.0 equiv.) in EtOH (10 mL) to 80 °C in a sealed scintillation vial for 8 hours. The reaction mixture was then cooled to ambient temperature and concentrated and the residue was purified by silica gel chromatography (EtOAc) to afford the desired amide, 2.73 g (93% yield) as a white solid. *R*_f (SiO₂, EtOAc): 0.36; ¹H NMR (400 MHz, CDCl₃): 1.32 (s, 6H), 3.41 (s, 2H), 3.58 (d, 2H, 5.9 Hz), 3.80 (t, 1H, *J* = 5.9 Hz), 6.48 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): 23.7, 26.7, 56.6, 69.0, 115.1, 161.8; HRMS (ESI⁺) Calc.: 155.0826, *M/z* found: 155.0825 (M-H)⁺.

2-cyano-*N*-(1-hydroxy-2-methylpropan-2-yl)-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide (12). To a solution of 3-(3,4,5-trimethoxyphenyl)-1*H*-indazole-5-carbaldehyde **14** (48mg, 0.16 mmol) and 2-cyano-*N*-(1-hydroxy-2-methylpropan-2-yl)acetamide **15** (23 mg, 0.16 mmol) in THF (1 mL) was added DBU (16 μ L, 0.16 mmol). The colorless solution slowly became bright orange after 1 hour. The reaction mixture was concentrated and purified by preparative TLC, eluting with EtOAc, to afford 44 mg (61%) of 2-cyano-*N*-(1-hydroxy-2-methylpropan-2-yl)-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide a bright yellow solid. ^1H NMR (DMSO- d_6): δ 1.31 (s, 6H), 3.47 (br d, 2H, $J = 4$ Hz), 3.74 (s, 3H), 3.92 (s, 6H), 5.04 (br t, 1H, $J = 4$ Hz), 7.22 (s, 2H), 7.38 (s, 1H), 7.75 (d, 1H, $J = 8.8$ Hz), 7.95 (d, 1H, $J = 8.8, 1.6$ Hz), 8.33 (s, 1H), 8.83 (s, 1H). ^{13}C NMR: δ 23.18, 55.29, 56.03, 60.12, 67.17, 104.19, 104.29, 111.55, 117.58, 120.29, 124.25, 125.13, 128.37, 137.77, 142.61, 145.04, 151.05, 153.38, 160.66. HPLC purity: 98%, $R_t = 1.35$ min. HRMS (ESI $^-$) Calc.: 449.1830, M/z found: 449.1841 ($M-H$) $^-$.

	RSK2 WT: 1	RSK2 WT: 8	RSK2 T493M: 12
Data Collection			
space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2
unit-cell parameters	a=47.7; b=47.7; c=292.7; a=90°; b=90°; g=90°	a=47.7; b=47.7; c=292.1; a=90°; b=90°; g=90°	a=47.8; b=47.8; c=291.4; a=90°; b=90°; g=90°
resolution range (Å)	2.60 – 46.7	3.00 – 47.0	3.10 – 45.8
highest resolution shell (Å)	2.60 – 3.19	3.00 – 3.18	3.10 – 3.29
measured reflections	96462 (15548)	44647 (6522)	113910 (16164)
unique reflections	11032 (1715)	7395 (1164)	6457 (992)
redundancy	8.7 (9.1)	6.0 (5.6)	17.64 (16.29)
completeness (%)	99.9 (99.7)	99.3 (99.3)	99.9 (99.3)
average I/s	22.07 (4.6)	21.1 (5.38)	35.24 (15.5)
R _{sym} (%)	6.5 (36.1)	9.2 (32.5)	3.4 (9.1)
Refinement			
resolution range (Å)	2.60 – 46.7	3.0 – 47.0	3.10 – 45.8
highest resolution shell (Å)	2.60 – 2.86	3.0 – 3.23	3.10 – 3.91
R _{work}	25.21 (36.68)	25.86 (32.56)	23.66 (26.93)
R _{free}	29.86 (44.17)	29.79 (35.94)	29.78 (35.12)
no. protein atoms	4748	4611	4802
no. water molecule	42	22	20
Wilson B-value (Å ²)	52.97	39.6	46.2
Average B factors (Å ²)	74.6	37.6	60.5
Protein	66.2	34.9	55.2
Solvent	38.9	30.0	25.0
Ligand	69.8	34.1	58.1
Ramachandran (%)			
favoured	97.3	95.1	95.3
outliers	0	0.3	0.3
rms deviations			
bond lengths (Å)	0.013	0.007	0.002
bond angles (°)	0.660	0.437	0.452

Table 2-5: Crystallographic Data Collection and Refinement Statistics for Compounds **1**, **8** and **12** bound to RSK CTD kinases.

2.4 References

- (1) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. *J. Med. Chem.* **2004**, 47, 3363–3482.
- (2) Howard, S.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Curry, J.; Devine, L. A.; Early, T. R.; Fazal, L.; Gill, A. L.; Heathcote, M.; Maman, S.; Matthews, J. E.; McMenamin, R. L.; Navarro, E. F.; O'Brien, M. A.; O'Reilly, M.; Rees, D. C.; Reule, M.; Tisi, D.; Williams, G.; Vinković, M.; Wyatt, P. G. *J. Med. Chem.* **2009**, 52, 379–388.
- (3) Bamborough, P.; Brown, M. J.; Christopher, J. A.; Chung, C.; Mellor, G. W. *J. Med. Chem.* **2011**, 54, 5131–5143.
- (4) Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 856–864.
- (5) Potashman, M. H.; Duggan, M. E. *J. Med. Chem.* **2009**, 52, 1231–1246.
- (6) Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trump-Kallmeyer, S.; Dobrusin, E. M. *Proc. Natl. Acad. Sci.* **1998**, 95, 12022–12027.

(7) Hagel, M.; Niu, D.; St Martin, T.; Sheets, M. P.; Qiao, L.; Bernard, H.; Karp, R. M.; Zhu, Z.; Labenski, M. T.; Chaturvedi, P.; Nacht, M.; Westlin, W. F.; Petter, R. C.; Singh, J. *Nat. Chem. Biol.* **2011**, 7, 22–24.

(8) Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, 9, 28–39.

(9) Leproult, E.; Barluenga, S.; Moras, D.; Wurtz, J. M.; Winssinger, N. *J. Med. Chem.* **2011**, 54, 1347–1355.

(10) Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. *Nat. Chem. Biol.* **2012**, 8, 471–476.

(11) Deak, M.; Clifton, A. D.; Lucocq, J. M.; Alessi, D. R. *EMBO J.* **1998**, 17, 4426–4441.

(12) McCoy, C. E.; Campbell, D. G.; Deak, M.; Bloomberg, G. B.; Arthur, J. S. C. *Biochem. J.* **2005**, 387, 507–517.

(13) Wiggan, G. R.; Soloaga, A.; Foster, J. M.; Murray-Tait, V.; Cohen, P.; Arthur, J. S. C. *Mol. Cell. Biol.* **2002**, 22, 2871–2881.

- (14) Odgerel, T.; Kikuchi, J.; Wada, T.; Shimizu, R.; Kano, Y.; Furukawa, Y. *Leukemia* **2010**, *24*, 1087–1090.
- (15) Pérez-Cadahía, B.; Drobic, B.; Espino, P. S.; He, S.; Mandal, S.; Healy, S.; Davie, J. R. *J. Biol. Chem.* **2010**, *286*, 42–49.
- (16) van der Heide, L. P.; van Dinther, M.; Moustakas, A.; ten Dijke, P. *J. Biol. Chem.* **2011**, *286*, 5003–5011.
- (17) Naqvi, S.; Macdonald, A.; McCoy, C. E.; Darragh, J.; Reith, A. D.; and Arthur, J. S. C. *Biochem. J.* **2012**, *441*, 347–357.
- (18) Ligand efficiency is calculated according to ref. 2, with IC₅₀ as a surrogate for K_d at 293 K, using $\Delta G = -RT\ln(IC_{50})$ and $LE = \Delta G/(\# \text{ heavy atoms})$.
- (19) Zhu, G.; Gandhi, V. B.; Gong, J.; Thomas, S.; Woods, K. W.; Song, X.; Li, T.; Diebold, R. B.; Luo, Y.; Liu, X.; Guan, R.; Klinghofer, V.; Johnson, E. F.; Bouska, J.; Olson, A.; Marsh, K.; Stoll, V. S.; Mamo, M.; Polakowski, J.; Campbell, T. J.; Martin, R. L.; Gintant, G. A.; Penning, T. D.; Li, Q.; Rosenberg, S. H.; Giranda, V. L. *J. Med. Chem.* **2007**, *50*, 2990–3003.
- (20) McTigue, M.; Murray, B. W.; Chen, J. H.; Deng, -L; Solowiej, J.; Kania, R. S. *Proc. Natl. Acad. Sci.* **2012**, *109*, 18281–18269.

- (21) Patel, R. Y.; Doerksen, R. J. *J. Proteome Res.* **2010**, *9*, 4433–4442.
- (22) Guimarães, C. R. W.; Rai, B. K.; Munchof, M. J.; Liu, S.; Wang, J.; Bhattacharya, S. K.; Buckbinder, L. *J. Chem. Inf. Model.* **2011**, *51*, 1199–1204.
- (23) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. *Nature* **2010**, *468*, 790–795.
- (24) Nonoo, R. H.; Armstrong, A.; Mann, D. J. *ChemMedChem.* **2012**, *7*, 2082–2086.
- (25) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *Proc. Natl. Acad. Sci.* **2000**, *97*, 9367–9372.
- (26) Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. *Science* **2005**, *308*, 1318–1321.
- (27) Rellos, P.; Ivins, F. J.; Baxter, J. E.; Pike, A.; Nott, T. J.; Parkinson, D-M.; Das, S.; Howell, S.; Fedorov, O.; Shen, Q. Y.; Fry, A. M.; Knapp, S.; Smerdon, S. *J. J. Biol. Chem.* **2006**, *282*, 6833–6842.

(28) Henise, J.; Taunton, J. *J. Med. Chem.* **2011**, *54*, 4133–4146.

(29) Knight, Z. A.; Feldman, M. E.; Shokat, K. M. *Nature Protocols*, **2007**, *2*, 2459-660.

(30) Kabsch, W. *Acta Cryst.* **2010**, *D66*, 125-132.

(31) Adams, P. D. ; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger T. C.; Zwart, P. H. *Acta Cryst.* **2010**, *D66*, 213-221.

(32) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Cryst.* **2010**, *D66*, 486–501.

(33) Nanosyn Single Point and IC₅₀ Profiling Services (<http://www.nanosyn.com>)

(34) Donald, A.; McHardy, T.; Rowlands, M. G.; Hunter, L. K.; Davies, T. G.; Berdini, V.; Boyle, R. G.; Aherne, G. W.; Garrett, M. D.; Collins, I. *J. Med. Chem.* **2007**, *50*, 2289-2292.

Chapter 3. Tuning and Stabilization of Doubly-Activated Olefins by Variation of the Non-Nitrile Electron- Withdrawing Substituent

3.1 Abstract

In spite of the success of our electrophilic fragment-based approach for the design of inhibitors, the unsubstituted cyanoacrylamide electrophile proved to be far from optimal for chemical probe applications. Specifically, these highly reactive electrophiles were found to be both chemically and metabolically unstable. We sought to explore beyond simple, unsubstituted cyanoacrylamide electrophiles for several reasons: 1) The amide nitrogen provides a facile diversification point, enabling rapid access to an SAR series on the same scaffold, 2) substitution of the amide nitrogen affords compounds with improved stability, both thermally and metabolically, and 3) the reactivity of the electrophile can be “tuned” by varying the non-nitrile substituent. Replacement of the amide substituent with heteroaromatic activating groups dramatically improved the thermal and metabolic stability of our optimized electrophilic inhibitors. Finally, further optimization of the indazole scaffold while retaining the beneficial heteroaryl-activated electrophile resulted in the stable, potent, soluble and orally bioavailable inhibitor **29** (RMM-64), capable of inhibiting RSK and MSK in live animals.

Note: Portions of this chapter were co-written with Dr. Shyam Krishnan and Prof. Jack Taunton (UC San Francisco), and have been prepared for submission to a peer-reviewed journal. Krishnan, S.; Miller, R. M.; Tian, B.; Serafimova, I. M.; Mullins, R. D.; Jacobson, M. P.; Taunton, J. Rational Design of Activated Acrylonitriles for Reversible Cysteine Targeting Applications, *prepared manuscript*, **2013**.

3.2 Introduction

To be genuinely useful as a chemical probe for biological studies, a compound must not only be potent and selective for its intended target, but must also be stable over the duration of the experiment. Compound integrity must be maintained in solution, and it must also be resistant to cellular metabolism. As we transitioned from acute biochemical assays to more extended experiments in tissue culture and animal models, we became concerned about the stability of the reactive cyanoacrylamide electrophile.

We noted that olefins activated by different electron withdrawing groups has variable stability in buffered solution at 37 °C in the context of a series of pyrrolopyrimidine RSK2 inhibitors.^{1,2} By monitoring the amount of compound present by LC-MS, an unsubstituted cyanoacrylamide analog of the pyrrolopyrimidine scaffold decomposes with a half life of about five days when stored at 37 °C. When the amide nitrogen was functionalized with an isopropyl group, the compound was significantly more stable, with only 10% decomposition observed after five days. In contrast, an analog in which the olefin was activated by a t-Bu ester was degraded after only one day. These preliminary data suggested that doubly activated olefins may be generally but variably unstable, motivating further variation of the non-nitrile electron withdrawing group.

Knoevenagel condensation of doubly activated methylenes like malononitrile, 2-cyanoacetamides and 2-cyanoacrylates with aldehydes is well established. For this reason, we were first interested in developing inhibitors based on amide- and ester-functionalized vinyl nitriles. This approach was particularly appealing because of the ease of amide and ester bond synthesis,

which would enable rapid access to an SAR series on the same scaffold. The drastic differences in hydrolytic stability observed with simple amides and esters were further motivation to vary this portion of the molecule. However, while amides and esters are synthetically accessible, the range of potential electron withdrawing substituents is far greater and in need of exploration.

Electrophilic olefins such as α -cyanoenones³, alkylidene rhodanines⁴, and alkylidene thiazolidine diones⁵ have been found to form reversible covalent adducts with their target proteins, but little is understood about the relationship between the chemical structure of a Michael acceptor activated by two electron withdrawing groups and its ability to react reversibly with thiols. In our efforts to stabilize our electrophilic inhibitors, we wished to retain the nitrile group as one of the electron withdrawing groups because of its low molecular weight, metabolic stability, and the ability to serve as a hydrogen bond acceptor in sterically congested environments.^{6,7} We conjectured that aryl and heteroaryl groups would serve as ideal electron-withdrawing groups for activating the olefin to react reversibly with thiols. Not only can heteroaromatic rings form non-covalent interactions with their protein targets and provide favorable pharmacological propertiesⁱ, but the enormous diversity in the positions and nature of the substituents and heteroatoms provides highly tunable electronic character. Apart from isolated reports⁸, the reaction of cysteine thiols with vinyl heteroaryl groups under physiological conditions has not been studied. Gratifyingly, replacement of the amide electron-withdrawing group with a variety of aromatic heterocycles in the context of both model olefins and our pyrrolopyrimidine inhibitors resulted in

predictable, reversible covalent bond formation with thiols. Furthermore, heteroaromatic activation of these vinyl nitrile inhibitors dramatically improved the metabolic stability in some cases (unpublished data, Dr. Shyam Krishnan). Given the unstable nature of the cyanoacrylamide electrophile when appended to certain inhibitor scaffolds, we sought to investigate the stability of this class of electrophile across a variety of inhibitor scaffolds. In particular, we sought to apply the heteroaryl-activated electrophile to our successful 3-arylindazole pharmacophore to enable the development of RSK/MSK inhibitors (Chapter 2) with suitable stability for use in prolonged cellular experiments and animal models.

3.3 Results and Discussion

As anticipated based on the variable instability observed in the context of the pyrrolopyrimidine RSK inhibitor series^{1,2}, isomeric indazole fragments **1** and **2** (Chapter 2, Figure 2-2) rapidly decomposed in buffered solution at 37 °C with a half-life of about 24 hours (Figure 3-1). The primary mechanism of decomposition was retro-Knoevenagel of the electrophile to afford the precursor aldehyde.

Figure 3-1: Cyanoacrylamides **1** and **2** are unstable when incubated at 37 °C in PBS at pH 7.6.

Starting with a relatively conservative N-methyl group, increasing the substitution and steric bulk at the amide nitrogen of cyanoacrylamide electrophiles was sufficient to dramatically increase their stability to hydrolytic decomposition in buffer (Figure 3-2).

Figure 3-2: Increasing steric bulk and substitution of the amide nitrogen of the cyanoacrylamide electrophile improves their thermal stability dramatically in buffered solution.

Encouraged by the thermal stability engendered by alkyl substitution at the amide nitrogen, and eager to take advantage of the ability to rapidly generate SAR at the same position, we synthesized a panel of N-alkyl cyanoacrylamide derivatives on a variety of inhibitor scaffolds, including isomeric indazoles **16** - **21**, benzoyl azaindoles **22** and **23**, and 3-arylindazoles **24** and **25** (**Figures 3-2 and 3-3**). These compounds were screened against recombinant RSK2 kinase domain, and we observed moderate variation in their inhibitory profile (**Figure 3-3**).

Figure 3-3: Facile substitution at the amide nitrogen with a variety of substituents enables the rapid generation of an SAR series on the same scaffold. Shown are *in vitro* kinase assay IC₅₀ values (nM) for (a) 5-substituted indazoles, (b) 6-substituted indazoles, (c) benzoyl azaindoles and (d) 3-arylandazoles.

Typically, monomethylation of the amide nitrogen affords a compound with slightly improved potency, but larger substitution resulted in varied potency and selectivity among a set of compounds based on the same scaffold. A striking

example of the potential improvement obtained by amide substitution is compound **12** (discussed in Chapter 2), which was the most RSK/MSK-selective cyanoacrylamide we synthesized based on the indazole pharmacophore. We surmised that interaction of the hydroxy-t-Bu amide substituent with the glycine-rich loop may be the source of this selectivity (see Figure 2-5).

Emboldened by the seemingly improved stability, selectivity and potency of N-alkyl cyanoacrylamides, we tested their potency in MDA-MB-231 breast cancer cells with varied incubation time in the presence of cells. Several of these inhibitors potently competed a fluorescent, irreversible RSK1/2 inhibitor after two hours, indicating complete occupancy of RSK1 and RSK2 (Figure 3-4). However, after 24 hours in the presence of cells, none of these inhibitors were able to compete the probe at concentrations up to 10 μ M. We attribute this decrease in observed potency to the metabolic decomposition of the cyanoacrylamide electrophile.

Figure 3-4: Varied amide nitrogen substitution fails to stabilize inhibitors to the cellular environment. Top panels: each inhibitor is capable of competing the fluorescent **BODIPY-FMK** probe after a 2 hour incubation, indicating occupancy of RSK1/2. Bottom panels: none of the compounds are capable of occupying RSK1/2 after 24 hours in the presence of cells.

Replacement of the Amide Electron Withdrawing Group with a Heteroaromatic Electron-Withdrawing Group Stabilizes Inhibitors

As mentioned in the introduction, activation of acrylonitriles with heteroaryl groups led to increased hydrolytic stability and tunable reactivity in a series of model olefins, as well as pyrrolopyrimidine-based RSK inhibitors. We surmised that we could use this approach to stabilize our problematic indazoles by virtue of their relatively slow (and tunable) reactivity. Therefore, we synthesized 1,2,4-triazole-activated indazole **26** from 5-formylindazole, and compared it directly to the cyanoacrylamide analog (Figure 3-5).

Figure 3-5: Synthesis of triazole-activated indazole **26**. ^aReagents and conditions: (a) NaOH, I₂, Dioxane. (b) 3,4,5-trimethoxyphenyl boronic acid, K₂CO₃, Pd(PPh₃)₄, 5:1 Dioxane:H₂O. (c) DBU, 1,2,4-triazolylacetonitrile, THF, rt.

Triazole **26** was slightly less potent than **12** against RSK2 *in vitro* under stringent conditions containing 0.1 mM ATP and 10 mM GSH (IC₅₀ values of 47 and 3 nM respectively, Figure 3-6). However, **26** was somewhat more selective for RSK2 over NEK2 and PLK1, two divergent kinases that contain a homologous active site cysteine.

Figure 3-6: Compounds **12** and **26** potently inhibit the RSK2 CTD *in vitro*, but are much less potent against NEK2 and PLK1 despite the presence of a homologous cysteine. ^adata for **22** reproduced from Chapter 2.

Additionally, **26** proved to be more stable than its cyanoacrylamide predecessor, making it potentially more useful for cell-based experiments. When MDA-MB-231 cells were treated for two hours with 1 μ M inhibitor, both **12** and **26** fully blocked labeling by the fluorescent **FMK-BODIPY** probe, indicating complete occupancy of RSK1/2. However, upon incubation for 24 hours in cells, cyanoacrylamide **12** was inactive while triazole **26** showed sustained occupancy of RSK1/2 (Figure 3-7).

Figure 3-7: Triazole **26** competes with **BODIPY-FMK** after 24 hours in cells, while cyanoacrylamide **12** is inactive after the same duration.

We attribute this prolonged occupancy to the improved metabolic stability and the slower off-rate engendered by the triazole-activated electrophile relative to the cyanoacrylamide. As predicted by our model studies of 1,2,4-triazole-activated electrophiles (unpublished data, Dr. Shyam Krishnan), the covalent bond between **26** and the RSK2 kinase domain was reversed upon denaturation of the folded protein with guanidine, with a $t_{1/2}$ of about one hour (Figure 3-8).

Figure 3-8: Triazole **26** dissociates from RSK2 CTD upon denaturation with a $t_{1/2}$ of about 60 minutes.

An x-ray co-crystal structure of **26** bound to the RSK2 T493M kinase (PDB: 4M8T) confirmed the presence of a covalent bond with C436 and a network of specific noncovalent interactions between the scaffold and the kinase (Figure 3-9).

Figure 3-9: A co-crystal structure of triazole **26** bound to RSK2 CTD reveals the covalent and noncovalent interactions required for potent inhibition

The indazole nitrogens form two hydrogen bonds to the backbone of hinge residues E494 and M496, and the 3-arylindazole core packs against a number of hydrophobic residues including M493, L546 and I428. The triazole portion of the adduct projects downward toward the floor of the ATP binding site, making polar contacts with the sidechains of N544 and D561 and with the backbone carbonyl of S543.

Additionally, an alignment of our RSK2 co-crystal structure with several NEK2 and PLK1 structures suggests a ligand binding mode that is incompatible with these other kinases, in spite of the presence of a homologous cysteine, providing an explanation for the selectivity shown in Figure 3-6. In this RSK2 structure, the triazole resides about 5.3 Å from L546, which forms part of the floor of the ATP binding site. Both NEK2 and PLK1 contain a large phenylalanine at this position that would lie within 2 Å of the triazole, potentially creating a steric clash (Figure 3-10). The trimethoxyphenyl substituent at the 3-position of the indazole also clashes with an arginine (R136) at the mouth of the ATP pocket of PLK1, a site occupied by a glycine residue in RSK2 and MSK1 that readily accommodates this aryl substituent. These structural insights show that even a highly thiol-reactive electrophilic inhibitor can differentiate between three kinases with a cysteine at the same position. Furthermore, while the tunable electronics of heteroaryl-activated olefins modulates cysteine reactivity, the wide variety of heteroaromatic shapes and sizes may allow the aryl electron withdrawing group to serve as a noncovalent selectivity element. Thus, incorporation of tunable,

heteroaryl-activated olefins may be a general method for the design of electrophilic chemical probes suitable for use in cells.

Figure 3-10: Selectivity may arise because **26** is accommodated by L546 of RSK2 (blue), but is incompatible with a phenylalanine at this position in NEK2 and PLK1 (yellow) as well as R136 in PLK1 (red).

Our lab has begun investigating a method for the rational design of novel aryl and heteroaryl-activated Michael acceptors that react reversibly with thiols. The mechanistic insights provided by both computational analysis and kinetic measurements have allowed us to quantify the effect of varied aryl and heteroaryl substituents on the stability of thiol Michael adducts and the relative rate of reversible bond formation (unpublished data, Dr. Shyam Krishnan). We anticipate that the trends observed in this study will motivate the computational prediction of thiol-reactive, reversible covalent inhibitors with tunable off-rates. In this chapter, we have demonstrated the application of this heteroaryl activation strategy to the improvement of the versatile indazole scaffold, which will enable

the study of RSK- and MSK-mediated signaling. We envision that the diversity of aromatic heterocycles will be exploited to design other reversible, covalent electrophilic inhibitors of desired reactivity and we envision the extension of this strategy to target cysteines residues of proteins beyond kinases.

Optimization of Heteroaryl-Activated Indazole Inhibitors Affords Orally

Bioavailable MSK and RSK Inhibitors

The improved selectivity, stability and potency of compound **26** made it an attractive chemical probe with which to investigate the role of MSK and RSK CTD activity in biological processes. Unfortunately, despite the incorporation of the triazole-activated electrophile, **26** remained unsuitable for use in long-term biological studies. For example, **26** had lower aqueous solubility relative to its cyanoacrylamide precursors **11** and **12**. In fact, we were unable to formulate this triazole-activated compound in an aqueous solvent system and were forced to dose mice with compound formulated in 100% DMSO. While formulation in DMSO was acceptable for short, one-hour treatment times, extended dosing with this vehicle is known to be toxic. Thus, we sought to develop an inhibitor that retained the positive attributes of compound **26** but with higher aqueous solubility for formulation in non-toxic solutions.

The crystal structures of compounds **12** and **26** in complex with RSK2 T493M CTD allow unambiguous placement of the trimethoxyphenyl substituent of the ligand in a partly solvent-exposed hydrophobic pocket at the mouth of the ATP binding site. Notably, the methoxy groups project away from the protein and

out toward solvent. The para-methoxy substituent appears least congested by the remainder of the protein and was selected as a candidate position for functionalization with an appropriate solubilizing group.

Based on this structural information, we sought to incorporate an N-Me-piperazine moiety at the para position of a phenyl ring at the 3-position of the indazole core. From the outset, we envisioned following a route similar to the synthesis of compounds **12** and **26**, starting from 5-formylindazole (Figure 3-11).

Figure 3-11: Initial, linear synthesis of triazole-activated indazole **29**. ^aReagents and conditions: (a) NaOH, I₂, Dioxane. (b) 3,4,5-trimethoxyphenyl boronic acid, K₂CO₃, Pd(PPh₃)₄, 5:1 Dioxane:H₂O. (c) DBU, 1,2,4-triazolylacetonitrile, THF, rt.

However, while unsubstituted cyanoacrylamides like those found in Chapter 2 are relatively easy to synthesize in moderate to high yields, modification of the amide withdrawing group (and in particular, replacement of the amide with heteroaromatic activating groups) invariably led to a multitude of side products in the final step and resulted in prohibitively low yields of the purified inhibitor.

These problems persisted regardless of solvent, base, and temperature variation when reactions were performed with a model indazole aldehyde. For example, the best yield for the Knoevenagel condensation of triazole **26** was 17% (Figure 3-5). In the context of the N-methylpiperazyl compound **29**, we were able to

isolate a small quantity of the desired condensation product from the final synthetic transformation; however, the initial 7% yield was unsatisfactory and prohibited the isolation of quantities sufficient for animal studies. Reorganization of the synthetic route improved both the scale and the isolated yield of the final compound (Figure 3-12). The new synthetic scheme began with N1-BOC protection of 3-iodoindazole, which proceeded in high yield on a gram scale to afford the corresponding protected iodide.

Figure 3-12: Reorganized synthetic route to solubilized triazole-activated indazole **29**, with the installation of the highly reactive electrophile moved to the beginning of the synthesis. ^aReagents and conditions: (a) BOC₂O, DMAP, MeCN (b) DBU, 1,2,4-triazolylacetonitrile, THF, rt. (c) sat. Na₂CO₃, PdCl₂dppf, Toluene:EtOH, 4-(N-Me-piperazyl)phenylboronic pinacol ester. (d) aq. 3N HCl.

More importantly, Knoevenagel condensation of this BOC-protected intermediate with 1,2,4-triazole-substituted acetonitrile resulted in rapid precipitation of the desired olefin **28** from the reaction mixture with THF as a solvent. The rapid evolution of this insoluble product subverted the formation of the traditional side products and enabled isolation of the pure olefin product by filtration in an acceptable 36% yield. It was not clear that the highly reactive doubly-activated olefin product would be stable to the conditions required for palladium-mediated cross coupling with an appropriate aryl boronic acid or boronic ester. However, an advanced intermediate like **28** would be amenable to late stage diversification

and thus we pursued it nonetheless. In spite of the highly reactive nature of the triazole activated starting material, we observed smooth conversion of iodide **30** to the corresponding N-Me-piperazylphenyl analog **31** by Suzuki coupling with the requisite phenylboronic pinacol ester. This compound was readily deprotected in acidic solution to afford **29** in 63% combined yield (150 mg) over the final two steps.

In spite of the crystallographic evidence that the N-Me-piperazine would be accommodated by the kinase, we were concerned that our placement of this bulky group might still interfere with ligand binding. However, the optimized inhibitor **29** was found to be equipotent to **26** *in vitro* and in cell-based assays, and maintained its durability in the presence of cells (Figure 3-13).

Figure 3-13: Compound **29** remains extremely potent in spite of the bulky N-Me-piperazyl substituent. (A) Compounds **26** and **29** are equipotent against RSK2 T493M kinase in an *in vitro* kinase assay. (B) N-Me-piperazine **29** potently competes the fluorescent **BODIPY-FMK** at low nM concentrations. (C) Both triazole-activated compounds **26** and **29** fully occupy RSK1/2 after 24 hours in the presence of cells.

Gratifyingly, this compound could be formulated at concentrations of up to 25 mg/mL in a 50% PEG400 : water solution, indicating that our attempt to solubilize the problematic, triazole-activated 3-arylindazole was successful. Furthermore, mice treated with this 50% PEG400 vehicle and a 40 mg/kg dose of compound **29** for 10 days suffered no discernable toxicity and were indistinguishable from one another. This indicates that prolonged exposure to moderate doses of an

extremely potent, reactive electrophilic inhibitor of MSK and RSK kinases does not lead to general toxicity in animals.

In conclusion, we were able to identify the hydrolytic and metabolic instability of the unsubstituted cyanoacrylamide electrophile as an impediment to their widespread use as chemical probes. Extensive variation of the non-nitrile electron withdrawing group led to the adoption of tunable, heteroaryl-activated electrophiles. Final optimization of the physical properties of this inhibitor scaffold resulted in **29**, a highly selective compound that is amenable to a broad spectrum of biological assays, including prolonged animal studies. Finally, installation of the bulky N-Me-piperazyl substituent at the para position of the 3-phenylindazole scaffold did not diminish its potency *in vitro* or in cellular assays. This suggests the potential for further substitution at this position with propargyl “click” handles, a biotin tag, or a fluorescent label. Such modifications will open the door to even broader application of the indazole scaffold for use as occupancy probes.

3.4 Methods

X-Ray Crystallography:

RSK2 T493M CTD was co-crystallized with compound **26** by hanging drop vapor diffusion as described previously. Briefly, diffraction data was collected at the ALS beamline 8.2.1, and diffraction images integrated and indexed using the XDS suite. Molecular replacement and refinement was carried out using the Phaser and Phenix.Refine modules of the Phenix program suite. Data collection and refinement statistics can be found in Table 3-1.

Data Collection

space group
unit-cell parameters

resolution range (Å)
highest resolution shell (Å)
measured reflections
unique reflections
redundancy
completeness (%)
average I/s
 R_{sym} (%)

RSK2 T493M: 26

$P4_12_12$
 $a=47.5$; $b=47.5$; $c=288.5$; $a=90^\circ$;
 $b=90^\circ$; $g=90^\circ$
47.5 – 2.6
2.74 – 2.60
93608 (12711)
10976 (1587)
8.5
97.5 (92.1)
28.2
6.1

Refinement

resolution range (Å)
highest resolution shell (Å)
 R_{work}
 R_{free}
no. protein atoms
no. water molecule
Wilson B-value (Å²)
Average B factors (Å²)
 Protein
 Solvent
 Ligand
Ramachandran (%)
 favoured
 outliers
RMS deviations
 bond lengths (Å)
 bond angles (°)

45.1 – 2.6
2.9 – 2.6
23.5 (29.7)
29.2 (35.4)
4671
41
44.3
57.2
52.7
39.7
54.2
94.9
0.3
0.009
0.99

Table 3-1: Crystallographic Data Collection and Refinement Statistics for compound **26** to RSK2 T493M CTD kinase.

In-cell Fluorescent Competition Assay:

MB-MDA-231 cells were seeded into six-well plates at an initial density of 300,000 cells/well in 2 mL media and grown overnight. Each well was treated with 1 μ M inhibitor (or DMSO) for 2 or 24 hours, followed by treatment with 0.6 mL of 2 μ M BODIPY-fmk for 1 hour. Cells were washed with cold PBS and frozen at -80°C prior to lysis in 60 μ L of lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% Triton X100 supplemented with phosphatase and protease inhibitor cocktails (Roche)). Lysates were clarified by centrifugation at 14,000 rpm, normalized by Bradford assay, denatured in SDS and separated by 7.5% SDS-PAGE. Gels were scanned with a Typhoon imaging system (Amersham Biosciences) and transferred to nitrocellulose with a semi-dry transfer apparatus. Membranes were blocked for 1 hour using Odyssey Blocking Buffer and incubated with 1:500 dilutions of antibodies against RSK1 (rabbit) and RSK2 (mouse) (Santa Cruz Biotechnology, Inc.) at 4°C overnight. Membranes were washed with TBST and incubated for 1 hour with 1:10,000 dilutions of fluorescently labeled secondary antibodies (Odyssey LiCOR) in TBST supplemented with 5% milk. Blots were washed thoroughly and scanned using an Odyssey LiCOR instrument.

Reversibility of compound 26 upon guanidine denaturation:

Compound **26** (4 μ M) in PBS (pH 7.4) was incubated at room temperature for one hour with RSK2 CTD (8 μ M) in a final volume of 100 μ L. The sample was then diluted with 100 μ L of 6M guanidine HCl or PBS, followed immediately by

the addition of 200 μ L of MeCN. This solution was injected onto a Waters Micromass UPLC-MS at varying time points and the amount of **26** recovered monitored by MS-MS analysis. The fraction of compound present relative to the input was calculated from a standard curve made in identical buffer.

Chemical Synthesis:

All purchased chemicals were used as received without further purification. Solvents were dried by passage through columns (either alumina or activated molecular sieves) on a Glass Contour solvent system. NMR spectra were obtained on a Varian Inova 400 MHz spectrometer and referenced to the residual solvent peak. LC-MS analysis was performed on a Waters Acquity LCT UPLC equipped with a TUV detector (monitored at 254 nm) and a Waters Acquity UPLC 1.7 μ m C-18 column, eluting at 0.6 mL/min with a 5 minute water:MeCN (with 0.1% formic acid) gradient method: 0 – 0.1 min, 5% MeCN; 0.1 – 3.8 min, 5 – 95% MeCN; 3.8 – 4.5 min, 95% MeCN; 4.5 – 4.7 min, 95 – 5% MeCN; 4.7 – 5.0 min, 5% MeCN. High resolution mass spectrometry (HRMS) was performed at the QB3 Mass Spectrometry Facility at UC Berkeley.

2-cyano-N-methylacetamide (**15**) To a solution of methyl cyanoacrylate (443 μ L, 5 mmol) was added 10 mL of methylamine (40% in water). After five minutes of stirring, the reaction was complete by TLC. Reaction mixture was evaporated

under reduced pressure to afford 410 mg (80%) of 2-cyano-N-methylacetamide as a white solid that was used without further purification. ^1H NMR (400 MHz, CDCl_3) δ 6.13 (br s, 1H), 3.63 (s, 2H), 2.89 (d, 3H, $J = 4.8$ Hz).

2-cyano-3-(1H-indazol-5-yl)-N-methylacrylamide (**16**) To a solution of 5-formyl-1H-indazole (96 mg, 0.68 mmol) in THF (1 mL) was added 2-cyano-N-methylacetamide (70 mg, 0.68 mmol) and piperidine (67 μL , 0.7 mmol). The solution was stirred at room temperature for 18 hours, monitored by TLC. The reaction was concentrated, taken up in ether and the resulting precipitate filtered to afford 37 mg (23%) of 2-cyano-3-(1H-indazol-5-yl)-N-methylacrylamide as a white solid. ^1H NMR (400 MHz, DMSO) δ 8.42 (s, 1H), 8.30 (br s, 1H), 8.28 (d, 1H, $J = 0.8$ Hz), 8.27 (s, 1H), 8.01 (dd, 1H, $J = 8.8, 1.6$ Hz), 7.70 (d, 1H, $J = 8.8$ Hz), 2.77 (d, 3H, $J = 4.4$ Hz).

2-cyano-3-(1H-indazol-5-yl)-N-isopropylacrylamide (**17**) To a solution of 5-formyl-1H-indazole (107 mg, 0.7 mmol) in THF (2 mL) was added 2-cyano-N-

isopropylacetamide (89 mg, 0.7 mmol) and DBU (130 μ L, 0.87 mmol). The yellowish reaction was stirred overnight with evolution of a yellow precipitate, which was filtered, washed twice with 0.5 mL water, dried and lyophilized from benzene:MeOH to afford 24 mg (13%) of 2-cyano-3-(1H-indazol-5-yl)-N-isopropylacrylamide as a yellow powder. ^1H NMR (400 MHz, DMSO) δ 8.41 (s, 1H), 8.29 (s, 1H), 8.23 (s, 1H), 8.16 (br d, 1H), 7.99 (d, 1H, J = 8.8 Hz), 7.71 (d, 1H, J = 8.8 Hz), 4.02 (m, 1H), 1.17 (d, 6H, J = 6.8 Hz).

3-(1H-indazol-5-yl)-2-(pyrrolidine-1-carbonyl)acrylonitrile (**18**) To a solution of 5-formyl-1H-indazole (119 mg, 0.8 mmol) in THF (2 mL) was added 2-cyano-N-isopropylacetamide (96 mg, 0.62 mmol) and DBU (130 μ L, 0.87 mmol). Reaction was stirred for 18 hours and the resulting precipitate filtered and washed with THF to afford 56 mg (30%) of 3-(1H-indazol-5-yl)-2-(pyrrolidine-1-carbonyl)acrylonitrile as a tan solid. ^1H NMR (400 MHz, DMSO) δ (2 isomers) 8.39 (s, 1H), 8.25 (s, 1H), 8.20 (s, 1H), 8.02 (s, 1H), 7.99 (dd, 1H, J = 8.8, 1.6 Hz), 7.95 (s, 1H), 7.75 (s, 1H), 7.67 (d, 1H, J = 8.8 Hz), 7.58 (d, 1H, J = 8.8 Hz), 7.39 (dd, 1H, J = 8.8, 1.6 Hz), 3.65 (m, 2H), 3.45 (m, 4H), 3.16 (m, 2H), 1.87 (m, 4H), 1.76 (m, 2H), 1.68 (m, 2H). Exact mass: 266.12, M/z found: 267.5 (M+H) $^+$.

2-cyano-3-(1H-indazol-6-yl)-N-methylacrylamide (19) To a solution of 6-formyl-1H-indazole (103.5 mg, 0.7 mmol) in THF (1 mL) was added 2-cyano-N-methylacetamide (70 mg, 0.7 mmol) and DBU (130 μ L, 0.87 mmol). Immediately upon DBU addition, the yellow slurry became a clear orange solution, followed by rapid evolution of an orange precipitate. After an hour of stirring at room temperature, the reaction was concentrated, taken up in a small volume of 1:1 EtOAc:H₂O and sonicated vigorously to break up the residue. The resulting solids were filtered, washed with EtOAc and H₂O and dried to afford 21 mg (13%) of 2-cyano-3-(1H-indazol-6-yl)-N-methylacrylamide as a tan solid. ¹H NMR (400 MHz, DMSO) δ 8.39 (br d, 1H), 8.31 (s, 1H), 8.20 (s, 1H), 8.18 (s, 1H), 7.92 (d, 1H, *J* = 8.8 Hz), 7.68 (d, 1H, *J* = 8.8 Hz), 2.77 (d, 3H, *J* = 4.4 Hz). Exact mass: 226.09, *M/z* found: 227.4 (*M*+H)⁺.

2-cyano-3-(1H-indazol-6-yl)-N-isopropylacrylamide (20) To a solution of 6-formyl-1H-indazole (101.5 mg, 0.7mmol) in THF (1 mL) was added 2-cyano-N-isopropylacetamide (99.9 mg, 0.8 mmol) and DBU (130 μ L, 0.87 mmol). After 1 hour, complete consumption of starting material was observed, monitored by

TLC. The reaction was diluted with EtOAc and washed with H₂O and brine.

Hexanes was added slowly to the remaining organic layer until a light precipitate formed, which was collected by filtration to afford 48 mg (27%) of 2-cyano-3-(1*H*-indazol-6-yl)-*N*-isopropylacrylamide as a white solid. ¹H NMR (400 MHz, DMSO) δ (major isomer) 8.49 (br d, 1H, *J* = 8 Hz), 8.27 (s, 1H), 8.19 (s, 1H), 8.19 (s, 1H, overlapping), 7.93 (d, 1H, *J* = 8.4 Hz), 7.67 (d, 1H, *J* = 8.8 Hz), 4.01 (m, 1H), 1.17 (d, 6H, *J* = 6.4 Hz).

3-(1*H*-indazol-6-yl)-2-(pyrrolidine-1-carbonyl)acrylonitrile (**21**) To a solution of 6-formyl-1*H*-indazole (115 mg, 0.8 mmol) in THF (2 mL) was added 2-cyano-*N*-isopropylacetamide (100 mg, 0.63 mmol) and DBU (130 μ L, 0.87 mmol).

Reaction was stirred for 36 hours, at which point the reaction was concentrated and purified by preparative TLC, eluting with EtOAc, to afford 43 mg (23%) of 3-(1*H*-indazol-6-yl)-2-(pyrrolidine-1-carbonyl)acrylonitrile as a tan solid. ¹H NMR (400 MHz, DMSO) δ (major isomer) 8.20 (s, 1H), 8.19 (s, 1H), 8.09 (s, 1H), 7.92 (d, 1H, *J* = 8.4 Hz), 7.69 (dd, 1H, *J* = 8.4, 1.2 Hz), 3.68 (m, 2H), 3.45 (m, 2H), 1.90 (m, 4H). Exact mass: 266.12, *M/z* found: 267.5 (*M*+H)⁺.

3-(3-(1H-pyrrolo[2,3-b]pyridine-3-carbonyl)phenyl)-2-cyano-N-methylacrylamide

(22) To a solution of 3-(1H-pyrrolo[2,3-b]pyridine-3-carbonyl)benzaldehyde (47 mg, 0.2 mmol) in THF (1 mL) was added 2-cyano-N-methylacetamide (19 mg, 0.2 mmol) and DBU (20 μ L, 0.2 mmol) and the slurry was stirred at room temperature for five minutes. A few drops of MeOH were added to the reaction to keep a gummy precipitate from sticking to the walls, and the reaction was stirred overnight with evolution of a yellowish precipitate. The precipitate was collected by filtration to afford 8 mg (12%) of 3-(3-(1H-pyrrolo[2,3-b]pyridine-3-carbonyl)phenyl)-2-cyano-N-methylacrylamide as a white solid. ^1H NMR (400 MHz, DMSO) δ 8.55 (dd, 1H, J = 8.4, 1.6 Hz), 8.43 (br m, 1H), 8.37 (dd, 1H, J = 8.8, 1.6 Hz), 8.36 (m, 1H), 8.30 (s, 1H), 8.20 (s, 1H), 8.16 (d, 1H, J = 8.4 Hz), 7.99 (m, 1H), 7.74 (t, 1H, J = 8.0 Hz), 7.31 (dd, 1H, J = 8.0, 1.6 Hz), 2.76 (d, 3H, J = 4.4 Hz). Exact mass: 330.11, M/z found: 331.05 ($M+H$) $^+$.

3-(3-(1H-pyrrolo[2,3-b]pyridine-3-carbonyl)phenyl)-2-cyano-N,N-

dimethylacrylamide (23) To a solution of 3-(1H-pyrrolo[2,3-b]pyridine-3-

carbonyl)benzaldehyde (47 mg, 0.2 mmol) in THF (1 mL) was added 2-cyano-*N*-dimethylacetamide (19 mg, 0.2 mmol) and DBU (20 μ L, 0.2 mmol) and the slurry was stirred at room temperature for several hours, at which point the brown slurry became a clear yellow solution. The following day, the reaction was concentrated to a brown, waxy solid, resuspended in EtOAc:MeOH (2:1), and the resultant precipitate was collected by filtration to afford 12 mg (17%) of 3-(3-(1*H*-pyrrolo[2,3-*b*]pyridine-3-carbonyl)phenyl)-2-cyano-*N,N*-dimethylacrylamide as a white solid. δ Exact mass: 330.11, *M/z* found: 331.05 (*M*+*H*)⁺. Exact mass: 344.13, *M/z* found: 345.09 (*M*+*H*)⁺.

2-cyano-*N*-methyl-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide (**24**)

To a solution of 3-(3,4,5-trimethoxyphenyl)-1*H*-indazole-5-carbaldehyde (30 mg, 0.100 mmol) and 2-cyano-*N*-methylacetamide (9.6 mg, 0.100 mmol) in THF (1 mL) was added DBU (10 μ L, 0.07 mmol). The colorless solution became bright yellow upon addition of DBU, and slowly turned bright red over the course of 15 minutes. After four hours, the reaction mixture was taken up in EtOAc (10 mL), and washed with 1M HCl (10 mL), NaHCO₃ (10 mL) and brine (10 mL). The organic layer was concentrated and purified by preparative TLC, eluting with 9:1 CH₂Cl₂:MeOH, to afford 17 mg (43%) of 2-cyano-*N*-methyl-3-(3-(3,4,5-

trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide as a bright yellow solid. ¹H NMR (400 MHz, DMSO) δ 8.86 (s, 1H), 8.39 (s, 1H), 8.33 (br d, 1H, *J* = 4.0 Hz), 7.96 (d, 1H, *J* = 8.8 Hz), 7.75 (d, 1H, *J* = 8.8 Hz), 7.22 (s, 2H), 3.92 (s, 6H), 3.74 (s, 3H), 2.77 (d, 3H, *J* = 4.4 Hz). Exact mass: 392.15, *M/z* found: 393.07 (*M*+*H*)⁺.

2-cyano-N,N-dimethyl-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide

(25) To a solution of 3-(3,4,5-trimethoxyphenyl)-1*H*-indazole-5-carbaldehyde (50 mg, 0.16 mmol) and 2-cyano-N-methylacetamide (17.6 mg, 0.16 mmol) in THF (1 mL) was added DBU (16 μ L, 0.10 mmol). The colorless solution became bright yellow upon addition of DBU, and slowly turned bright red over the course of 15 minutes. After four hours, the reaction mixture was taken up in EtOAc (10 mL), and washed with 1M HCl (10 mL), NaHCO₃ (10 mL) and brine (10 mL). The organic layer was concentrated and purified by preparative TLC, eluting with 9:1 CH₂Cl₂:MeOH, to afford 17 mg (26%) of 2-cyano-*N,N*-dimethyl-3-(3-(3,4,5-trimethoxyphenyl)-1*H*-indazol-5-yl)acrylamide as a bright yellow solid. ¹H NMR (400 MHz, CDCl₃) δ 8.72 (s, 1H), 7.93 (s, 1H), 7.85 (dd, 1H, *J* = 8.8, 1.6 Hz), 7.53 (d, 1H, *J* = 8.8 Hz), 7.18 (s, 2H), 3.96 (s, 6H), 3.92 (s, 3H), 3.15 (br d, 6H). Exact mass: 406.16, *M/z* found: 407.05(*M*+*H*)⁺.

2-(1H-1,2,4-triazol-1-yl)-3-(3-(3,4,5-trimethoxyphenyl)-1H-indazol-5-yl)acrylonitrile (**26**) To a solution of 3-(3,4,5-trimethoxyphenyl)-1H-indazole-5-carbaldehyde (400 mg, 1.28 mmol) and 2-(1H-1,2,4-triazol-1-yl)acetonitrile (140 mg, 1.28 mmol) in THF (5 mL) was added DBU (192 μ L, 1.28 mmol). The yellowish solution became bright orange as the reaction progressed. After stirring overnight, the reaction mixture was concentrated and the crude residue was purified by silica gel chromatography, eluting with 49:49:2 EtOAc:Hexanes:MeOH, to afford 85 mg (17%) of (2-(1H-1,2,4-triazol-1-yl)-3-(3-(3,4,5-trimethoxyphenyl)-1H-indazol-5-yl)acrylonitrile as a bright yellow solid. ^1H NMR (400 MHz, DMSO) δ 9.13 (s, 1H), 8.72 (s, 1H), 8.33 (s, 1H), 8.31 (s, 1H), 7.90 (d, 1H, J = 8.8 Hz), 7.79 (d, 1H, J = 8.8 Hz), 7.22 (s, 2H), 3.92 (s, 6H), 3.74 (s, 3H). ^{13}C NMR (125 MHz, DMSO) δ 153.37, 152.63, 144.78, 143.78, 143.75, 142.25, 138.68, 137.73, 128.44, 127.22, 123.50, 123.35, 120.27, 114.45, 111.68, 105.43, 104.23, 60.13, 56.03. HRMS (ESI) found 401.1362, $\text{C}_{21}\text{H}_{17}\text{N}_6\text{O}_3^-$ ($\text{M}-\text{H}^-$) requires 401.1368. HRMS (ESI) found 401.1362, $\text{C}_{21}\text{H}_{17}\text{N}_6\text{O}_3^-$ ($\text{M}-\text{H}^-$) requires 402.1368.

3-(4-(4-methylpiperazin-1-yl)phenyl)-1*H*-indazole-5-carbaldehyde (**27**) A

microwave reaction vessel was charged with a stirbar and 5-formyl-3-iodo-1*H*-indazole (400 mg, 1.47 mmol), which was dissolved in 6 mL of 5:1 Dioxane:H₂O. This solution was purged with argon for 30 minutes followed by addition of 4-(*N*-Me-piperazyl)phenylboronic pinacol ester (525 mg, 1.75 mmol), K₂CO₃ (460 mg, 4.4 mmol) and Pd(PPh₃)₄ (174 mg, 0.18 mmol). The reaction vessel was purged with argon again, then microwaved at 150 °C for 15 minutes. This reaction was diluted with EtOAc (40 mL) and washed twice with water (20 mL) and once with brine (20 mL). The organic layer was concentrated and the crude residue purified by column chromatography, eluting with EtOAc + 1% triethylamine to afford 212 mg (55%) of 3-(4-(4-methylpiperazin-1-yl)phenyl)-1*H*-indazole-5-carbaldehyde as a bright yellow solid. ¹H NMR (400 MHz, DMSO) δ 10.10 (s, 1H), 8.66 (s, 1H), 7.95 (d, 2H, *J* = 8.8 Hz), 7.91 (d, 1H, *J* = 8.4 Hz), 7.70 (d, 1H, *J* = 8.4 Hz), 7.07 (d, 2H, *J* = 8.8 Hz), 3.26 (m, 4H), 2.52 (m, 4H).

tert-butyl 5-(2-cyano-2-(1H-1,2,4-triazol-1-yl)vinyl)-3-iodo-1H-indazole-1-carboxylate (28) To a solution of tert-butyl 5-formyl-3-iodo-1H-indazole-1-carboxylate **30** (834 mg, 2.24 mmol) was added 1,2,4-triazolylacetonitrile (242 mg, 2.24 mmol) and DBU (335 μ L, 2.24 mmol). This reaction almost immediately turned yellow and precipitate began forming. After one hour of stirring, the precipitate was filtered to afford 344 mg (33%) of tert-butyl 5-(2-cyano-2-(1H-1,2,4-triazol-1-yl)vinyl)-3-iodo-1H-indazole-1-carboxylate as a slightly yellow solid. ^1H NMR (400 MHz, DMSO) δ (major isomer) 9.25 (s, 1H), 8.15 (s, 1H), 8.02 (d, 1H, J = 8.4 Hz), 7.58 (d, 1H, J = 8.8 Hz), 7.45 (s, 1H), 6.29 (s, 1H), 1.61 (s, 9H).

3-(3-(4-(4-methylpiperazin-1-yl)phenyl)-1H-indazol-5-yl)-2-(1H-1,2,4-triazol-1-yl)acrylonitrile (29) First procedure: To a solution of 3-(4-(4-methylpiperazin-1-

yl)phenyl)-1*H*-indazole-5-carbaldehyde (150 mg, 0.468 mmol) in THF (2 mL) was added 1,2,4-triazolylacetonitrile (50 mg, 0.468 mmol). DBU (70 μ L, 0.468 mmol) was slowly added and the yellow solution became more bright yellow. The reaction was left stirring overnight, at which point it was concentrated and purified by preparatory TLC, eluting with 9:9:2 EtOAc:Hexanes:MeOH + 1% TEA to afford 13 mg (7%) of 3-(3-(4-(4-methylpiperazin-1-yl)phenyl)-1*H*-indazol-5-yl)-2-(1*H*-1,2,4-triazol-1-yl)acrylonitrile as a brilliant yellow solid.

Second procedure: A microwave reaction vessel was charged with tert-butyl 5-(2-cyano-2-(1*H*-1,2,4-triazol-1-yl)vinyl)-3-iodo-1*H*-indazole-1-carboxylate (410 mg, 0.887 mmol), 4-(*N*-Me-piperazyl)phenylborinic pinacol ester (294 mg, 0.975 mmol) and saturated Na₂CO₃ (2 mL) in 10:1 Toluene:EtOH (6 mL) and the slurry purged with argon for 30 minutes. PdCl₂dppf (65 mg, 0.088 mmol) was added, the vessel was purged with argon again, and the reaction was microwaved at 80 °C for 30 minutes. The crude reaction was diluted in EtOAc (40 mL), washed with water (20 mL) and brine (20 mL), the organics dried with sodium sulfate and concentrated, and the residue was chromatographed in 4:1 EtOAc:MeOH +1% triethylamine to afford 357 mg (79%) of the BOC-**29** product. This was directly deprotected by stirring for four hours in 3*M* HCl, which was then diluted into large volumes of EtOAc, washed with NaHCO₃, and the organics concentrated to afford 229 mg (80%, 63% over the two steps) of 3-(3-(4-(4-methylpiperazin-1-yl)phenyl)-1*H*-indazol-5-yl)-2-(1*H*-1,2,4-triazol-1-yl)acrylonitrile as a brilliant yellow solid. ¹H NMR (400 MHz, DMSO) δ 9.13 (s, 1H), 8.64 (s, 1H), 8.32 (s, 1H),

8.29 (s, 1H), 7.97 (d, 1H, $J = 8.4$ Hz), 7.86 (d, 2H, $J = 8.8$ Hz), 7.74 (d, 1H, $J = 8.4$ Hz), 7.08 (d, 2H, $J = 8.8$ Hz), 3.24 (m, 4H), 2.50 (m, 4H), 2.24 (s, 3H).

tert-butyl 5-formyl-3-iodo-1H-indazole-1-carboxylate (**30**) To a solution of 3-iodo-5-formyl indazole (500 mg, 1.84 mmol) in MeCN (10 mL) was added DMAP (11 mg, 0.092 mmol) and triethylamine (384 μ L, 2.76 mmol), which was stirred for 15 minutes. Then, Boc₂O was added slowly (602 mg, 2.76 mmol). The solution was stirred for 36 hours at room temperature, at which point the reaction was diluted with 40 mL EtOAc, washed twice with 20 mL 1M HCl, once with 20 mL saturated NaHCO₃, and once with 20 mL brine. The organic layer was concentrated and the crude residue purified by column chromatography, eluting with 4:1 Hexanes:EtOAc to afford 507 mg of tert-butyl 5-formyl-3-iodo-1H-indazole-1-carboxylate as a white solid (98%). ¹H NMR (400 MHz, DMSO) δ 10.12 (s, 1H), 8.08 – 8.17 (m, 3H), 1.66 (s, 9H).

tert-butyl 5-formyl-3-(4-(4-methylpiperazin-1-yl)phenyl)-1H-indazole-1-carboxylate (31) To a solution of 3-(4-(4-methylpiperazin-1-yl)phenyl)-1H-indazole-5-carbaldehyde (160 mg, 0.5 mmol) in MeCN (2 mL) was added DMAP (3 mg, 0.085 mmol) and triethylamine (104 μ L, 0.75 mmol), which was stirred for 15 minutes. Then, Boc₂O was added slowly (164 mg, 0.75 mmol). The solution was stirred for 24 hours at room temperature, at which point the reaction was diluted with 20 mL EtOAc, washed twice with 10 mL 1M HCl, once with 10 mL saturated NaHCO₃, and once with 10 mL brine. The organic layer was concentrated and the crude residue purified by column chromatography, eluting with 9:9:2 Hexanes:EtOAc:MeOH +1% triethylamine to afford 86 mg (40%) of tert-butyl 5-formyl-3-(4-(4-methylpiperazin-1-yl)phenyl)-1H-indazole-1-carboxylate as a yellow solid. ¹H NMR (400 MHz, CDCl₃) δ 10.07 (s, 1H), 8.45 (s, 1H), 8.27 (d, 1H, *J* = 8.4 Hz), 8.02 (d, 1H, *J* = 8.4 Hz), 7.89 (d, 2H, *J* = 8.8 Hz), 7.01 (d, 2H, *J* = 8.8 Hz), 3.30 (m, 4H), 2.58 (m, 4H), 2.33 (s, 3H), 1.72 (s, 9H).

3.5 References

1. Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. *Nat. Chem. Biol.* **2012**, 8, 471–476.
2. Serafimova, I. M. Ph.D. Thesis entitled “Reversible Targeting of Noncatalytic Cysteines with Chemically Tuned Electrophiles.” **2012**
3. Couch, R. D. et al. Studies on the reactivity of CDDO, a promising new chemopreventive and chemotherapeutic agent: Implications for a molecular mechanism of action. *Bioorg. Med. Chem. Lett.* 15, 2215–2219 (2005).
4. Powers, J. P. et al. SAR and Mode of Action of Novel Non-Nucleoside Inhibitors of Hepatitis C NS5b RNA Polymerase. *J. Med. Chem.* 49, 1034–1046 (2006).
5. Patch, R.J. et al. Identification of Diaryl Ether-Based Ligands for Estrogen-Related Receptor α as Potential Antidiabetic Agents. *J. Med. Chem.* 54, 788–808 (2011).

6. Fleming, F. F., Yao, L., Ravikumar, P. C., Funk, L. & Shook, B. C. Nitrile-Containing Pharmaceuticals: Efficacious Roles of the Nitrile Pharmacophore. *J. Med. Chem.* 53, 7902–7917 (2010).
7. Meanwell, N. A. Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *J. Med. Chem.* 54, 2529–2591 (2011).
8. Friedman, M., Krull, L. H. & Cavins, J. F. The Chromatographic Determination of Cystine and Cysteine Residues in Proteins as S- β -(4-Pyridylethyl)cysteine. *J. Biol. Chem.* 245, 3868-3871 (1970).

**Chapter 4. Computational, Fragment-Based Design of
Reversible Covalent, Cell-Active Cyanoacrylamide
Inhibitors of RSK2 and MSK1**

4.1 Abstract

Chapter 2 describes the development and application of a cyanoacrylamide fragment-based method for the development of reversible covalent MSK and RSK kinase inhibitors. However, our implementation of this method suffered from several limitations. First, the panel of compounds used in this study was small in size. Scaling to a larger library requires the chemical synthesis of each member, consuming time and resources. Second, the chemical diversity of an experimentally derived library is inherently limited and biased by the chemist's selection process. We sought to improve the throughput and diversity of a covalent fragment based approach by adapting the virtual screening methodology, DOCK. Such an approach would enable the computational assembly of extremely large and diverse compound libraries, and facilitate the rapid virtual screening of these compounds against a target of interest. The selection of the best candidate inhibitors from such a screen significantly reduces the number of compounds that must be synthesized and tested experimentally. Using DOCKoValent, a covalent virtual screening strategy developed by Dr. Nir London and Prof. Brian Shoichet (UC San Francisco), we were able to screen a 12,000-member library of cyanoacrylamides computationally to identify novel potent, cell-active inhibitors of MSK1 and RSK2 C-terminal kinase domains.

Note: This chapter was co-written with Dr. Nir London and Profs. Brian Shoichet and Jack Taunton (UC San Francisco). This work has been submitted and is under review as of publication of this thesis: London, N.; Miller, R. M.; Irwin, J. J.; Eidam, O.; Gibold, L.; Bonnet, R.; Shoichet, B. K.; Taunton, J. Covalent Docking of Large Libraries for the Discovery of Chemical Probes, *submitted manuscript*, **2013**.

4.2 Introduction

Small-molecule modulators of protein activity are useful as tools for investigating biological questions and as leads for drug discovery. Development of genuinely useful probes typically involves iterative rounds of medicinal chemistry to optimize the potency and selectivity of initial active molecules. A strategy for enhancing both properties is via covalent bond formation with a nucleophilic residue that is specific to a target of interest and ideally absent from off-targets. Such covalent-acting chemical probes have increasingly been used in proteome-wide target identification¹, visualization² and for finding inhibitors with high specificity among related enzymes and enzyme isoforms.^{3,4} A challenge in developing covalent probes is identifying reactive functional groups ("warheads") that do not make the molecule so reactive as to be promiscuous. Less recognized is the challenge of screening a wide variety of scaffolds for optimal presentation of such reactive functionality.

The most widely used technique for ligand discovery in drug development and chemical biology is high-throughput screening (HTS), and one could potentially screen extant libraries for new small molecules that react covalently. However, protein-reactive compounds are typically shunned in high-throughput screening efforts⁵ or flagged as likely artifacts due to concerns about promiscuous activity.^{6,7} Whereas this is often sensible in the context of certain drug discovery campaigns, it removes from consideration potential starting points for developing useful covalent chemical probes, which have a proven track record in chemical biology and drug discovery.

Covalent ligands can be divided into two classes: those targeting catalytically essential nucleophiles in enzymes (e.g., serine and cysteine hydrolases), and those targeting noncatalytic nucleophiles (usually cysteine), which are frequently found in small-molecule binding sites on proteins, including proteins without enzymatic activity (e.g., GPCRs and nuclear receptors).⁸ Each class can be further subdivided into ligands that bind reversibly or irreversibly, depending on the electrophile and nucleophile. In all cases, specific noncovalent interactions contributed by the scaffold portion of the ligand are critical for orienting the electrophilic warhead relative to the protein nucleophile, thereby increasing the rate (and stability, in the case of reversible covalent ligands) and selectivity of covalent bond formation. A key unsolved problem in the discovery of reversible and irreversible covalent probes is how to identify a protein-binding scaffold that optimally orients the warhead, while minimizing the number of compounds that must be synthesized and tested.

In principle, structure-based docking screens⁹⁻¹¹ can address the gap left by HTS and its large libraries of synthesized small molecules. Given the structure of a protein target with a defined binding site, docking programs¹² computationally screen large libraries for small molecules predicted to form favorable binding interactions. Docking has been widely used for the discovery of reversible, non-covalent ligands.¹³⁻¹⁵ To date, there have been few attempts to perform covalent virtual screens with electrophilic small molecules,¹⁶ and no prospective screens have been conducted with more than a few hundred compounds.¹⁷ Moreover, covalent docking has yet to be applied prospectively to

noncatalytic cysteines, targeted by an important and growing class of covalent drugs.¹⁹

Here, we adapt the non-covalent docking program DOCK to enable large-scale, covalent virtual screening of electrophilic small molecules, including low-molecular weight electrophilic fragments (Figure 4-1). Notably, we report the first computational screen of electrophilic fragments against a non-catalytic, rare cysteine residue. Prospective covalent docking against non-catalytic residues will result in the identification of more selective compounds relative to those targeting conserved, catalytic nucleophiles. The ability to target non-conserved residues makes covalent docking generally applicable to any target that contains a nucleophilic residue near a suitable binding site. By screening large virtual libraries comprising two unrelated sets of electrophiles, we discover novel, ligand-efficient, reversible covalent inhibitors that exhibit potent on-target effects in cells.

4.3 Results and Discussion

Our method begins with the construction of large virtual libraries of commercially available (or easily synthesized) small molecules, all of which contain an electrophilic warhead suitable for a covalent probe. All stereoisomers, protonation states and conformations of the covalent adduct are pre-generated for each ligand, enabling rapid docking of the library to any target of interest (Figure 4-1).

Figure 4-1: Experimental outline for covalent docking. (A) A library of commercially available or easily synthesized small molecules containing a specific warhead is constructed virtually. In this example, the cyanoacrylamide warhead is shown in red. (B) All stereoisomers, protonation states and conformations of each ligand are pre-generated. (C) Conformational space is exhaustively sampled around the covalent bond for each pre-generated ligand state, and each pose is scored using a physics-based energy function. (D) Each molecule is represented by its best scoring pose, and high-ranking candidates are manually selected for experimental validation.

For each ligand, DOCKoValent exhaustively samples all poses and ligand conformations with respect to the covalent bond to the target nucleophile, constrained by ideal bond lengths and angles (Figure 4-1C). Each sampled conformation is scored using DOCK's physics-based scoring function¹⁹ and the best scoring pose for each ligand is saved. The library is then ranked by this docking score (Figure 4-1D). The top 1-3% of the list is curated to remove misdocked ligands inevitable in any large scale automated docking, compounds with incorrect ionization states, or overly strained conformations.²⁰ The remaining compounds are filtered for novelty, diversity, and availability. Molecules that pass these filters are then selected for testing. An advantage of DOCKoValent is that it plugs seamlessly into the existing DOCK infrastructure, enabling the rapid construction and docking of large virtual libraries containing tens or hundreds of thousands (even millions) of chemical entities.

The C-terminal kinase domain (CTD) of p90 ribosomal protein S6 kinase-2 (RSK2) contains a noncatalytic active-site cysteine shared by only 11 of the 518 human protein kinases. Starting with an established kinase inhibitor scaffold, we previously designed irreversible³ and reversible²¹ covalent inhibitors that target this cysteine (Cys436). To achieve reversible covalent inhibition of a noncatalytic cysteine, we exploited the atypical reactivity of cyanoacrylamide Michael acceptors, which react rapidly and reversibly with cysteine thiols at physiological pH. RSK2, and the related paralog MSK1, are attractive therapeutic targets implicated in tumor metastasis,²²⁻²⁴ neurodegeneration,²⁵ and atherosclerosis²⁶ among other pathological conditions.

By testing a focused set of low molecular weight cyanoacrylamides, followed by structure-based merging and optimization of the hits, we recently developed the first inhibitor of the MSK1 CTD.²⁷ Given the limited number of compounds in our collection, along with their strong bias toward known kinase binders, we were motivated to test whether DOCKovalent could identify novel RSK2/MSK1 inhibitors by screening thousands of cyanoacrylamide fragments computationally.

As an initial blind test, we used our covalent docking method to predict the poses of two cyanoacrylamide fragments bound to RSK2, prior to obtaining the x-ray coordinates. The predicted binding modes closely recapitulated the experimental structures (Figure 4-2, 1.93 Å and 1.56 Å RMSD, respectively).

Figure 4-2: Blind docking predictions of two cyanoacrylamide fragments (**1** and **8**) from Chapter 2 covalently bound to RSK2 (magenta) recapitulate their crystallographic poses (yellow, PDB: 4JG7,4JG6). Performed by Dr. Nir London and Prof. Brian Shoichet.

Retrospective docking of two larger cyanoacrylamides produced similar results (Figure 4-3, 0.66Å and 1.52Å RMSD). In each of the predictions, the scaffold portion of the molecule, which forms critical non-covalent interactions

with RSK2, closely matched the x-ray structures (0.91–1.36 Å RMSD). The nitrile and carboxamide substituents, which are poorly ordered in the crystal structures, were not predicted as accurately.

Figure 4-3: DOCKoValent predictions (magenta) accurately recapitulate the crystallographic poses (yellow) of optimized cyanoacrylamide inhibitors of RSK2. (A) Recapitulation of ligand binding in PDB: 4D9T with an RMSD of 0.66Å overall (0.48Å over the scaffold alone). (B) Recapitulation of ligand binding in PDB: 4JG8 with an RMSD of 1.52Å overall (0.91Å for the scaffold alone). Performed by Dr. Nir London and Prof. Brian Shoichet.

Encouraged by these initial pose predictions, we sought to screen for novel cyanoacrylamide inhibitors of these kinases. Cyanoacrylamide fragments are rare in commercial collections (602 out of 474,770 of the “fragments in-stock” in ZINC.²⁸ However, β -substituted cyanoacrylamides can be synthesized in one step by condensing aldehyde fragments with cyanoacetamide. We therefore assembled ~12,000 aldehyde fragments from ZINC (molecular weight < 250 Da). These were converted *in silico* to generate a virtual library of cyanoacrylamide fragments. Because our method can generate and screen thousands of cyanoacrylamides, it drastically reduces the number of compounds that must be chemically synthesized prior to experimental testing.

We docked this library against Cys436 of RSK2. Given that nearly all kinase inhibitors form one or more hydrogen bonds with the hinge region (backbone of Met496 and Glu494 in RSK2), we increased the polarity of these groups to favor electrostatic interactions, a technique we have long used.²⁹ We manually inspected the top-ranking compounds and pursued eight virtual cyanoacrylamide fragments ranked between 96 and 391 (top 3%, Figure 4-4).

Figure 4-4: Chemical structures of cyanoacrylamide fragments selected for synthesis and testing.

The corresponding aldehydes were purchased and converted to the cyanoacrylamides, which were tested against wild-type RSK2 and the T493M gatekeeper mutant (Figure 4-5). We have previously used this mutant as a biochemical surrogate for MSK1.²⁷ Five of the eight high-ranking cyanoacrylamides inhibited RSK2 with IC₅₀ values less than 10 μ M under

stringent assay conditions in the presence of competing ATP (0.1 mM) and reduced glutathione (10 mM). Pyridine **34** and pyrazole **37** were the most potent against wild-type RSK2 and inhibited the T493M mutant with submicromolar potency (IC₅₀ 430 nM and 370 nM, respectively).

Compound	DOCK rank	IC ₅₀ (μM)	
		RSK2 WT	RSK2 T493M
32	66	50.4	27.9
33	96	7	5.2
34	122	1.1	0.43
35	132	3.3	6.8
36	142	12.7	6.4
37	200	1.2	0.37
38	368	17.9	12.8
39	391	6	7.1

Figure 4-5: Docking rank and *in vitro* IC₅₀ values for cyanoacrylamides **32** – **39** against RSK2 WT and T493M mutant C-terminal kinase domain.

We determined the co-crystal structure of compound **37** bound to T493M RSK2 at 3.0 Å resolution. Even at this modest resolution, the electron density allowed unambiguous modeling of the phenylpyrazole fragment and the covalent bond to Cys436 (Figure 4-6). The crystallographic structure superposed well with the docking prediction, with an RMSD of 0.86 Å over the phenylpyrazole fragment. The pyrazole moiety forms two hydrogen bonds with the hinge region (N1 to the carbonyl of Glu494 and N2 to the amide NH of Met496). The pyrazole ring of **37** also packs edgewise against the methionine gatekeeper, and the additional van der Waals contacts may explain the enhanced potency for the T493M mutant.

Figure 4-6: Docking predictions for cyanoacrylamide hits **34** and **37**. (A) Docking prediction for the most potent fragment **37** corresponds well to the experimental structure. The $F_o - F_c$ omit map is shown in green. (B) Docking prediction of the binding mode of compound **34**.

As further validation of the utility of our virtual screening method, we tested cyanoacrylamides **34** and **37** for activity in mammalian cells stimulated with phorbol myristate acetate, which activates kinase cascades upstream of RSK2 and MSK1. Both compounds inhibited the activating autophosphorylation of wild-type RSK2 and MSK1, with somewhat greater potency toward MSK1

(EC₅₀ < 5 mM; Figure 4-7). A cysteine to valine mutation in the C-terminal kinase domain of MSK1 (C458V) conferred complete resistance to both inhibitors, providing genetic proof for on-target efficacy. Hence, the unoptimized cyanoacrylamide fragments identified by covalent docking inactivate the desired targets without affecting upstream kinases such as PKC, RAF, MEK, and ERK.

Figure 4-7: Compounds **37** and **34** inhibit the activating autophosphorylation of RSK2 and MSK1 in stimulated mammalian cells. Neither compound inhibits the cysteine to valine mutant of MSK1 at concentrations up to 20 μM.

Covalent probes play a crucial role in chemical biology,¹⁻⁴ but there are few ways to screen for them. Hence, the principal goals of this study were: (1) construct large, easily accessible libraries of electrophilic fragments, and (2)

devise a facile way to screen these fragments for structural complementary to a target of interest using covalent molecular docking.

Established docking programs such as GOLD and AutoDock allow users to model small molecules as covalent adducts to proteins, and several papers have reported the adaptation of these rudimentary covalent bond parameters. One study employed GOLD to dock 343 known molecules against Cathepsin K, 39 of which were tested experimentally.¹⁷ Three inhibitors displayed IC_{50} values < 1 μ M against Cathepsin K by targeting the conserved catalytic cysteine, and two of them were known ligands for the related cysteine protease, Cruzain. Another purely retrospective computational study improved the covalent modeling module of AutoDock, which resulted in better crystallographic pose recapitulation relative to both AutoDock and GOLD.¹⁶

DOCKoValent represents several advancements over these previous covalent docking methods. Our method was designed specifically to compile and screen very large digital compound libraries. We report five retrospective and two prospective screens comprised of more than 420,000 molecules, a scale made possible by the inherent efficiency of DOCK and the integration of rapid library generation using ZINC. We employ a broad spectrum of warhead libraries, each with tens to hundreds of thousands of members. These include composite electrophile libraries like the cyanoacrylamides, which contain unprecedented inhibitors that are synthetically accessible from commercially available starting materials. Practically, such libraries dramatically reduce the number of compounds that must be chemically synthesized and tested. Additionally, we

report the first computational screen of covalent fragments (MW < 250 da), enabling the high-throughput identification of ligand-efficient compounds suitable for extensive optimization. DOCKoValent is the only covalent virtual screening method is available through an automated web-server, and is therefore accessible to researchers in diverse fields. Finally, our method is not merely a proof that covalent docking can be used to select known binders from a pool of decoys, but provides practical evidence that it can be used to identify novel molecules with cellular activity.

Our confidence in the utility of our method is supported by the discovery of new chemotypes in two wholly prospective docking screens. The hit rate was high, and the most potent kinase inhibitors were highly selective in cells, being inactive against the C458V mutant of MSK1, as well as the upstream kinases, MEK and ERK, which also have non-catalytic cysteines in their active sites.

Whereas most of the initial hits do not have the potency typically associated with optimized chemical probes, the best examples from each screen were sufficiently potent and selective to be genuinely useful in cell-based assays. Hits from our digital screen were remarkably ligand efficient, with LE values as high as 0.66, and these hits were optimizable. The origins of the selectivity and affinities of these ligands are captured by their docking poses, which superpose well with the subsequently determined x-ray structures. The high fidelity of the docking to the x-ray structures suggests that the covalently docked poses alone may guide ligand optimization in cases where obtaining experimental structures is challenging.

One of our concerns at the outset was that, in mixing covalent and non-covalent scoring terms, the former would overwhelm the latter.³⁰ To avoid this, we did not explicitly score the covalent bond energy of the adducts, but rather imposed distance and angle constraints on bond formation and then scored the docked poses by our standard physics-based scoring function. An advantage is that the scoring function does not become dominated by the covalent bond to the warhead, and instead favors structural complementarity and specificity provided by noncovalent interactions with the scaffold. The approach is especially well-suited to the reversible, and hence thermodynamically driven, inhibitors discovered for β -lactamase and RSK2. Nevertheless, this approach can be extended to model high-energy intermediate states of the covalent bond-forming reaction,³¹ which may also be treated as a pseudo-equilibrium event. Preliminary studies suggest that the method may be suited to irreversible Michael acceptors, such as acrylamide-based inhibitors of the EGFR kinase.

The failure to explicitly model the energy of covalent bond formation is one of several methodological gaps that merit mention. Incorporating quantum-level approaches to quantitatively predict the reactivity of specific scaffold/warhead combinations may be a direction forward.³² Even within our current constraint-based approach, sampling can be improved by more efficient minimization of ligand dihedral degrees of freedom, and perhaps by accounting for bonded energy terms that can be derived from established force fields. Lastly, accounting for receptor flexibility may be essential for accurate modeling of receptor/ligand interactions. This remains an area of active research in the field.^{9, 33,34}

These caveats should not overshadow the potential benefits offered by this and related covalent docking methods.^{16,17} We hope it will begin to address what has been a substantial gap in our ability to screen widely for useful covalent chemical probes.

4.4 Methods

Ligand Generation (Performed by Dr. Nir London and Prof. Brian Shoichet).

Following DOCK's methodology, ligand flexibility is sampled by generating ligand conformations prior to docking. Given a compound SMILES string of a ligand with a specific warhead, we use the OEChem library (OpenEye Software, Santa Fe NM)³⁵ to convert the ligand to its final, reacted form (either the product or high energy intermediate). This process is particular to each class of warhead. The receptor's nucleophilic atom involved in the covalent bond is represented by a dummy atom (silicon, for technical reasons). Following the generation of the 'reacted' warhead, the ligand's 3D structures and stereoisomers are built by Corina³⁶ (Molecular Networks, Bechtesgarten, Germany), and then protonated and tautomerized by EPIK³⁷ (Schrodinger software, Catsville NY). Partial atomic charges and solvation energies are calculated for each of these structures with AMSOL.³⁸ The warhead serves as a starting rigid body fragment, and conformations are generated for the rest of the ligand's degrees of freedom using Omega.³⁹ Owing to faster sampling compared to regular docking—the ligand is anchored by the covalent bond, eliminating pose sampling—we are able to sample more ligand conformations (Omega parameters: EnergyWindow=30.0; MaxConfs=10,000; RMSThreshold=0.5). The collection of pre-generated ligand conformations in the reacted state is saved to a DOCK-readable flexibase format file using a modified version of the program mol2db.⁴⁰

Sampling (Performed by Dr. Nir London and Prof. Brian Shoichet). Sampling of ligand poses within the protein binding site is restricted to exhaustive ligand placement with respect to the covalent bond. The covalent attachment point is sampled in steps of 20° around the terminal dihedral of the nucleophilic side chain. Based on the covalent warhead geometry determined during ligand generation, as well as user provided parameters, the vectors of the covalent bond from the ligand and receptor sides are aligned and the ligand is exhaustively rotated around this vector in 20° steps. For each such placement, all of the pre-generated ligand conformations are scored and the score for the best pose is saved. This process is repeated for different values of the covalent bond length and angles, centered on ideal values. The magnitude of deviation from the ideal values, as well as the step size, is user specified.

Scoring (Performed by Dr. Nir London and Prof. Brian Shoichet). Scoring is performed as previously described, using pre-calculated van der-Waals, electrostatic, and ligand solvent-excluded desolvation grids, correcting for ligand desolvation.⁴¹ Receptor structures were prepared using an automated procedure as described in ref. 42 using DELPHI⁴³ for electrostatics grid generation. The ligand's electrophilic atom participating in the bond is omitted from the overall ligand score. As mentioned in the Discussion, we anticipate exploring a harmonic restraint-based energy term for this bond in the future.

Virtual ligand libraries used in this study (Performed by Dr. Nir London and Prof. Brian Shoichet). For the curation of ligand libraries, the warheads were represented by SMARTS regular expressions (Supplementary Table 6). The Full, Lead-like, and Fragment-like subsets of the ZINC database of commercially available molecules were filtered using these patterns to identify warhead-bearing molecules.⁴⁴

Prospective Covalent Virtual Screening (Performed by Dr. Nir London and Prof. Brian Shoichet). The RSK2 screen used PDB: 4D9T for the receptor structure. The backbone amides of the kinase hinge residues (Glu494 and Met496) were polarized to emphasize hinge-binding hydrogen bonds. Covalent bond sampling parameters were set to $d=2.0\text{\AA}$ $a=109.5^\circ\pm10^\circ$ and $b=109.5^\circ\pm10^\circ$ (Step size of 2.5° ; Supplementary Fig. 1). This bond length was chosen based on the two available RSK2 cyanoacrylamide complexes (PDBs: 4D9U, 4D9T). While longer than a typical thioether by about $0.2\text{\AA}$, this helps ensure that the scores are dominated by the non-covalent terms of the scoring function, minimizing hard van der Waals repulsion in the region of the new bond. We have not attempted to optimize this term. Calculation time for the docking screen was 103 CPU hours, elapsed wall time was less than 1 hour owing to use of a cluster.

RMSD calculations (Performed by Dr. Nir London and Prof. Brian Shoichet). RMSD values were calculated using the Hungarian matching algorithm as implemented in DOCK6.⁴⁵ For comparison purposes, RMSDs for the β -lactam

pose recovery benchmark were calculated using software generously provided by Xuchang Ouyang.⁴⁶ For AmpC docking predictions, receptors were superimposed based on chain A before calculations. For RSK2, receptors were superimposed based on C436, M496, and C560.

Crystallography. Co-crystals of RSK2 T493M in complex with compound **37** (PDB: 4M8T) were grown by hanging drop diffusion as described in Chapter 2.3,²⁷ and diffraction data were collected at the ALS on beamline 8.2.2. Diffraction images were indexed and integrated using XDS. Molecular replacement was performed using apo-RSK2 CTD as a starting model (PDB: 2QR8) using Phaser, and TLS refinement was carried out using the Phenix.Refine module of the Phenix suite.

	RSK2 T493M/ 37
PDB Code	4M8T
Space Group	P4 ₁ 2 ₁ 2
Cell dimensions a, b, c [Å] α , β , γ [°]	46.99, 46.99 291.10 90, 90, 90
Resolution [Å] ^a	46.38 – 3.0
R _{merge} [%] ^a	18.5 (20.9)
Completeness [%] ^a	98.7 (99.9)
I/ σ I ^a	13.1 (6.0)
Redundancy ^a	23.4 (26.1)
Resolution [Å] ^a	46.39 – 3.00 (3.78 – 3.00)
No. of reflections / Size of test set	7265 / 348
R _{work} / R _{free} [%] ^a	25.3 (29.8) / 31.3 (38.0)
No. atoms Protein Ligand/Ions Water	2370 20 0
Avg. B-factors [Å ²] ^b Protein Ligand/Ions Water	59.5 55.3 55.1
R.m.s. deviations bond lengths [Å] bond angles [°]	0.011 0.91

Table 4-1: Crystallographic Data Collection and Refinement Statistics for compound **37** bound to RSK2 T493M.

Synthetic Chemistry. All purchased chemicals were used as received without further purification. Solvents were dried by passage through columns (either alumina or activated molecular sieves) on a Glass Contour solvent system. NMR spectra were obtained on a Varian Inova 400 MHz spectrometer and referenced to the residual solvent peak. LC-MS analysis was performed on a Waters Acquity LCT UPLC equipped with a TUV detector (monitored at 254 nm) and a Waters Acquity UPLC 1.7 μ m C-18 column, eluting at 0.6 mL/min with a 5 minute water:MeCN (with 0.1% formic acid) gradient method.

3-(4-((1H-1,2,4-triazol-1-yl)methyl)thiophen-2-yl)-2-cyanoacrylamide (**32**)

4-((1H-1,2,4-triazol-1-yl)methyl)thiophene-2-carbaldehyde (25 mg, 0.129 mmol) was dissolved in THF (0.5 mL) in a vial with a stirbar, to which was added 2-cyanoacetamide (11 mg, 0.129 mmol) and piperidine (12 μ L, 0.129 mmol). The reaction was stirred for 7 hours. The precipitate was collected by filtration and dried *in vacuo* to afford cyanoacrylamide **32** (17 mg, 51%) as a tan solid. ^1H NMR (400 MHz, DMSO) δ 8.64 (s, 1H), 8.32 (s, 1H), 8.01 (s, 1H), 7.94 (s, 1H), 7.79 (br s, 1H), 7.74 (s, 1H), 7.69 (br s, 1H, overlaps peak at 7.74), 5.47 (s, 2H). ^{13}C NMR (125 MHz, DMSO) δ 162.52, 151.82, 144.18, 143.09, 137.99, 137.00, 136.31, 132.46, 116.46, 102.73, 46.93 LRMS. (ESI) Exact Mass: 259.05, Found: 260.3 ($\text{M} + \text{H}^+$).

Compounds **33** - **39** were synthesized analogously and purified by filtration or preparative TLC, eluting with hexanes/EtOAc. Yields refer to chromatographically and spectroscopically pure compounds.

4-(3-amino-2-cyano-3-oxoprop-1-en-1-yl)benzamide (**33**)

Yield: 47 mg (65%) ^1H NMR (400 MHz, DMSO) δ 8.22 (s, 1H), 8.12 (br s, 1H), 7.95 – 8.03 (m, 5H), 7.82 (br s, 1H), 7.56 (br s, 1H). ^{13}C NMR (125 MHz, DMSO) δ 167.0, 162.5, 149.7, 137.1, 134.4, 129.8, 128.2, 116.3, 108.2. LRMS (ESI) Exact Mass: 215.07, Found: 216.3 ($\text{M} + \text{H}^+$).

2-cyano-3-(3-(pyridin-4-yl)phenyl)acrylamide (**34**)

Yield: 33 mg (46%). ^1H NMR (400 MHz, DMSO) δ 8.69 (m, 2H), 8.34 (m, 1H), 8.31 (s, 1H), 8.05 - 8.60 (m, 2H), 7.96 (br s, 1H), 7.83 (br s, 1H), 7.70 – 7.76 (m, 3H). ^{13}C NMR (125 MHz, DMSO) δ 162.47, 150.44, 150.36, 145.98, 138.0,

132.88, 130.46, 130.16, 128.59, 121.25, 116.43, 107.65 (one carbon not observed). LRMS (ESI) Exact Mass: 249.09, Found: 250.2 ($M + H^+$).

3-(3-amino-2-cyano-3-oxoprop-1-en-1-yl)benzamide (**35**)

Yield: 36 mg (50%). ^1H NMR (400 MHz, DMSO) δ 8.37 (m, 1H), 8.23 (s, 1H), 8.09 (br s, 1H), 8.07 (s, 1H), 8.02 – 8.06 (m, 1H), 7.97 (br s, 1H), 7.81 (br s, 1H), 7.65 (t, 1H, $J = 8\text{Hz}$), 7.54 (br s, 1H). ^{13}C NMR (125 MHz, DMSO) δ 167.1, 162.6, 150.1, 135.3, 132.1, 132.0, 130.8, 129.7, 129.3, 116.2, 107.7. LRMS (ESI) Exact Mass: 215.07, Found: 216.3 ($M + H^+$).

2-cyano-3-(isoquinolin-6-yl)acrylamide (**36**)

Yield: 86 mg (60%). ^1H NMR (400 MHz, DMSO) δ 9.42 (s, 1H), 8.63 (d, 1H, $J = 5.6\text{ Hz}$), 8.62 (d, 1H, $J = 1.6\text{ Hz}$), 8.38 (s, 1H), 8.31 (dd, 1H, $J = 1.6, 8.4\text{ Hz}$), 8.14 (d, 1H, $J = 8.4\text{ Hz}$), 8.01 (br s, 1H), 7.91 (d, 1H, $J = 5.6\text{ Hz}$), 7.85 (br s, 1H). ^{13}C NMR (125 MHz, DMSO) δ 162.5, 153.4, 149.8, 145.0, 136.5, 131.7, 131.0,

129.3, 127.7, 127.6, 120.3, 116.4, 107.9. LRMS (ESI) Exact Mass: 223.07,
Found: 224.3 ($M + H^+$).

3-(3-(1*H*-pyrazol-4-yl)phenyl)-2-cyanoacrylamide (37)

Yield: 44 mg (64%). ^1H NMR (400 MHz, DMSO) δ 8.23 (br s, 1H), 8.21 (s, 1H), 8.14 (m, 1H), 7.94 (br d, 2H), 7.78 – 7.83 (m, 3H), 7.55 (t, 1H, $J = 8$ Hz). ^{13}C NMR (125 MHz, DMSO) δ 162.62, 150.82, 136.13 (br), 133.86, 132.47, 129.70, 128.84, 127.11, 126.65, 125.63 (br), 120.19, 115.52, 106.78. LRMS (ESI) Exact Mass: 238.09, Found: 239.3 ($M + H^+$).

2-cyano-3-(4-(2-fluorobenzoyl)-1-methyl-1*H*-pyrrol-2-yl)acrylamide (38)

Yield: 42 mg (33%). ^1H NMR (400 MHz, DMSO) δ 8.00 (s, 1H), 7.89 (br s, 1H), 7.83 (s, 1H), 7.69 (d, 1H, $J = 2$ Hz), 7.54 – 7.67 (m, 3H), 7.32 – 7.40 (m, 2H), 3.83 (s, 3H). ^{13}C NMR (125 MHz, DMSO) δ 185.41, 162.92, 158.81 ($J = 247$ Hz),

136.64, 135.38, 132.88 ($J = 8.4$ Hz), 129.74 ($J = 2.8$ Hz), 128.30, 127.702 ($J = 15.3$ Hz), 125.35, 124.66 ($J = 3.2$ Hz), 117.11, 116.28 ($J = 25$ Hz), 115.37, 100.99, 34.53. LRMS (ESI) Exact Mass: 297.09, Found: 298.3 ($M + H^+$).

2-cyano-3-(1,3-dimethyl-2-oxo-2,3-dihydro-1H-benzo[d]imidazol-5-yl)acrylamide
(39)

Yield: 19 mg (57%). 1H NMR (400 MHz, DMSO) δ 8.19 (s, 1H), 7.80 (br s, 1H, overlaps peak at 7.77), 7.77 (s, 1H), 7.74 (d, 1H, $J = 8$ Hz), 7.67 (br s, 1H), 7.34 (d, $J = 8$ Hz), 3.38 (s, 3H), 3.35 (s, 3H). ^{13}C NMR (125 MHz, DMSO) δ 162.52, 153.39, 149.78, 145.02, 136.45, 131.67, 130.99, 129.30, 127.73, 127.66, 120.32, 116.34, 107.88. LRMS (ESI) Exact Mass: 256.10, Found: 257.3 ($M + H^+$).

RSK2 kinase assays. Wild-type and T493M RSK2 kinase activity were assayed as reported previously.³ Briefly, ERK2-activated RSK2 CTD (5 nM) was incubated with varying concentrations of each inhibitor for 30 minutes in the presence of 100 μ M ATP and 10 mM GSH. Each reaction was initiated by the addition of 167 μ M substrate peptide (RRQLFRGFSFVAK) and 0.3 μ Ci/ μ L γ - ^{32}P -ATP in a final volume of 25 μ L for an additional 30 minutes. Reactions were spotted on phosphocellulose membranes, washed once with 10% AcOH, twice with 0.1% H_3PO_4 , and once with MeOH prior to drying. Blots were exposed to a

phosphor storage plate (Amersham Biosciences), imaged with a Typhoon scanner (GE Healthcare), and quantitated using the SPOT program.⁴⁷ IC₅₀ values were calculated using a sigmoidal dose response fitting in the Prism program (Graphpad). Full IC₅₀ curves are presented in Supplementary Figure 9.

Cell-based Assay and Western Blotting. Confluent COS-7 cells were transfected overnight with HA-tagged full length RSK2, MSK1, and MSK1 C458V as reported previously.²⁷ Transfected cells were seeded into 6-well plates at 500,000 cells/well in 2 mL DMEM supplemented with 10% FBS (Axenia), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco) and allowed to adhere for 4 hours. Cells were serum starved for 18 hours, then treated with inhibitor or DMSO for an additional 2 hours. Cells were stimulated with PMA (100 ng/mL) for 30 minutes, washed with cold PBS, and frozen. Cells were thawed into 60 µL of lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% Triton X100 supplemented with Roche phosphatase and protease inhibitor cocktails). Lysates were clarified by centrifugation at 14,000 rpm, normalized by Bradford assay, denatured in SDS and separated by 7.5% acrylamide SDS-PAGE. Gels were transferred to nitrocellulose, blocked with Odyssey LiCOR blocking buffer for one hour, and probed with 1:1000 HA, 1:1000 pS380 RSK2 or 1:1000 pS376 MSK1 antibody dilutions. After thorough washing with TBST, blots were incubated with 1:10,000 dilutions of fluorescent secondary antibodies (Odyssey) for one hour, washed with TBST, and scanned on an Odyssey LiCOR instrument.

4.5 References

1. Weerapana, E., G.M. Simon, and B.F. Cravatt, Disparate proteome reactivity profiles of carbon electrophiles. *Nat Chem Biol*, 2008. 4(7): p. 405-7.
2. Blum, G., et al., Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Nat Chem Biol*, 2007. 3(10): p. 668-77.
3. Cohen, M.S., et al., Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science*, 2005. 308(5726): p. 1318-21.
4. Chang, J.W., D.K. Nomura, and B.F. Cravatt, A potent and selective inhibitor of KIAA1363/AADACL1 that impairs prostate cancer pathogenesis. *Chem Biol*, 2011. 18(4): p. 476-84.
5. Sirois, S., et al., Assessment of chemical libraries for their druggability. *Comput Biol Chem*, 2005. 29(1): p. 55-67.
6. Jadhav, A., et al., Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. *J Med Chem*, 2010. 53(1): p. 37-51.

7. Baell, J.B. and G.A. Holloway, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem*, 2010. 53(7): p. 2719-40.
8. Weerapana, E., et al., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature*, 2010. 468(7325): p. 790-5.
9. Totrov, M. and R. Abagyan, Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol*, 2008. 18(2): p. 178-84.
10. Jorgensen, W.L., The many roles of computation in drug discovery. *Science*, 2004. 303(5665): p. 1813-8.
11. Shoichet, B.K., Virtual screening of chemical libraries. *Nature*, 2004. 432(7019): p. 862-5.
12. Kellenberger, E., et al., Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins*, 2004. 57(2): p. 225-42.
13. Zhu, T., et al., Hit Identification and Optimization in Virtual Screening: Practical Recommendations Based on a Critical Literature Analysis. *J Med Chem*, 2013.

14. Kitchen, D.B., et al., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*, 2004. 3(11): p. 935-49.
15. Glick, M. and E. Jacoby, The role of computational methods in the identification of bioactive compounds. *Curr Opin Chem Biol*, 2011.
16. Ouyang, X., et al., CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J Comput Chem*, 2013. 34(4): p. 326-36.
17. Schroder, J., et al., Docking-Based Virtual Screening of Covalently Binding Ligands: An Orthogonal Lead Discovery Approach. *J Med Chem*, 2013.
18. Singh, J., et al., The resurgence of covalent drugs. *Nat Rev Drug Discov*, 2011. 10(4): p. 307-17.
19. Mysinger, M.M. and B.K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J Chem Inf Model*, 2010. 50(9): p. 1561-73.

20. Mysinger, M.M., et al., Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc Natl Acad Sci U S A*, 2012. 109(14): p. 5517-22.
21. Serafimova, I.M., et al., Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol*, 2012. 8(5): p. 471-6.
22. Doehn, U., et al., RSK is a principal effector of the RAS-ERK pathway for eliciting a coordinate promotile/invasive gene program and phenotype in epithelial cells. *Mol Cell*, 2009. 35(4): p. 511-22.
23. Kang, S., et al., p90 ribosomal S6 kinase 2 promotes invasion and metastasis of human head and neck squamous cell carcinoma cells. *J Clin Invest*, 2010. 120(4): p. 1165-77.
24. Smolen, G.A., et al., A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. *Genes Dev*, 2010. 24(23): p. 2654-65.
25. Park, J., et al., RAS-MAPK-MSK1 pathway modulates ataxin 1 protein levels and toxicity in SCA1. *Nature*, 2013. 498(7454): p. 325-31.

26. Le, N.T., et al., A crucial role for p90RSK-mediated reduction of ERK5 transcriptional activity in endothelial dysfunction and atherosclerosis. *Circulation*, 2013. 127(4): p. 486-99.
27. Miller, R.M., et al., Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J Am Chem Soc*, 2013. 135(14): p. 5298-301.
28. Irwin, J.J., et al., ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model*, 2012.
29. Carlsson, J., et al., Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat Chem Biol*, 2011. 7(11): p. 769-78.
30. Smith, A.J., et al., Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J Med Chem*, 2009. 52(2): p. 225-33.
31. Hermann, J.C., et al., Structure-based activity prediction for an enzyme of unknown function. *Nature*, 2007. 448(7155): p. 775-9.
32. Schwobel, J.A., et al., Prediction of michael-type acceptor reactivity toward glutathione. *Chem Res Toxicol*, 2010. 23(10): p. 1576-85.

33. Osguthorpe, D.J., W. Sherman, and A.T. Hagler, Exploring protein flexibility: incorporating structural ensembles from crystal structures and simulation into virtual screening protocols. *J Phys Chem B*, 2012. 116(23): p. 6952-9.
34. Morris, G.M., et al., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 2009. 30(16): p. 2785-91.
35. OEChem TK 1.7.4; Openeye Scientific Software: Santa Fe, NM.
<http://www.eyesopen.com>.
36. Gasteiger, J., C. Rudolph, and J. Sadowski, Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Computer Methodology*, 1990. 3(6): p. 537-547.
37. Shelley, J.C., et al., Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des*, 2007. 21(12): p. 681-91.
38. Li, J., et al., Extension of the platform of applicability of the SM5. 42R universal solvation model. *Theoretical Chemistry Accounts*, 1999. 103(1): p. 9-63.

39. Hawkins, P.C., et al., Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model*, 2010. 50(4): p. 572-84.
40. Lorber, D.M. and B.K. Shoichet, Hierarchical docking of databases of multiple ligand conformations. *Curr Top Med Chem*, 2005. 5(8): p. 739-49.
41. Mysinger, M.M. and B.K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J Chem Inf Model*, 2010. 50(9): p. 1561-73.
42. Mysinger, M.M., et al., Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem*, 2012. 55(14): p. 6582-94.
43. Gilson, M.K., K.A. Sharp, and B.H. Honig, Calculating the electrostatic potential of molecules in solution: method and error assessment. *Journal of computational chemistry*, 1988. 9(4): p. 327-335.
44. Irwin, J.J., et al., ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model*, 2012.
11. Gaulton, A., et al., ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res*, 2012. 40(Database issue): p. D1100-7.

45. Brozell, S.R., et al., Evaluation of DOCK 6 as a pose generation and database enrichment tool. *J Comput Aided Mol Des*, 2012. 26(6): p. 749-73.
46. Ouyang, X., et al., CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J Comput Chem*, 2013. 34(4): p. 326-36.
47. Knight, Z.A., et al., A membrane capture assay for lipid kinase activity. *Nat Protoc*, 2007. 2(10): p. 2459-66.

Chapter 5: Targeting Protein Kinases with Electrophilic Inhibitors

5.1 Abstract

The development of small molecule inhibitors of protein kinases presents a significant challenge. Covalent bond formation between a ligand and its kinase target has proven to be a powerful method to address this challenge by improving inhibitor potency, selectivity, and residence time. In this chapter, we present several approaches to the design of covalent chemical probes and the application of these compounds to study protein kinase function. We begin with the bioinformatics and structure-based design of selective, irreversible fluoromethylketone inhibitors of the RSK1/2 C-terminal kinase domain, which exploit a rare cysteine and a small gatekeeper residue. We discuss the rational design of reactive electrophilic compounds capable of forming reversible covalent bonds with noncatalytic cysteines, and the application of this type of electrophile to the design of novel inhibitors of the RSK and MSK CTD kinases. We present the adaptation of a known electrophilic natural product, hypothemycin, as a chemical probe for the identification of therapeutic kinase targets in the parasite *T. brucei* for the treatment of African sleeping sickness. Finally, we discuss the use of two different electrophiles on the same, promiscuous inhibitor scaffold to report on the endogenous kinase selectivity of an approved drug in living cells.

Note: This chapter was co-written with Prof. Jack Taunton, and has been prepared for publication. Miller, R.M.; Taunton, J. Targeting Protein Kinases with Electrophilic Inhibitors, *Meth. Enz.*, prepared manuscript, **2014**.

5.2 Introduction

Protein kinases are enzymes that catalyze the transfer of the γ -phosphate from ATP to the hydroxyl of serine, threonine or tyrosine residues of their substrate proteins. This post-translational modification is capable of altering protein function and is a ubiquitous mediator of intracellular signaling¹. Protein phosphorylation is frequently associated with cell growth, proliferation, and differentiation, and the misregulation of kinase signaling is implicated in many diseases. For this reason, protein kinases have remained a highly sought target class for small molecule inhibition for decades. There are 518 protein kinases in the human proteome, but only a fraction of these has been targeted with small molecule inhibitors. By virtue of their architectural similarity and sequence homology, particularly in the active site targeted by competitive inhibitors, selective inhibition of individual protein kinases remains a major challenge.

Covalent targeting of noncatalytic, nonconserved cysteine residues with electrophilic compounds has emerged as a powerful strategy for the development of potent and selective kinase inhibitors,^{2,3} Sequence and structural alignment has predicted that about 200 of the 518 protein kinases contain at least one solvent-exposed cysteine residue within the ATP binding site that may be targeted with reactive electrophilic inhibitors.⁴ In spite of this potential for improved selectivity, reactive electrophilic compounds have been shunned by the pharmaceutical industry out of fear of toxic side effects^{5,6}. However, covalent inhibitors are presently enjoying resurgence in academia and in the

pharmaceutical industry.⁷ The free energy of covalent bond formation between a ligand and its target typically leads to increased potency, and irreversible covalent inhibitors theoretically exhibit infinite potency given a long enough incubation time. However, in order to attain maximal inhibition, the electrophile must be positioned appropriately, and the reversible binding affinity of the scaffold must be high enough to bring the electrophilic center into close proximity with the targeted cysteine. Thus, medicinal chemists must optimize not only the electrophile but also the noncovalent scaffold to which it is attached.

The first rationally designed covalent kinase inhibitors were based on adenosine as inhibitors of EGFR, with a 2' thiol modification intended to form a disulfide bond with C797.⁸ This bond was predicted to be possible based on the proximity of the 2' of the ribose ring to the analogous cysteine-bearing position in a reported structure of the CDK kinase bound to ATP analogs. In related studies, the placement of an irreversible acrylamide electrophile on a general anilinoquinazoline core resulted in potent and selective EGFR inhibitors, which spawned decades of development of this scaffold.⁹ In spite of the success of this rational approach to covalent inhibitor design, few covalent inhibitors of protein kinases have been approved by the FDA. One is Afatinib, an inhibitor of EGFR based on the anilinoquinazoline scaffold discussed above, approved for the treatment of non-small cell lung cancer (NSCLC).¹⁰ Another recently approved drug is Ibrutinib, a Bruton's tyrosine kinase (BTK) inhibitor based on a pyrazolopyrimidine scaffold, approved for the treatment of mantle cell lymphoma.¹¹ Both of these compounds target a homologous cysteine found at

the “mouth” of the ATP binding site, present in about 2% of human kinases, and therefore may suffer from alkylation of off-target kinases.

The field of chemical biology has recently experienced a push for the development of “quality” chemical probes in response to the increase in overlap between the fields of chemistry and biology.¹² A set of five required characteristics has been proposed for quality chemical probes: 1. potency and selectivity for the desired target, 2. a known mechanism of action, 3. a known active chemical species, 4. proof of biological utility and 5. public availability. Covalent bond formation has the potential to aid both in the development of quality chemical probes, and as tools to help assess whether a new small molecule adheres to these stringent requirements. In this chapter, we discuss our laboratory’s design and characterization of covalent, cysteine-targeted inhibitors of the RSK and MSK C-terminal kinase domains, including a discussion of the rational design of tuned, reversible covalent inhibitors. We then present the synthesis and application of tagged derivatives that contain a fluorescent label or affinity handle. Finally, we demonstrate that even promiscuous covalent inhibitors targeting rare cysteines or a conserved catalytic lysine may be used for the identification of therapeutic kinase targets and to report on the selectivity and of chemical probes *in vivo*.

5.3 Irreversible, Covalent Inhibitors of RSK1, 2 and 4 C-terminal Kinase

Domains

The RSK family of protein kinases is comprised of a set of four proteins (RSK1-4), which lie downstream of ERK as part of the mitogen activated signaling cascade.¹³⁻¹⁵ Each of these proteins contains two kinase domains: an AGC-like N-terminal kinase domain (NTD) and a CAMK-like C-terminal kinase domain (CTD). These two kinase domains are connected by a short, hydrophobic regulatory linker peptide. Following phosphorylation of T577 (human RSK2 numbering) by ERK, the activated C-terminal kinase domain phosphorylates its own linker peptide on S386. This serves as a docking site for PDK1, which phosphorylates the NTD on its activation loop at S227, leading to full activation of this kinase domain and downstream signaling.¹⁶ The RSK2 NTD is known to phosphorylate transcription factors like CREB¹⁷ and NF- κ B,¹⁸ immediate-early gene products like Fos,¹⁹ and histone H3S10.²⁰ The RSK kinases help mediate cell growth and proliferation, and have been implicated in cognitive function²¹ tumor invasion,²² and leukemia.²³ We sought to develop a potent and selective small molecule inhibitor of the RSK kinases to further investigate their role in normal cellular signaling and disease.

A sequence alignment of all human protein kinases revealed a rare cysteine residue at the “top” of the ATP binding site that is present in only 11 kinases, including the CTDs of RSK1-4 and MSK1 and MSK2.²⁴ Of those 11 kinases, only RSK1, 2 and 4 contain a small threonine “gatekeeper” residue that

can be accessed by bulky inhibitor scaffolds (see Chapter 1, Figure 1-2).

Because the RSK CTD is required for activation of the NTD and downstream signaling, we wanted to design an inhibitor that could access the hydrophobic pocket behind the threonine gatekeeper and project an electrophile toward this rare cysteine residue. A compound that exploits these two elements would likely inhibit only those three kinases, while the remainder of the 518 kinases lacking one or both “selectivity filters” would be resistant.

Co-crystal structures of ATP analogs bound to several kinases indicates that the C8 position of the adenine ring forms Van der Waals contacts with the valine that typically occupies the site of C436 in RSK2. Additionally, the pyrrolopyrimidine-based inhibitor PP1 projects a bulky tolyl substituent beyond the small gatekeeper found in several Src-family kinases (Figure 5-1A).^{25, 26} We surmised that an electrophile positioned at C6 of a similar scaffold would covalently target the RSK2 cysteine and access the pocket behind the small threonine gatekeeper in RSK1, 2, and 4. Thus, we prepared **fmk**, a PP1-like pyrrolopyrimidine derivative containing a fluoromethylketone electrophile at the C6 position (analogous to C8 of adenine) and a hydroxypropyl substituent on the N7 position, predicted to point toward solvent (Figure 5-1B).²⁷

As anticipated, this inhibitor potently blocks the activation of the RSK NTD by its CTD only when both selectivity filters are present. Mutagenesis of either the cysteine (to valine) or the threonine gatekeeper (to methionine) abolishes drug binding *in vitro*. A kinome-wide screen of **fmk** shows that this requirement for both selectivity filters extends across the enzyme family (data not shown).

Indeed, **fmk** is among the most selective kinase inhibitors reported to date. This compound is potently cell-active as well, abolishing not only RSK2 autophosphorylation at S386 by the CTD, but also downstream phosphorylation of histone H3S10, a known RSK NTD substrate.

In order to adapt this inhibitor as a probe for interrogation of RSK function in cells, we prepared several tagged derivatives for use as occupancy probes or affinity handles (Figure 5-1C and D).²⁸

Figure 5-1: Chemical structures of (a) **PP-1**, a reversible Src inhibitor, (b) **fmk**, with proposed binding mode, (c) **BODIPY-fmk**, and (d) **fmk-pa**.

Based on the proposed binding mode of the PP1-like scaffold, we surmised that the hydroxypropyl portion of the molecule would be solvent-accessible and could accommodate larger tags. We first prepared a biotinylated derivative for use as a streptavidin affinity reagent, but this compound suffered from reduced *in vitro* potency as well as poor cell permeability. We also prepared a derivative functionalized with a cell-permeable BODIPY fluorophore at the same position (**BODIPY-fmk**); while this compound is 100-fold less potent *in vitro* than the unsubstituted precursor, it selectively labels RSK1 and RSK2 at micromolar concentrations in living cells. **BOIPY-fmk** is used routinely to assess the potency and stability of RSK inhibitors in both cells and animals. However, the cellular potency of this compound is reduced by the presence of serum in the tissue culture media, and the hydrophobic fluorophore may interact nonspecifically with a variety of proteins.

While **BODIPY-fmk** proved useful as an occupancy probe, we required a more versatile and potent compound. We sought to incorporate an alkyne functional group amenable to bioorthogonal Cu(I)-mediated azide-alkyne cycloaddition, or “click chemistry.” This functional group is small and is therefore less likely to interfere with ligand binding than bulky fluorophores or biotin tags. These tags can be installed by click chemistry after the inhibitor has irreversibly bound its target kinases. We synthesized the propargylated derivative **fmk-pa**, which was five-fold more potent than the unsubstituted **fmk** in cells (Figure 5-1D). After labeling in intact cells or cell lysates by **fmk-pa**, the click handle enables the installation of a fluorescent azide or a biotin azide, which can be

used to pull down the cellular targets of the inhibitor. Using **fmk-pa**, we demonstrated for the first time that RSK2 NTD activation is context-dependent.²⁸ Under phorbol myristate acetate (PMA) stimulation in bone marrow-derived macrophages, 300 nM **fmk-pa** is sufficient to block autophosphorylation of S386, whereas upon lipopolysaccharide (LPS) stimulation even μ M concentrations of **fmk-pa** has no effect.

Finally, we wanted to make a molecule with high aqueous solubility for extended dosing of animals. We synthesized the more stable and soluble derivative **fmk-mea** (fmk-methoxyethylamine, Figure 5-2A).²⁹

Figure 5-2: Solubilized compound **fmk-mea** is selective for RSK CTDs. (a) chemical structure of **fmk-mea**. (b) Screen of **fmk-mea** against 443 kinases (Ambit) at 1 μ M shows that only RSK1 and RSK4 CTD are inhibited > 50%.

When screened against a 443-kinase panel (Ambit/DiscoverX) at 1 μ M, only the RSK1 and RSK4 CTDs were inhibited greater than 50% (Figure 5-2B, RSK2 CTD was not included in the panel). The bis-tosylate salt of **fmk-mea** can

be formulated at concentrations greater than 25 mg/mL in PBS. The high aqueous solubility of **fmk-mea** enabled us to dose mice up to 60 mg/kg, both by intraperitoneal (IP) injection and by oral gavage.

The versatility of the **fmk** inhibitor scaffold is best revealed by the simultaneous application of several derivatives. In collaboration with the Abe laboratory (University of Rochester), **fmk-mea** was used to show that pharmacological blockade of RSK1/2 CTD signaling is sufficient to block atherosclerosis.³⁰ During this study, we used **BODIPY-fmk** as a probe to assess the extent of RSK1/2 occupancy in the heart by **fmk-mea** in living mice (Figure 5-3).

Figure 5-3: Using **BODIPY-fmk** as a probe, the occupancy of RSK1 and RSK2 by **fmk-mea** was assessed in the heart tissue of living animals. Competition of the fluorescently-labeled probe is indicative of occupancy, and maximal occupancy was assessed by incubation of vehicle-treated lysates with **fmk** prior to fluorescent probe labeling. Figure adapted from the supporting information for reference XX.

Mice were dosed by IP injection with varying concentrations of **fmk-mea** for 1 hour, organs were harvested and lysed by mechanical homogenization, and these lysates treated with 5 μM of **BODIPY-fmk** for 1 hour to assess RSK1/2 occupancy. Maximum potential occupancy was assessed by treatment of a

vehicle lysate with 3 μ M **fmk** prior to incubation with the fluorescent competitor. At all doses of **fmk-mea**, we observed partial occupancy of both RSK1 and RSK2. RSK1 appeared more sensitive to **fmk-mea**, as complete occupancy was attained with a 35 mg/kg dose, whereas RSK2 required a 50 mg/kg dose to achieve approximately the same level of occupancy. This experiment highlights the power of an irreversible occupancy probe, and we have included a detailed procedure for assaying compound occupancy in living animals using an irreversible, fluorescent probe.

*Representative procedure for assessing **fmk-mea** occupancy using the fluorescently labeled derivative **BODIPY-fmk***

Protocol:

1. Dose mice with inhibitor as desired.
2. Sacrifice animals, harvest desired tissues, and freeze tissues until ready to process.
3. Chop tissue into small pieces with a clean razor blade, while the tissue is still slightly frozen.
4. Place each tissue sample into a 2 mL Eppendorf tube. Add 0.75 mL PBS supplemented with protease and phosphatase inhibitor cocktails (Roche).
5. Prepare lysates using a mechanical rotor-stator style tissue homogenizer (Janke & Kunkel/IKA T25-Ultra-Turrax Homogenizer),

- keeping tubes on ice. Begin by homogenizing for 30 seconds on low speed, allowing samples to rest for 30 seconds, and then homogenizing at high speed for 30 seconds.
6. Transfer crude lysates to 1.5 mL ultracentrifuge tubes. Centrifuge samples at 4 °C in a benchtop ultracentrifuge for 1 hour at 30,000 x g (Beckman Coulter Optima TLX Ultracentrifuge).
 7. Transfer clarified supernatants to new 1.5 mL Eppendorf tubes on ice and snap freeze in liquid nitrogen if desired. Typically, lysate total protein will be in the 3 - 5 mg/mL range.
 8. Quantify lysate protein concentration by Bradford assay, normalize each sample by adding cold PBS, and transfer 100 µL of lysate to new Eppendorf tubes for labeling with fluorescent competitor (in this case, **BODIPY-fmk**).
 9. Add 5 µM **BODIPY-fmk** (5.25 µL of a 100 µM DMSO stock) and incubate for 1 hour at room temperature.
 10. Remove 40 µL of sample and quench with 10 µL of 5x Laemmli sample buffer as a pre-IP lysate sample. Note: Depending on the abundance of the protein and the efficiency of fluorescent probe labeling, these samples may suffice to visualize and quantitate compound occupancy. Competition of the fluorescent probe indicates occupancy of the protein in the live mouse.
 11. To the remaining 60 µL of lysates freshly labeled with **BODIPY-fmk**, add 300 µL of cold PBS +1 % NP40. Keep samples on ice.

12. Add 10 uL of α -Rsk1 and 10 uL of α -Rsk2 antibodies (Santa Cruz Biotechnology, sc-231 and sc-9986). Place samples in a rotator at 4 °C for 2 – 4 hours.
13. Before opening tubes, spin them down to collect liquid that may be trapped in the cap. Add 30 uL of ProteinG Dynabeads (Invitrogen 100-04D) to each tube, place in a rotator, and incubate overnight at 4 °C.
14. Place tubes in magnetic racks to separate the beads. Remove 40 uL and quench with 5x Laemmli sample buffer as a post-IP lysate sample.
15. Discard the remaining supernatant, and add 300 uL of cold PBS + 1% NP40 to wash. Place in a rotator for 5 minutes.
16. Place tubes in magnetic racks to separate beads, and discard the supernatant. Add 300 uL of cold PBS + 1% NP40, mix thoroughly, and transfer each sample to a new pre-chilled 1.5 mL Eppendorf tube. Place on a rotator at 4 °C for 5 minutes.
17. Wash once more with 300 uL of cold PBS + 1% NP40 for 5 minutes, separate the beads with a magnetic rack, and discard the supernatant.
18. To elute the proteins from the beads, add 50 uL of 2x Laemmli sample buffer and 10 uL of freshly made 1M DTT to the beads. Vortex briefly and boil for 1 minute at 90 °C. Samples may be snap frozen in liquid nitrogen prior to use.
19. Collect all samples (pre-IP lysate, post-IP lysate, and immunoprecipitated samples) and resolve proteins by SDS-PAGE. Visualize competitive labeling by fluorescent scanning on a Typhoon

imager. Transfer proteins to nitrocellulose and assess even loading of RSK1 and RSK2 by Western blot.

5.4 Reversible Covalent Inhibitors

During the development of RSK CTD inhibitors based on **fmk**, we synthesized several derivatives containing modified thiol-reactive electrophiles. We began by increasing the reactivity of a series of vinyl nitrile Michael acceptors with geminal amide or ester electron withdrawing groups (Figure 5-4A).³¹ In studies with model olefins, we discovered that this modification not only increased the rate of the forward reaction (1,4-conjugate addition), but also accelerates retro-Michael reaction at physiological pH, presumably by lowering the pK_a of the α -proton by about 10 log units (Figure 5-4B).

Figure 5-4: Cyanoacrylamides and cyanoacrylates form reversible covalent bonds with thiols at physiological pH. (a) chemical structures of **fmk**-based cyanoacrylate and cyanoacrylamide inhibitors. (b) Geminal electron withdrawing groups lower the adduct α -proton pK_a and promote the reverse reaction. Figure adapted from reference 32.

Reactive electrophilic compounds have been avoided by the pharmaceutical industry for fear of non-specific alkylation, off target toxicity and immunogenicity^{5,6}. When electrophilic compounds *are* implemented, it is common practice to use the least reactive electrophile possible to avoid spurious alkylation of off-target nucleophiles. Reversible covalent bond formation has the potential to address many of the problems associated with reactive electrophiles. Because the covalent bond between a highly reactive cyanoacrylamide is rapidly reversible, it can be defined by thermodynamic equilibrium dissociation constant (K_d). The K_d for reaction of most cyanoacrylamides with β -mercaptoethanol (BME) or glutathione (GSH) lies in the range of tens to hundreds of mM. Thus, reaction with the mM concentrations of intracellular small molecule thiols like GSH or surface nucleophiles on off-target proteins will likely be transient and, in theory, nontoxic. In contrast, this covalent bond is durable when the scaffold portion of the inhibitor is stabilized by noncovalent interactions with a binding site. These stable on-target adducts will then fall apart upon disruption of the noncovalent interactions by unfolding or degradation, minimizing the risk of haptenization.

Modification of the **fmk** scaffold to incorporate a series of these activated cyanoacrylamides or cyanoacrylates at the C6 position afforded compounds that inhibited RSK2 *in vitro* with an IC_{50} of 5 nM, similar to that of the irreversible compound **fmk** (Figure 5-4A).³¹ However, while **BODIPY-fmk**, a fluorescent isopropyl cyanoacrylamide analog covalently labels no detectable proteins in cell lysates, confirming that the reversible bond formation reduces the number of

nonspecific interactions. A co-crystal structure of **CN-OtBu** shows a clear covalent bond to C436 of RSK2 and a network of noncovalent interactions that stabilize the complex. These compounds block binding of **BODIPY-fmk** in a competitive labeling experiment, indicating full occupancy of RSK1 and RSK2 in cells. Finally, these reversible covalent RSK CTD inhibitors abolish the invasive phenotype of Ras-transformed epithelial cells in a model of aggressive cancer. Taken together, these data suggest that rational reversible covalent bond formation is a viable method for the development of selective, cell-active kinase inhibitors.

As an orthogonal approach to inhibitor design, we harnessed the cyanoacrylamide electrophile to reversibly and covalently capture small electrophilic fragments targeting noncatalytic cysteines.³² The geometric constraint of the covalent bond enables the identification of low molecular weight fragments capable of forming selective noncovalent interactions with the kinase active site. Additionally, the potency increase provided by the covalent bond enables the biochemical screening of typically insoluble heteroaromatic fragments at much lower concentrations than previously possible. MSKs are dual-kinase proteins similar to RSKs that have been implicated in leukemia,³³ malignant growth of Ras-transformed fibroblasts,³⁴ and inflammation.³⁵ However, the CTD of each of these kinases has a bulky methionine gatekeeper and is resistant to our **fmk**-like compounds. Using our fragment-based method, we were able to identify several ligand efficient fragments that selectively inhibited the RSK kinases without requiring access to the gatekeeper pocket. We used co-

crystal structures of these fragments to design **RMM-46** (compound **12** from Chapter 2), the first reported inhibitor of the MSK1, MSK2 CTDs (Figure 5-5).

Figure 5-5: Cyanoacrylamide fragment screening enables the identification of selective, ligand efficient compounds. (a) indazole fragment **1** identified ligand efficient, covalent scaffold. (b) RMM-46 (**12**) was designed by analysis of co-crystal structures of several cyanoacrylamide fragments and is the first reported inhibitor of the MSK1 and MSK2 CTDs.

Our previous studies adapted known inhibitor scaffolds and improved their potency and selectivity with a covalent bond. As a complementary approach to inhibitor design, this cyanoacrylamide fragment-based approach exploits the covalent bond to identify novel scaffolds as starting points for inhibitor development.

5.5 Using Promiscuous Covalent Inhibitors as Chemical Probes

Covalent bond formation is frequently employed to make an inhibitor as potent and selective as possible for a single kinase target. However, a question asked by our laboratory (and others) is whether promiscuous covalent inhibitors can be applied to investigate novel biology. It is becoming apparent that many successful approved drugs inhibit many other proteins than their intended targets, and that this polypharmacology is necessary for their function. Our laboratory has repurposed several promiscuous inhibitors as chemical tools to identify novel therapeutic targets for the treatment of disease, and to report the selectivity of known approved drugs in cells.

*Adaptation of a fungal antibiotic natural product to identify therapeutic kinase targets in *T. Brucei**

African sleeping sickness is an orphan disease caused by the parasite *Trypanosoma brucei*.^{36,37} There are few drugs for this disease, and those that are widely used exhibit suffer from poor bioavailability and numerous side-effects. Hypothemycin is a resorcylic acid lactone natural product that is known to covalently inhibit human protein kinases that contain a cysteine proximal to the conserved DXG motif (CDXG), which comprise about 10% of the human kinome (Figure 5-6A).³⁸ We showed that hypothemycin potently inhibits proliferation of the bloodstream form (BSF) of the *T. brucei* parasite with an EC₅₀ of 170 nM, and partially clears infection in mice after daily administration of a 10 mg/kg dose.³⁹

Trypanosomes contain 182 protein kinases, 21 of which contain a CDXG cysteine. We attribute the potent trypanocidal behavior of hypothemycin to covalent inhibition of one or more of these kinases, and sought to adapt the natural product to identify these targets.

In a crystal structure of hypothemycin covalently bound to ERK2, the C4 methyl ether of the natural product is solvent exposed, suggesting that this position may be amenable to functionalization (Figure 5-6B).

Figure 5-6: The natural product hypothemycin may be adapted as a chemical probe. (a) chemical structure of hypothemycin. (b) crystal structure of hypothemycin bound to ERK suggests chemical modification at C4. (c) chemical structure of the semisynthetic chemical probe, propargylhypothemycin. Figure adapted from ref. 39.

Selective installation of a propargyl group to 4-desmethoxyhypothemycin⁴⁰ was achieved by alkylation with propargyl bromide mediated by Cs_2CO_3 to afford propargylhypothemycin (Figure 5-6C). Consistent with our structural analysis, this propargyl functionality did not interfere with ligand binding or cellular permeability. Propargylhypothemycin was equipotent to the natural product in a cell proliferation assay against the BSF parasite ($\text{IC}_{50} = 170 \text{ nM}$). Treatment of intact trypanosomes with propargylhypothemycin results in irreversible covalent modification of its targets, and subsequent Cu(I) -mediated cycloaddition was used to functionalize the adduct with either a rhodamine-azide fluorophore or a

biotin azide affinity handle. Affinity purification and quantitative mass spectrometry enabled the identification of 11 potential kinase targets of this natural product. Selective RNAi knockdown of tbCLK1 and tbGSK3 (both CDXG kinases) resulted in loss of cell viability, suggesting these two kinases are the relevant targets of the natural product. Hypothemycin and its derivatives are toxic at high doses, presumably due to inhibition of human CDXG kinases, but this study demonstrates that selective inhibition of tbCLK1 and tbGSK3 may be a viable therapy for *T. brucei* infection.

Targeting conserved nucleophilic residues enables the cellular selectivity profiling of approved drugs

Thus far we have discussed inhibitors designed to target nonconserved cysteine residues in order to gain specificity. However, protein kinases contain other nucleophilic amino acids. Active kinases contain a conserved, nucleophilic active site lysine (K295 in human Src) suitable for electrophilic targeting. While this lysine is not involved in covalent catalysis, it is required to coordinate the negatively charged phosphates of ATP to help catalyze phosphotransfer. We merged elements of the promiscuous, lysine-targeted inhibitor **FSBA** with those of **PP1** to afford tolyl-substituted adenosine fluorosulfonate **40**, designed to covalently and irreversibly inactivate kinases with a small gatekeeper (Figure 5-7A, B).⁴¹ This compound is also functionalized with a propargyl ether which allowed us to monitor its occupancy of even low-abundance endogenous proteins in cells. Following sequential affinity purification and western blotting, **40**

labels at least five endogenous Src-family kinases in cells: Src, Yes, Lck, Blk and Lyn. This promiscuous electrophilic compound was used in a competitive labeling experiment to show that an inhibitor of BCR-ABL approved for treatment of chronic myeloid leukemia, ponatinib, also potently occupies the important Src-family kinases Yes, Blk, and Lck in living cells.

Figure 5-7: An adenosine-like scaffold can target two different nucleophiles to interrogate inhibition of Src family kinases in cells. (a) Crystal structures of Src indicate the presence of two active site nucleophiles. (b) design of lysine-targeted fluorosulfonyl **40**. (c) chemical structure of cysteine targeted inhibitor **41**. Figure adapted from ref. 41.

To highlight the versatility of this tolyl-adenosine scaffold, we replaced the amine-reactive fluorosulfonyl electrophile with a thiol-reactive vinylsulfonate predicted to lie in close proximity to a rare cysteine (C277 in human Src) at the tip of the glycine-rich loop found in only 9 human kinases (compound **41**, Figure 5-7A, C). In contrast to **40**, cysteine-targeted inhibitor **41** only labels the two kinases with a glycine-rich loop cysteine, Src and Yes, but not the other Src-family kinases.

This study demonstrates that appropriate electrophile selection is a key determinant in the selectivity of a covalent inhibitor. We also show that we can vary the selectivity of even a promiscuous scaffold like adenosine by the exploitation of multiple different selectivity filters. Importantly, we anticipate that the workflow presented in this study will motivate the use of promiscuous, covalent chemical probes as reporters of the selectivity of inhibitors of other kinase families.

5.6 Conclusions and Future Directions

An understanding of the intricate roles of protein kinases in the maintenance of cellular function requires continued development of potent and selective small molecules. Additionally, protein kinases remain an important drug target class for the treatment of disease. Covalent inhibition has emerged as an enticing method for the design of selective molecules for use as drugs and as chemical probes. In this chapter, we have presented our laboratory's approach to the design of such inhibitors, the improvement of their physical properties, and their application as chemical probes.

There are many crystal structures of small molecule inhibitors bound to kinases deposited in the PDB, and the position of nearly any nucleophilic amino acid side chain can be predicted by sequence alignment or homology modeling. The integration of these data makes it possible to design inhibitors for covalent targeting of almost any kinase, even in the absence of experimentally determined structures. Furthermore, in collaboration with N. London and B. Shoichet (UCSF), our laboratory has pioneered the virtual screening of large, virtual libraries of electrophiles targeted to noncatalytic cysteines (see Chapter 4). This computational approach allows the transition from intuition-based electrophile placement on known inhibitor scaffolds to the unbiased identification of probable binders based on relative binding energy calculations.

Thus far, we have only discussed the design of inhibitors that form a covalent interaction with a wild-type cysteine. Targeting endogenous cysteines in this manner limits chemical biologists to the roughly 200 kinases known to

contain a solvent exposed cysteine in the ATP binding site. However, genetic introduction of a selectivity filter into the active site of a kinase is a viable method for sensitizing a kinase with known inhibitors or for developing new inhibitors altogether. The genetic introduction of sensitizing mutations, originally applied to the gatekeeper residue, is known broadly as “chemical genetics.”⁴² Chemical genetic mutations can be introduced easily into recombinant proteins for biochemical assays, but it remains a challenge in living organisms. Chemical genetic sensitization of the yeast homolog of the Polo kinase domain (Cdc5) to the **fmk** scaffold by introduction of a small gatekeeper allowed us to interrogate substrates of Polo kinase in yeast.⁴³ Mutagenesis of the gatekeeper residue of Src to a cysteine enabled the development of gatekeeper-targeted covalent inhibitors based on a pyrrolopyrimidine scaffold that inhibit only two off-target kinases with endogenous cysteine gatekeepers.⁴⁴ Recently, a zinc-finger nuclease technology has been reported for highly efficient and targeted genome editing with single stranded oligonucleotides in several different cell types.⁴⁵ The CRISPR-Cas system provides yet another method for the rapid introduction of multiple, targeted mutations in animals.⁴⁶ Technologies like these open the door for the chemical genetic introduction of reactive cysteines and small gatekeepers in any kinase of interest, and even desensitization of other kinases in the same organism.

In spite of these recent advances, rational covalent inhibitor design remains nontrivial. Covalent inhibitor design requires the selection of an appropriate noncovalent scaffold, an electrophile with reactivity matched to the

nucleophilic residue of interest, and an optimized linkage between the two. Here we have presented several strategies to address these challenges, and we hope this chapter will serve as a framework for covalent inhibitor development. We anticipate that, as fears of off-target toxicity associated with reactive electrophiles continue to be addressed, companies and academic labs will harness this type approach for the design of better probes and drugs.

5.7 References

1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S.
Science, **2002**, 298, 1912 -- 1934.
2. Potashman, M. H.; Duggan, M. E. *J. Med. Chem.* **2009**, 52, 1231–1246.
3. Hagel, M.; Niu, D.; St Martin, T.; Sheets, M. P.; Qiao, L.; Bernard, H.; Karp, R. M.; Zhu, Z.; Labenski, M. T.; Chaturvedi, P.; Nacht, M.; Westlin, W. F.; Petter, R. C.; Singh, J. *Nat. Chem. Biol.* **2011**, 7, 22–24.
4. Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, 9, 28–39.
5. Uetrecht, J. *Chem. Res. Toxicol.* **2008**, 21, 84-92.
6. Evans, D. C., Watt, A. P., Nicoll-Griffith, D. A.; Baillie, T. A. *Chem. Res. Toxicol.* **2004**, 17, 3 -16.
7. Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Discov.* **2011**, 10, 307 – 317.
8. Singh, J.; Dobrusin, E. M.; Fry, D. W.; Haske, T.; Whitty, A.; McNamara, D. J. *J. Med. Chem.* **1997**, 40, 1130 - 1135.

9. Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trump-Kallmeyer, S.; Dobrusin, E. M. *Proc. Natl. Acad. Sci.* **1998**, *95*, 12022–12027.
10. Sequist, L. V.; Yang, J. C.; Yamamoto, N.; O'Byrne, K.; Hirsh, V.; Mok, T.; Geater, S. L.; Orlov, S.; Tsai, C. M.; Boyer, M.; Su, W. C.; Bennouna, J.; Kato, T.; Gorbunova, V.; Lee, K. H.; Shah, R.; Massey, D.; Zazulina, V.; Shahidi, M.; Schuler, M. *J. Clin. Oncol.* **2013**, *31*, 3327-3334.
11. Pan, Z.; Scheerens, H.; Li, S. J.; Schultz, B. E.; Sprengeler, P. A.; Burrill, L. C.; Mendonca, R. V.; Sweeney, M. D.; Scott, K. C.; Grothaus, P. G.; Jeffery D. A.; Spoerke, J.M.; Honigberg, L. A.; Young, P. R.; Dalrymple, S. A.; Palmer J. T. *ChemMedChem*, **2007**, *2*, 58 – 61.
12. Frye, S. *Nat. Chem. Biol.* **2010**, *6*, 159 – 161.
13. Hauge, C.; Frodin, M. RSK and MSK in MAPK signalling. *Journal of Cell Science*, **2006**, *119*, 3021 – 3023.
14. Fisher, T. L.; Blenis, J. *Mol. Cell. Biol.*, **1996**, *16*, 1212-1219.
15. Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J. & Cohen, P. *J Biol*

Chem **1998**, 273, 1496-1505.

16. Frodin, M., Jensen, C. J., Merienne, K. & Gammeltoft, S. *Embo J* **2000**, 19, 2924-2934.

17. Ginty, D. D; Bonni, A.; Greenberg, M. E. *Cell*, **1994**, 77, 713-725.

18. Ghoda, L.; Lin, X.; Greene, W. C. *J Biol Chem*, **1997**, 272, 21281-21288.

19. Chen, R. H.; Abate, C.; Blenis, J. *Proc. Natl. Acad. Sci.* **1993**, 90, 10952-10956.

20. Sassone-Corsi, P.; Mizzen, C. A.; Cheung, P.; Crosio, C.; Monaco, L.; Jacquot, S.; Hanauer, A.; Allis, C. D. *Science*, **1999**, 285, 886-891.

21. Trivier, E. *et al. Nature*. **1996**, 384, 567-570.

22. Doehn, U.; Hauge, C.; Frank, S. R.; Jensen, C. J.; Duda, K.; Nielsen, J. V.; Cohen, M. S.; Johansen, J. V.; Winther, B. R.; Lund, L. R.; Winther, O.; Taunton, J.; Hansen, S. H.; Frödin, M. *Mol. Cell*, **2009**, 511-522.

23. Elf, S. *et al. Blood* **2011**, 117, 6885-6894.
24. Buzko, O.; Shokat, K. M. *Bioinformatics* **2002**, 18, 1274.
25. Liu, Y.; Bishop, A.; Witucki, L.; Kraybill, B.; Shimizu, E.; Tsien, J.; Ubersax, J.; Blethrow, J.; Morgan, D. O.; Shokat, K. M. *Chem. Biol.* **1999**, 6, 671-678.
26. Schindler, T.; Sicheri, F.; Pico, A.; Gazit, A.; Levitzki, A.; Kuriyan, J. *Mol. Cell.* **1999**, 3, 639 – 648.
27. Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J.. *Science* **2005**, 308, 1318-1321.
28. Cohen, M. S.; Hadjivassiliou, H.; Taunton, J. *Nat. Chem. Biol.* **2007**, 3, 156 – 160.
29. Cohen, M. S. *Ph.D. Thesis*, **2007**
30. Le, N.T.; Heo, K.S.; Takei, Y.; Lee, H.; Woo, C.H.; Chang, E.; McClain, C.; Hurley, C.; Wang, X.; Li, F.; Xu, H.; Morrell, C.; Sullivan, M. A.; Cohen, M. S.; Serafimova, I. M.; Taunton, J.; Fujiwara, K.; Abe, J. *Circulation*, **2013**, 127, 486 - 499.
31. Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.;

- Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. *Nat. Chem. Biol.* **2012**, *8*, 471–476.
32. Miller, R.M.; Paavilainen, V. O.; Krishnan, S.; Serafimova, I. M.; Taunton, J. *J. Am. Chem. Soc.* **2013**, *135*, 5298 - 5301.
33. Odgerel, T.; Kikuchi, J.; Wada, T.; Shimizu, R.; Kano, Y.; Furukawa, Y. *Leukemia* **2010**, *24*, 1087–1090.
34. Pérez-Cadahía, B.; Drobic, B.; Espino, P. S.; He, S.; Mandal, S.; Healy, S.; Davie, J. R. *J. Biol. Chem.* **2010**, *286*, 42–49.
35. Ananieva, O.; Darragh, J.; Johansen, C.; Carr, J. M.; McIlrath, J.; Park, J. M.; Wingate, A.; Monk, C. E.; Toth, R.; Santos, S. G.; Iversen, L.; Arthur, J. S. *Nat Immunol.* **2008**, *9*, 1028-1036.
36. Simarro, P. P.; Jannin, J.; Cattand, P. *PLoS Medicine*, **2008**, *5*, 174 – 180.
37. Fairlamb, A. H. *Trends Parasitol.* **2003**, *19*, 488 - 494.
38. Leproult, E.; Barluenga, S.; Moras, D.; Wurtz, J. M.; Winssinger, N. *J. Med. Chem.* **2011**, *54*, 1347–1355.

39. Nishino, M.; Choy, J. W.; Gushwa, N. N.; Oses-Prieto, J. A.; Koupparis, K.; Burlingame, A. L.; Renslo, A. R.; McKerrow, J. H.; Taunton, J. *eLife*, **2013**, 2, 1 -15.
40. Wee, J.L.; Sundermann, K.; Licari, P.; Galazzo, J. *J. Nat. Prod.* **2006**, 69, 1456–1459.
41. Gushwa, N. N.; Kang, S.; Chen, J.; Taunton, J. *J. Am. Chem. Soc.* **2012**, 134, 20214 – 20217.
42. Bishop, A. C.; Shah, K.; Liu, Y.; Witucki, L.; Kung, C.; Shokat KM. *Curr Biol.* **1998**, 8, 257 - 266.
43. Paulson, J. L.; Sullivan, M.; Lowery, D. M.; Cohen, M. S.; Zhang, C.; Randle, D. H.; Taunton, J.; Yaffe, M. B.; Morgan, D. O.; Shokat, K. M. *Chem. Biol.* **2007**, 14, 1261 – 1272.
44. Garske, A. L.; Peters, U.; Cortesi, A. T.; Perez, J. L.; Shokat, K. M. *Proc. Natl. Acad. Sci.* **2011**, 108, 15046 -15052.
45. Chen, F.; Pruett-Miller, S. M.; Huang, Y.; Gjoka, M.; Duda, K.; Taunton, J.; Collingwood, T. N.; Frodin, M.; Davis, G. D. *Nat Methods.* **2011**, 8, 753-755.
46. Wang, H.; Yang, H.; Shivalila, C. S.; Dawlaty, M. M.; Cheng, A. W.; Zhang, F.; Jaenisch, R. *Cell*, **2013**, 153, 910-918.

Chapter 6: Conclusions and Future Directions

The continued development of novel small molecule modulators of protein activity is of fundamental importance for the elucidation of biological processes, as well as in the treatment of disease. However, as the productivity of traditional high-throughput screening efforts has begun to diminish, the need for novel methods to identify druglike leads has increased. In this thesis, we report an electrophilic fragment-based method for the identification of inhibitors of proteins containing nonconserved, noncatalytic cysteines. We applied this method to the design of the first reported inhibitors of the MSK1 CTD kinase, which will allow us to begin studying the role of MSK1 in the maintenance of normal cellular function and disease pathogenesis.

Fragment-based design of inhibitors is a well-established method for the design of selective small molecules. Additionally, cysteines have been exploited for the identification of fragment hits in the context of “tethered” disulfide fragments. Here, we report a complementary electrophilic fragment-based approach to inhibitor design based on the spontaneous, reversible covalent bond formation between noncatalytic cysteines and the cyanoacrylamide electrophile. The reversible bond formation between cyanoacrylamide fragments and a noncatalytic cysteine has several advantages over other methods for inhibitor identification. First, the potency gained from the free energy of covalent bond formation enables biochemical screening of often insoluble heteroaromatic fragments at much lower concentrations than previously possible (low micromolar as opposed to mM). The ligand efficiency of the electrophilic fragment hits can be extremely high (in the range of 0.6 kcal/mol per heavy

atom). Second, the cyanoacrylamide electrophile is compatible with the nucleophile-rich cellular environment, and can be retained throughout the inhibitor optimization process. Third, cyanoacrylamides are capable of reacting with noncatalytic, unactivated cysteines and may be used to target any exposed cysteine within reach of a binding pocket, potentially even low-affinity pockets like those found in protein-protein interfaces. Reversible covalent fragment hits identified from a screen are more selective than their noncovalent counterparts due to the lack of conservation of targeted noncatalytic cysteines across a large protein family.

Adaptation of the virtual screening program DOCK to include covalent bond parameters enabled us simplify the process of electrophilic fragment screening (performed by Dr. Nir London and Prof. Brian Shoichet, UC San Francisco). Computational assembly and screening of large virtual fragment libraries enables the focused synthesis and testing of only the highest ranking compounds. In this work, we have shown that we can computationally assemble and screen a 12,000 member cyanoacrylamide library from a simple pool of commercially available aldehyde fragments. This approach led to the identification of several potent, cell-active inhibitors of the RSK and MSK kinases. We envision the extension of this technique to more complicated “synthetic” libraries assembled computationally from commercially available starting materials, further directing the chemistry required to identify novel, complex inhibitor scaffolds.

The work presented here will likely impact the field of MSK1 biology by enabling the rapid, selective and reversible blockade of these important signaling molecules. With the help of collaborators, we are currently engaged in a variety of studies to investigate the role of the MSK1 CTD in mediating tumor angiogenesis, leukemia proliferation, and immune cell function, among others. Preliminary experiments suggest a complex role for MSK1 in the expression of cytokines required for the proper development and activation of B- and T-cells, and we look forward to exploring the consequences of MSK inhibition on adaptive immunity. Finally, little is known about the ability of the RSK and MSK kinases to compensate for one another in the event of inhibition or knockout of a single member. We anticipate that our potent pan-RSK/MSK inhibitors may provide the first tools to study whether cellular processes are dependent on partially redundant functions of these proteins.

Looking forward, we hope that this reversible covalent fragment method will be applied both experimentally and computationally to the design of reversible covalent fragment libraries targeting nearly any solvent-exposed cysteine in the proteome.

Appendix A: Reproduced NMR Spectra of Synthesized Compounds

3

6

7

9

10

19

20

23

24

26

29

30

33

34

35

36

37

39

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

Date