Title
Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles

Permalink
https://escholarship.org/uc/item/0cj6m628

Author
Serafimova, Iana Margaritova

Publication Date
2013

Peer reviewed|Thesis/dissertation

Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles

by

Iana M. Serafimova

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

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgments

I am incredibly grateful for all my experiences at UCSF and especially thankful to my advisor, Professor Jack Taunton, for his guidance and support, and for always demanding the best of us.

I would like to thank my thesis committee members, Prof. Kevan Shokat and Prof. John Gross, for all of their helpful comments and advise.

In addition, I would like to thank all the members of the Taunton lab who have been a great source of help, and without whom my graduate school experience would not have been the same.

Part of this dissertation is a reproduction of material that has been previously published, and contains contributions from collaborators. Chapter 2 is reproduced in part with permission from Serafimova, I.M., Pufall, M.O., Krishnan, S., Duda, K., Cohen, M.S., Maglathlin, R.L., McFarland, J.M., Miller, R.M., Frodin, M., and Taunton, J. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nature Chemical Biology.* 2012, 8: 471-476. Chapter 4 contains experimental results (Figure 4.4b) that were obtained by Ville Paavilainen and are reproduced with his permission.

To my family.

Abstract

Targeting non-catalytic cysteine residues with irreversible electrophiles improved the selectivity and potency of various inhibitors. This strategy has been applied to scaffolds targeting diverse proteins and in all cases results in an irreversible bond between the cysteine thiol and the electrophile. Recently, several acrylamide-based irreversible kinase inhibitors have been advanced to clinical trials, and the potency and selectivity of these compounds is greatly dependent on the covalent modification of a cysteine residue found in the enzyme active site. However, concerns about off-target modification motivated us to seek out strategies to develop reversible cysteine-targeting small molecule inhibitors.

Here, we describe the development of cysteine-targeting inhibitors that form a covalent bond with their kinase target, but do so in a reversible manner. The covalent bond to a cysteine thiol is long-lived in the context of the folded active-site, but readily broken with off target proteins. We show that electron-deficient olefins, including acrylamides, can be chemically tuned to react with cysteine thiols in a rapidly reversible manner. By installing a nitrile group, the intrinsic reactivity of the olefins towards thiols increases while at the same time the formation of irreversible adducts is eliminated. These properties result in an inhibitor with sustained occupancy of the active site of the kinase due to covalent bond formation as well as high selectivity for its target and minimal off-target

effects. We specifically combined 2-cyanoacrylates and 2-cyanoacrylamides with a scaffold that binds promiscuously to protein kinases and to obtain highly selective and reversible, yet covalent inhibitors of the C-terminal domain of p90 ribosomal S6 kinases (RSK-CTD).

Our results establish a chemical framework for engineering sustained covalent inhibition without accumulating permanently modified proteins and peptides. Future studies will seek to define the druggable cysteinome and apply this technology to generate novel small molecules inhibitors for several cysteine containing therapeutic targets. We believe that the application of reversible covalent technology to drug discovery and development will yield compounds with both dramatically increased potency and minimal toxicity.

Table Of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Protein kinases and signaling pathways

Protein kinases catalyze the transfer of the γ-phosphate of adenosine triphosphate (ATP) to serine, threonine or tyrosine amino acids of their protein substrates. This phosphorylation event can have a profound effect on the substrates' function and controls a multitude of cellular processes, including cell growth and proliferation[1]. Aberrant kinase activity is the cause or consequence of many diseases, especially cancer, and multitude of drugs have been developed with the aim of inhibiting protein kinase activity and thereby killing tumor cells and preventing metastasis[2]. The Ras-mitogen activated protein kinase (Ras/MAPK) pathway plays a central role in converting extracellular stimuli into a wide range of cellular responses and this pathway is dysregulated in many types of cancers[3].
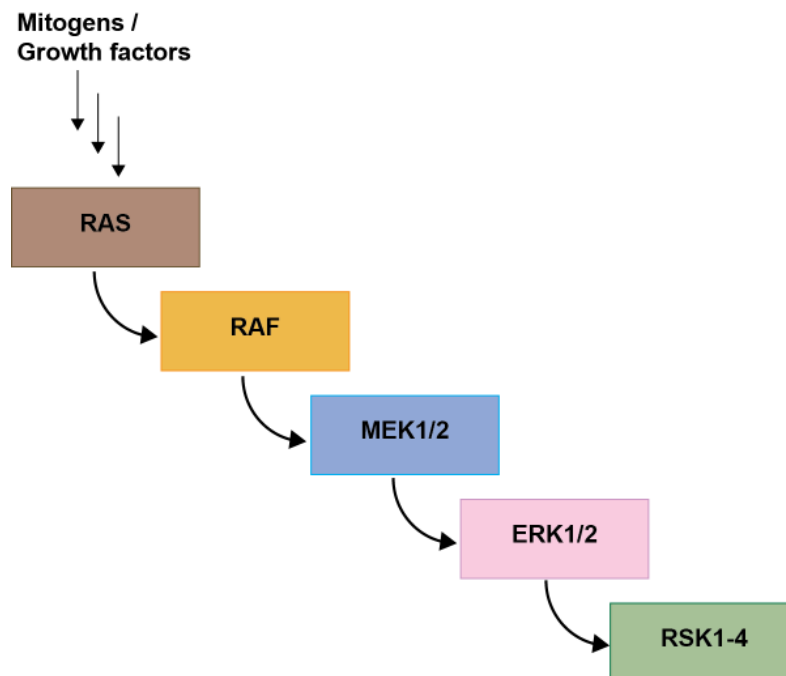
Figure 1.1. Schematic of RSK in the MAPK signaling pathway

Members of the p90 ribosomal S6 kinase (RSK) family are serine-threonine protein kinases that act as downstream effectors of the Ras/MAPK signaling pathway[4,5]. The four RSK isoforms in humans, RSK1 – RSK4 are ubiquitously expressed in cell lines and tissues[6-8], and regulate various cellular processes such as cell growth, proliferation and survival, and cell cycle progression. The RSK isoforms are activated by the extracellular signal-regulated kinases 1 and 2 (ERK 1/2) in response to variety of signaling molecules (growth factors, cytokines, hormones and neurotransmitters) that stimulate the Ras–Raf–Mek–Erk pathway[9] (Figure 1.1). RSKs play an important role in the MAPK signaling cascade and their misregulation has been implicated in multiple diseases, including breast and prostate cancers[10,11].

The human RSK family members are highly identical to each other[12] (75 – 80% sequence identity) and a characteristic feature of this family of kinases is that they all contain two distinct and catalytically functional kinase domains[13,14], an N-terminal kinase domain (NTD) connected by a hydrophobic linker region to a C-terminal kinase domain (CTD) (Figure 1.2). The NTD belongs to the AGC (protein kinase A, G and C) family, while the CTD belongs to the CAMK (Ca^{2+}/calmodulin-dependent protein kinase) family of protein kinases. The only known function of the CTD is activation of the NTD by phosphorylation of a serine residue (S386, Rsk2 human numbering) in a hydrophobic motif of the linker region[14]. Once activated, the NTD phosphorylates a variety of cellular substrates that play a role in diverse biological processes[12,15,16] (REF*10-16o*).
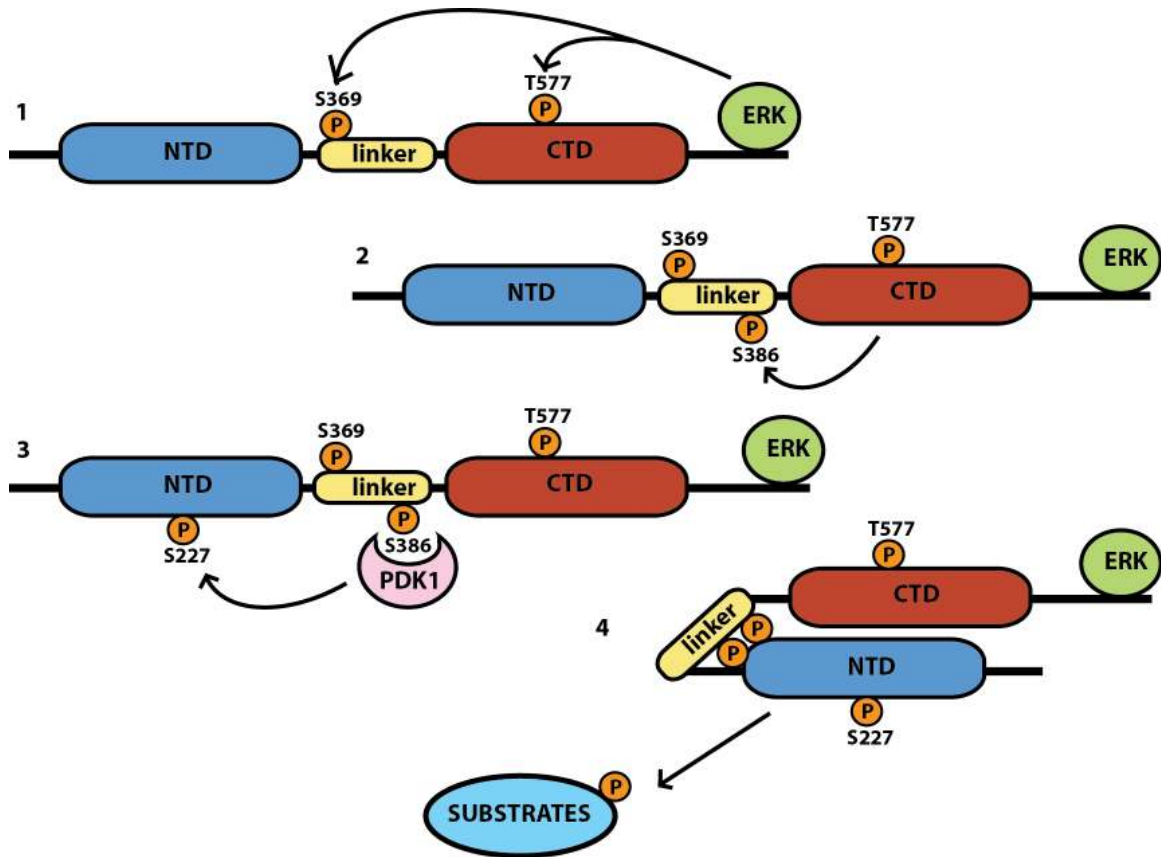
Figure 1.2. Model for ERK1/2 activation of RSK kinases

1.2 RSK activation

The mechanism of RSK activation is complex and the current model involves sequential phosphorylation of six different residues, of which four have been found to be essential for biological RSK activity (Ser227, Ser369, Ser386 and Thr577, human RSK2 numbering)[9,17,18]. Upon stimulation of the RAS-ERK pathway, activated ERK 1/2 binds the carboxy-terminus of RSK and phosphorylates Thr577 in the activation loop of the CTD, as well as Ser369 and Thr365 in the linker region between the two kinase domains[17] (Figure 1.2, step 1). Phosphorylation of Thr577 activates the CTD, resulting in auto-

phosphorylation of Ser386 in the hydrophobic motif of the linker region[18] (Figure 1.2, step 2). The pSer386 serves as a docking site for PDK1, which in turn phosphorylates the activation loop of the NTD at Ser227 and results in complete RSK activation[19,20] (Figure 1.2, step 3). Once PDK1 dissociates, RSK is stabilized in an active conformation and is capable of phosphorylating its downstream cellular targets[21] (Figure 1.2, step 4). Subsequent phosphorylation of RSK at its C-terminus (Ser737) by the NTD promotes release of ERK1/2, resulting in inactivation of the kinase[22]. The NTD of RSK phosphorylates multiple cellular and nuclear proteins and requires the minimum consensus phosphorylation motifs of RRx**S/T** or R/KxRxx**S/T**[23].

1.3 RSK role in biological processes

RSK1 and RSK2 are the most studied isoforms and were originally identified in *Xenopus laevis* extracts[24,25]. They were thought to be responsible for phosphorylating the ribosomal protein S6 (rpS6), but in later experiments the predominant rpS6 kinases were revealed to be the 70 kDa ribosomal S6 kinases (S6K1 and S6K2)[26]. Since their discovery in the late eighties, great strides have been made toward understanding RSK's role in cellular processes. RSK can regulate gene expression via association and phosphorylation of transcriptional regulators including c-Fos[27,28], cAMP response element-binding protein (CREB)[29-31], estrogen receptor-α (ERα)[32], and NFκB/IκBα[33,34] and play an important role in cell survival, migration and proliferation. Loss of RSK2 activity was shown to be the cause of Coffin-Lowry syndrome (CLS)[35], an X-linked

disorder characterized by severe psychomotor retardation[36]. Patients with CLS have multiple mutations in the RSK2 gene, which result in truncated or inactive RSK2 proteins[35].

RSK1 and RSK2 have been shown to positively regulate cell survival in different cell types[37]. Both isoforms phosphorylate the proapoptotic protein Bad, causing its inactivation by cytosolic 14-3-3 proteins[37,38]. The tumor suppressor BimEL is also degraded upon phosphorylation by RSK1 and RSK2[39]. In addition, RSK1 and RSK2 phosphorylate and inactivate the death-associated protein kinase (DAPK), inhibiting its proapoptotic activity and resulting in increased cell survival in response to mitogenic stimulation[40]. Phosphorylation of the transcription factor CREB was shown to promote survival of primary neurons through increased transcription of survival-promoting genes, including those from Bcl-2, Bcl-XL and Mcl-1 [29-31].

Expression profiling analysis of cells treated with a highly selective RSK1/2/4 inhibitor developed in our lab had shown that RSK1 and RSK2 are important ERK effectors of Raf1-dependent migration in epithelial cells and overexpression of constitutively active RSK2 elicits a motile phenotype in MDCK cells[41]. RSK1 and RSK2 were found to stimulate motility and invasion by activating a transcriptional program that enables RSK to coordinately modulate the extracellular environment, the intracellular motility proteins, and receptors mediating communication between the two. A genome-wide mRNA expression analyses revealed RSK-regulated genes with established roles in motility and invasion that are responsible for a diverse forms of motility, including cell

scattering, wound healing, cell multilayering, chemotaxis and 3D EMC invasion[41].

Additionally, a genome wide RNAi screen revealed a broad functional dependence of cell motility phenotypes on RSK activation[42].

RSK2 appears to be a key regulator of cell transformation induced by tumor promoters such as epidermal growth factor (EGF) or 12-*O*-tetraecanoylphorbol-13-acetate (TPA)[43]. Ectopic expression of RSK2 was found to increase proliferation as well as anchorage-independent transformation of mouse epidermal cells[43]. In addition, phospho-proteomics studies revealed that RSK2 is a substrate and a downstream effector of the oncogenic FGFR3 receptor-tyrosine kinase and plays an important role in FGFR3-mediated hematopoietic transformation. Using FGFR3-expressing t(4;14)-positive myeloma cells, FGFR3 was shown to activate RSK2 by facilitating both the ERK-RSK2 interaction and the subsequent phosphorylation of RSK2 by ERK[44]. RSK2 was also found to be activated in diverse leukemia cell lines expressing different leukemogenic tyrosine kinases, including BCR-ABL and FLT3-ITD[45]. Inhibition of RSK2 by the small molecule inhibitor **FMK** induced significant apoptosis in FLT3-ITD positive leukemia cell lines, but not the BCR-ABL induced cells, suggesting that RSK2 is required for the pathogenesis in FLT3-IDT induced myeloid leukemia and may represent an alternative therapeutic target in the treatment of hematologic cancers[45].

1.4 Pharmacological RSK inhibitors

To date, several RSK inhibitors have been identified with promising *in vivo* properties (Figure 1.3).
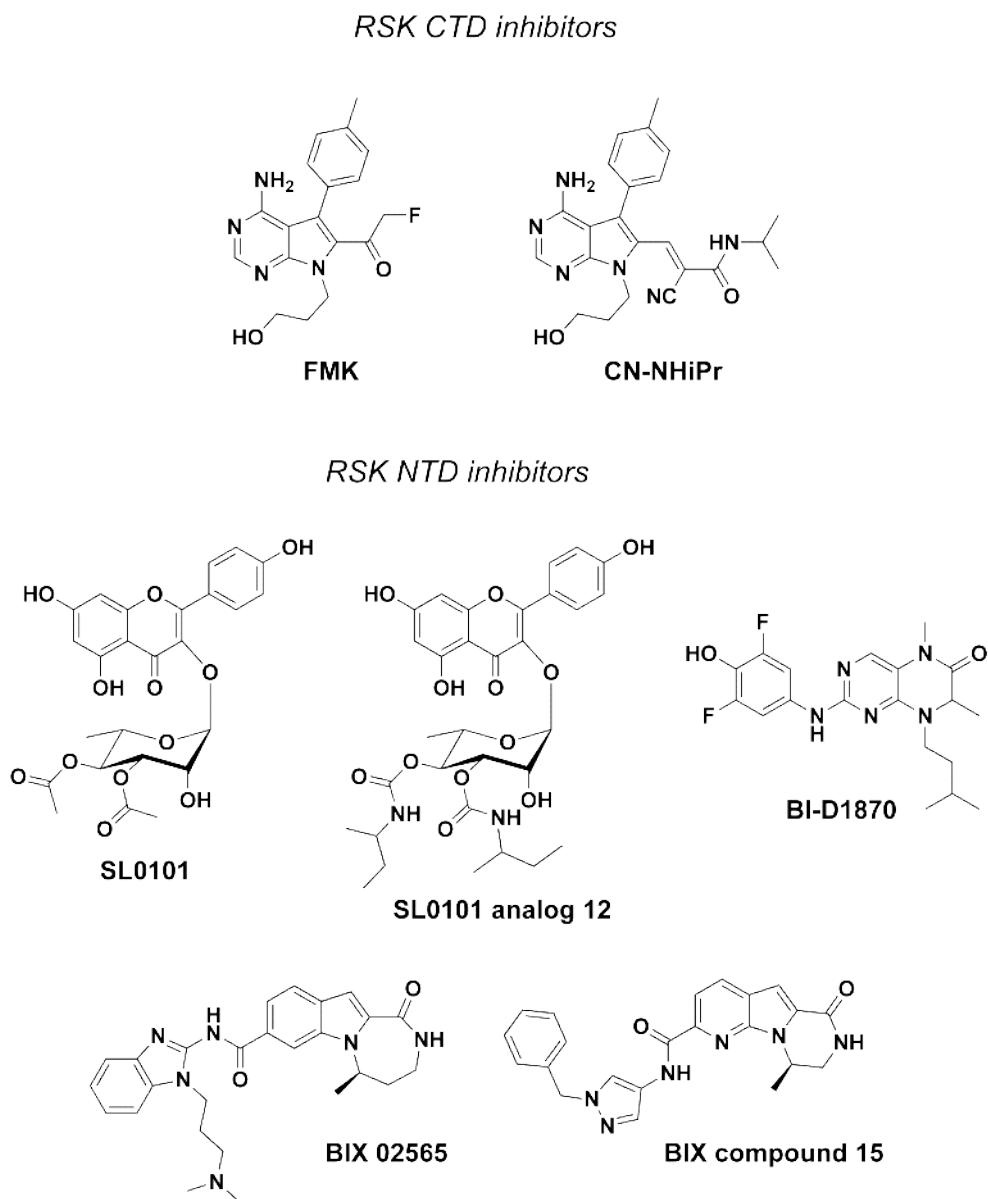
Figure 1.3. Chemical structures of RSK specific small molecule inhibitors

Two of these inhibitors (**FMK** and **CN-NHiPr**) are highly potent and selective and target the CTD of RSK1, RSK2 and RSK4 by covalently modifying a cysteine residue found in the RSK CTD ATP binding site[46,47], while the other three (**CL0101** and its analogs, **BI-D1870**, and the indole-based **BIX02565** and its analogs) target the NTD of RSK1-4 in a reversible, ATP-competitive manner[11,48-50], but have other off-target effects on other proteins.

C-Terminal kinase domain (CTD) RSK inhibitors

FMK is the first irreversible small molecule inhibitor that was rationally designed to selectively inhibit the C-terminal domains of RSK1, 2 and 4 (Figure 1.3). A bioinformatics approach identified two selectivity filters that could be exploited for the design of RSK CTD specific inhibitors – a cysteine residue (Cys436, RSK2 numbering) at the top of the ATP-binding site, as well as a small threonine (Thr493, RSK2 numbering) gatekeeper[46]. All of the RSK isoforms contain that cysteine, while in contrast to RSK1, 2 and 4, RSK3 had a sterically demanding methionine gatekeeper. **FMK** is based on a pyrrolo-pyrimidine scaffold with an appended fluoromethyl ketone electrophile that covalently reacts with Cys436. It also has a *para*-tolyl substituent that is accommodated by the small threonine gatekeeper and therefore fits only in the binding sites of RSK1, RSK2 and RSK4. **FMK** is a highly potent and selective RSK CTD inhibitor with an IC$_{50}$ value of 15 nM against RSK2 in the presence of 100 μM ATP and its exclusive selectivity for RSK1 and RSK2 was demonstrated in human epithelial cell lysates using a biotinylated derivative (**biotin-fmk**) as a probe as well as

from commercial kinase profiling screens. **FMK** has been successfully used as a chemical tool to illuminate RSK's role in tumorigenesis as well as cancer invasion and metastasis[41,44,45]. The specific inhibition of RSKs by **FMK** induced apoptosis in FGFR3 expressing myeloma cells and in FLT3-ITD leukemia cells[45].

We recently developed a closely related analog of **FMK** with a novel mode of covalent inhibition[47]. **CN-NHiPr** (Figure 1.3) contains a doubly substituted Michael acceptor electrophile and potently and selectively inhibits RSK CTD in a covalent, yet reversible manner. **CN-NHiPr** displayed single digit nanomolar potency in enzymatic kinase activity assays against RSK2 CTD (IC_{50} = 4 nM in the presence of 100 μM ATP) and demonstrated longer residence time in cell-based assays, despite the fact that it is a reversible inhibitor. A crystal structure of RSK2 CTD with a *tert*-butyl cyanoacrylate derivative revealed that RSK2 forms a covalent adduct by 1,4 nuclephilic conjugate addition of Cys436 to the electrophilic β-carbon of the cyanoacrylate. **CN-NHiPr** was tested against a large panel of purified human kinases and was found to be highly specific to RSK1 and RSK4 CTDs. **CN-NHiPr** was shown to react rapidly and reversibly with cellular thiols such as glutathione (GSH), yet was able to abolish RAF-induced epithelial cell migration and invasion, consistent with the recently established role of RSK signaling in these processes [41,47].

N-Terminal kinase domain (NTD) RSK inhibitors

The kaempferol glycoside **SL0101** (Figure 1.3) was the first identified specific RSK inhibitor that binds the NTD[11,51]. **SL0101** is a natural product

isolated from the tropical plant *Fosteronia refracta* and was identified by screening botanical extracts in biochemical assays for RSK inhibition. **SL0101** is an ATP-competitive inhibitor and was reported to reversibly inhibit RSK2 with an IC_{50} of 90 nM in the presence of 10 μM ATP. Its activity in intact cells was almost 500-fold lower, with reported EC_{50} of 50 μM. Treatment with **SL0101** inhibits proliferation of human prostate and breast tumors, but has no effect on healthy breast tissue cells[11]. The anti-proliferative effect of **SL0101** is sustained for 48 hours and is due to RSK regulation of cyclin D1 levels, which leads to cell cycle block in G1 phase. However, after 48 hours of treatment cells resume proliferation, indicating that **SL0101** has poor *in vitro* stability. Recently, a crystal structure of RSK2 NTD with **SL0101** revealed that the N-lobe, the hinge region, and the a-D helix of the kinase undergo a conformational change that results in a formation of a hydrophobic pocket that accommodates the inhibitor[52]. Further analysis and chemical substitution at the 3' and 4' positions of the sugar moiety with carbamate analogues yielded **SL0101 analog 12** with improved *in vitro* stability and retained RSK potency and specificity[53,54].

The dihydropteridinone **BI-D1870** is another reversible and ATP competitive RSK NTD inhibitor[48]. It is relatively selective and inhibits all four RSK isoforms *in vitro* with an IC_{50} of 15 – 30 nM, but does not significantly inhibit ten other closely related AGC kinase members and over 40 other protein kinases tested at 100-fold higher concentrations. The main exception is PLK1, which is inhibited with an IC_{50} of 100 nM. **BI-D1870** is cell permeable and was able to inhibit the RSK-mediated phorbol ester- and EGF-induced phosphorylation of

11

glycogen synthase kinase-3β (GSK3β) and LKB1 in human embryonic kidney 293 cells, however complete inhibition of these RSK substrates required a concentration of 10 μM **BI-D1870**. In contrast, **BI-D1870** had little effect on the PMA-induced activation of ERK1/2, which is catalyzed by Raf and MKK1. Its effect on cellular proliferation and survival has not been reported. In addition, we have observed in our lab that the formation of monopolar spindles is induced at 10 μM **BI-D1870** (unpublished data), so any conclusions from cellular experiments using such high concentrations of inhibitor need to be scrutinized for off-target effects.

Most recently, a series of novel RSK NTD inhibitors with an indole lactam scaffold have been developed[49,50]. An X-ray crystal structure of the NTD of RSK2 as well as molecular docking were used to guide the design of new compounds. The lead hits were tested in *in vitro* RSK2 activity assay and were further optimized for cellular potency and kinase selectivity. **BIX 02565** was identified as novel RSK2 inhibitor with high potency (IC_{50} = 1nM), but unfortunately it bound multiple adrenergic receptors in subsequent selectivity assays, suggesting a potential for cardiovascular liabilities. Additional efforts and structure activity relationship (SAR) studies around the off-target adrenergic receptor binding resulted in the identification of a new potent RSK NTD inhibitor (**BIX compound 15**, *in vitro* RSK2 IC_{50} = 0.2 nM and cellular IC_{50} = 0.32 nM; see Figure 1.3 for structure) with no adverse cardiovascular effects *in vivo* at high plasma concentrations, but its utility as a RSK inhibitor has yet to be described[55].

1.5 REFERENCES

1   Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934 (2002).

2   Melnikova, I. & Golden, J. Targeting protein kinases. *Nat Rev Drug Discov* **3**, 993-994 (2004).

3   Sebolt-Leopold, J. S. & Herrera, R. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat Rev Cancer* **4**, 937-947 (2004).

4   Frodin, M. & Gammeltoft, S. Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Mol Cell Endocrinol* **151**, 65-77 (1999).

5   Cargnello, M. & Roux, P. P. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev* **75**, 50-83 (2011).

6   Zeniou, M., Ding, T., Trivier, E. & Hanauer, A. Expression analysis of RSK gene family members: the RSK2 gene, mutated in Coffin-Lowry syndrome, is prominently expressed in brain structures essential for cognitive function and learning. *Hum Mol Genet* **11**, 2929-2940 (2002).

7   Moller, D. E., Xia, C. H., Tang, W., Zhu, A. X. & Jakubowski, M. Human rsk isoforms: cloning and characterization of tissue-specific expression. *Am J Physiol* **266**, C351-359 (1994).

8   Dummler, B. A. *et al.* Functional characterization of human RSK4, a new 90-kDa ribosomal S6 kinase, reveals constitutive activation in most cell types. *J Biol Chem* **280**, 13304-13314 (2005).

9   Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J. & Cohen, P. Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK. *J Biol Chem* **273**, 1496-1505 (1998).

10   Clark, D. E. *et al.* The serine/threonine protein kinase, p90 ribosomal S6 kinase, is an important regulator of prostate cancer cell proliferation. *Cancer Res* **65**, 3108-3116 (2005).

11   Smith, J. A. *et al.* Identification of the first specific inhibitor of p90 ribosomal S6 kinase (RSK) reveals an unexpected role for RSK in cancer cell proliferation. *Cancer Res* **65**, 1027-1034 (2005).

12   Carriere, A., Ray, H., Blenis, J. & Roux, P. P. The RSK factors of activating the Ras/MAPK signaling cascade. *Front Biosci* **13**, 4258-4275 (2008).

13   Fisher, T. L. & Blenis, J. Evidence for two catalytically active kinase domains in pp90rsk. *Mol Cell Biol* **16**, 1212-1219 (1996).

14   Chrestensen, C. A. & Sturgill, T. W. Characterization of the p90 ribosomal S6 kinase 2 carboxyl-terminal domain as a protein kinase. *J Biol Chem* **277**, 27733-27741 (2002).

15   Romeo, Y., Zhang, X. & Roux, P. P. Regulation and function of the RSK family of protein kinases. *Biochem J* **441**, 553-569 (2012).

16   Anjum, R. & Blenis, J. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* **9**, 747-758 (2008).

17   Gavin, A. C. & Nebreda, A. R. A MAP kinase docking site is required for phosphorylation and activation of p90(rsk)/MAPKAP kinase-1. *Curr Biol* **9**, 281-284 (1999).

18   Vik, T. A. & Ryder, J. W. Identification of serine 380 as the major site of autophosphorylation of Xenopus pp90rsk. *Biochem Biophys Res Commun* **235**, 398-402 (1997).

19   Frodin, M., Jensen, C. J., Merienne, K. & Gammeltoft, S. A phosphoserine-regulated docking site in the protein kinase RSK2 that recruits and activates PDK1. *Embo J* **19**, 2924-2934 (2000).

20   Jensen, C. J. *et al.* 90-kDa ribosomal S6 kinase is phosphorylated and activated by 3-phosphoinositide-dependent protein kinase-1. *J Biol Chem* **274**, 27168-27176 (1999).

21   Hauge, C. *et al.* Mechanism for activation of the growth factor-activated AGC kinases by turn motif phosphorylation. *Embo J* **26**, 2251-2261 (2007).

22   Roux, P. P., Richards, S. A. & Blenis, J. Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Mol Cell Biol* **23**, 4796-4804 (2003).

23   Leighton, I. A., Dalby, K. N., Caudwell, F. B., Cohen, P. T. & Cohen, P. Comparison of the specificities of p70 S6 kinase and MAPKAP kinase-1 identifies a relatively specific substrate for p70 S6 kinase: the N-terminal

kinase domain of MAPKAP kinase-1 is essential for peptide phosphorylation. *FEBS Lett* **375**, 289-293 (1995).

24   Erikson, E. & Maller, J. L. A protein kinase from Xenopus eggs specific for ribosomal protein S6. *Proc Natl Acad Sci U S A* **82**, 742-746 (1985).

25   Jones, S. W., Erikson, E., Blenis, J., Maller, J. L. & Erikson, R. L. A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proc Natl Acad Sci U S A* **85**, 3377-3381 (1988).

26   Sturgill, T. W., Ray, L. B., Erikson, E. & Maller, J. L. Insulin-stimulated MAP-2 kinase phosphorylates and activates ribosomal protein S6 kinase II. *Nature* **334**, 715-718 (1988).

27   Chen, R. H., Abate, C. & Blenis, J. Phosphorylation of the c-Fos transrepression domain by mitogen-activated protein kinase and 90-kDa ribosomal S6 kinase. *Proc Natl Acad Sci U S A* **90**, 10952-10956 (1993).

28   Chen, R. H., Juo, P. C., Curran, T. & Blenis, J. Phosphorylation of c-Fos at the C-terminus enhances its transforming activity. *Oncogene* **12**, 1493-1502 (1996).

29   Bonni, A. *et al.* Cell survival promoted by the Ras-MAPK signaling pathway by transcription-dependent and -independent mechanisms. *Science* **286**, 1358-1362 (1999).

30   Ginty, D. D., Bonni, A. & Greenberg, M. E. Nerve growth factor activates a Ras-dependent protein kinase that stimulates c-fos transcription via phosphorylation of CREB. *Cell* **77**, 713-725 (1994).

31   Xing, J., Ginty, D. D. & Greenberg, M. E. Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. *Science* **273**, 959-963 (1996).

32   Joel, P. B. *et al.* pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. *Mol Cell Biol* **18**, 1978-1984 (1998).

33   Ghoda, L., Lin, X. & Greene, W. C. The 90-kDa ribosomal S6 kinase (pp90rsk) phosphorylates the N-terminal regulatory domain of IkappaBalpha and stimulates its degradation in vitro. *J Biol Chem* **272**, 21281-21288 (1997).

34   Schouten, G. J. *et al.* IkappaB alpha is a target for the mitogen-activated 90 kDa ribosomal S6 kinase. *Embo J* **16**, 3133-3144 (1997).

35   Trivier, E. *et al.* Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome. *Nature* **384**, 567-570 (1996).

36   Pereira, P. M., Schneider, A., Pannetier, S., Heron, D. & Hanauer, A. Coffin-Lowry syndrome. *Eur J Hum Genet* **18**, 627-633 (2010).

37   Shimamura, A., Ballif, B. A., Richards, S. A. & Blenis, J. Rsk1 mediates a MEK-MAP kinase cell survival signal. *Curr Biol* **10**, 127-135 (2000).

38   Tan, Y., Ruan, H., Demeter, M. R. & Comb, M. J. p90(RSK) blocks bad-mediated cell death via a protein kinase C-dependent pathway. *J Biol Chem* **274**, 34859-34867 (1999).

39   Dehan, E. *et al.* betaTrCP- and Rsk1/2-mediated degradation of BimEL inhibits apoptosis. *Mol Cell* **33**, 109-116 (2009).

40   Anjum, R., Roux, P. P., Ballif, B. A., Gygi, S. P. & Blenis, J. The tumor suppressor DAP kinase is a target of RSK-mediated survival signaling. *Curr Biol* **15**, 1762-1767 (2005).

41   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* **35**, 511-522 (2009).

42   Smolen, G. A. *et al.* A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. *Genes Dev* **24**, 2654-2665 (2010).

43   Cho, Y. Y. *et al.* Ribosomal S6 kinase 2 is a key regulator in tumor promoter induced cell transformation. *Cancer Res* **67**, 8104-8112 (2007).

44   Kang, S. *et al.* FGFR3 activates RSK2 to mediate hematopoietic transformation through tyrosine phosphorylation of RSK2 and activation of the MEK/ERK pathway. *Cancer Cell* **12**, 201-214 (2007).

45   Elf, S. *et al.* p90RSK2 is essential for FLT3-ITD- but dispensable for BCR-ABL-induced myeloid leukemia. *Blood* **117**, 6885-6894 (2011).

46   Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **308**, 1318-1321 (2005).

47   Serafimova, I. M. *et al.* Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol* **8**, 471-476 (2012).

48   Sapkota, G. P. *et al.* BI-D1870 is a specific inhibitor of the p90 RSK (ribosomal S6 kinase) isoforms in vitro and in vivo. *Biochem J* **401**, 29-38 (2007).

49   Boyer, S. J. *et al.* Indole RSK inhibitors. Part 1: discovery and initial SAR. *Bioorg Med Chem Lett* **22**, 733-737 (2012).

50   Kirrane, T. M. *et al.* Indole RSK inhibitors. Part 2: optimization of cell potency and kinase selectivity. *Bioorg Med Chem Lett* **22**, 738-742 (2012).

51   Maloney, D. J. & Hecht, S. M. Synthesis of a potent and selective inhibitor of p90 Rsk. *Org Lett* **7**, 1097-1099 (2005).

52   Utepbergenov, D. *et al.* Insights into the inhibition of the p90 ribosomal S6 kinase (RSK) by the flavonol glycoside SL0101 from the 1.5 A crystal structure of the N-terminal domain of RSK2 with bound inhibitor. *Biochemistry* **51**, 6499-6510 (2012).

53   Smith, J. A., Maloney, D. J., Hecht, S. M. & Lannigan, D. A. Structural basis for the activity of the RSK-specific inhibitor, SL0101. *Bioorg Med Chem* **15**, 5018-5034 (2007).

54   Hilinski, M. K., Mrozowski, R. M., Clark, D. E. & Lannigan, D. A. Analogs of the RSK inhibitor SL0101: optimization of in vitro biological stability. *Bioorg Med Chem Lett* **22**, 3244-3247 (2012).

55   Fryer, R. M. *et al.* Mitigation of off-target adrenergic binding and effects on cardiovascular function in the discovery of novel ribosomal S6 kinase 2 inhibitors. *J Pharmacol Exp Ther* **340**, 492-500 (2012).

CHAPTER 2

Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles[1]

2.1 Abstract

Targeting noncatalytic cysteine residues with irreversible acrylamide-based inhibitors is a powerful approach for enhancing pharmacological potency and selectivity. Nevertheless, concerns about off-target modification motivate the development of reversible cysteine-targeting strategies. Here we show that electron-deficient olefins, including acrylamides, can be tuned to react with cysteine thiols in a rapidly reversible manner. Installation of a nitrile group increased the olefins' intrinsic reactivity, yet paradoxically eliminated the formation of irreversible adducts. Incorporation of these electrophiles into a noncovalent kinase recognition scaffold produced slowly dissociating, covalent inhibitors of the p90 ribosomal protein S6 kinase, RSK. A cocrystal structure revealed specific noncovalent interactions that stabilize the complex by positioning the electrophilic carbon near the targeted cysteine. Disruption of these interactions by protein unfolding or proteolysis promoted instantaneous cleavage of the covalent bond. Our results establish a chemistry-based framework for engineering sustained covalent inhibition without accumulating permanently modified proteins and peptides.

3.2 Introduction

Cysteine displays rich chemistry through its nucleophilic thiol group. It is also one of the least common amino acids in proteins. Together, these properties make cysteine residues ideal for targeting with covalent drugs, which have the potential to exhibit high levels of target specificity and a prolonged duration of action[1-3]. Although frequently designed to inactivate conserved, catalytically essential nucleophiles (e.g., in Ser, Thr, and Cys proteases), covalent inhibitors can achieve maximal selectivity among related targets by exploiting the intrinsic nucleophilicity of poorly conserved, noncatalytic cysteines[4]. This strategy, guided by structural bioinformatics analysis, has led to the design of selective, irreversible inhibitors of protein kinases[5-9], and more recently, the NS3/4A serine protease from hepatitis C virus[10].

Protein kinases are challenging therapeutic targets from the standpoint of achieving sustained inhibition of the desired kinase without affecting structurally related kinases. A majority of the 518 human kinases have an accessible noncatalytic cysteine within reach of the active site[11,12], and at least four cysteine-targeted kinase inhibitors are in clinical trials for advanced cancer indications. They all rely on an acrylamide electrophile to form an irreversible covalent bond with the kinase[4]. Acrylamide-based kinase inhibitors react irreversibly with glutathione[13] and therefore may react with proteins other than the desired target, especially proteins with hyper-reactive cysteines[14]. Although the risk may be low and more relevant to chronic diseases than advanced cancer, there are currently no preclinical models that can accurately predict the

22

toxicological potential of chemically reactive drugs and drug metabolites[15-17]. Thus, current drug discovery efforts mostly aim to avoid the formation of irreversible covalent adducts.

Based on these considerations, we sought reversible electrophilic inhibitors that would retain the advantages of covalent cysteine targeting (prolonged duration of action and high selectivity) without the potential liabilities associated with irreversible adduct formation. The few known covalent inhibitors that reversibly target noncatalytic cysteines were discovered by random high-throughput screening[18,19], and the chemical basis of their reversibility is not clear. In this study, we elucidate specific structural features underlying reversible thiol addition to electron-deficient olefins and apply these principles to the design of reversible, cysteine-targeted kinase inhibitors.

2.3 Results

Reversibility of thiol addition to activated olefins

Experiments in the 1960s revealed that simple thiols react instantaneously with 2-cyanoacrylates at physiological pH, but the products could not be isolated or structurally characterized[20]. A potential explanation for these results is that the reaction, possibly a Michael-type conjugate addition, is a rapid-equilibrium process. To test this hypothesis and define the structural requirements for rapid reversibility, we compared three simple Michael acceptors, activated by a methyl ester (**1**), a nitrile (**2**), or both electron-withdrawing groups (**3**) (Figure 2.1a).
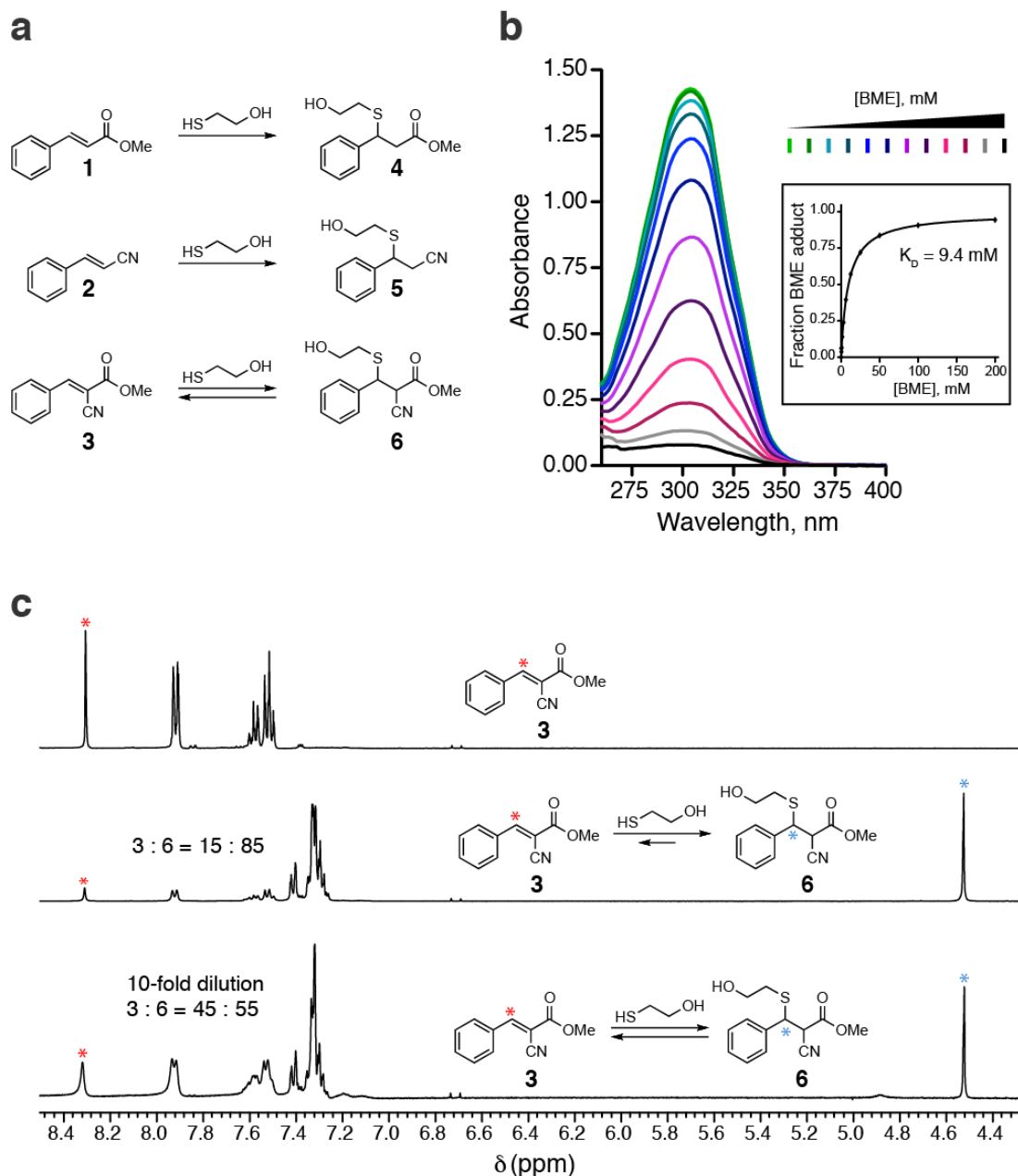
Figure 2.1. Thiol reactivity of electron-deficient olefins. a, Conjugate addition reactions of beta-mercaptoethanol (BME) with olefins **1**–**3**. **b,** Cyanoacrylate **3** (100 µM) was treated with increasing concentrations of BME (0.2–200 mM in two-fold increments) at pH 7.4 and monitored by UV-visible absorption spectroscopy. The inset shows the fraction of BME adduct vs. [BME], from which the equilibrium dissociation constant (K_D) was derived. **c,** ^{1}H NMR spectra showing reversible addition of BME to cyanoacrylate **3**. Cyanoacrylate **3** (90 mM, top spectrum) was treated with 100 mM BME in DMSO-d_6:PBS-d (3:1), affording an 15:85 mixture that favored the BME adduct **6** (middle). Upon 10-fold dilution, the equilibrium shifted to the left (bottom).   Red and blue asterisks

indicate protons used to determine ratios of **3:6.** The NMR dilution experiment was carried out by Dr. Shyam Krishnan .

Reactions of acrylate **1** and acrylonitrile **2** with the model thiol, beta-mercaptoethanol (BME), produced the stable thioether adducts **4** and **5**, which were easily isolated and characterized (Figure 2.2).
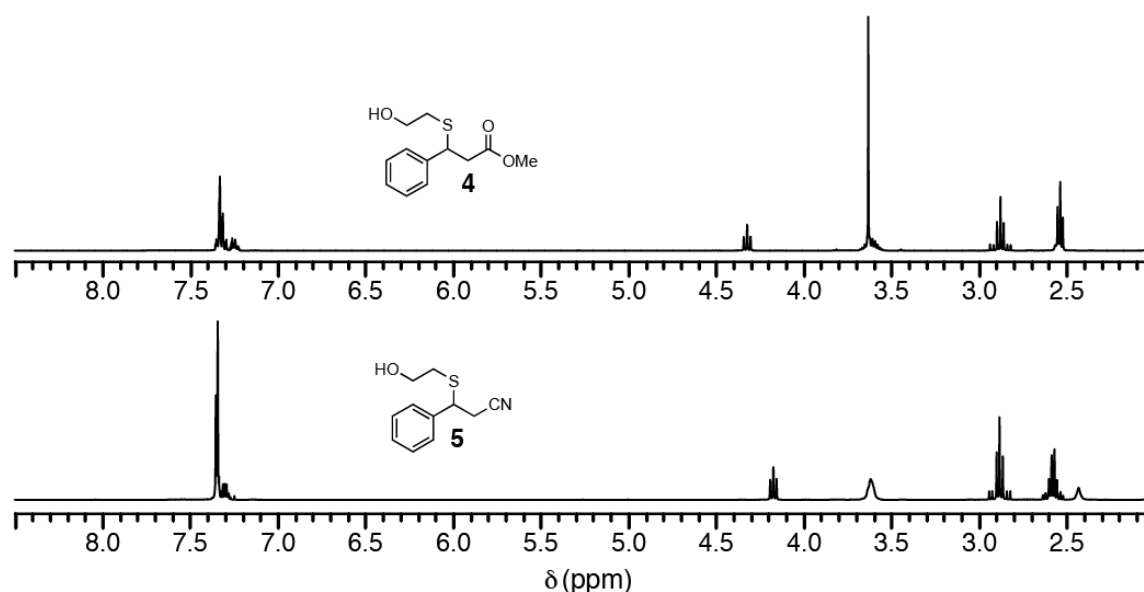
Figure 2.2. ^{1}H NMR spectra (CDCl$_3$) of stable BME adducts **4** and **5** (see Methods for additional characterization data).

By contrast, when the doubly activated Michael acceptor **3** was treated with BME (Figure 2.1a), only the starting cyanoacrylate was recovered. Addition of increasing concentrations of BME caused a stepwise reduction in the prominent UV-visible absorption band of cyanoacrylate **3** (l$_{max}$ 304 nm), and fitting these titration data provided an apparent equilibrium dissociation constant (K$_D$) of 9.4 mM (Figure 2.1b). ^{1}H NMR provided further spectroscopic evidence for the formation of an adduct corresponding to thioether **6**, and dilution

experiments confirmed that the reaction was rapidly reversible (Figure 2.1c). The facile reversion of thioether adduct **6** to the starting cyanoacrylate likely derives from its enhanced kinetic and thermodynamic acidity, estimated[21] to be 10–15 orders of magnitude greater than the carbon acidity of stable adducts **4** and **5**. Hence, the combined influence of a nitrile and an ester on the acidity of the alpha C–H bond facilitates rapid elimination of thiol adducts at physiological pH, while at the same time accelerating the rate of thiol addition.

To test whether this chemistry is applicable to structurally related Michael acceptors, we synthesized five cyanoacrylamides with diverse substituents attached to the electrophilic beta-carbon (**7**-**11**, Figure 2.3a).

	BME K_D (mM)
7	1.29 ± 0.03
8	6.23 ± 1.19
9	0.18 ± 0.004
10	11.86 ± 0.92
11	32.82 ± 0.57
entacapone	4.21 ± 0.13
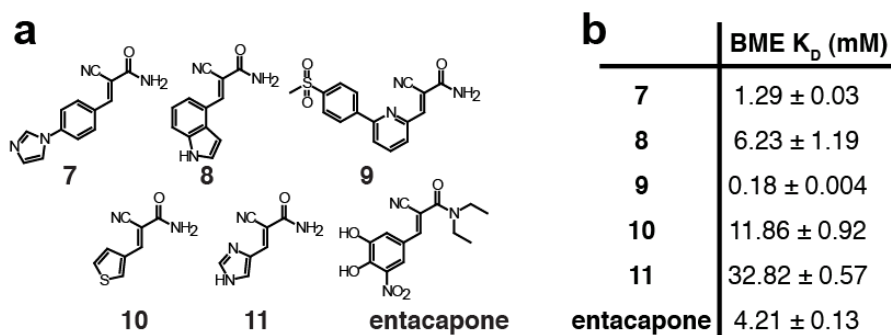
Figure 2.3. BME reacts rapidly and reversibly with diverse cyanoacrylamides. a, Chemical structures of cyanoacrylamides tested for BME reactivity (for synthesis and characterization data, see Methods). **b**, Equilibrium dissociation constants of BME/cyanoacrylamide adducts determined by BME titration and UV-visible spectroscopy.

We included entacapone, a cyanoacrylamide-based drug used for the treatment of Parkinson's disease; to our knowledge, the reactivity of entacapone toward thiols has not been reported previously. Similar to cyanoacrylate **3**, the cyanoacrylamides reacted with millimolar concentrations of BME in a manner

that was reversed within seconds upon dilution (Figure 2.4). Depending on the beta-substituent, the equilibrium dissociation constants ranged from 0.2 to 33 mM, corresponding to standard free energy changes (DG_o) of –2 to –5 kcal/mol. The ability of beta-substituted cyanoacrylamide derivatives, including a widely used drug, to form rapidly reversible adducts with thiols suggested an attractive approach for targeting noncatalytic cysteines in small-molecule binding sites, without the use of irreversible electrophiles.
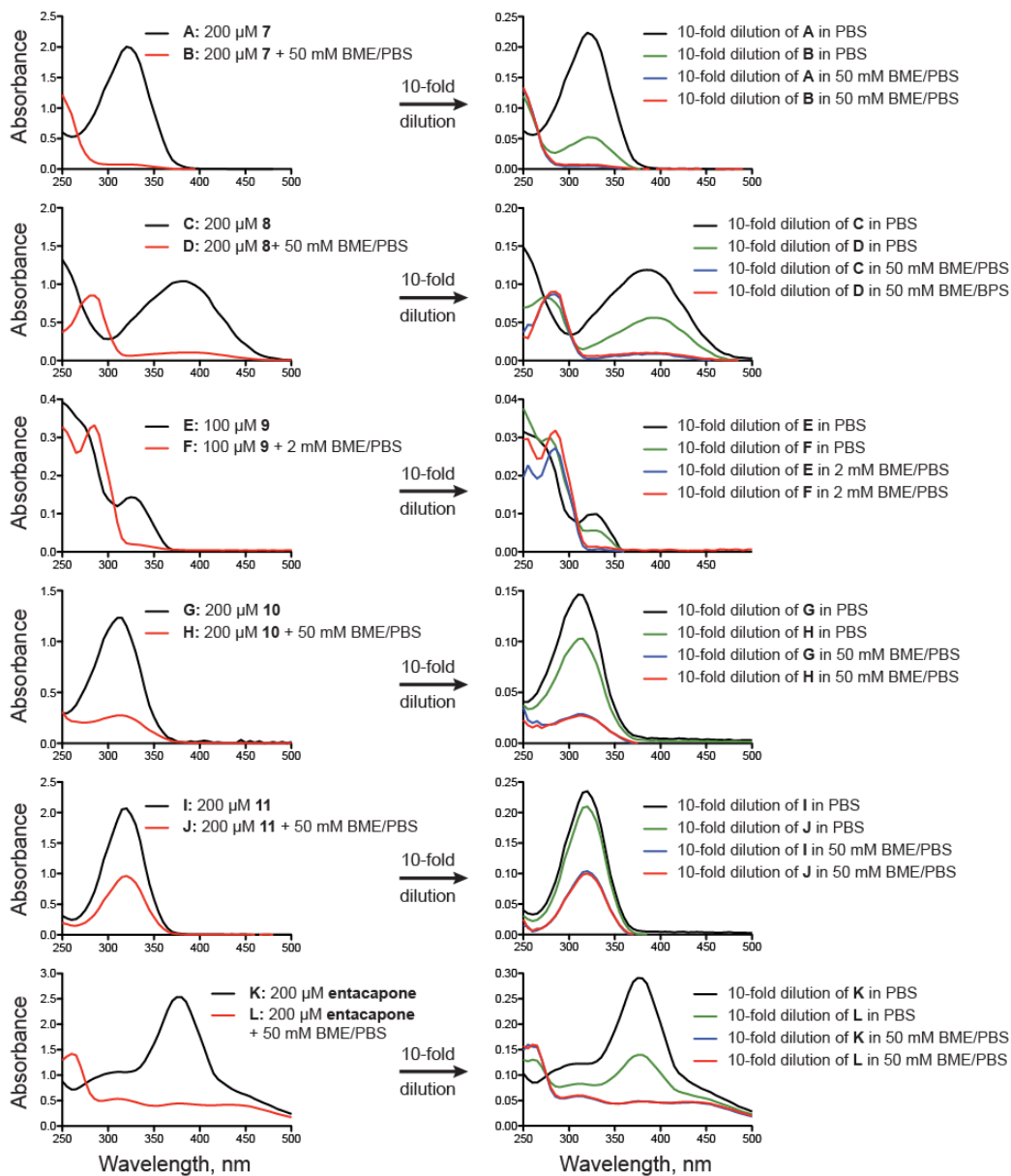
Figure 2.4. Reversibility of BME addition to activated olefins. UV-visible spectroscopy assay showing reversible formation of BME adducts with cyanoacrylamides. Cyanoacrylamides in PBS (1-2% DMSO) were treated with or without the indicated concentration of BME for 10 min, after which UV-visible absorption spectra were recorded (left panels). Each reaction was then diluted 10-fold into PBS or PBS/BME as indicated, and the absorption spectra were recorded again (right panels). Reversibility is demonstrated by the reappearance or increase of the cyanoacrylamide absorption peak after dilution of BME-containing samples into PBS (compare red and green curves in right panels). The extent of cyanoacrylamide recovery varies depending on the equilibrium constant and the concentration of BME after dilution.

Rational design of reversible covalent kinase inhibitors

We prepared a series of electrophilic pyrrolopyrimidines (**12**-**15**, Figure 2.5a), analogous to the model Michael acceptors in Figure 2.1, and compared their biochemical activities against the C-terminal kinase domain (CTD) of p90 ribosomal protein S6 kinase, RSK2. The electrophilic beta-carbon was designed to be proximal to Cys436 in the RSK2 active site based on the predicted binding orientation of the pyrrolopyrimidine scaffold, similar to the irreversible fluoromethylketone-based inhibitor developed previously ('**FMK**'; see Fig. 7 for structure) [6,22]. Treatment of the RSK2-CTD with five equivalents of acrylate **12** or acrylonitrile **13** led to the slow formation of a covalent adduct, as revealed by liquid chromatography–mass spectrometry (LC-MS) (Figure 2.5b). By contrast, the doubly activated Michael acceptors **14** and **15** failed to produce adducts detectable by LC-MS (Figure 2.5b), despite being more than 200-fold more potent than the singly activated Michael acceptors **12** and **13** in kinase activity assays (Figure 2.5c). Mutation of Cys436 to valine conferred ~1000-fold resistance to inhibitors **14** and **15** (Figure 2.5d), suggesting that potent RSK2 inhibition requires covalent bond formation. Cyanoacrylamide **15** reacted rapidly and reversibly with BME and glutathione (K_D = 5.1 and 7.3 mM, Figure 2.6a-c), yet pretreatment with 10 mM glutathione had no effect on its ability to inhibit RSK2 (IC_{50} = 4 nM, Figure 2.6d). Cyanoacrylamide **15** did not react significantly with lysine or ethanolamine (up to 100 mM, Figure 2.6e-f).
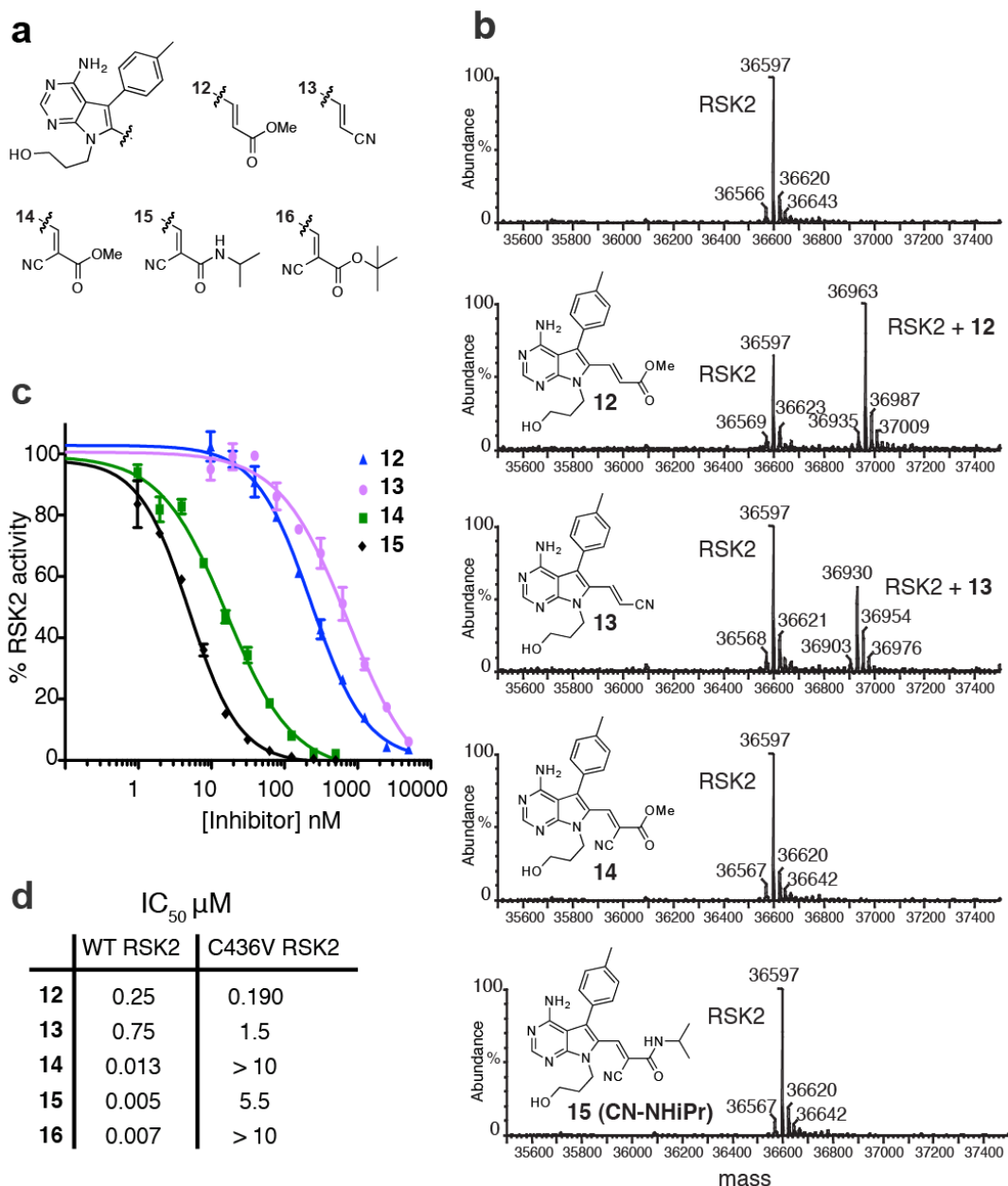
Figure 2.5. Inhibition of RSK2-CTD by doubly activated Michael acceptors.
a, Chemical structures of pyrrolopyrimidine Michael acceptors **12**–**15** . **b,** RSK2-CTD (5 µM) was treated with pyrrolopyrimidines **12**–**15** (25 µM) for 1 h, followed by LC-MS analysis. Observed molecular masses of unmodified RSK2-CTD and the 1:1 adducts with **12** and **13** are consistent with the predicted values. **c,** In vitro kinase assays of RSK2-CTD. Shown are mean values from duplicate measurements (± range). Data were fit with PRISM 4.0 to provide IC_{50} values. **d,** Half-maximum inhibitory concentrations (IC_{50} in µM) against wild type (WT) or C436V.
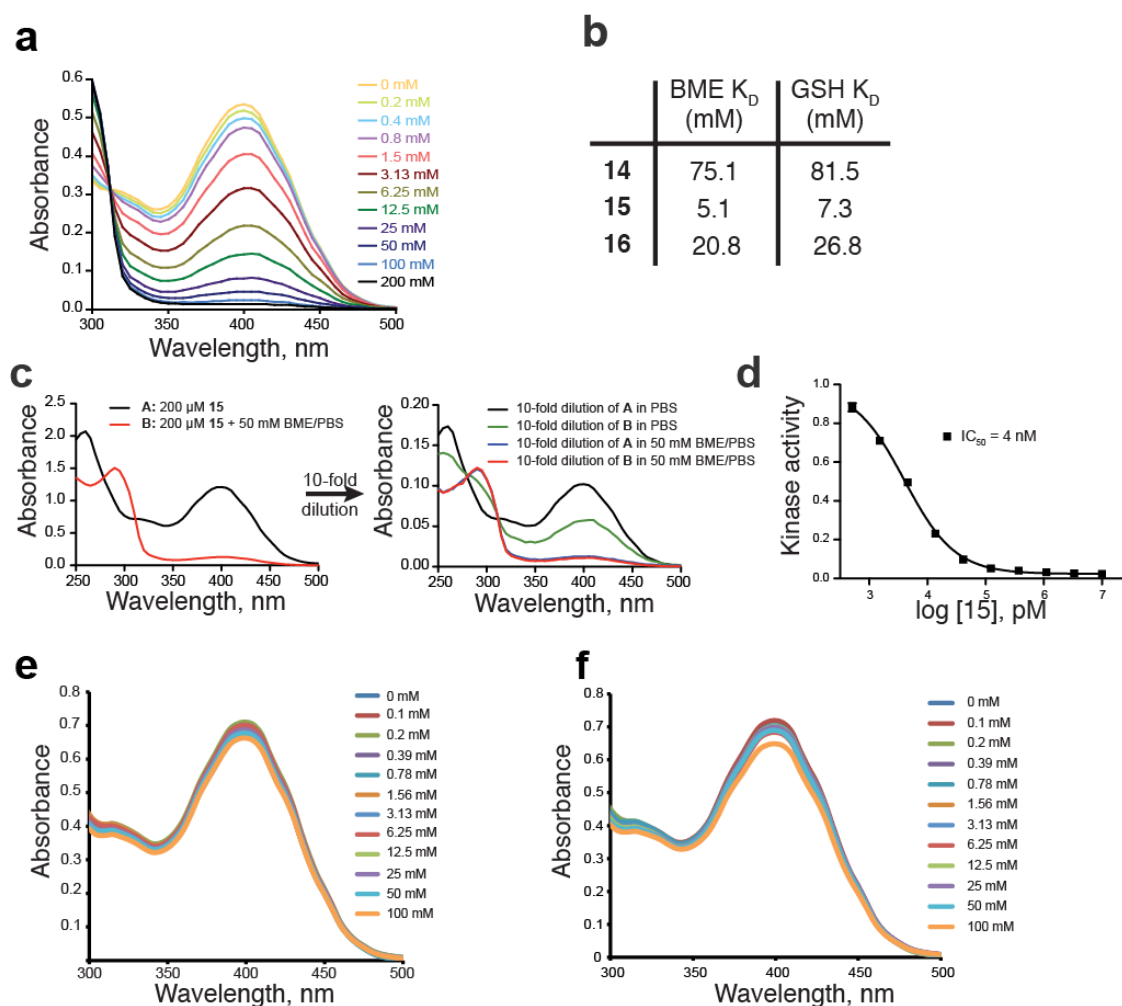
Figure 2.6. Reactivity of electrophilic pyrrolopyrimidines with thiols and amines. a, Absorption spectra of cyanoacrylamide **15** (100 µM in PBS, pH 7.4) in the presence of increasing concentrations of BME (0.2–200 mM, PBS, pH 7.4). Similar spectra were obtained by titration with glutathione (GSH, data not shown). **b**, Equilibrium dissociation constants determined by titrating BME or GSH vs. the doubly activated Michael acceptors **14–16**. **c**, UV-visible absorption assay depicting the rapid reversion upon dilution of the BME/**15** adduct. **d**, Kinase assay inhibition curve showing that pretreatment of cyanoacrylamide **15** with 10 mM GSH does not affect its potency toward RSK2-CTD. This experiment was performed by Rand Miller identically to the one shown in Figure 2.5c, except that 10 mM GSH was included during the 30 min pre-incubation period. **e**, Absorption spectra of cyanoacrylamide **15** (100 µM in PBS, pH 7.4) in the presence of increasing concentrations of ethanolamine (0.1–100 mM, PBS, pH 7.4). **f**, Absorption spectra of cyanoacrylamide **15** (100 µM in PBS, pH 7.4) in the presence of increasing concentrations of lysine (0.1–100 mM, PBS, pH 7.4). The absorbance assays depicted in **e** and **f** were performed by Dr. Shyam Krishnan.

RSK2-CTD kinase activity recovered from inhibition by cyanoacrylate **14** after dialysis for three days at 4°C, whereas recovery from cyanoacrylamide **15** was much slower under these conditions (Figure 2.7a). To estimate their biochemical dissociation rates, we developed a kinetic trapping assay in which RSK2-CTD was first saturated with a reversible inhibitor, followed by treatment with a large excess of FMK. In the absence of inhibitor pretreatment, FMK (100 µM) reacted with Cys436 of RSK2-CTD to form a 1:1 covalent adduct ($t_{1/2}$ < 1 min), as determined by MS analysis (Figure 2.7b). Pretreatment of RSK2-CTD (5 µM) with cyanoacrylate **14** or cyanoacrylamide **15** (10 µM) caused a dramatic reduction in the rate of FMK labeling (Figure 2.7c), from which we could estimate dissociation half-times of 42 and 245 min, respectively. Thus, inhibition of RSK2-CTD by cyanoacrylamide **15** is long-lived yet fully reversible, with complete dissociation occurring on a time scale of several hours.
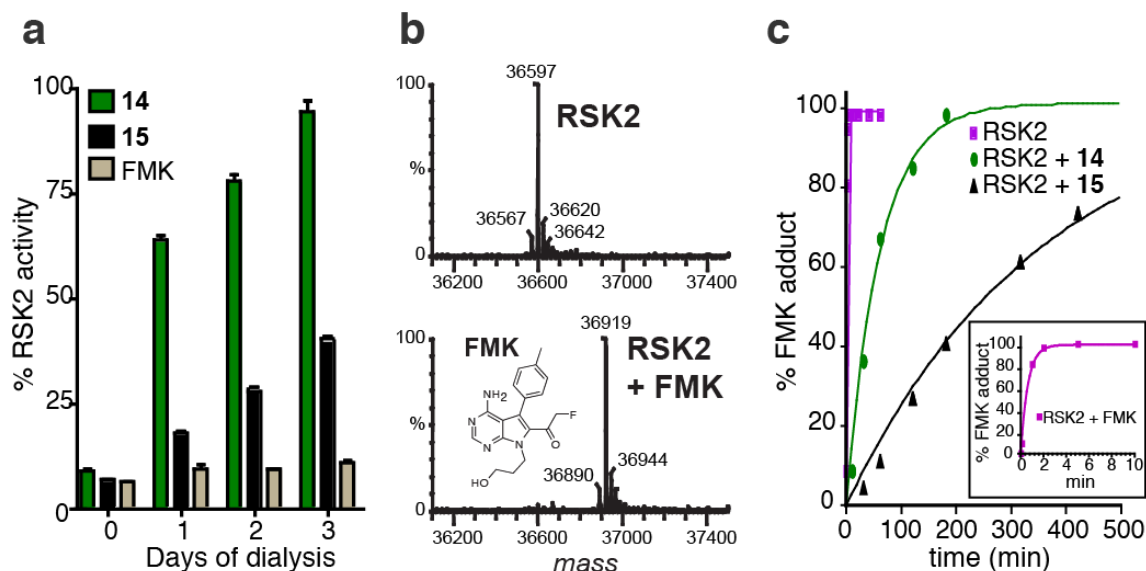
Figure 2.7. Sustained, reversible inhibition of RSK2-CTD by doubly activated Michael acceptors. a, RSK2-CTD (50 nM) was treated with the indicated inhibitors (1 µM) or DMSO for 1 h at room temp, prior to dialyzing at 4˚C. Aliquots were tested daily for kinase activity, normalized to DMSO control (mean ± SD, triplicate values). **b,** Mass spectra showing irreversible modification of RSK2-CTD by **FMK** (+322 Da, with loss of HF). RSK2-CTD (5 µM) was incubated with **FMK** (100 µM) for 10 min, followed by LC-MS analysis of the intact protein. **c,** RSK2-CTD (5 µM) was pretreated with inhibitors **14** or **15** (10 µM) for 1 h, followed by addition of excess FMK (100 µM final). At each time point, an aliquot was removed and quenched by diluting 1:1 into 0.4% formic acid. Relative amounts of RSK2-CTD and the FMK adduct were quantified by LC-MS. Progress curve fits to a single exponential were used to estimate dissociation half-times of **14** and **15**. Inset: progress curve detail showing reaction of RSK2-CTD with 100 µM **FMK**.

We selected the *N*-isopropyl cyanoacrylamide **15** (hereafter referred to as '**CN-NHiPr**') for further biochemical and cellular experiments due to its greater potency and slower off-rate compared to cyanoacrylate **14**. Kinase profiling[23] revealed that **CN-NHiPr** is highly selective for the C-terminal kinase domains of RSK1 and RSK4 (RSK2-CTD was not on the panel of profiled kinases). Besides RSK1/4-CTD, only 6 of the 442 kinases profiled showed greater than 70% inhibition by 1 µM **CN-NHiPr** relative to the DMSO control (Figure 2.8a). Subsequent K_D determinations demonstrated that **CN-NHiPr** bound RSK1-CTD (K_D = 540 pM) with ~80-fold higher affinity than MAP3K1 and at least 400-fold higher affinity than the remaining five kinases (STK16, RIPK2, RET, MEK5, and PDGFRB) (Figure 2.8b).

Kinase	CN-NHiPr K_D (nM)
RSK1	0.54
RSK4	1.2
MAP3K1	45
STK16	230
RIPK2	310
RET	360
MEK5	2300
PDGFRB	2800
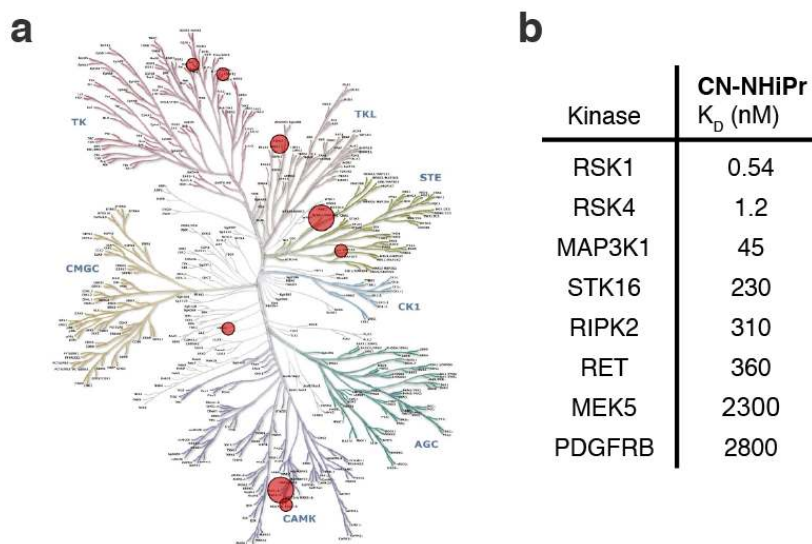
Figure 2.8. Kinase profiling of CN-NHiPr using the Ambit competitive binding platform (now administered by DiscoverRx). a, In the initial screen, CN-NHiPr (1 µM) was profiled against the entire 442-kinase panel. Only 8 kinases (including RSK1/4 CTD) showed greater than 70% inhibition by **CN-NHiPr** in this single-point determination. **b,** K_D values, depicted above, were determined from dose-response curves for these 8 kinases. RSK3-CTD, which has a Met gatekeeper but is otherwise 85% identical to RSK1/2-CTD, was not significantly inhibited by 1 µM CN-NHiPr in the initial screen. RSK2-CTD was not on the panel.

Sustained, quasi-irreversible occupancy of RSK in cells

We compared the cellular activity of **CN-NHiPr** with **FMK**, previously shown to be an irreversible inhibitor of endogenous RSK1 and RSK2[6,22]. Treatment of MDA-MB-231 breast cancer cells with **CN-NHiPr** for 2 h led to potent inhibition of RSK2-CTD-mediated autophosphorylation of Ser386 (Figure 2.9a, EC_{50} < 10 nM), a key regulatory site for the N-terminal effector kinase domain[24]. With a similar dose response, **CN-NHiPr** prevented an irreversible fluorescent probe, **FMK-BODIPY**[22], from labeling RSK1/2 in cells (Figure 2.9b). Despite its irreversible binding mode, **FMK** was less potent than **CN-NHiPr** in both the cellular occupancy and autophosphorylation assays, suggesting that **CN-NHiPr** enters cells and binds to endogenous RSK1/2 with faster kinetics. Washout experiments revealed a striking duration of RSK1/2 occupancy by **CN-NHiPr**, indistinguishable from **FMK**. After washout of **FMK** or **CN-NHiPr**, reappearance of inhibitor-free RSK1/2 required at least 48 h, presumably the time required for RSK1/2 resynthesis (Figure 2.9c). Similar to **FMK**, **CN-NHiPr** abolished RAF-induced epithelial cell migration and invasion (Figure 2.9d–f), consistent with the recently established role of RSK signaling in these processes[25-27].
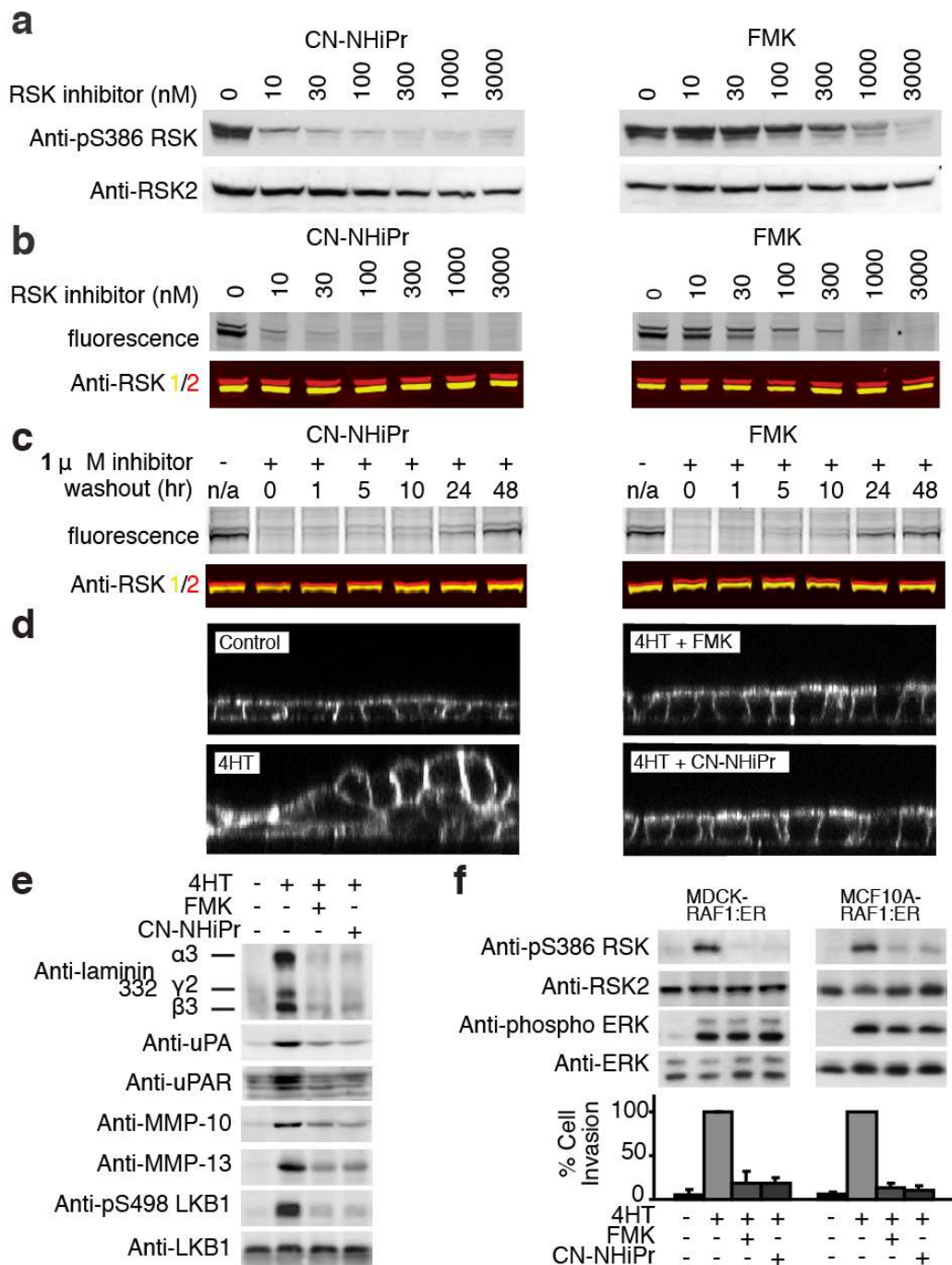
Figure 2.9. Sustained inhibition of cellular RSK1/2 by CN-NHiPr. a, MDA-MB-231 cells were treated with **CN-NHiPr** or **FMK** (2 h). Cell lysates were immunoblotted for pS386-RSK or RSK2. **b,** MDA-MB-231 cells were treated with **CN-NHiPr** or **FMK** (2 h), followed by **FMK-BODIPY** (3 µM, 1 h). Cell lysates were resolved by SDS-PAGE and detected by in-gel fluorescence scanning or immunoblotting for RSK1 (yellow) and RSK2 (red). **c,** MDA-MB-231 cells were

treated with **CN-NHiPr** or **FMK** (1 µM, 1 h). The media was exchanged and cells were harvested at the indicated times after inhibitor washout. Cell lysates were treated with **FMK-BODIPY** (5 µM, 1 h). Proteins were resolved on SDS-PAGE and detected by in-gel fluorescence or immunoblotting for RSK1 (yellow) and RSK2 (red). **d,** Polarized MDCK-RAF:ER cell monolayers were treated with **CN-NHiPr** or **FMK** (2 µM) and stimulated with 4-hydroxytamoxifen (4HT, 1 mM) to activate RAF-dependent epithelial cell multilayering. After 24 h, cells were imaged by confocal fluorescence microscopy (XZ plane). **e,** Polarized MDCK-RAF:ER cell monolayers were treated as in **d**. After 24 h, cell lysates and conditioned media were analyzed by immunoblotting for the indicated proteins. **f,** MDCK-RAF:ER and MCF10A-RAF:ER cells were treated with **CN-NHiPr** or **FMK** (2 µM) and 4HT (1 µM) as indicated for 24 h. Cells were then either lysed for immunoblot analysis or trypsinized and subjected to Matrigel invasion assays (see Methods). After 24 h, cell invasion was quantified and expressed as a percentage of the maximum value (mean ± SD from three experiments). The experiments shown in panels **d** – **f** were performed by Katarzyna Duda under the guidance of Morten Frodin.

Cocrystal structure of the RSK2/cyanoacrylate complex

A cocrystal structure provided insight into the molecular basis of sustained RSK2 inhibition by doubly activated Michael acceptors (Figure 2.10). We screened several related cyanoacrylate and cyanoacrylamide inhibitors and obtained the best diffraction data (2.4 Å resolution) with a *tert*-butyl cyanoacrylate derivative **16** (For structure see Figure 2.5a) that has similar potency and dissociation kinetics ($t_{1/2}$ = 163 min) as **CN-NHiPr**. The structure confirms the covalent nature of the complex (PDB code 4D9U), with strong electron density connecting Cys436 to the electrophilic beta-carbon of the cyanoacrylate (Figure 2.10a).
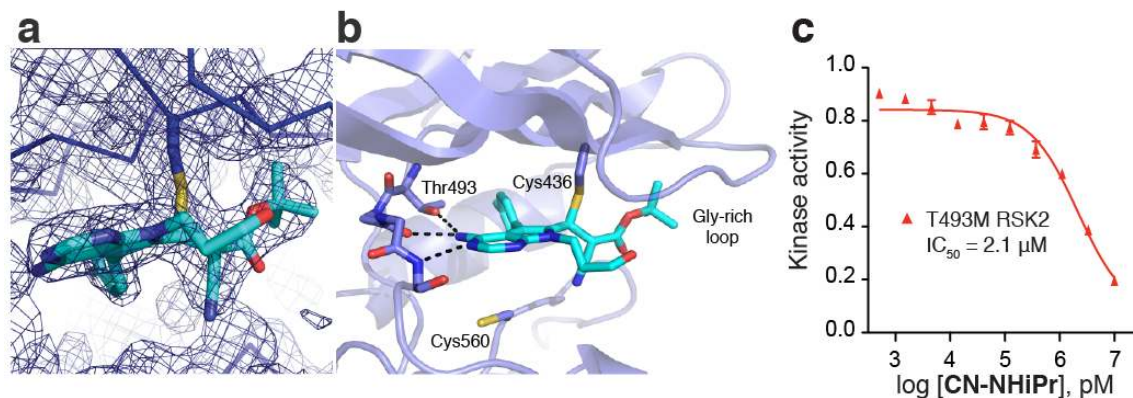
Figure 2.10. **Specific noncovalent interactions drive covalent bond formation.** **a,** $2F_0$-F_c electron density map (contoured at 1.0 s) showing the covalent bond between Cys436 and the electrophilic beta-carbon of compound **16**. **b,** Cocrystal structure of RSK2-CTD bound to *tert*-butyl cyanoacrylate **16**. Ribbon representation of RSK2-CTD with cyanoacrylate **16** in green highlights hydrogen bonds to the hinge and the gatekeeper, Thr493. Note the close proximity of Cys560, which does not form a covalent bond with the electrophilic beta-carbon. This crystal structure was solved and refined by Dr. Miles Pufall. **c,** Mutation of the RSK2-CTD gatekeeper (T493M) confers ~1000-fold resistance to **CN-NHiPr** as evident from kinase assays. Kinase assays with T493M RSK2-CTD were performed as described for WT RSK2-CTD.

A second cysteine (Cys560) within ~7 Å of the electrophilic carbon is incapable of forming a stable covalent bond, as indicated by the ~1000-fold decrease in potency of **CN-NHiPr** toward C436V RSK2 (Figure 2.5d). Hydrogen bonds between the pyrrolopyrimidine scaffold, the RSK2 hinge region, and the side chain of Thr493 position the electrophilic beta-carbon beneath Cys436 and are likely essential for driving covalent bond formation (Figure 2.10b). An additional anchor is provided by the *p*-tolyl group, which packs against the gatekeeper side chain (Thr493) and extends into a hydrophobic pocket. Mutation of Thr493 to Met in RSK2-CTD conferred ~1000-fold resistance to **CN-NHiPr** (Figure 2.10c).

RSK3-CTD and the related kinase domains of MSK1/2 have a Met gatekeeper, explaining their insensitivity to **CN-NHiPr**.

Protein unfolding promotes covalent bond dissociation

If the combined nonbonded interactions between RSK2 and the pyrrolopyrimidine are essential for stabilizing the covalent bond, disruption of the kinase domain's tertiary fold should promote rapid elimination of Cys436 and regeneration of the original Michael acceptor. We tested this hypothesis by monitoring covalent bond formation in solution with UV-visible spectroscopy. Treatment of **CN-NHiPr** (20 µM in pH 7.5 buffer) with a slight excess of RSK2-CTD (25 µM) caused disappearance of the strong 400 nm absorption peak (Figure 2.11a and 2.11b), consistent with nucleophilic attack on the cyanoacrylamide. Mutation of Cys436 abolished this effect (Figure 2.11b), demonstrating that none of the remaining cysteines (including Cys560 in the ATP binding site) detectably react with **CN-NHiPr**, even at millimolar concentrations. When the covalent RSK2-CTD/**CN-NHiPr** complex was unfolded by the addition of sodium dodecylsulfate or guanidine, the cyanoacrylamide absorption peak reappeared within seconds (Figure 2.11a and 2.11b), and LC-MS analysis indicated quantitative recovery of CN-NHiPr (Figure 2.11). Proteolytic digestion of the complex with trypsin or proteinase K also promoted complete reversal of the covalent bond (Figure 2.11c), suggesting that cyanoacrylamide-modified peptides derived from cellular RSK turnover would be short-lived.
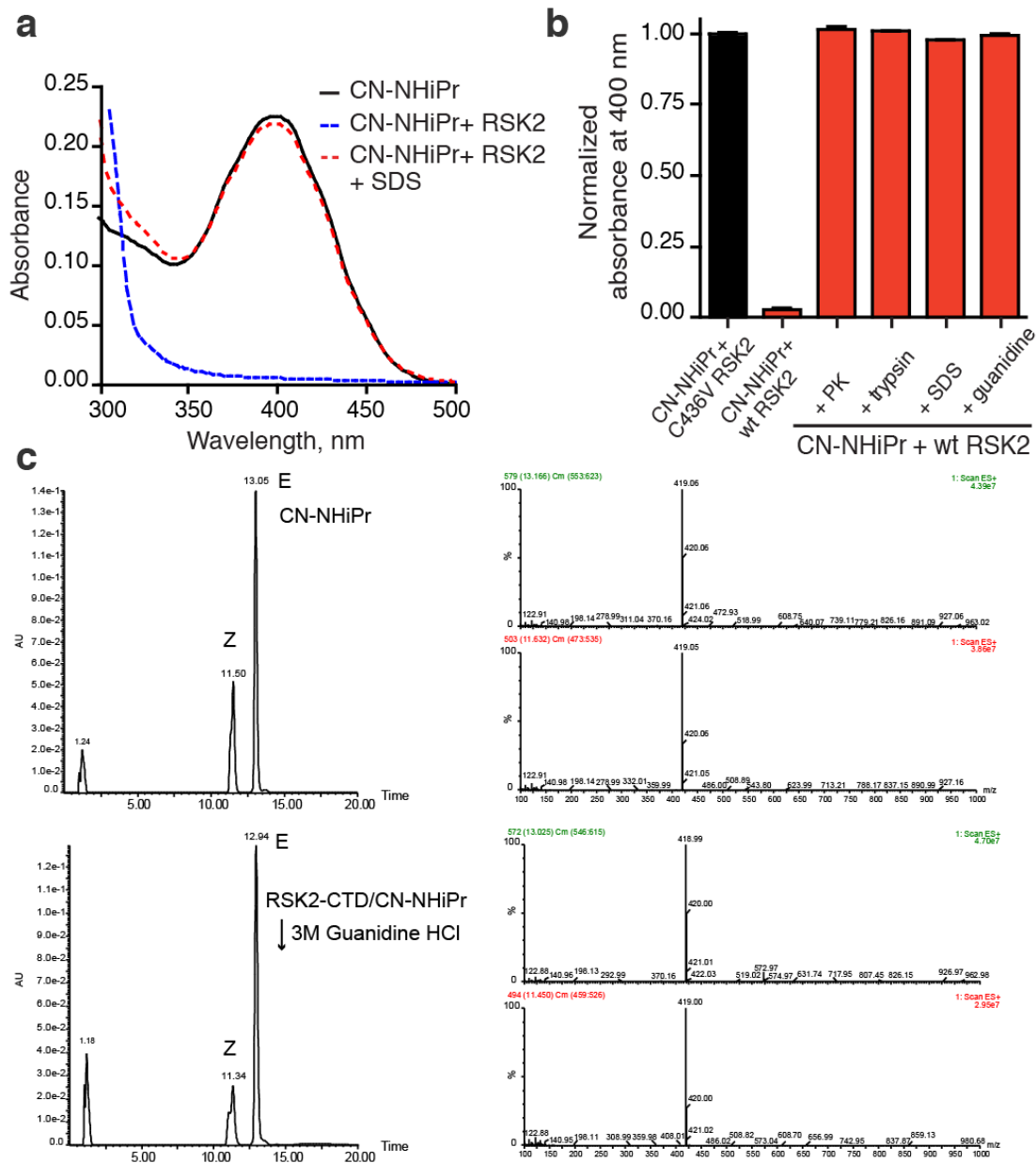
Figure 2.11. Protein unfolding results in dissociation of CN-NHiPr. a, , UV-visible spectra of **CN-NHiPr** showing the 400 nm cyanoacrylamide absorption peak. **CN-NHiPr** (20 µM, black curve) was treated with wild-type RSK2-CTD (25 µM), resulting in a loss of the cyanoacrylamide peak (blue curve). Addition of SDS (2% final) caused instantaneous recovery of the cyanoacrylamide (red curve). **b,** Effect of protein unfolding (2% SDS; 3 M guanidine HCl) and proteolysis (proteinase K, PK; trypsin) on the covalent complex derived from CN-NHiPr (20 µM) and RSK2-CTD (25 µM). Recovery of free **CN-NHiPr** was determined by quantifying the absorbance at 400 nm (triplicate measurements, mean ± SD). C436V RSK2-CTD (25 µM) had no effect on the **CN-NHiPr** absorption peak. **c**, LC-MS chromatogram and mass spectra showing quantitative recovery of CN-NHiPr after unfolding the RSK2-CTD/**CN-NHiPr** complex with 3 M guanidine-HCl.

We have shown that cyanoacrylamides react reversibly with model thiols and recombinant RSK2-CTD. However, formation of irreversible adducts with endogenous full-length RSK or other cellular proteins remained possible, especially given the apparent irreversible binding to RSK1/2 (resistant to washout) in cells (Figure 2.9c). We addressed this possibility by treating cells with fluorescent BODIPY conjugates derived from **FMK**, **CN-NHiPr**, and the pyrrolopyrimidine scaffold as a non-electrophilic control (Figure 2.12a). Analysis of cell lysates by SDS-PAGE revealed several prominent bands labeled by the irreversible probe, **FMK-BODIPY**, including endogenous RSK1 and RSK2[22] (Figure 2.12b). By contrast, BODIPY conjugates of **CN-NHiPr** and the non-electrophilic scaffold failed to label RSK1/2. Instead, both BODIPY conjugates bound weakly to an identical subset of denatured proteins, presumably in a noncovalent manner. We conclude that specific interactions provided by the folded kinase domain, in concert with a precisely positioned cysteine, are required to cooperatively stabilize a covalent complex with **CN-NHiPr**. The lack of irreversible adduct formation by **CN-NHiPr** in a cellular context is remarkable given that it is more than 50,000 times more reactive toward thiols than **FMK** (reaction with 10 mM glutathione: CN-NHiPr $t_{1/2}$ < 1 sec; FMK $t_{1/2}$ = 16 h).
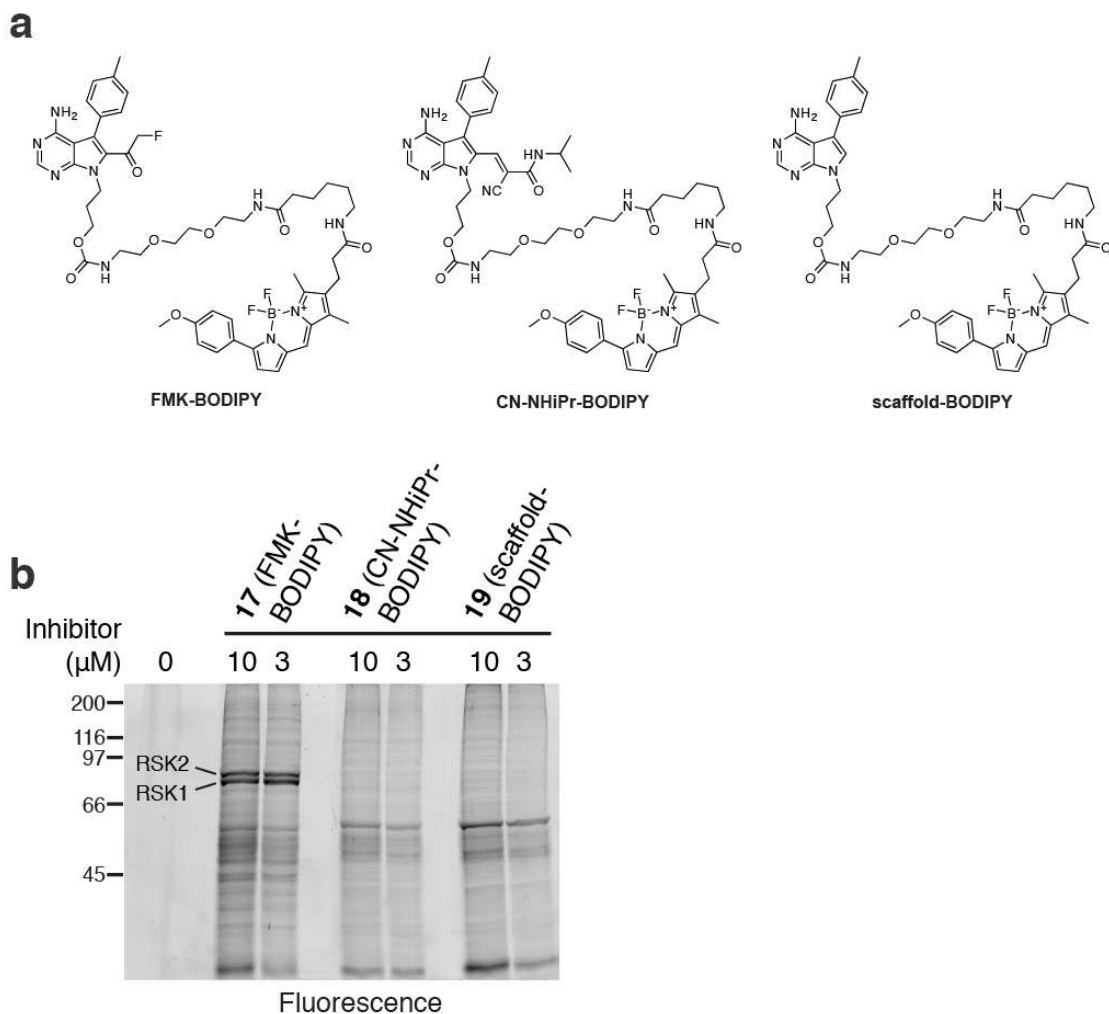
Figure 2.12. CN-NHiPr does not covalently modify cellular proteins. a, Chemical structures of the BODIPY conjugates **17**-**19**. These analogs were synthesized by Dr. Shyam Krishnan. **b,** Modification of cellular proteins by BODIPY conjugates of **FMK** (**17**), **CN-NHiPr** (**18**), and the pyrrolopyrimidine scaffold (**19**). MDA-MB-231 cells were treated for 1 h with the indicated BODIPY conjugates. Cell lysates were resolved by SDS-PAGE and BODIPY adducts detected by in-gel fluorescence scanning.

2.4 Discussion

In this study, we report a novel approach for targeting noncatalytic cysteines with reversible covalent inhibitors. The predominant strategy for targeting noncatalytic cysteines, exemplified by several acrylamide-based kinase inhibitors in clinical trials, has been to minimize the intrinsic chemical reactivity of the electrophile as much as possible[4]. This strategy derives from the chemically intuitive hypothesis that attenuated electrophiles have a reduced likelihood of reacting irreversibly with off-target proteins. Our results challenge this notion. Although attenuated Michael acceptors such as acrylamides may react slowly with most cysteine residues, the probability of modifying a hyper-reactive cysteine (comprising up to ~10% of solvent-exposed cysteines[14]) is unpredictable; such off-target reactions are clearly undesirable.

Our chemical and biochemical experiments demonstrate that addition of a nitrile group (with a molecular mass of only 26 Da) converts irreversible Michael acceptors (e.g., acrylate or acrylamide) into electrophiles with unanticipated properties. The resulting electrophiles (e.g., cyanoacrylate or cyanoacrylamide) react with cysteine thiols under physiological conditions in a manner that is energetically favorable yet rapidly reversible, thus minimizing the chance of producing irreversibly modified peptides. Adding the nitrile group to the alpha-carbon of an acrylamide increases the susceptibility of the beta-carbon to nucleophilic attack but also stabilizes the resultant carbanion. This leads to a kinetic regime in which thiol addition and elimination proceed on a subsecond/second time scale, a surprising result that, to our knowledge, has not

been demonstrated previously. Rapid thiol-addition/elimination chemistry appears to be a general property of cyanoacrylamides, and in preliminary work, extends to some, but not all, combinations of two electron-withdrawing groups in doubly activated Michael acceptors. We envision that the reversible covalent targeting strategy described herein can be applied to any kinase that has an exposed cysteine within striking distance of the active site, provided that the geometric requirements for covalent bond formation are satisfied with minimal strain. More generally, by using structural bioinformatics to identify solvent-exposed cysteines near potential small-molecule binding sites, it may be possible to apply this strategy to additional target classes outside of the protein kinase family.

2.5 Experimental

Determination of equilibrium dissociation constants (K_D) for thiol/Michael acceptor adducts

Reactions of cyanoacrylates and cyanoacrylamides with BME, GSH, ethanolamine, and lysine were monitored with a Spectramax M5 plate reader (Molecular Devices, Sunnyvale CA). Reactions were initiated by mixing equal volumes of the Michael acceptor (200–400 μM in PBS, pH 7.4) with the thiol (0–200 mM BME or GSH in PBS pH 7.4, two-fold dilution series). Final solutions (Costar flat-bottom clear 96-well plate, 100 μL per well), containing 100–200 μM Michael acceptor and increasing concentrations of BME/GSH, were incubated for 10 min at room temp prior to acquiring absorption spectra (250–500 nm). Equilibration of these doubly activated Michael acceptors with BME or GSH occurred within seconds or less (as determined by sequential absorbance measurements, data not shown). Formation of the thiol adduct was quantified by monitoring the disappearance of the absorbance peak (λ_{max}) relative to the no-thiol control sample. Data were fit using PRISM 4.0 to obtain equilibrium dissociation constants. Test reactions of **CN-NHiPr** with ethanolamine or lysine were also carried out. Reactions were initiated as described above by mixing equal volumes of **CN-NHiPr** (100 μM in PBS, pH 7.4) with solutions of ethanolamine or lysine (0.1–100 mM amine solution in PBS pH 7.4). Negligible reaction was observed after 5 hours of incubation

.

Reversible BME addition to cyanoacrylate 3 monitored by ^{1}H NMR
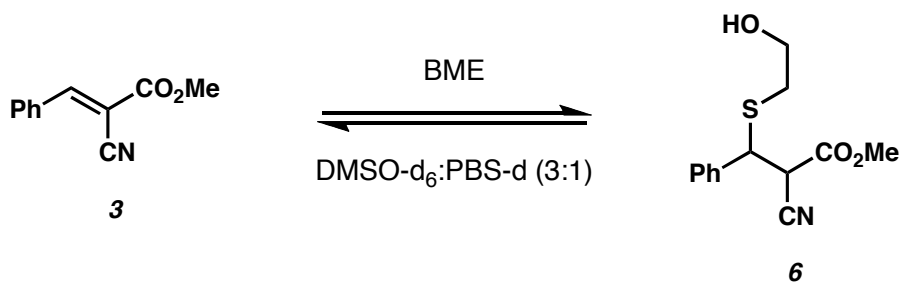
To a solution of cyanoacrylate **3** (17.1 mg, 91.4 μmol) in 0.75 mL of DMSO-d6 was added a solution of 400 mM BME in deuterated PBS (0.25 mL). Analysis of the reaction mixture by ^{1}H NMR after 30 minutes indicated an 85:15 ratio of the thiol adduct **6** to the starting cyanoacrylate **3**. The reaction mixture was then diluted 10-fold by addition of 100 μL to 900 μL of 3:1 v/v DMSO-d6:deuterated PBS. Analysis of the solution after 30 minutes by ^{1}H NMR indicated a 55:45 ratio of the thiol adduct **6** to the starting cyanoacrylate **3**.

In a second experiment, a solution of cyanoacrylate **3** (5.1 mg, 27.3 μmol) in 0.75 mL of DMSO-d6 was added to 0.25 mL of 400 mM BME in deuterated PBS. Analysis of the reaction mixture by ^{1}H NMR indicated a >95:5 ratio of the thiol adduct **6** to the starting cyanoacrylate **3**. The sample was then allowed to stand at ambient temperature. The ratio of adduct **6** to starting cyanoacrylate **3** changed with time as the excess BME in solution underwent air oxidation to the corresponding disulfide. After 17 days, the ratio of **6:3** was <5:95 (data not shown). This experiment was performed by Dr. Shyam Krishnan.

Characterization of the BME adduct 6.

For the purpose of characterization, a solution of cyanoacrylate 3 (13.5 mg) in DMSO-*d6* (0.75 mL) was treated with 1.2 M BME in deuterated PBS (0.25 mL) to drive the reaction towards the adduct 6 (final concentrations of cyanoacrylate 3 and BME were 72 mM and 300 mM respectively). ^{1}H and ^{13}C spectra were acquired under these conditions and are reported for the mixture of diastereomeric adducts in solution containing excess BME.

^{1}H NMR (400 MHz, 3:1 v/v DMSO-*d6*: deuterated PBS): 7.40 (m, 1H, minor diastereomer), 7.39 (m, 1H, major diastereomer), 7.33-7.23 (m, 4H), 4.51 (s, 1H), 3.63 (s, 3H, minor diastereomer), 3.54 (s, 3H, major diastereomer), 3.47 (m, 2H, minor diastereomer), 3.38 (m, 2H, major diastereomer), 2.64–2.40 (m, 2H); ^{13}C NMR (100 MHz, 3:1 v/v DMSO-*d6*: deuterated PBS): 166.2, 139.0, 138.1, 129.9, 129.8, 129.6, 128.92, 128.87, 116.77, 116.74, 61.4, 61.2, 54.8, 54.6, 49.2, 48.3, 34.9, 34.8. This was performed by Dr. Shyam Krishnan.

LC-MS detection of RSK2-CTD adducts with electrophilic pyrrolopyrimidines.

Purified RSK2-CTD (5 μM) was incubated with the indicated compounds (25 μM, 5 equiv) for 1 h at room temp in buffer (20 mM Hepes pH 8.0, 100 mM NaCl, 10 mM MgCl$_2$). The reaction was quenched by adding an equal volume of 0.4% formic acid, and the samples were analyzed by liquid chromatography (Microtrap C18 Protein column [Michrom Bioresources], 5% MeCN, 0.2% formic acid, 0.25 mL/min; eluted with 95% MeCN, 0.2% formic acid) and in-line ESI

mass spectrometry (LCT Premier, Waters). Mass/charge envelope data were deconvolved with MassLynx software to provide molecular masses. Experimentally determined molecular masses for RSK2-CTD and derived adducts were in accord with predicted values.

In vitro kinase assays

WT or mutant RSK2-CTD (10 μM) was first activated by His$_6$-ERK2 (10 μM in 20 mM Hepes pH 8.0, 10 mM MgCl$_2$, 2.5 mM TCEP, 0.2 mg/mL BSA, 200 μM ATP) for 30 min at room temp. Active RSK2-CTD (5 nM in 20 mM Hepes pH 8.0, 10 mM MgCl$_2$, 2.5 mM TCEP, 0.25 mg/mL BSA, 100 μM ATP) was then treated with inhibitors (ten concentrations, in duplicate) for 30 min. Kinase reactions were initiated by the addition of 5 μCi of [γ-^{32}P]ATP (6000 Ci/mmol, NEN) and 167 μM peptide substrate (RRQLFRGFSFVAK, CTD-tide) and incubated at room temp for 30 min. Kinase activity was determined by spotting 5 μL of each reaction onto P81 cation exchange paper (Whatman). Each blot was washed once with 1% AcOH solution, twice with 0.1% H$_3$PO$_4$ solution, and once with MeOH (5-10 minutes per wash). Dried blots were exposed for 30 min to a storage phosphor screen and scanned by a Typhoon imager (GE Life Sciences). The data were quantified using ImageQuant 5.2 software and fit using PRISM 4.0.

Dialysis of RSK2-CTD after treatment with electrophilic pyrrolopyrimidines

Activated RSK2-CTD (50 nM in 20 mM Hepes pH 8.0, 10 mM MgCl$_2$, 2.5

mM TCEP, 0.25 mg/mL BSA, 100 μM ATP) was treated with 1 μM inhibitor for 60 min at room temp. The reactions were transferred to a dialysis cassette (0.1–0.5 mL Slide-A-Lyzer, 10 kDa MWCO, Pierce) and dialyzed vs. 2 L buffer (20 mM Hepes pH 8.0, 10 mM MgCl$_2$, 1 mM DTT) at 4°C. Buffer was exchanged after 2 h of dialysis, and then daily until the end of the experiment. 50 μL aliquots were removed each day for kinase assays. Kinase reactions (in triplicate for each time point) were initiated by adding 10 μL of 5 μCi of [γ-^{32}P]ATP (6000 Ci/mmol, NEN) and 167 μM peptide substrate (RRQLFRGFSFVAK, CTD-tide) to 10 μL of each time point aliquot and were incubated at room temp for 30 min. Kinase activity was determined as described above.

Estimation of inhibitor dissociation rates using FMK as a kinetic trap

RSK2-CTD (5 μM in 20 mM Hepes pH 8.0, 10 mM MgCl$_2$, 100 mM NaCl, 2.5 mM TCEP, 0.2 mg/mL BSA) was preincubated with 10 μM of **14** or **15** (or DMSO) for 60 min at room temp. **FMK** (100 μM) was then added, and 50 μL aliquots were removed at various times after **FMK** addition. Each time point reaction was stopped by adding an equal volume of 0.4% formic acid, and the samples were analyzed by by liquid chromatography (Microtrap C18 Protein column [Michrom Bioresources], 5% MeCN, 0.2% formic acid, 0.25 mL/min; eluted with 95% MeCN, 0.2% formic acid) and in-line ESI mass spectrometry (LCT Premier, Waters). After charge envelope deconvolution, relative amounts of unmodified RSK2-CTD and the **FMK** adduct were determined by quantifying the mass peak intensities using MassLynx software (peak areas provided nearly

identical results). These data (percentage of **FMK** adduct vs. time) were fit to a single exponential function (assuming pseudo-first order kinetics) using PRISM 4.0 to estimate the dissociation half-times of **14** and **15**. This assumes that dissociation of **14** and **15** from RSK2-CTD is rate-limiting and that the rebinding rate of free **14** and **15** is negligible given the large excess of FMK competitor.

pS386 RSK inhibition in MDA-MB-231 cells

MDA-MB-231 cells were seeded into 6-well plates at a density of 300,000 cells/well in DMEM supplemented with 10% FBS. After 48 h, the media was exchanged with serum-free DMEM, and the cells were treated for 2 h with the indicated concentrations of inhibitors. Following inhibitor treatment, the cells were stimulated for 20 min with PMA (100 ng/mL), then washed with 2 mL of cold PBS and frozen onto the plate at –80˚C. The cells were thawed in the presence of 70 μL CelLytic M lysis buffer (Sigma) supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors, and mechanically scraped from the wells. The lysates were cleared by centrifugation at 14K rpm for 10 min at 4˚C, and normalized by Bradford assay. Laemmli sample buffer was added to the lysates and the proteins were separated by 7.5% SDS-PAGE and analyzed by immunoblot using phospho-Ser386 RSK (1:500 dilution, rabbit Ab, Cell Signaling #9341) and RSK2 (E-1) (1:500 dilution, mouse monoclonal, Santa Cruz cs-9986) antibodies. Immunoblots were developed and imaged as described [22].

Cellular RSK1/2 occupancy assay using FMK-BODIPY

MDA-MB-231 cells were seeded into 6-well plates at 300,000 cells/well. After 48 h, the media was exchanged with serum-free DMEM, and the cells were treated for 2 h with the indicated concentrations of inhibitors, followed by 1 h incubation with 3 μM of the **FMK-BODIPY** probe. The media was aspirated and the cells were washed with 2 mL of cold PBS and lysed with 70 μL CelLytic M lysis buffer (Sigma) supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors.   The lysates were cleared by centrifugation at 14K rpm for 10 min at 4°C and normalized by Bradford assay. Laemmli sample buffer was added to the lysates and the proteins were separated by 7.5% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by immunoblot analysis using RSK1 (C-21, Santa Cruz sc-231) and RSK2 (E-1, Santa Cruz sc-9986) antibodies. After incubation with primary antibodies, immunoblots were incubated with infrared dye-labeled secondary antibodies (IR680 or IR800) and visualized using the LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE).

Cellular RSK1/2 occupancy after inhibitor washout

MDA-MB-231 cells were seeded into 6-well plates at 300,000 cells/well. After 24 h, the cells were treated with 1 μM of the indicated inhibitors for 1 h in DMEM supplemented with 10% serum. The media was then exchanged with inhibitor-free media (with 10% serum).   At the indicated time points (post-

washout), the cells were trypsinized, pelleted and frozen in liquid nitrogen. The cell pellets were lysed with 70 μL PBS supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors. The lysates were cleared by centrifugation at 14K rpm for 10 min at 4°C, normalized by Bradford assay and treated with 5 μM **fmk-BODIPY** for 1 h. Laemmli sample buffer was added to the lysates and the proteins were separated by 7.5% SDS-PAGE and detected by in-gel fluorescence scanning (Typhoon 9400), followed by immunoblot analysis using RSK1 (C-21, Santa Cruz sc-231) and RSK2 (E-1, Santa Cruz sc-9986) antibodies. After incubation with primary antibodies, immunoblots were incubated with infrared dye-labeled secondary antibodies (IR680 or IR800) and visualized using the LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE).

Polarized epithelial cell multilayering assay

MDCK-RAF:ER cells [25] were plated onto 0.4 μm pore size Transwell polyester filters (Corning Inc., New York; #3460), and experiments were initiated 3 days after the cells reached confluency. Cells were treated for 24 h with the indicated inhibitors (2 μM) in the presence or absence of 4-hydroxytamoxifen (1 μM), after which time the cells were fixed for 10 min with Lillies fixative (4% phosphate-buffered formaldehyde [pH 7.0]) and thereafter permeabilized with 0.5% Triton X-100. Actin filaments were stained with rhodamine-phalloidin (Invitrogen) and imaged with a Zeiss LSM 510 confocal microscope. This experiment was performed by Katarzyna Duda.

Matrigel invasion assay

MDCK-RAF1:ER and MCF10A-RAF1:ER cells [25] were seeded at 8 × 10^4 cells/cm^2 in 9.6 cm^2 wells. The next day, medium was changed to complete medium with or without indicated treatment (2 μM RSK inhibitors, 1 μM 4HT). After 24 h, cells were trypsinized, and 100,000 cells were seeded in serum-free medium in the top chamber of 8 μm pore-size Transwell polyester filters (Corning, #3422) coated with growth factor-reduced Matrigel (BD Biosciences, #356231). The cells were allowed to invade against a chemotactic agent (MDCK-RAF1:ER cells, 8 nM HGF; MCF10A-RAF1:ER , 10% FBS) for 24 h with or without the indicated treatment (2 μM RSK inhibitors, 1 μM 4HT). Invading cells (attached to the filter facing the bottom chamber) were fixed in methanol, stained with 0.1% crystal violet, and quantified as described previously [28]. This experiment was performed by Katarzyna Duda.

RSK2-CTD expression and purification

RSK2-CTD (mouse RSK2 399-740) was expressed in E. coli strain BL21 (DE3)-RIL (pET-46 Ek/LIC His6-fusion vector was kindly provided by M. Malakhova and Zigang Dong, University of Minnesota). Cells were lysed in lysis buffer (50 mM Tris pH 8.0, 0.5 M NaCl, 10% glycerol, 15 mM imidazole) using a homogenizer (EmulsiFlex-C5, Avestin) operated at 15000 psi at 4°C for 15 min (continuous flow lysis). Soluble His-tagged RSK2-CTD was purified by Ni/NTA affinity chromatography (50 mM Tris pH 8.0, 0.5 M NaCl, 10 mM imidazole) using a 10 mL column at 2 mL/min flow rate with gradient elution (10–500 mM imidazole),

followed by cleavage of the His$_6$-tag with enterokinase (1U per 1 mg protein) (EMD Biosciences) at 4°C. The cleaved RSK2-CTD was further purified by HiLoad 16/60 Superdex-75 size exclusion chromatography (20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT) and the protein eluted as a single peak at 55 mL, corresponding to a monomer. Purified protein was concentrated and flash frozen in liquid nitrogen in 20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT, 12% glycerol and stored at -80°C. Protein concentrations were determined using the calculated extinction coefficients at 280 nm measured with a NanoDrop1000 spectrophotometer (Thermo Scientific).

RSK2-CTD crystallization and data collection

Initial RSK2-CTD crystallization conditions were determined through conventional screens. Purified RSK2-CTD was concentrated to 5–10 mg/mL in 20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT, and then incubated with one molar equivalent of *tert*-butyl cyanoacrylate **16**. Drops of 1 μl RSK2-CTD/**16** and 1 μL precipitant solution were deposited in hanging drop 24-well plates using the commercial Classics II Suite (Qiagen) crystallization screen. Several conditions produced crystals, the best quality of which were under condition 45 of the Classics II Suite (Qiagen) screen, composed of 0.1 M Tris pH 8.5, 25% PEG 3350. Crystals were cryoprotected in mother liquor with 30% ethylene glycol and flash frozen by plunging in liquid nitrogen. Crystallographic data acquisition was performed at the Advanced Light Source (ALS, Berkeley, CA) Beamline 8.3.1. Two data sets were collected for cyanoacrylate **16**, one with the detector at 300

mm and the other at 450 mm with an oscillation of 0.5 degrees per frame in order to maximize the resolution and minimize spot overlap. The data sets were indexed, integrated, and scaled using the XDS suite of programs [29]. All of the data sets fit the 4_12_12 space group, with unit cell dimensions of approximately 46 x 46 x 295 Å. The length of this unit cell caused two problems: significant overlapping or spots, and highly anisotropic data sets. Overlap was minimized as described above and by reorienting the crystal in the beamline. We were unable to improve on the anisotropy of the crystal, which subsequently resulted in higher than expected B-factors.

Structure determination

Prior to solving the liganded RSK2-CTD structure, we first re-solved the structure of apo-RSK2-CTD. We expressed the protein as described above, recrystallized it, and collected data sets on ALS Beamline 8.3.1, with the exception that we flash froze the crystals in Paratone oil (Hampton Research). These crystals exhibited significantly higher resolution (~1.9 Å), with less mosaicity. After processing the data sets using XDS, we solved the structure by molecular replacement using 2qr8 as a search model [30]. Interestingly, our data sets enabled us to build residues N-terminal to the reported structure out to residue 406 (data not shown). We then used this structure as a model for molecular replacement in solving the liganded crystal structures.

The liganded complexes were modeled by several rounds of manual rebuilding and restrained refinement with programs Coot [31] and Phenix.refine [32].

Ligands were then built in ChemDraw (CambridgeSoft), and exported in SMILES string format and read into PhenixELBOW to generate both a coordinate and restraint file that were used during model building and refinement. Before adding the ligand to the model, it was clear from the electron density map where the ligand would be placed in the pocket, and that it was covalently linked to Cys436. After ligand placement in this density, the ligand was redrawn with a tetrahedral carbon at the C11 attachment position. The previous ligand was then replaced with the newer one, and the structure refined with Phenix. The placement of the ligand was then confirmed by composite omit maps. The models were validated using Molprobity [33] from within Phenix. The crystal data was collected and analyzed by Dr. Miles Pufall.

Table 2.1: Crystallographic data collection and refinement statistics

	4d9u (Rsk2-CTD_**16**)
Data collection	
Space group	$P4_12_12$
Cell dimensions	
a, b, c (Å)	46.7, 46.7, 293.0
a, b, g (°)	90, 90, 90
Resolution (Å)	50-2.4 (2.5-2.4)*
R_{sym} or R_{merge}	0.110(0.753)
I/sI	14.7(2.06)
Completeness (%)	99.7(99.2)
Redundancy	8.0(3.7)
Refinement	
Resolution (Å)	50-2.4
No. reflections	13,695 (unique)
R_{work}/ R_{free}	20.4/27.6
No. atoms	2464
Protein	2403
Ligand/ion	32/1
Water	31
B-factors	
Protein	67.0
Ligand/ion	74.7
Water	52.1
R.m.s deviations	
Bond lengths (Å)	0.008
Bond angles (º)	1.10

*Highest resolution shell is shown in parenthesis.

Formation of covalent RSK2-CTD/CN-NHiPr complex; covalent bond dissociation upon RSK2-CTD unfolding or proteolysis

RSK2-CTD (25 μM) in PBS was added to 20 μM **CN-NHiPr** (or DMSO control) in a final reaction volume of 0.7 mL. UV-visible absorption spectra (250–500 nm) were obtained for CN-NHiPr before (control) and after addition of RSK2-CTD using Spectramax M5 Spectrophotometer (1 cm path). SDS (2% w/v final) or guanidine-HCl (3 M final) was added, followed by acquisition of absorption spectra within ~1 min. Alternatively, trypsin (32 μg total) or Proteinase K (32 μg

total) were added to the covalent RSK2-CTD/CN-NHiPr complex for 1 h at 37˚C, followed by acquisition of absorption spectra.

LC-MS detection of recovered CN-NHiPr after guanidine-induced unfolding of the RSK2-CTD/CN-NHiPr covalent complex

RSK2-CTD (300 μM, PBS) was added to **CN-NHiPr** (250 μM, PBS) in a total volume of 50 µL. The covalent complex was then unfolded by treatment with 50 µL of 6 M guanidine-HCl for 1 min, after which time acetonitrile was added to a final concentration of 50%. The mixture was centrifuged, filtered through a 0.2 µm filter, and finally filtered through a Microcon device (10 kDa MWCO). The filtrate was then analyzed by LC-MS (20 µL injection, Waters XTerra MS C18 column, 20 min gradient, 5–70% MeCN/0.1% formic acid; Waters 2695 Alliance Separations Module; Waters Micromass ZQ mass spectrometer), as shown in Figure 2.11c. A control CN-NHiPr sample without RSK2-CTD was processed identically (Figure 2.11b).

Treatment of MDA-MB-231 cells with electrophilic and non-electrophilic BODIPY conjugates

MDA-MB-231 cells were seeded into 6-well plates at 300,000 cells/well. After 24 h, cells were treated with 10 μM or 3 μM of the indicated BODIPY conjugates **17**-**19** for 1 h in serum-free media. The media was aspirated and the cells washed with 2 mL of cold PBS and lysed with 70 μL CelLytic lysis buffer (Sigma) supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors. The lysates were cleared by centrifugation at 14K

rpm for 10 min at 4°C and normalized by Bradford assay. Laemmli sample buffer was added to the lysates and the proteins were separated by 10% SDS-PAGE and detected by in-gel fluorescence scanning (Typhoon 9400), followed by Coomassie blue staining.

Kinase profiling

Kinase profiling of **CN-NHiPr** (1 µM) by competitive binding assays was performed at Ambit Biosciences (kinase profiling with this platform is now administered by DiscoveRx) [23]. The eight kinases showing >70% inhibition relative to the DMSO control are highlighted in the Figure 2.8a. K_D values were subsequently determined for these kinases.

Synthetic chemistry methods

Except as otherwise indicated, reactions were carried out under dry argon or nitrogen with dry solvents. Dry solvents were dispensed from a delivery system that passes the solvents through packed columns (tetrahydrofuran, acetonitrile, and methylene chloride: dry neutral alumina; dimethylformamide: activated molecular sieves). All other reagents were purchased from commercial sources and used as received.

Yields of reactions refer to chromatographically and spectroscopically pure compounds. Reactions were monitored by thin layer chromatography (TLC) using glass plates precoated with Merck silica gel 60 F_{254}. Visualization was by the quenching of UV fluorescence (λ_{max} = 254 nm) or by staining with ceric ammonium molybdate, *p*-anisaldehyde, potassium permanganate and Iodoplatinate stain. Retention factors (R_f) are quoted to 0.01. Proton magnetic resonance spectra were recorded on Varian Inova (400 MHz) spectrometer. Chemical shifts (δ_H) are quoted in ppm and are referenced to tetramethylsilane (internal). Coupling constants (*J*) are reported in Hertz to the nearest 0.1 Hz. Data are reported as follows: chemical shift, integration, multiplicity (br, broad; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet), coupling constant(s) and number of protons. Carbon magnetic resonance spectra were recorded on a Varian Inova (100 MHz) spectrometer. Chemical shifts (δ_C) are quoted in ppm to the nearest 0.1 ppm, and are referenced to tetramethylsilane (internal). Low resolution electron impact (EI) mass spectra were obtained with a Waters AutoSpec mass spectrometer with an Agilent 7890A gas chromatograph; JEOL

AX-505H, SX-102A (CI/EI), Micromass Platform II and LCT (APCI/ES/LCMS) spectrometers. Only molecular ions, fractions from molecular ions and other major peaks are reported. High-resolution mass spectra (HRMS) was obtained with a Thermo Electron Corporation LTQFT spectrometer using electrospray ionization with FT resolution set to 30000, and reported mass values are within the error limits of ± 5 ppm mass units.
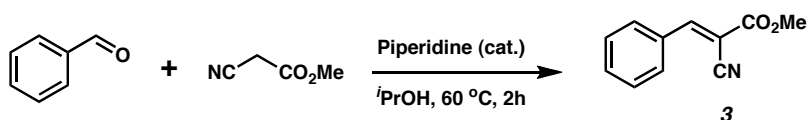
Methyl 2-cyano-3-phenylacrylate 3 (Synthesized by Dr. Shyam Krishnan)

To a solution of benzaldehyde (1.28 g, 12.06 mmol) and methyl cyanoacetate (1.21 g, 12.21 mmol, 1.01 equiv) in 2-propanol (5 mL) was added piperidine (150 μL, 1.52 mmol, 0.13 equiv). The reaction mixture was heated to 60°C for 2 h and DI water (15 mL) was added dropwise. The resulting slurry was cooled to 0–5°C for 15 min and filtered. The filter cake was washed with 2-propanol:DI water (1:3 v/v, 30 mL) and dried in vacuo to afford a white solid (ca. 7:1 of the desired methyl ester to the corresponding isopropyl ester formed by transesterification. Purification by silica gel chromatography (5–10% EtOAc in Hexanes) afforded 1.54 g (68% yield) of the desired methyl ester **3** as a white solid [34]. R_f 0.27 (9:1 Hexanes:EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.22 (s, 1H), 7.95 (m, 2H), 7.53–7.44 (m, 3H), 3.89 (s, 3H) ; ^{13}C NMR (100 MHz, CDCl$_3$): δ 162.8, 155.1, 133.3, 131.3, 131.0, 129.2, 115.3, 102.4, 53.3, MS (EI): 187.2 (M$^+$)
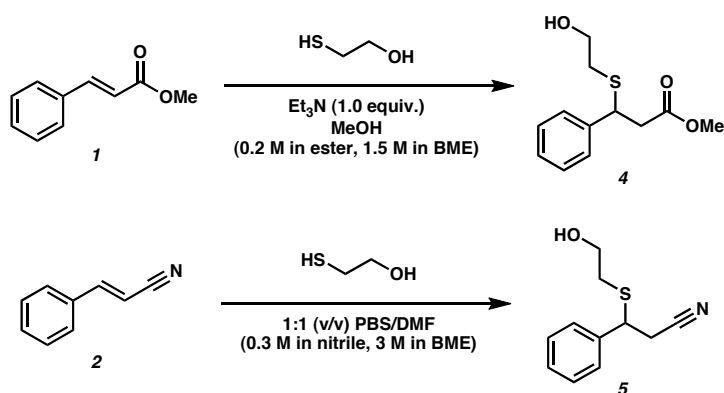
Methyl 3-(2-hydroxyethylthio)-3-phenylpropanoate 4 (Synthesized by Dr. Shyam Krishnan)

To a solution of *trans*-methylcinnamate **1** (133.5 mg, 0.823 mmol) and triethylamine (115 μL, 1.0 equiv) in MeOH (3 mL) was added BME (173 μL, 3.0 equiv). The reaction mixture was maintained at 20–25°C and after 2.5 h additional BME (200 μL, 3.5 equiv) was added. The reaction mixture was maintained at 20–25 °C for an additional 45.5 h, at which point complete conversion of the cinnamate was observed. The reaction mixture was concentrated and the residue was purified by silica gel chromatography (3:1 Hexanes:EtOAc) to afford 197 mg (>99% yield) of ester **4** as a colorless oil. R_f 0.20 (3:1 Hexanes:EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.35–7.22 (m, 5H), 4.32 (t, *J* = 7.7 Hz, 1H), 3.62 (s, 3H), 3.61 (m, 2H), 2.88 (m, 2H), 2.56–2.52 (m, 3H) ; ^{13}C NMR (100 MHz, CDCl$_3$): δ 171.3, 141.2, 128.7, 127.6, 127.5, 60.6, 51.9, 44.9, 41.3, 34.3; IR (film, cm^{-1}): 3429 (b, s), 3028 (s), 2952 (s), 1734 (s), 1436 (s), 1150 (s), 1044 (s); MS (EI): 240.1 (M$^+$).

3-(2-hydroxyethylthio)-3-phenylpropanenitrile 5 (Synthesized by Dr. Shyam Krishnan)

To a solution of cinnamonitrile **2** (204.6 mg, 1.58 mmol) in DMF (2 mL) was added 3 mL of a solution of 2-mercaptoethanol (BME, 5M) in phosphate buffered saline (PBS). The reaction mixture was maintained at 20–25°C for 11 days at which point complete conversion was observed. The reaction mixture was diluted with EtOAc (50 mL) and washed with DI water (4 x 25 mL). The combined aqueous washes were extracted with EtOAc (50 mL). The combined EtOAc extracts were washed with brine (50 mL), dried (Na_2SO_4) and concentrated to afford a colorless oil. Purification by silica gel chromatography (3:1→5:2 Hexanes:EtOAc) afforded 201.7 mg (61% yield) of nitrile **5** as a colorless solid. R_f 0.10 (3:1 Hexanes:EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ ; 7.39–7.28 (m, 5H), 4.19 (t, J = 7.3 Hz, 1H), 3.63 (m, 2H), 2.89 (m, 2H), 2.59 (m, 2H), 2.44 (bs, 1H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 139.1, 129.0, 128.4, 127.3, 117.3, 60.9, 45.2, 34.5, 25.7; IR (film, cm^{-1}): 3426 (b, s), 2251 (s), 1494 (m), 1291 (m), 1046 (s); MS (EI): 207.1 (M$^+$).
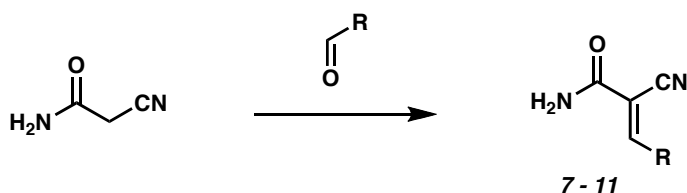
General procedure for the preparation of 2-cyanoacrylamides 7-11

(Synthesized by Robert Gutierrez and Dr. Jesse McFarland)

In a 2 mL vial fitted with a magnetic stir bar was combined the aldehyde (0.1 mmol, purchased from Aldrich), 2-cyanoacetamide (16 mg, 0.2 mmol), piperidine (0.01 mL, 0.1 mmol), and 2-propanol (0.3 mL). The mixture was stirred at room temp or 60°C for 1–18 h. Products that precipitated from the reaction mixture were isolated by filtration, washed with water, and dried *in vacuo*. Products that did not precipitate were extracted with EtOAc and water. The organic layer was then dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The resulting residue was then purified by silica gel chromatography (elution with EtOAc).

3-(4-(1*H*-Imidazol-1-yl)phenyl)-2-cyanoacrylamide (7). Yield: 23 mg (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.44 (s, 1H), 8.22 (s, 1H), 8.08 (d, 2H, *J* = 7.5), 7.95-7.88 (m, 4H), 7.79 (broad s, 1H), 7.16 (s, 1H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 162.6, 149.4, 139.3, 135.7, 131.8, 130.9, 130.4, 130.0, 120.2, 119.6, 117.6, 116.5, 106.3; HRMS (ESI) found 239.0943, calcd for $C_{13}H_{11}N_4O$ (MH$^+$) 239.0927.

2-Cyano-3-(1*H*-indol-4-yl)acrylamide (8).   Yield: 15 mg (67%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.52 (s, 1H), 8.52 (s, 1H), 8.04 (broad s, 1H), 7.96 (d, 1H, *J* = 7.7), 7.72 (broad s, 1H), 7.63 (d, 1H, *J* = 8.0), 7.57 (t, 1H, *J* = 2.2), 7.26 (t, 1H, *J* = 7.5 Hz), 6.81 (s, 1H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 163.3, 147.6, 136.2, 129.1, 127.7, 122.8, 121.0, 118.9, 117.1, 115.9, 105.4, 99.5; HRMS (ESI) found 234.0649, calcd for $C_{12}H_9N_3O$ (M Na$^+$) 234.0638.

2-Cyano-3-(6-(4-(methylsulfonyl)phenyl)pyridin-2-yl)acrylamide (9).   Yield: 12 mg (37%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.59 (d, 2H, *J* = 8.5), 8.30 (d, 1H, *J* = 8.5), 8.25 (s, 1H), 8.16 (t, 1H, *J* = 7.8), 8.04 (d, 2H, *J* = 8.5), 7.84 (d, 1H, *J* = 7.8), 3.30 (s, 3H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 162.6, 154.5, 149.9, 148.3, 142.3, 141.5, 139.2, 127.8, 127.39, 127.37, 123.4, 116.5, 109.8, 43.3; HRMS (ESI) found 328.0767, calcd for $C_{16}H_{14}N_3O_3S$ (MH$^+$) 328.0750.

2-Cyano-3-(thiophen-3-yl)acrylamide (10).

In a 20 mL vial fitted with a magnetic stir bar was dissolved 2-cyanoacetamide (80 mg, 1 mmol) in 10% NaHCO$_3$ (3 mL). To the solution was added 3-thiophene carboxaldehyde (101 mg, 0.9 mmol) and the resulting mixture was stirred vigorously for 3 h at room temp. The product was isolated by filtration, washed with water and dried *in vacuo*. Yield: 123 mg (69%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.37 (broad s, 1H), 8.19 (s, 1H), 7.82 (broad s, 1H), 7.79 (m, 2H), 7.69 (broad s, 1H). ^{13}C NMR (100 MHz, CDCl$_3$): δ 162.8, 146.9, 135.4, 127.9, 127.4, 117.5, 101.6; MS (ES$^+$): 179.1 (MH$^+$), calcd for $C_8H_6N_2OS$ 178.0.

**2-Cyano-3-(1*H*-imidazol-5-yl)acrylamide (11).**

In a 20 mL vial fitted with a magnetic stir bar was dissolved 2-cyanoacetamide (80 mg, 1 mmol) in 10% $NaHCO_3$ (3 mL). To the solution was added 4(5)-imidazole carboxaldehyde (86 mg, 0.9 mmol) and the resulting mixture was stirred vigorously for 3 h at room temp. The product was isolated by filtration, washed with water and dried *in vacuo*. Yield: 58 mg (39%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.78 (s, 1H), 8.03 (s, 1H), 7.95 (s, 1H), 7.92 (s, 1H), 7.64 (broad s, 1H), 7.52 (broad s, 1H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 163.4, 143.5, 138.1, 134.7, 124.6, 117.0, 100.6; HRMS (ESI) found 163.0622, calcd for $C_7H_7N_4O$ (MH[+]) 163.0614
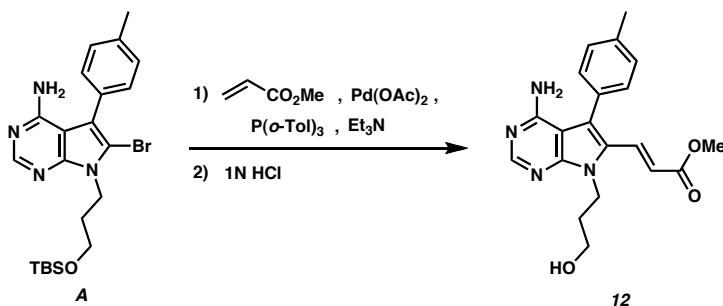
(*E*)-methyl 3-(4-amino-7-(3-hydroxypropyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-6-yl)acrylate 12

Bromide **A** was synthesized as previously reported [22]. In a 10 mL glass microwave tube containing 2mL DMF were added bromide **A** (50 mg, 0.1 mmol), $Pd(OAc)_2$ (7 mg, 0.3 equiv), tri-*o*-tolyl phosphine (20 mg, 0.6 equiv), Et_3N (47 μL, 3 equiv) and methyl acrylate (95 μL, 10 equiv). The reaction mixture was heated by microwave irradiation at 100˚C with stirring for two 10 min cycles at a power of

250W. The crude reaction was diluted with 5 mL EtOAc and the organic layer was washed with saturated aqueous sodium bicarbonate (2 x 5 mL), brine (2 x 5 mL), dried over MgSO$_4$. The mixture was then concentrated to a residue and absorbed on silica and purified by column chromatography (50% EtOAc/hexanes isocratic) to afford the TBS-protected methyl ester. The TBS-protected methyl ester (28 mg) was dissolved in THF (2 mL) and cooled to 0°C, then 1N HCl (0.3 mL) was added and the reaction was allowed to warm to room temperature (20–25°C). After 2.5 h the deprotection was complete and the reaction was diluted with 10 mL EtOAc. The organic phase was washed with saturated aqueous sodium bicarbonate (2 x 15 mL), brine (2 x 15 mL), dried over MgSO$_4$ and concentrated. The residue was absorbed on silica and purified by column chromatography (3% MeOH in EtOAc, isocratic) to yield 12.3 mg (33%) of methyl acrylate **12**. ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.29 (s, 1H), 7.61 (d, *J* = 17.6 Hz, 1H), 7.28—7.21 (m, 4H), 6.05 (d, *J* = 17.6 Hz, 1H), 4.53 (t, *J* = 5.4 Hz, 2H), 3.73 (s, 3H), 3.479 (t, *J* = 5.1 Hz, 2H), 2.44 (s, 3H), 2.01 (m, 2H), 1.67 (s, 2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 167.5, 157.7, 157.2, 153.8, 151.9, 138.8, 131.7, 130.3, 128.1, 126.9, 121.2, 118.1, 102.4, 57.4, 51.9, 38.7, 32.9, 21.5 HRMS (ESI) found 367.1775, calcd for C$_{20}$H$_{22}$N$_4$O$_3$ (MH$^+$) 367.1765
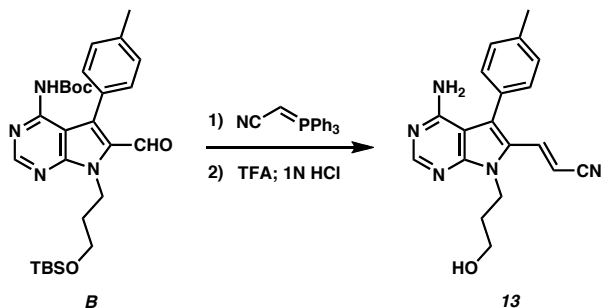
(*E*)-3-(4-amino-7-(3-hydroxypropyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-6-yl)acrylonitrile 13

A flame dried 50 mL round bottom flask was charged with aldehyde **B** (218 mg, 0.41 mmol; prepared under contract by Albany Molecular Research Inc., Albany, NY; details available upon request) and (triphenylphosphoranylidine)acetonitrile (500 mg, 4 equiv). CH_2Cl_2 (10 mL) was added and the reaction mixture was stirred at room temp. After 16 h the reaction was complete and the crude mixture after concentration was adsorbed on silica and purified by column chromatography (20% EtOAc/hexanes isocratic) to yield the protected nitrile, which was dissolved in CH_2Cl_2 (2 mL) and was cooled to 0˚C. TFA (2 mL) was added dropwise and the reaction mixture was allowed to warm to room temp. After 12 h the reaction mixture was concentrated and the residue redissolved in THF (3 mL). The solution was cooled to 0˚C and 1N HCl (1 mL) was added after which the reaction mixture was allowed to warm to room temp. After 4.5 h the deprotection was complete and the reaction mixture was diluted with 10 mL of EtOAc. The organic phase was washed with saturated aqueous sodium bicarbonate (2 x 15 mL), brine (2 x 15 mL), dried over $MgSO_4$ and concentrated. The residue was adsorbed on silica and purified by column

chromatography (3% MeOH in EtOAc, isocratic) to yield 9.1 mg (11.4%) of acrylonitrile **13**. ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.30 (s, 1H), 7.36—7.29 (m, 4H), 7.24 (d, *J* = 16.8 Hz, 1H), 5.42 (d, *J* = 16.8 Hz, 1H), 5.15 (s, broad, 2H), 4.80 (s, broad, 1H), 4.48 (t, *J* = 5.8 Hz, 2H), 3.48 (m, 2H), 2.46 (s, 3H), 1.96 (m, 2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 157.9, 154.4, 151.9, 139.4, 136.9, 130.7, 129.9, 127.2, 124.4, 121.6, 118.7, 102.5, 96.4, 57.4, 38.7, 32.9, 21.5 HRMS (ESI) found 334.1677, calcd for C$_{19}$H$_{20}$N$_5$O (MH$^+$) 334.1662.
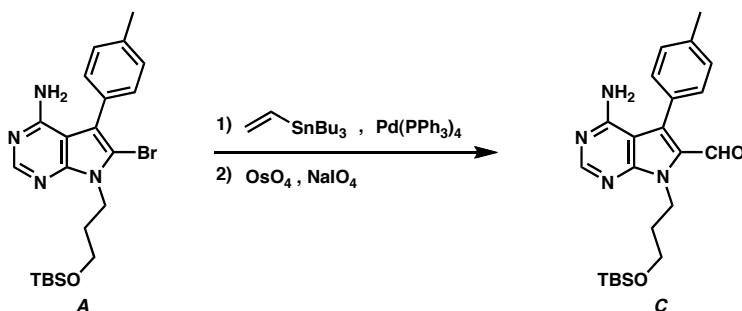
4-amino-7-(3-(*tert*-butyldimethylsilyloxy)propyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidine-6-carbaldehyde C

To a solution of bromide **A** [22] (1.0 g, 2.1 mmol) in toluene (30 ml) was added tributylvinyltin (0.8 ml, 2.73 mmol). The solution was sparged with argon for 10 min. Tetrakis(triphenylphosphine)palladium (244 mg, 0.21 mmol) was quickly added and the reaction mixture was sparged with argon for a further 10 min, then heated to reflux for 3 h. The reaction mixture was filtered through celite and the filtrate was concentrated. The residue was purified by flash column chromatography in 50% EtOAc/hexanes to give a light yellow residue that was lyophilized from benzene to afford the corresponding vinyl pyrrolopyrimidine as a light yellow powder.

To a solution of the vinyl pyrrolopyrimidine (760 mg, 1.8 mmol) in 3:1 v/v THF:H$_2$O (11.3 ml) under argon was added, dropwise, osmium tetroxide solution (1.75 ml, 0.18 mmol, 2.5% in *t*-BuOH). The reaction was stirred at room temperature (20–25 °C) under argon for 20 min. Sodium periodate (860 mg, 3.6 mmol, dissolved in 2.4 mL of warm water) was added dropwise to the reaction over a period of 30 min. The reaction mixture was stirred for 1.5 h at room temp and was then diluted with ethyl acetate. The organic layer was washed with saturated aqueous sodium thiosulfate and the aqueous layer was back extracted with ethyl acetate. The combined organic extracts were dried over sodium sulfate and concentrated. The residue was purified by flash column chromatography in 50% EtOAc/hexanes to give 417 mg (55% yield) of aldehyde **C** as a yellowish oil that solidified upon standing. ^{1}H NMR (400 MHz, CDCl$_3$): δ 9.63 (s, 1H), 8.34 (s, 1H), 7.39—7.32 (m, 4H), 4.72 (m, 2H), 3.72 (m, 2H), 2.45 (s, 3H), 2.043 (m, 2H), 0.88 (s, 9H), 0.04 (s, 6H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 182.1, 158.8, 156.0, 151.8, 139.1, 130.5, 130.3, 129.7, 128.9, 128.4, 101.8, 60.9, 41.2, 33.7, 25.9, 21.3, 18.3, -5.4 MS (ES$^+$): 425.0 (MH$^+$).
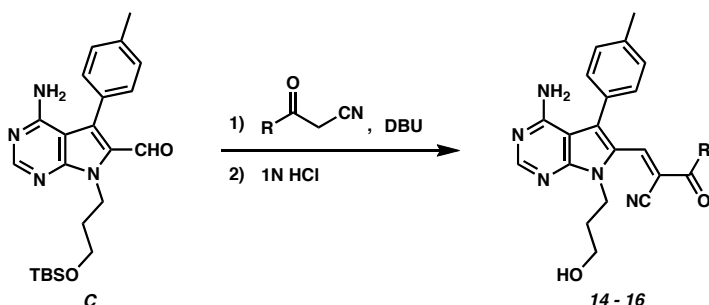
General procedure for the preparation of cyanoacrylates and cyanoacrylamides 14 – 16

To a solution of aldehyde **C** (0.3 mmol) in THF (2 ml) were added DBU (1.2 equiv) and the appropriate cyanoacrylate or cyanoacetamide (1.2 equiv). The reaction was stirred at room temp until all the starting material had been consumed. The reaction mixture was concentrated and the residue was purified by flash column chromatography to give the TBS-protected products.

To a solution of the TBS-protected cyanoacrylates or cyanoacrylamides (0.1 mmol) in THF (1 ml) was added 1N aqueous HCl. The reaction mixture was stirred at room temp for 1.5 h and then diluted with ethyl acetate. The organic layer was washed with saturated aqueous sodium bicarbonate followed by brine and then dried over sodium sulfate, filtered and concentrated. The residues were purified by flash column chromatography in 100% EtOAc → 5% MeOH/EtOAc to give the desired compounds **14–16**.

**Methyl 3-(4-amino-7-(3-hydroxypropyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-6-yl)-2-cyanoacrylate 14**

Cyanoacrylate **14** was prepared in 22% yield from methyl 2-cyanoacetate and aldehyde **C**. ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.32 (s, 1H), 8.24 (s, 1H), 7.35—7.29 (m, 4H), 5.47 (s, broad, 2H), 4.49 (t, *J* = 6.0 Hz, 2H), 3.86 (s, 3H), 3.44 (t, *J* = 5.0 Hz, 2H), 2.47 (s, 3H), 1.95 (m, 2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 162.8, 158.5, 154.9, 152. 8, 142.1, 139.1, 130.9, 130.3, 129.6, 125.4, 114.2, 102.7, 57.6, 53.7, 39.3, 32.9, 21.6 HRMS (ESI) found 392.1723, calcd for C$_{21}$H$_{22}$N$_5$O$_3$ (MH$^+$) 392.1717.

3-(4-amino-7-(3-hydroxypropyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-6-yl)-2-cyano-*N*-isopropylacrylamide 15 (CN-NHiPr)

Cyanoacrylamide **16** was prepared in 13% yield from *N*-isopropylcyanoacetamide [35] and aldehyde **C**. ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.32 (s, 1H), 8.30 (s, 1H), 7.35—7.29 (m, 4H), 5.94 (d, *J* = 7.7 Hz, 1H), 5.51 (broad s, 1H), 4.46 (t, *J* = 5.8 Hz, 2H), 4.12 (m, 1H), 3.45 (t, *J* = 5.5 Hz, 2H), 2.43 (s, 3H), 1.95 (m, 2H), 1.20 (d, *J* = 6.6, 6H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 158.9, 158.3, 154.4, 152.4, 139.4, 138.7, 130.8, 130.6, 129.7, 125.8, 123.3, 115.6, 108.1, 102.5, 57.5, 43.1, 39.1, 33.0, 22.6, 21.6 HRMS (ESI) found 419.2193, calcd for C$_{23}$H$_{27}$N$_6$O$_2$ (MH$^+$) 419.2190.

tert-butyl 3-(4-amino-7-(3-hydroxypropyl)-5-*p*-tolyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-6-yl)-2-cyanoacrylate 16

Cyanoacrylate **16** was prepared in 16% yield from *tert*-butyl 2-cyanoacetate and aldehyde **C**. ^{1}H NMR (400 MHz, CDCl$_3$):δ 8.37 (s, 1H), 8.15 (s, 1H), 7.34—7.28 (m, 4H), 5.34 (broad s, 2H), 4.95 (t, *J* = 5.8 Hz, 2H), 3.43 (m, 2H), 2.43 (s, 3H), 1.95 (m, 2H), 1.52 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 161.1, 158.4, 156.7, 154.7, 152.1, 141.1, 138.9, 136.4, 130.7, 129.7, 124.4, 114.4, 107.9, 102.6, 57.6, 39.3, 33.0, 28.1, 27.7, 21.6 HRMS (ESI) found 434.2181, calcd for C$_{24}$H$_{28}$N$_5$O$_3$ (MH$^+$) 434.2187.
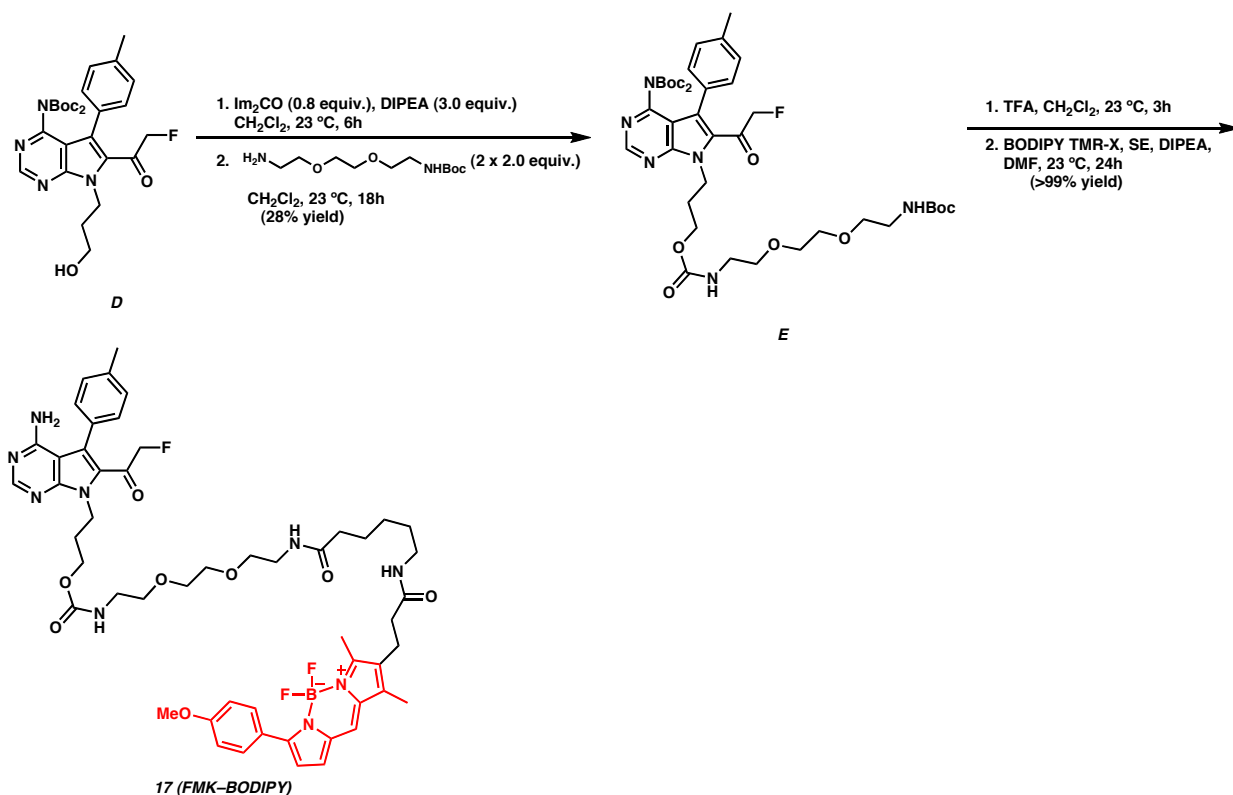
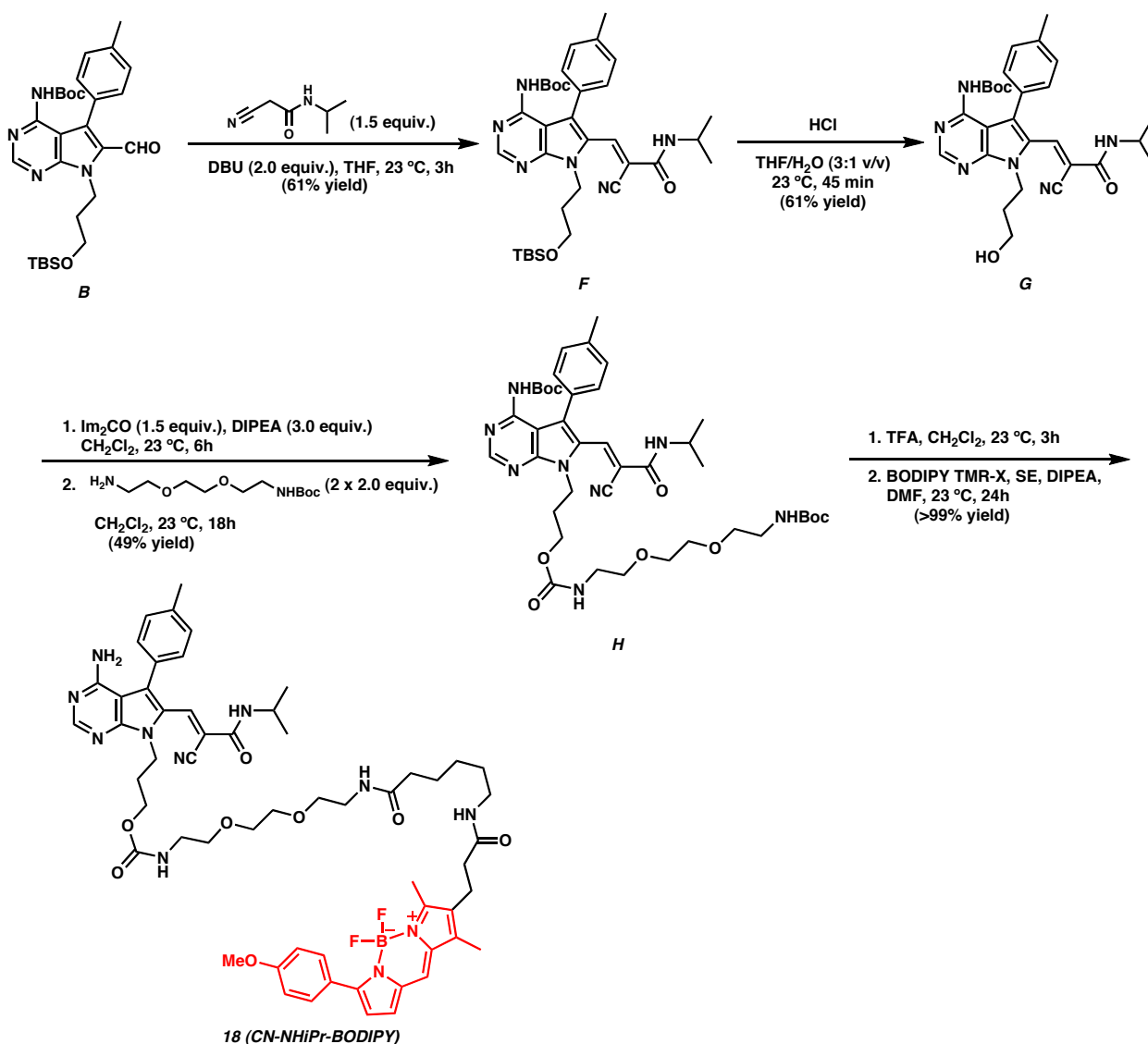**Carbamate E** (Synthesized by Dr. Shyam Krishnan)

To a solution of alcohol **D** [22] (65.5 mg, 0.121 mmol) and dicarbonylimidazole (15.7 mg, 0.8 equiv) in CH_2Cl_2 (1 mL) was added DIPEA (16.8 µL, 0.8 equiv). The reaction mixture was maintained at 20–25°C for 2 h and *N*-Boc-2,2'-(ethylenedioxy)diethylamine (37.3 µL, 1.3 equiv) was added. After 3 h, additional diamine (37.3 µL, 1.3 equiv) was added. After 2 h, the reaction mixture was diluted with EtOAc (50 mL) and washed with 0.5 M aqueous HCl (20 mL) and brine (30 mL). The organic phase was dried (Na_2SO_4) and concentrated to afford a colorless oil. Purification by preparative TLC (1:2 Hexanes:EtOAc x 5 elutions) afforded unreacted alcohol **D** (15.7 mg, 24% yield), the acylimidazole of the alcohol (10.1 mg, 16% yield) and the desired carbamate

E (20.9 mg , 28% yield) as a colorless oil. R_f 0.32 (1:1 Hexanes/EtOAc x 3 elutions); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.95 (s, 1H), 7.25 (m, 4H), 5.17 (bs, 1H), 5.09 (bs, 1H), 4.80 (app. t, J = 7.0 Hz, 2H), 4.63 (d, J = 46.7 Hz, 2H), 4.14 (m, 2H), 3.61 (s, 4H), 3.54 (app. q, J = 5.3 Hz, 4H), 3.34 (m, 4H), 2.46 (s, 3H), 2.18 (m, 2H), 1.44 (s, 9H), 1.33 (s, 18H); MS (ESI): 817.4 (MH$^+$).

17 (FMK-BODIPY) (Synthesized by Dr. Shyam Krishnan)

To a solution of carbamate **L** (10.1 mg, 12.4 μmol, 1.5 equiv) in CH$_2$Cl$_2$ (0.9 mL) was added TFA (0.9 mL) at 20–25°C. The reaction mixture was maintained at ambient temperature for 3 h, then concentrated under reduced pressure to afford a colorless oil. A solution of BODIPY-TMR-X, SE (5.0 mg, 8.2 μmol, 1.0 equiv) in DMF (1.5 mL) was then added to the oil followed by DIPEA (14.3 μL, 20.0 equiv). The reaction mixture was maintained at ambient temperature for 24 h while protected from light, then concentrated under reduced pressure to afford a purple solid. Purification by preparative TLC (10:1 CH$_2$Cl$_2$:MeOH x 2 elutions) afforded FMK-BODIPY **17** (7.6 mg, 92% yield based on BODIPY-TMR-X, SE; 94.4:5.6 of fluoromethyl ketone: methyl ketone) as a purple solid. R_f 0.47 (10:1 CH$_2$Cl$_2$:MeOH); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.35 (s, 1H), 7.86 (d, J = 9.0 Hz, 2H), 7.36 (d, J = 8.0 Hz, 2H), 7.30 (d, J = 7.8 Hz, 2H), 7.08 (s, 1H), 7.00–6.94 (m, 3H), 6.53 (d, J = 4.2 Hz, 1H), 6.21 (bs, 1H), 5.93 (bs, 1H), 5.36 (bs, 1H), 4.89 (bs, 2H), 4.73 (t, J = 6.9 Hz, 2H), 4.61 (d, J = 46.8 Hz, 2H), 4.09 (m, 2H), 3.86 (s, 3H), 3.64–3.51 (m, 8H), 3.41 (q, J = 5.3 Hz, 2H), 3.36 (m, 2H), 3.20 (q, J = 6.4 Hz, 2H), 2.75 (t, J = 7.5 Hz, 2H), 2.51 (s, 3H), 2.47 (s,

3H), 2.28 (t, *J* = 7.4 Hz, 2H), 2.20 (s, 3H), 2.15–2.08 (m, 4H), 1.57 (m, 2H), 1.25

(m, 2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 171.9, 159.1, 156.1, 140.2, 130.90,

130.86, 130.82, 130.7, 130.5, 129.6, 128.0, 126.9, 125.7, 123.0, 113.9, 84.9,

83.1, 70.5, 70.47, 70.38, 70.36, 70.2, 70.17, 70.15, 70.12, 70.1, 62.6, 55.5, 41.5,

41.0, 39.5, 39.4, 39.3, 36.6, 36.3, 30.1, 29.1, 26.4, 25.1, 21.6, 20.4, 14.0, 13.3,

9.8; HRMS (ESI) found 1010.4952, calcd for C$_{52}$H$_{64}$BF$_3$N$_9$O$_8$ (MH$^+$): 1010.

18 (CN-NHiPr-BODIPY)

Cyanoacrylamide F (Synthesized by Dr. Shyam Krishnan)

To a solution of aldehyde **B** (247.3 mg, 0.471 mmol; prepared under contract by Albany Molecular Research Inc., Albany, NY; details available upon request) and *N*-isopropylcyanoacetamide [35] (89.2 mg, 0.707 mmol, 1.5 equiv) in THF (5 mL) was added DBU (141 μL, 0.943 mmol, 2.0 equiv). The reaction mixture was maintained at 20–25°C for 2.5 h, quenched with saturated aqueous NH_4Cl (5 mL) and extracted with EtOAc (4 x 10 mL). The combined organic extracts were dried ($MgSO_4$) and concentrated. The residue was purified by silica gel chromatography (4:1→3:1 Hexanes:EtOAc) to afford 181 mg (61% yield) of the cyanoacrylamide **F** as a yellow oil. R_f 0.38 (3:1 Hexanes:EtOAc); ^1H NMR (400 MHz, $CDCl_3$): δ 8.74 (s, 1H), 8.35 (s, 1H), 7.36 (d, *J* = 8.0 Hz, 2H), 7.29 (d, *J* = 8.0 Hz, 2H), 5.95 (d, *J* = 7.7 Hz, 1H), 4.49 (t, *J* = 7.3 Hz, 2H), 4.14 (sextet, *J* = 6.6 Hz, 1H), 3.63 (t, *J* = 5.9 Hz, 2H), 2.44 (s, 3H), 2.01 (m, 2H), 1.42 (s, 9H), 1.21 (d, *J* = 6.6 Hz, 6H), 0.88 (s, 9H), 0.36 (s, 6H) ; ^{13}C NMR (100 MHz, $CDCl_3$): δ 158.1, 153.8, 152.71, 152.67, 149.4, 140.0, 138.8, 130.5, 130.0, 129.6, 127.8, 120.9, 115.5, 109.7, 105.4, 81.6, 60.1, 42.9, 41.0, 33.1, 28.0, 25.9, 22.5, 21.3, 18.3, –5.4; MS (ESI): 633 (MH^+)

Alcohol G (Synthesized by Dr. Shyam Krishnan)

To a solution of Boc pyrrolopyrimidine **F** (181 mg, 0.286 mmol) in THF (3 mL) was added 1 M aqueous HCl (1 mL). The reaction mixture was maintained at 20–25°C for 45 min, then quenched with saturated aqueous $NaHCO_3$ (5 mL) and extracted with EtOAc (3 x 5 mL). The combined organic extracts were dried

(MgSO₄) and concentrated to afford a colorless oil, which was purified by preparative TLC (EtOAc) to afford the alcohol **G** as an orange-yellow foam, 89.7 mg (61% yield). R_f 0.24 (EtOAc); ^{1}H NMR (400 MHz, CDCl₃): δ 8.70 (s, 1H), 8.35 (s, 1H), 7.35 (d, *J* = 7.9 Hz, 2H), 7.28 (d, *J* = 7.9 Hz, 2H), 5.95 (d, *J* = 7.7 Hz, 1H), 4.48 (t, *J* = 6.1 Hz, 2H), 4.11 (m, 1H), 4.05 (t, *J* = 6.9 Hz, 2H), 3.45 (q, *J* = 5.8 Hz, 2H), 2.43 (s, 3H), 1.96 (m, 2H), 1.40 (s, 9H), 1.19 (d, *J* = 6.6 Hz, 6H); ^{13}C NMR (100 MHz, CDCl₃): δ 158.2, 153.9, 153.0, 152.6, 149.3, 139.2, 139.0, 130.7, 129.9, 129.6, 127.5, 121.0, 115.1, 110.2, 105.0, 81.9, 57.7, 43.1, 39.3, 32.8, 28.0, 22.4, 21.4; MS (ESI): 519.6 (MH⁺).
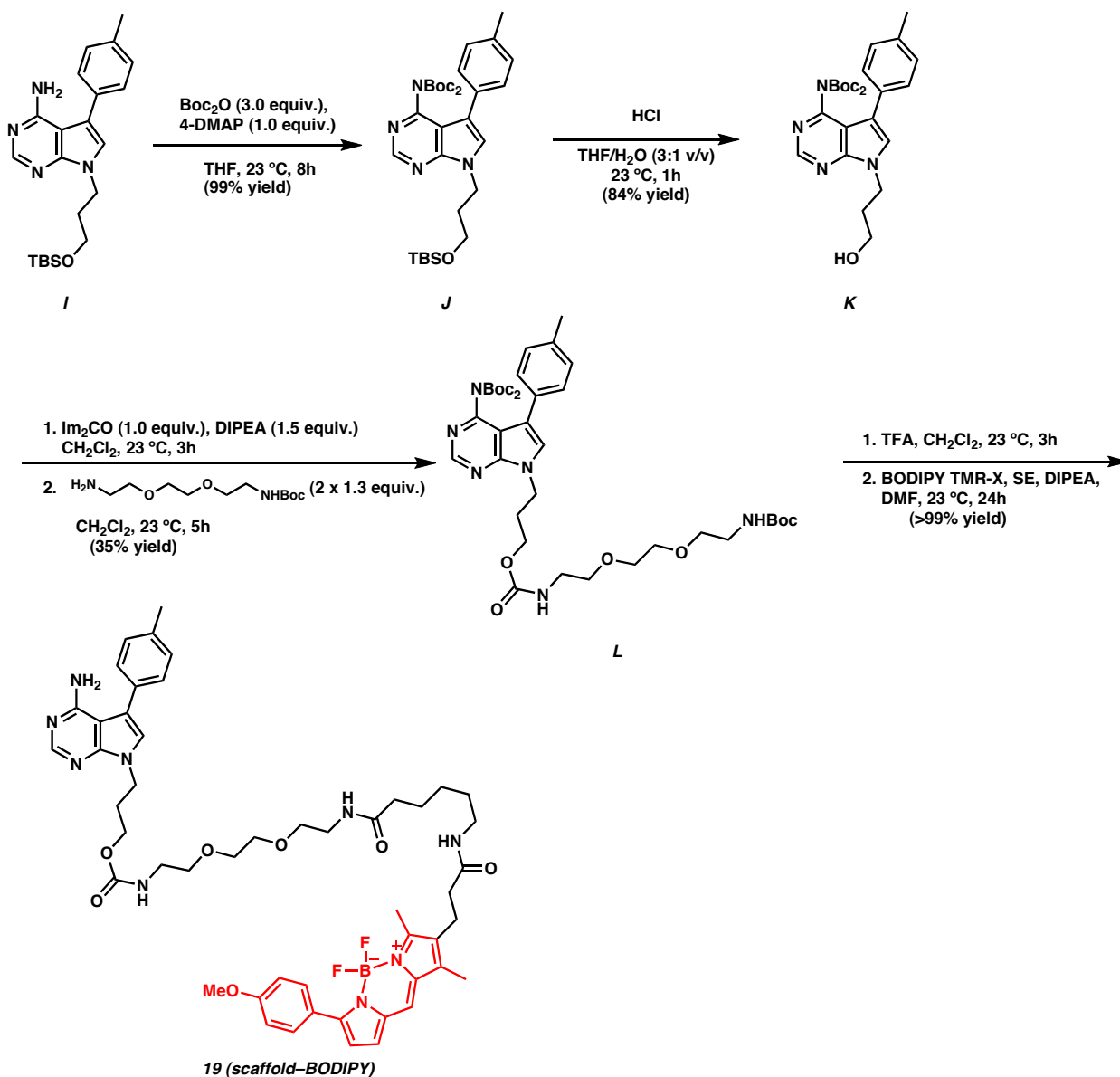
Carbamate H (Synthesized by Dr. Shyam Krishnan)

To a solution of alcohol **G** (72.5 mg, 0.140 mmol) and dicarbonylimidazole (34 mg, 0.21 mmol, 1.5 equiv) in CH₂Cl₂ (2 mL) was added DIPEA (73 μL, 0.42 mmol, 3.0 equiv). The reaction mixture was maintained at 20–25°C for 6 h and *N*-Boc-2,2'-(ethylenedioxy)diethylamine (66.4 μL, 0.28 mmol, 2.0 equiv) was added. After 6 h, additional diamine (66.4 μL, 0.28 mmol, 2.0 equiv) was added and the reaction mixture was maintained at 20–25°C for 12 further hours, after which it was diluted with EtOAc (10 mL) and saturated aqueous NH₄Cl (5 mL). The phases were separated and the aqueous phase was extracted with EtOAc (2 x 5 mL). The combined organic extracts were concentrated to afford a colorless oil. Purification by preparative TLC (EtOAc x 2 elutions) afforded carbamate **H** (53.9 mg , 49% yield, *E:Z* = 85:15) as a yellow foam. Characterized as a mixture of geometric isomers (chemical shifts in italics represent peaks due to the minor

isomer), R_f 0.30 (1:2 Hexanes:EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.71 (s, 1H), *8.69* (s, 1H), 8.57 (s, 1H), 7.40–7.21 (m, 4H), 6.06 (bs, 1H), 5.99 (d, *J* = 7.6 Hz, 1H), 5.06 (bs, 1H), 4.44 (t, *J* = 6.7 Hz, 2H), *4.32* (t, *J* = 6.8 Hz, 2H), 4.11 (m, 1H), 4.05 (m, *J* = 5.6 Hz, 2H), 3.63–3.51 (m, 8H), 3.38 (q, *J* = 5.6 Hz, 2H), 3.31 (m, 2H), 2.43 (s, 3H), *2.41* (s, 3H), 2.18 (m, 2H), 1.85 (bs, 1H), 1.42 (s, 9H), 1.40 (s, 9H), *1.37* (s, 9H), 1.20 (d, *J* = 6.6 Hz, 6H), *1.04* (d, *J* = 6.6 Hz, 6H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 158.7, 156.3, *156.0*, 154.0, *153.2*, 152.8, *152.53*, 152.47, 152.1, *149.6*, 149.4, 140.2, *139.5*, 138.9, 130.6, *130.1*, 130.0, 129.9, *129.7*, 127.6, 120.9, 115.1, 109.2, 105.5, 81.7, 70.34, 70.29, 70.25, 70.1, 60.9, 43.1, *42.6*, 40.8, *40.6*, 40.4, 40.1, 29.4, 22.4, 22.1, 21.4, *21.3*; MS(ESI): 793.4 (MH$^+$).

18 (CN-NHiPr-BODIPY) (Synthesized by Dr. Shyam Krishnan)

To a solution of carbamate **H** (5.2 mg, 6.6 μmol, 1.5 equiv) in CH$_2$Cl$_2$ (0.5 mL) was added TFA (0.5 mL) at 20–25°C. The reaction mixture was maintained at ambient temperature for 3 h, then concentrated under reduced pressure to afford a yellow oil. A solution of BODIPY-TMR-X, SE (2.5 mg, 4.1 μmol, 1.0 equiv) in DMF (1 mL) was then added to the oil followed by DIPEA (7.2 μL, 10.0 equiv). The reaction mixture was maintained at ambient temperature for 24 h while protected from light, then concentrated under reduced pressure to afford a purple oil. Purification by preparative TLC (10:1 CH$_2$Cl$_2$:MeOH x 2 elutions) afforded CN-NHiPr-BODIPY **18** (4.5 mg, E:Z = 3.9:1, >99% yield based on BODIPY-TMR-X, SE) as a purple solid. R_f 0.50 (9:1 CH$_2$Cl$_2$:MeOH); ^{1}H NMR (400 MHz, CDCl$_3$), chemical shifts of peaks from the minor isomer are in italics: δ

8.53 (s, 1H), 8.32 (s, 1H), *8.31* (s, 1H), 7.86 (d, *J* = 9.0 Hz, 2H), 7.33 (d, *J* = 8.4

Hz, 2H), 7.29 (d, *J* = 8.4 Hz, 2H), 7.08 (s, 1H), 7.00–6.94 (m, 3H), 6.53 (d, *J* = 4.0

Hz, 1H), 6.17 (bs, 1H), 6.12 (t, *J* = 5.6 Hz, 1H), 5.99 (d, *J* = 7.7 Hz, 1H), 5.84 (t, *J*

= 5.7 Hz, 1H), 5.23 (bs, 2H), *5.12* (bs, 2H), 4.41 (t, *J* = 6.8 Hz, 2H), *4.28* (t, *J* =

6.9 Hz, 2H), 4.11 (m, 1H), 4.03 (t, *J* = 5.5 Hz, 2H), *3.95* (t, *J* = 6.2 Hz, 2H), 3.85

(s, 3H), 3.61–3.56 (m, 6H), 3.52 (t, *J* = 5.1 Hz, 2H), 3.42–3.34 (m, 4H), 3.20 (m,

2H), 2.75 (t, *J* = 7.3 Hz, 2H), 2.52 (s, 3H), 2.43 (s, 3H), *2.41* (s, 3H), 2.28 (t, *J* =

7.5 Hz, 2H), 2.21 (s, 3H), 2.17–2.09 (m, 4H), 1.59 (m, 2H), 1.44 (m, 2H), 1.25 (m,

2H), 1.20 (d, *J* = 6.6 Hz, 6H), *1.05* (d, *J* = 6.5Hz, 6H); ^{13}C NMR (100 MHz,

CDCl$_3$): δ 174.8, 171.6, 154.5, 140.3, 130.6, 130.5, 130.0, 129.6, 127.8, 122.7,

118.2, 113.7, 70.3, 70.0, 69.90, 69.88, 55.3, 43.0, 39.2, 39.1, 39.05, 36.4, 36.0,

29.7, 29.4, 28.9, 26.2, 24.8, 22.4, 21.4, 20.1, 13.2, 9.6; HRMS (ESI) found

1086.5583, calcd for C$_{57}$H$_{71}$BF$_2$N$_{11}$O$_8$ (MH$^+$): 1086.5543.

DiBoc pyrrolopyrimidine J (Synthesized by Dr. Shyam Krishnan)

To a solution of pyrrolopyrimidine **I** [22] (116.4 mg, 0.293 mmol) and DMAP (7.2 mg, 0.2 equiv) in MeCN/THF (2:1 v/v, 6 mL) was added Boc$_2$O (160.1 mg, 2.5 equiv). The reaction mixture was maintained at 20–25°C for 2 h, at which point additional Boc$_2$O (160.1 mg, 2.5 equiv) and DMAP (36 mg, 1.0 equiv) was added. The reaction mixture was maintained at 20–25°C for an additional 8 h,

then concentrated to afford a brown solid. Purification by preparative TLC (3:1 Hexanes:EtOAc) afforded the DiBoc pyrrolopyrimidine **J** as a colorless foam, 173.8 mg (99% yield). R_f 0.38 (3:1 Hexanes:EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.81 (s, 1H), 7.35 (d, J = 8.1 Hz, 2H), 7.27 (s, 1H), 7.18 (d, J = 8.1 Hz, 2H), 4.43 (t, J = 6.9 Hz, 2H), 3.65 (t, J = 5.8 Hz, 2H), 2.38 (s, 3H), 2.09 (app. q, J = 6.4 Hz, 2H), 1.27 (s, 18H), 0.91 (s, 9H), 0.06 (s, 6H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 153.4, 151.4, 150.9, 150.7, 136.5, 130.6, 129.2, 128.6, 127.8, 116.0, 113.3, 83.0, 59.7, 41.8, 32.8, 27.7, 26.0, 21.2, 18.3, –5.3; MS (ESI): 597.3 (MH$^+$).

Alcohol K (Synthesized by Dr. Shyam Krishnan)

To a solution of DiBoc pyrrolopyrimidine **H** (113.1 mg, 0.190 mmol) in THF (3 mL) was added 1M aqueous HCl (1 mL). The reaction mixture was maintained at 20–25°C for 1 h, then quenched with saturated aqueous NaHCO$_3$ (5 mL) and extracted with EtOAc (3 x 5 mL). The combined organic extracts were dried (MgSO$_4$) and concentrated to afford a colorless oil, which was purified by preparative TLC (1:2 Hexanes:EtOAc) to afford the alcohol **K** as a colorless oil, 76.2 mg (84% yield). R_f 0.17 (1:1 Hexanes/EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.80 (s, 1H), 7.35 (d, J = 7.9 Hz, 2H), 7.24 (s, 1H), 7.19 (d, J = 7.9 Hz, 2H), 4.46 (t, J = 6.2 Hz, 2H), 3.77 (t, J = 6.3 Hz, 1H), 3.49 (m, 2H), 2.38 (s, 3H), 2.03 (m, 2H), 1.29 (s, 18H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 153.6, 151.8, 150.9, 150.6, 136.8, 130.2, 129.3, 128.6, 127.5, 116.7, 113.2, 83.2, 58.0, 41.1, 33.0, 27.6, 21.2; MS (ESI): 483 (MH$^+$).

Carbamate L (Synthesized by Dr. Shyam Krishnan)

To a solution of alcohol **I** (54 mg, 0.112 mmol) and dicarbonylimidazole (18.1 mg, 1.0 equiv) in CH_2Cl_2 (2 mL) was added DIPEA (29.2 μL, 1.5 equiv). The reaction mixture was maintained at 20–25°C for 3 h and *N*-Boc-2,2'-(ethylenedioxy)diethylamine (34.5 μL, 1.3 equiv) was added. After 3 h, further diamine (34.5 μL, 1.3 equiv) was added. The reaction mixture was maintained at ambient temperature for a further 2 h and then diluted with EtOAc (10 mL) and saturated aqueous NH_4Cl (5 mL). The phases were separated and the aqueous phase was extracted with EtOAc (2 x 5 mL). The combined organic extracts were concentrated to afford a colorless oil. Purification by preparative TLC (1:2 Hexanes:EtOAc x 5 elutions) afforded carbamate **L** (30 mg , 35% yield) as a colorless oil. R_f 0.30 (1:2 Hexanes/EtOAc); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.80 (s, 1H), 7.35 (d, *J* = 8.1 Hz, 2H), 7.26 (s, 1H), 7.18 (d, *J* = 8.1 Hz, 2H), 5.20 (bs, 1H), 5.05 (bs, 1H), 4.42 (t, *J* = 6.9 Hz, 2H), 4.14 (t, *J* = 5.7 Hz, 2H), 3.59 (s, 4H), 3.53 (app. q, *J* = 4.7 Hz, 4H), 3.33 (m, 4H), 2.38 (s, 3H), 2.22 (m, 2H), 1.43 (s, 9H), 1.28 (s, 18H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 156.4, 156.1, 153.5, 151.5, 151.1, 150.8, 136.7, 130.5, 129.3, 128.7, 127.3, 116.4, 113.4, 83.2, 70.33, 70.28, 70.1, 62.1, 42.3, 40.8, 40.4, 29.7, 28.5, 28.1, 27.8, 21.2; MS (ESI): 757.9 (MH[+]).

19 (scaffold-BODIPY) (Synthesized by Dr. Shyam Krishnan)

To a solution of carbamate **L** (10.8 mg, 14.3 μmol, 3.5 equiv) in CH_2Cl_2 (1 mL) was added TFA (1 mL) at 20–25°C. The reaction mixture was maintained at ambient temperature for 3 h, then concentrated under reduced pressure to afford

a colorless oil. A solution of BODIPY-TMR-X, SE (2.5 mg, 4.1 μmol, 1.0 equiv) in DMF (1 mL) was then added to the oil followed by DIPEA (7.2 μL, 10.0 equiv). The reaction mixture was maintained at ambient temperature for 24 h while protected from light, then concentrated under reduced pressure to afford a purple oil. Purification by preparative TLC (10:1 CH_2Cl_2:MeOH x 2 elutions) afforded scaffold-BODIPY **19** (4.1 mg, >99% yield based on BODIPY-TMR-X, SE) as a purple oil. R_f 0.40 (10:1 CH_2Cl_2:MeOH); ^{1}H NMR (400 MHz, CDCl$_3$): δ 8.29 (s, 1H), 7.86 (d, J = 8.9 Hz, 2H), 7.35 (d, J = 8.0 Hz, 2H), 7.25 (d, J = 8.0 Hz, 2H), 7.08 (s, 1H), 7.00–6.93 (m, 4H), 6.53 (d, J = 4.1 Hz, 1H), 6.19 (bs, 1H), 5.91 (bs, 1H), 5.33 (bs, 1H), 5.14 (bs, 2H), 4.30 (t, J = 6.5 Hz, 2H), 4.08 (t, J = 5.8 Hz, 2H), 3.85 (s, 3H), 3.59 (bs, 4H), 3.55–3.50 (m, 4H), 3.41 (q, J = 5.2 Hz, 2H), 3.34 (m, 2H), 3.19 (q, J = 6.5 Hz, 2H), 2.75 (t, J = 7.4 Hz, 2H), 2.51 (s, 3H), 2.40 (s, 3H), 2.27 (t, J = 7.3 Hz, 2H), 2.19 (s, 3H), 2.18 (m, 2H), 2.09 (t, J = 7.2 Hz, 2H), 1.56 (m, 2H), 1.42 (m, 2H), 1.25 (m, 2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 171.6, 160.4, 157.0, 151.8, 139.9, 137.0, 134.9, 134.4, 133.3, 131.7, 130.7, 129.7, 128.7, 127.8, 125.5, 122.7, 122.5, 118.3, 118.2, 116.4, 116.3, 113.7, 70.2, 70.0, 69.9, 61.9, 55.3, 41.5, 40.8, 39.9, 39.6, 36.4, 36.0, 29.7, 28.8, 26.2, 24.8, 21.1, 20.1, 14.1, 13.2, 9.6; HRMS (ESI) found 950.4940, calcd for $C_{50}H_{63}BF_2N_9O_7$ (MH$^+$): 950.4906.

2.6 REFERENCES

1   Copeland, R. A., Pompliano, D. L. & Meek, T. D. Opinion - Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* **5**, 730-739 (2006).

2   Potashman, M. H. & Duggan, M. E. Covalent Modifiers: An Orthogonal Approach to Drug Design. *J Med Chem* **52**, 1231-1246 (2009).

3   Smith, A. J. T., Zhang, X. Y., Leach, A. G. & Houk, K. N. Beyond Picomolar Affinities: Quantitative Aspects of Noncovalent and Covalent Binding of Drugs to Proteins. *J Med Chem* **52**, 225-233 (2009).

4   Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10**, 307-317 (2011).

5   Fry, D. W. *et al.* Specific, irreversible inactivation of the epidermal growth factor receptor and erbB2, by a new class of tyrosine kinase inhibitor. *P Natl Acad Sci USA* **95**, 12022-12027 (1998).

6   Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **308**, 1318-1321 (2005).

7   Honigberg, L. A. *et al.* The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. *P Natl Acad Sci USA* **107**, 13075-13080 (2010).

8   Zhou, W. J. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070-1074 (2009).

9   Zhou, W. J. *et al.* A Structure-Guided Approach to Creating Covalent FGFR Inhibitors. *Chem Biol* **17**, 285-295 (2010).

10   Hagel, M. *et al.* Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine. *Nat Chem Biol* **7**, 22-24 (2011).

11   Leproult, E., Barluenga, S., Moras, D., Wurtz, J. M. & Winssinger, N. Cysteine Mapping in Conformationally Distinct Kinase Nucleotide Binding Sites: Application to the Design of Selective Covalent Inhibitors. *J Med Chem* **54**, 1347-1355 (2011).

12   Zhang, J. M., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* **9**, 28-39 (2009).

13   Wissner, A. *et al.* Synthesis and structure-activity relationships of 6,7-disubstituted 4-anilinoquinoline-3-carbonitriles. The design of an orally active, irreversible inhibitor of the tyrosine kinase activity of the epidermal growth factor receptor (EGFR) and the human epidermal growth factor receptor-2 (HER-2). *J Med Chem* **46**, 49-63 (2003).

14   Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790-797 (2010).

15   Uetrecht, J. Idiosyncratic drug reactions: Past, present, and future. *Chem Res Toxicol* **21**, 84-92 (2008).

16   Evans, D. C., Watt, A. P., Nicoll-Griffith, D. A. & Baillie, T. A. Drug-protein adducts: An industry perspective on minimizing the potential for drug bioactivation in drug discovery and development. *Chem Res Toxicol* **17**, 3-16 (2004).

17   Park, B. K. *et al.* Managing the challenge of chemically reactive metabolites in drug development. *Nat Rev Drug Discov* **10**, 292-306 (2011).

18   Lee, G. *et al.* Novel inhibitors of hepatitis C virus RNA-dependent RNA polymerases. *J Mol Biol* **357**, 1051-1057 (2006).

19   Patch, R. J. *et al.* Identification of Diaryl Ether-Based Ligands for Estrogen-Related Receptor alpha as Potential Antidiabetic Agents. *J Med Chem* **54**, 788-808 (2011).

20   Pritchard, R. B., Lough, C. E., Currie, D. J. & Holmes, H. L. Equilibrium Reactions of N-Butanethiol with Some Conjugated Heteroenoid Compounds. *Can J Chemistry* **46**, 775-781 (1968).

21   Pearson, R. G. & Dillon, R. L. Rates of Ionization of Pseudo Acids. IV. Relation between Rates and Equilibria. *Journal of the American Chemical Society* **75**, 2439-2443 (1953).

22   Cohen, M. S., Hadjivassiliou, H. & Taunton, J. A clickable inhibitor reveals context-dependent autoactivation of p90 RSK. *Nat Chem Biol* **3**, 156-160 (2007).

23   Fabian, M. A. *et al.* A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol* **23**, 329-336 (2005).

24   Frodin, M. & Gammeltoft, S. Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Mol Cell Endocrinol* **151**, 65-77 (1999).

25   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* **35**, 511-522 (2009).

26   Kang, S. M. *et al.* p90 ribosomal S6 kinase 2 promotes invasion and metastasis of human head and neck squamous cell carcinoma cells. *J Clin Invest* **120**, 1165-1177 (2010).

27   Smolen, G. A. *et al.* A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. *Gene Dev* **24**, 2654-2665 (2010).

28   Shaw, L. M., Rabinovitz, I., Wang, H. H. F., Toker, A. & Mercurio, A. M. Activation of phosphoinositide 3-OH kinase by the alpha 6 beta 4 integrin promotes carcinoma invasion. *Cell* **91**, 949-960 (1997).

29   Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J Appl Crystallogr* **26**, 795-800 (1993).

30   Malakhova, M. *et al.* Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2. *Nat Struct Mol Biol* **15**, 112-113 (2008).

31   Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126-2132 (2004).

32   Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213-221 (2010).

33   Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12-21 (2010).

34   Erichsen, M. N. *et al.* Structure-Activity Relationship Study of First Selective Inhibitor of Excitatory Amino Acid Transporter Subtype 1: 2-Amino-4-(4-methoxyphenyl)-7-(naphthalen-1-yl)-5-oxo-5,6,7,8-tetrahydro-4H-chromene-3-carbonitrile (UCPH-101). *J Med Chem* **53**, 7180-7191 (2010).

35   Basheer, A., Yamataka, H., Ammal, S. C. & Rappoport, Z. Enols of substituted cyanomalonamides. *J Org Chem* **72**, 5297-5312 (2007).

CHAPTER 3

Inhibitor stability and *in vivo* RSK-CTD active site occupancy

studies in mice

3.1 Abstract

Reversible covalent Michael acceptor RSK CTD inhibitors were analyzed for their stability *in vitro* as well as in cultured cells by a competitive cell based labeling assay to pre-screen compounds and select the most stable inhibitors for RSK CTD occupancy assays in mice. This study identified several RSK CTD inhibitors that have improved potency, stability and oral bioavailability over the originally reported **CN-NHiPr** analog.

3.2 Introduction

Electrophile-bearing small molecule inhibitors that can covalently modify their targets display enhanced selectivity and potency, but also have the potential to be neutralized by reaction with off-target nucleophiles. Recently, we reported the discovery of a new class or RSK2 CTD inhibitors that act in a reversible covalent manner[1]. These compounds contain a reactive Michael acceptor electrophile that forms reversible adducts with various thiols (β-mercaptoethanol (BME), glutathione (GSH) or cysteines found in proteins). We were interested in characterizing the chemical stability of the electrophiles, and more importantly whether these compounds can survive the complex environment of animal tissues and reach their target protein *in vivo*.

3.3 Results

The first generation reversible covalent RSK CTD inhibitors contained a cyanoacrylamide or cyanoacrylate groups and were synthesized by a Knoevenagel condensation reaction from a common precursor aldehyde (Scheme 3.1). Liquid chromatography mass spectrometry (LC/MS) revealed that at 37°C in phosphate buffered saline (**PBS**), the *tert*-butyl ester analog **CN-OtBu** (Figure 3.1) was completely degraded after overnight incubation, while the amide analogs **CN-NH2** and **CN-NHiPr** were more stable (Figure 3.1a and 3.1b).
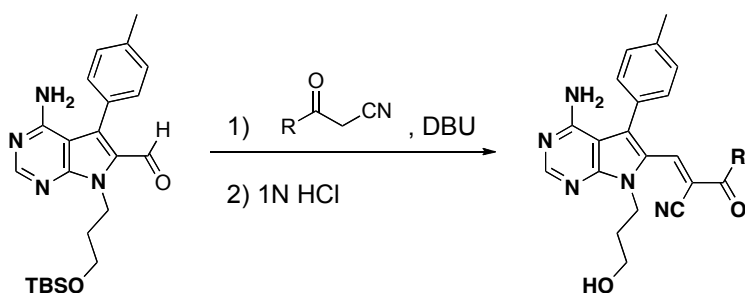
Scheme 3.1 Knoevenagel condensation for the synthesis of various Michael acceptor RSK inhibitors
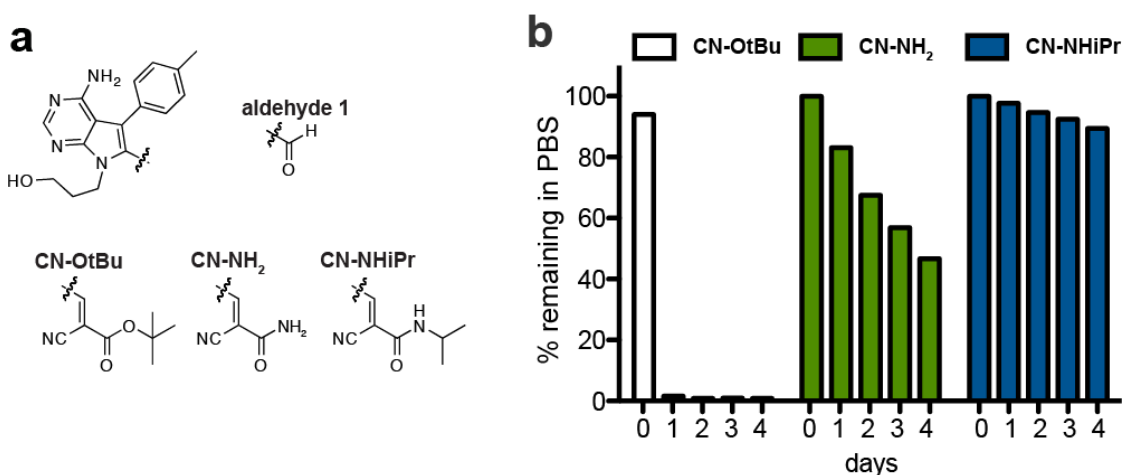
Figure 3.1. Stability of RSK CTD Michael acceptor inhibitors. a) Chemical structure of RSK inhibitors. **b)** Compound stability was assayed using LC-MS after incubating compounds (150 μM) in PBS with at 37°C for various times.

The isopropyl substituted amide CN-NHiPr was the most stable of the compounds and even after 16 days of incubation at 37°C, more than 60% remained in solution. Analysis by LC/MS revealed that the major decomposition product of these analogs was the precursor **aldehyde 1**, a product of a retro-Knoevenagel reaction (data not shown). Surprisingly, the stability trend was not correlated with cellular potency (Figure 3.2). We observed that these analogs all potently inhibited RSK2 CTD in *in vitro* kinase activity assays (IC_{50} values were between 5 – 7 nM, (REF) data not shown), but in cellular based assays, **CN-OtBu**, the least chemically stable analog, inhibited the phosphorylation of Ser380 more potently than the acrylamide analogs. **CN-NHiPr** was slightly less potent in cells, but because of its enhanced stability, we chose to move forward with this analog into more complex *in vitro* and *in vivo* active site occupancy experiments .
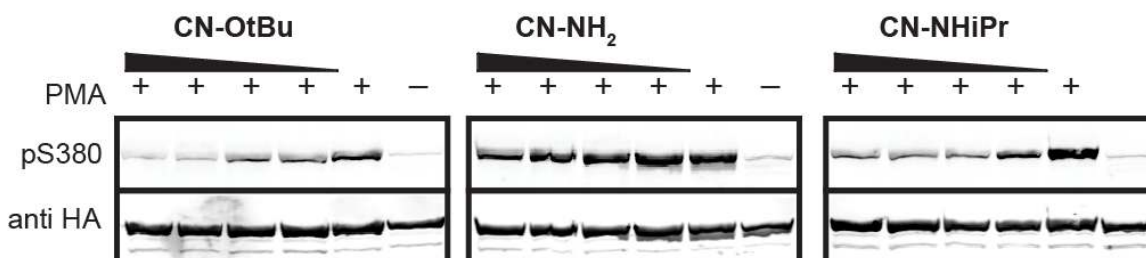
Figure 3.2 Inhibition of RSK-CTD in HEK293 cells. Cells were treated for 2 hours in serum free medium with inhibitors (1, 0.3, 0.1, 0.03 and 0 μM), stimulated with phorbol 12-myristate 13-acetate (PMA, 100 ng/mL for 30 min) and analyzed by western blot.

During the initial biochemical characterization of **CN-NHiPr**, we showed that it binds RSK CTD covalently by reacting with Cys436 (RSK2 numbering) and that the covalent bond is only stable in the context of the folded protein[1]. We were interested in measuring the rate of the covalent bond formation and used a

stopped flow apparatus to determine the kinetics of this interaction[2,3]. CN-NHiPr

has a distinct absorbance peak at 400 nm that is disrupted by the Michael

addition and loss of conjugation and we used that to measure the rate of covalent

modification with increasing amounts of RSK2 CTD protein (10 – 400 μM). A

non-linear regression fit[3] to the observed experimental data gave covalent bond

formation rate of 2.18 sec^{-1} and dissociation constant for the initial non-covalent

reversible binding (K_I) of 29 μM. (Fugure 3.3).

$$E + I \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E\cdot I \underset{k_4}{\overset{k_3}{\rightleftharpoons}} E-I$$

RSK [μM]	k_{obs} (sec^{-1})
10	0.4728
25	0.9330
50	1.3397
100	1.5830
200	1.7725
400	2.0099

k_3 = 2.1854 sec^{-1}
K_I = 29.04 μM
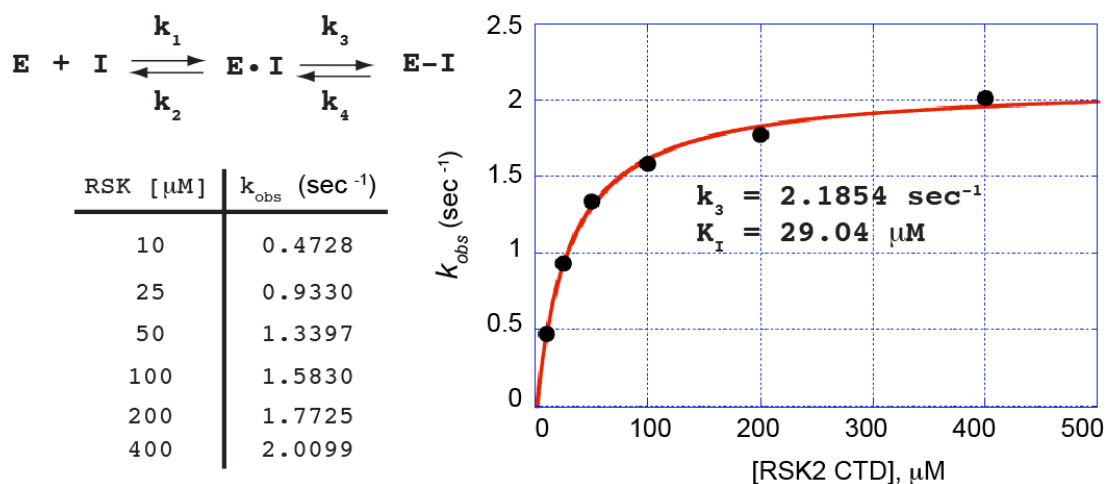
Figure 3.3. Stopped flow results. Experimental stopped flow data and non-linear regression fit for determining the rate of the covalent bond formation (k_3) between **CN-NHiPr** (10 μM) and RSK2-CTD (10 – 400 μM).
See experimental section for further details.

Next we wanted to test if the **CN-NHiPr** can inhibit RSK in mouse tissues.

We used a competition based occupancy assay[4] in which various tissue lysates

derived from mice treated with the inhibitor were labeled with a fluorescent

irreversible RSK specific probe (**FMK-BPDIPY**)[4]. We also used the irreversible

inhibitor **fmk-mea**, a very closely related analogue of **FMK** with improved

pharmacological properties, as a positive control for achieving complete inhibitor

occupancy[5]. BALB/c mice were treated (intraperitoneal injection, 100% DMSO solution) for 1 h with various concentrations of **CN-NHiPr** (100 mg/kg) or **fmk-mea** (10, 20, 40 or 60 mg/kg; see Figure 3.5 for chemical structure) and their hearts were harvested and processed. The clarified heart lysates were then treated with saturating amounts of **FMK-BODIPY**, followed by immunodepletion of the endogenous RSK1 protein. The proteins were resolved by denaturing gel electrophoresis and detected by in-gel fluorescence scanning (Figure 3.4). Quantification of the fluorescent bands relative to the vehicle control (0 mg/kg dose) revealed that even at 100 mg/kg dosing, **CN-NHiPr** cannot fully inhibit RSK1, while **fmk-mea** achieves full inhibition at about 40 mg/kg.
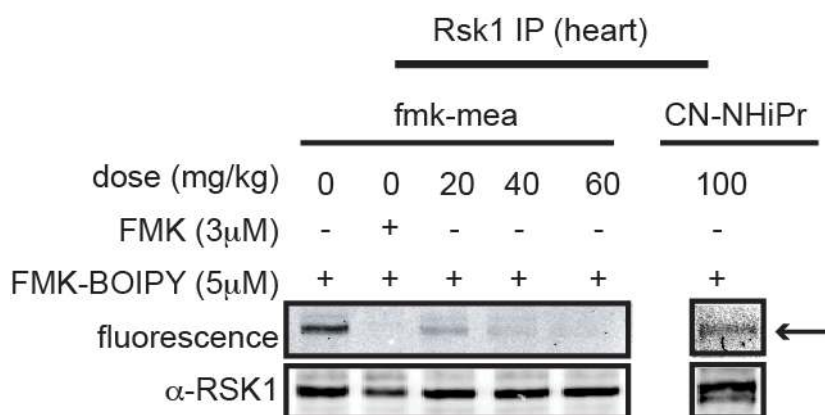
Figure 3.4. *In vivo* **RSK CTD occupancy of CN-NHiPr**. RSK CTD occupancy results from mouse experiments show complete RSK1 inhibition by **fmk-mea** at 60 mg/kg (Left panel), but not by **CN-NHiPr** at 100 mg/kg (Right panel). Mice were treated for 1 h (intraperitoneal injection) with indicated dose of inhibitors and their hearts were harvested and processed. The lysates were labeled with fluorescent **FMK-BODIPY** probe to reveal the degree of RSK CTD active site occupancy.

In efforts to improve the bioavailability of the reversible covalent RSK inhibitors, new compounds were synthesized with different Michael acceptor

electrophiles[6,7]. Some of the new analogs contained a hetero-aryl electron-withdrawing group, instead of an amide. In other analogs, N(7) of the pyrrolopyrimidine was alkylated with a cyclopropylmethyl group instead of the 3-hydroxypropyl group. A *p*-chlorophenyl substituent was also introduced instead of the *p*-tolyl group at C(5) of the pyrrolopyrimidine (Figure 3.5).
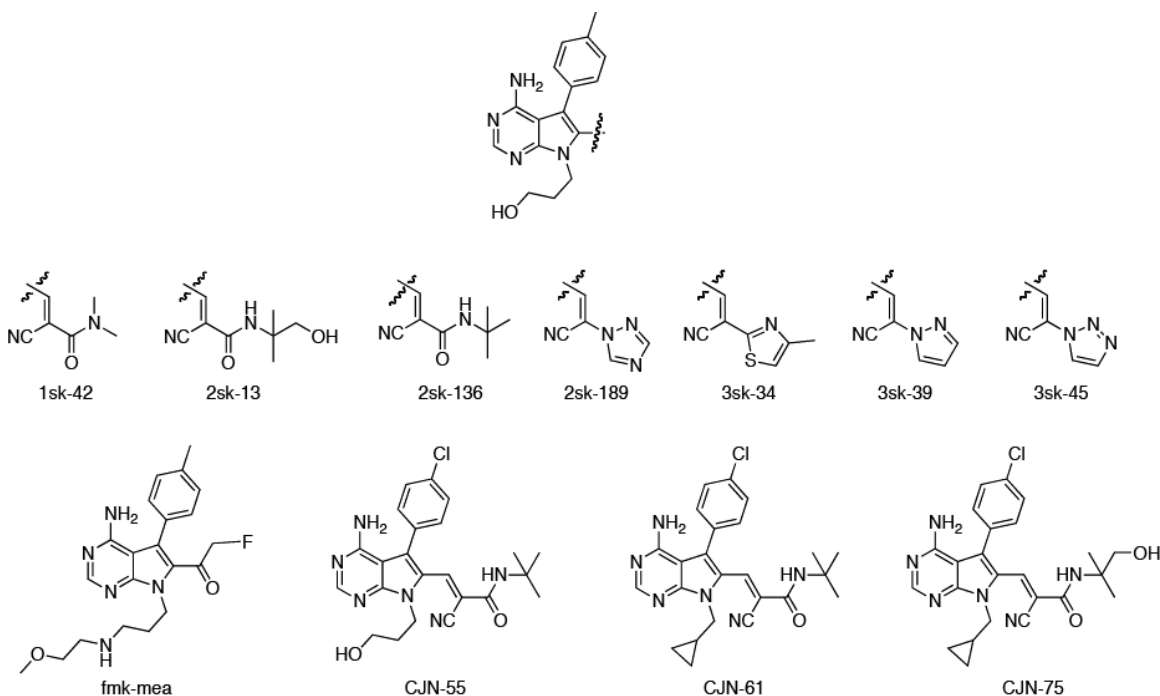

Figure 3.5. Chemical structures of RSK-CTD inhibitors (analogs prepared by Dr. Shyam Krishnan and Chris J Novotny)

We measured the active site occupancy of the new compounds with cellular RSK after a 24 h incubation using a competition-based assay. Human MDA-MB321 breast cancer cells were treated for 24 h with 1 μM of each inhibitor, then treated with saturating amount of **FMK-BODIPY**[1]. The endogenous RSK1 and RSK2 were immunoprecipitated from the clarified cellular lysates, followed by denaturing gel electrophoresis and the fluorescently labeled full

length RSK1 and RSK2 were detected by in-gel fluorescence scanning (Figure 3.6). Quantification of the fluorescent bands relative to the DMSO control revealed that **2sk-136**, **CJN-61** and **3sk-39** were as potent and stable as the irreversible **fmk-mea** in inhibiting RSK1 and RSK2 even after 24 h of incubation.
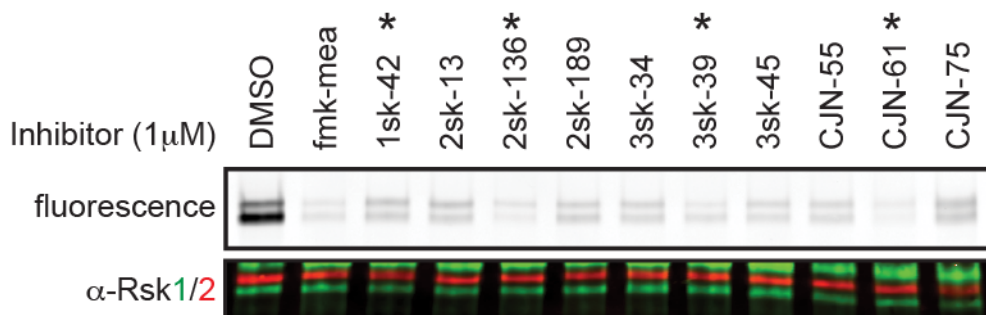
Figure 3.6. Cellular RSK CTD occupancy of various inhibitors. A competition-based assay was used to directly probe for RSK CTD active site occupancy. Intact MDA-MB231 cell were treated with indicated inhibitors (1 μM) for 24 hours in absence of serum, followed by 1 h treatment with **FMK-BODIPY** (3 μM) to reveal the degree of RSK CTD active site occupancy. Compounds marked with asterisk were selected for mouse studies.

Both **2sk-136** and **CJN-61** possess a *tert*-Butyl cyanoacrylamide Michael acceptor but different alkylations on the pyrrole nitrogen, as well as different substitutions on the benzyl group, while **3sk-39** contains a pyrazole as the electron withdrawing group. We chose these three analogs, together with **1sk-42** (the only doubly substituted amide in this series) to be tested in mouse occupancy studies.

We performed a dose response mouse occupancy experiment with **1sk-42** that was very similar to previous experiments with **FMK-mea** (Figure 3.4), but used lung tissue instead of heart (Figure 3.7a). BALB/c mice were treated (intraperitoneal injection) for 1 h with various concentrations of **1sk-42** (10, 20, 40

or 70 mg/kg). The lung tissues were processed and the clarified lysates labeled with **FMK-BODIPY.** After immunoprecipitaiton of the fluorescently labeled endogenous RSK1, we observed that only the highest dose (70 mg/kg) resulted in complete occupancy of RSK1 (Figure 3.7a). We next tested the other three analogs, **2sk-136**, **3sk-39** and **CJN-61** in the same assay using a single dose of 60 mg/kg (intraperitoneal injection). This time we immunoprecipitated both RSK1 and RSK2 to look at the isoform specific inhibition (Figure 3.7b). All three inhibitors were successful at potently inhibiting RSK1 and RSK2 with the pyrazole analog **3sk-39** being the most potent. These compounds were present in the serum after one hour, as evidenced by LC/MS (data not shown), with concentrations of 0,4 μM, 1.6 μM and 1.2 μM for **2sk-136**, **3sk-39** and **CJN-61**, respectively.
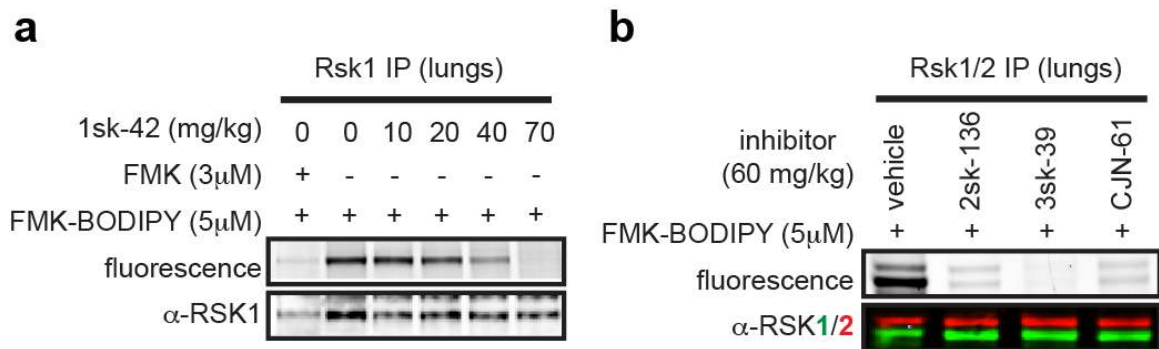
Figure 3.7. *In vivo* **RSK CTD occupancy of second generation RSK inhibitors. a)** RSK1 CTD occupancy results from mouse experiments show complete RSK1 inhibition by **1sk-42** at 70 mg/kg (intraperitoneal injection), as evidenced by a dose response in lung lysates. The vehicle lysates were pre-treated with **FMK** prior to the **FMK-BODIPY** labeling as control for maximum inhibition. **b)** RSK1 and RSK2 CTD occupancy in lung lysates from mice treated for 1 h with a single dose of inhibitors (60 mg/kg, intraperitoneal injection).

We performed a dose response with **2sk-136** and compared the results from both lung or spleen lysates. The minimum concentration needed to fully inhibit RSK in lungs was 80 mg/kg, while in spleens we observed full inhibition at significantly lower doses (40 mg/kg). This discrepancy could be due to differential expression of RSK in different tissues, as well as differential distribution of the inhibitor *in vivo*.
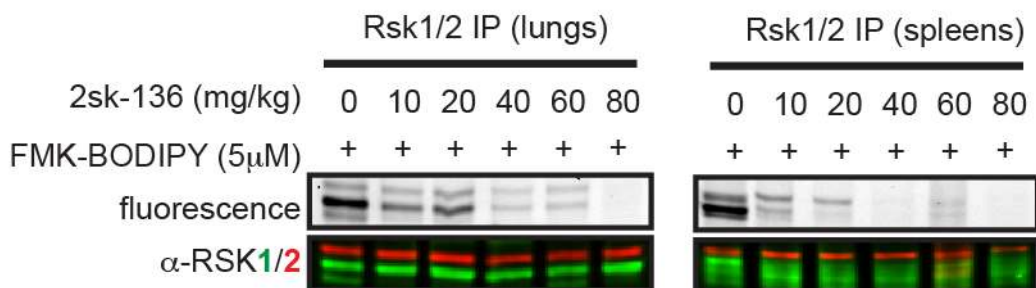

Figure 3.8. *In vivo* RSK CTD occupancy of 2sk136 in various mouse tissues. Higher dose of 2sk136 is needed to achieve full RSK inhibition in lungs (80 mg/kg) vs spleens (40 mg/kg). Mice were treated for 1 h (intraperitoneal injection) with indicated doses of **2sk136** and the lungs and spleens were harvested and processed. The clarified lysates were labeled with fluorescent **FMK-BODIPY** probe to reveal the degree of RSK CTD active site occupancy.

In order to evaluate the oral bioavailability, we performed oral gavage administration **of 2sk-136** and **CJN-61** and used the lung and spleen lysates in the occupancy assay (Figure 3.9). Complete RSK1 and RSK2 inhibition was observed when both compounds were administered at 100 mg/kg or 150 mg/kg (formulated in PHOSAL-50PG + 5% DMSO), suggesting that these scaffolds are quite stable and suitable for oral administration. Additionally, plasma analysis by LC/MS/MS detected both **2sk-136** (0.6 μM for the 100 mg/kg dose, 2.02 μM for

the 150 mg/kg dose) and **CJN-61** (4.5 μM for the 100 mg/kg dose, 2.4 μM for the 150 mg/kg dose) 1 hour after administration.
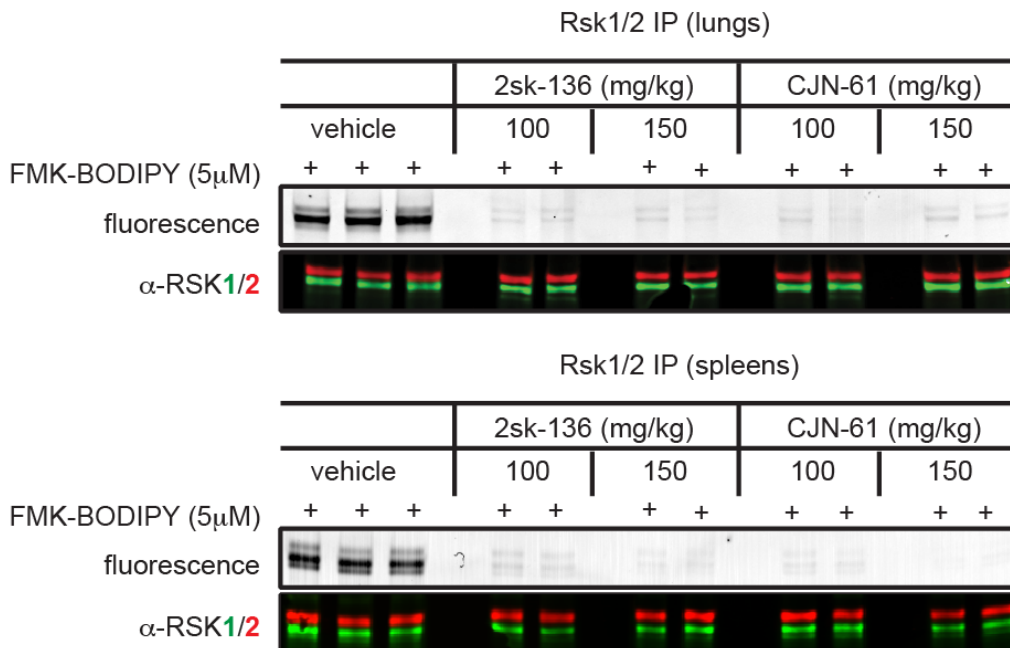

Figure 3.9. *In vivo* **RSK CTD occupancy of 2sk136 and CJN-61 after oral administration.** Mice were dosed by oral gavage with 100 mg/kg or 150 mg/kg inhibitors and their lungs and spleens were harvested 1 h post drug administration and processed. The clarified lysates were labeled with fluorescent **FMK-BODIPY** probe to reveal the degree of RSK CTD active site occupancy.

3.4 Conclusion

By changing the electron withdrawing groups of the Michael acceptor electrophile, were able to identify new RSK reversible covalent inhibitors with improved stability and oral bioavailability. A common feature of the first generation reversible covalent RSK inhibitors is that they all contain a vynilnitrile with either a carboxyl ester (CN-OtBu) or carboxamide (CN-NHiPr, CN-NH$_2$) substitution. Changing those substitutions to heterocycles or to different

substituted carboxamides afforded compounds with improved cellular stability and potency. We also tested these analogs in mice in *in vivo* RSK active site occupancy assay in order to assess their oral bioavailability. We compared two different routes of administration (intraperitoneal injection vs oral gavage) and determined that **2sk-136** can achieve complete RSK inhibition regardless of how it is administered. This finding is encouraging and suggests that inhibitors bearing Michael acceptor electrophiles can survive both the harsh acidic stomach environment as well as the high concentration of thiols in the cells and still retain potency.

3.5 Experimental

Compounds

aldehyde1, **CN-OtBu**, **CN-NH2** and **CN-NHiPr** were synthesized as previously described[1]

FMK, and **FMK-mea** were synthesized as described[4,5,8] by Michael S. Cohen.

1sk-42, **2sk-13**, **2sk-136**, **2sk-189**, **3sk-34**, **3sk-39**, **3sk-45** and **FMK-BODIPY** were synthesized as described[1,6] by Shyam Krishnan.

CJN-61 and **CJN-75** were synthesized by Chris Novotny[7].

Monitoring inhibitor stability at 37°C in PBS by LC-MS

A 150 µM solutions of different Michael acceptor RSK2-CTD inhibitors in PBS (pH 7.4) with 5% MeCN were incubated at 37°C. 100 µL aliquots were drawn daily and analyzed by LC-MS (20 µL injection, Waters XTerra MS C18 column, 25 min gradient, 5–70% MeCN/0.1% formic acid; Waters 2695 Alliance Separations Module; Waters Micromass ZQ mass spectrometer). The percent of inhibitor remaining was calculated as the ratio of the area under the peaks for the starting compound and the decomposition aldehyde product.

CN-NHiPr / RSK2 CTD kinetic measurements

Stopped-flow experiments were conducted using a Hi-Tech SF-61Dx2 instrument (TgK Scientific Ltd., Bradford on Avon, U.K.). Data were collected in dual beam mode using photomultiplier detection of absorbance at 400 nm. **CN-NHiPr** (20

μM) was mixed rapidly with a solution of RSK2 CTD (20 – 800 μM) and the disappearance of the **CN-NHiPr** was monitored over time (3 – 30 seconds). Data for **CN-NHiPr** reacting with RSK2 CTD was fit to a two-step, reversible inhibition model:

$$E \ + \ I \ \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \ E \cdot I \ \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \ E - I$$
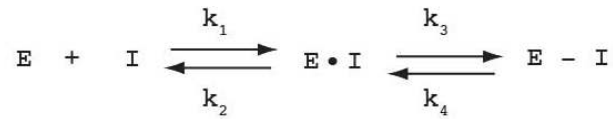
Where $K_I = k_2/k_1$ and $\ \ k_{obs} \ = \ (k_3[E] \ / \ ([E] \ + \ K_d)) \ + \ k_4$

Plots of k_{obs} vs RSK2 CTD concentration were fit using KaleidaGraph to obtain values for K_I and k_3. The value for k_4 could not be accurately determined and was assumed to be zero (calculated result for k_4 was $-0.079 \ \text{sec}^{-1}$).

RSK1 occupancy assay using heart or lung lysates from mice treated with FMK-mea ,CN-NHiPr or 1sk-42

Frozen heart or lung tissues from mice that were treated with **fmk-mea** (0, 20, 40 and 60 mg/kg), **CN-NHiPr** (0 and 100 mg/kg) or 1sk-42 (0, 10, 20, 40 and 70 mg/kg) were thawed on ice in presence of 2 mL PBS containing protease (Complete, Roche) and phosphatase (Coctail I and II, Sigma) inhibitors. The tissues were kept on ice and were homogenized three times for 30 sec at 13500 rpm with an electric tissue tearor (Janke & Kunkel/IKA T25-Ultra-Turrax Homogenizer). The homogenates were transferred to clean 4mL ultracentrifuge tubes (Beckman) and were clarified by centrifugation at 30000 x g for 60 min (Beckman Coulter Optima TLX Ultracentrifuge; TLA-100.3 rotor). Next, 250 μL each of the clarified lysates (normalized to 2.5 mg/mL) derived from the vehicle-treated mouse was pre-treated with **FMK** (3 μM) for 1 hour at room temperature,

and then reacted with **FMK-BODIPY** (5 μM) for 1 hour at room temperature. This sample was a control for the labeling specificity of **FMK-BODIPY** in the heart tissue lysates. Separately, 250 μL each of the clarified lysates (from mice treated with 0, 20, 40 and 60 mg/kg **fmk-mea**, or 0 and 100 mg/kg CN-NHiPr, or 0, 10, 20, 40 and 70 mg/kg 1sk-42; normalized to 2.5 mg/mL) were reacted with **FMK-BODIPY** (5 μM) for 1 hour at room temperature (without **FMK** pre-treatment). Following the **FMK-BODIPY** reaction, each sample (270 μL total) was diluted with 730 μL ice cold PBS + 1% NP40. Immunoprecipation was performed using 4 μg Rsk1 antibody (C-21, sc-231 rabbit polyclonal antibody, Santa Cruz Biotechnology) for 3 hr at 4°C, followed by overnight incubation at 4°C with 50 μL of Protein G Sepharose beads (GE Healthcare). The beads were washed three times for 10 min with 500 μL PBS + 1% NP40, pelleted and resuspended in 50 μL Laemmli sample buffer. The samples were resolved on 10% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by transfer to nitrocellulose membrane and western blot detection with antibodies against RSK1.

General procedure for RSK1/2 occupancy assay using lung or spleen lysates from mice treated with 2sk-136, 3sk-39 and CJN-61

Frozen mouse tissues (lungs or spleens) from mice that were treated with different doses of RSK-CTD inhibitors (0, varying mg/kg) were thawed on ice in presence of 2 mL PBS containing protease (Complete, Roche) and phosphatase (Coctail I and II, Sigma) inhibitors. The tissues were kept on ice and were

homogenized three times for 30 sec at 13500 rpm with an electric tissue tearor (Janke & Kunkel/IKA T25-Ultra-Turrax Homogenizer). The homogenates were transferred to clean 4mL ultracentrifuge tubes (Beckman) and were clarified by centrifugation at 30000 x g for 60 min (Beckman Coulter Optima TLX Ultracentrifuge; TLA-100.3 rotor). Next, 250 μL each of the clarified lysates (normalized to 2.5 mg/mL) were reacted with **FMK-BODIPY** (5 μM) for 1 hour at room temperature (without **FMK** pre-treatment). Following the **FMK-BODIPY** reaction, each sample (270 μL total) was diluted with 730 μL ice cold PBS + 1% NP40. Immunoprecipitation was performed using 4 μg Rsk2 antibody (E-1, sc-9986 mouse monoclonal antibody, Santa Cruz Biotechnology) and 4 μg Rsk1 antibody (C-21, sc-231 rabbit polyclonal antibody, Santa Cruz Biotechnology) for 3 hr at 4°C for each sample, followed by overnight incubation at 4°C with 50 μL of Protein G Sepharose beads (GE Healthcare). The beads were washed three times for 10 min with 500 μL PBS + 1% NP40, pelleted and resuspended in 50 μL Laemmli sample buffer. The samples were resolved on 10% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by transfer to nitrocellulose membrane and western blot detection with antibodies against RSK1 (C-21, Santa Cruz sc-231) and RSK2 (E-1, Santa Cruz sc-9986) antibodies. After incubation with primary antibodies, immunoblots were incubated with infrared dye-labeled secondary antibodies (IR680 or IR800) and visualized using the LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE).

Cellular RSK1/2 competition based occupancy assay

MDA-MB-231 cells were seeded into 6-well plates at 300,000 cells/well. After 48 h, the media was exchanged with serum-free DMEM, and the cells were treated for 24 h with the indicated concentrations of inhibitors, followed by 1 h incubation with 3 μM of the **FMK-BODIPY** probe. The media was aspirated and the cells were washed with 2 mL of cold PBS and lysed with 70 μL CelLytic M lysis buffer (Sigma) supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors. The lysates were cleared by centrifugation at 14K rpm for 10 min at 4°C and normalized by Bradford assay. Immunoprecipitation was performed using 4 μg Rsk2 antibody (E-1, sc-9986 mouse monoclonal antibody, Santa Cruz Biotechnology) and 4 μg Rsk1 antibody (C-21, sc-231 rabbit polyclonal antibody, Santa Cruz Biotechnology) for 3 hr at 4°C for each sample, followed by overnight incubation at 4°C with 50 μL of Protein G Sepharose beads (GE Healthcare). The beads were washed three times for 10 min with 500 μL PBS + 1% NP40, pelleted and resuspended in 50 μL Laemmli sample buffer. The samples were resolved by 7.5% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by transfer to nitrocellulose membrane and western blot analysis using RSK1 (C-21, Santa Cruz sc-231) and RSK2 (E-1, Santa Cruz sc-9986) antibodies. After incubation with primary antibodies, immunoblots were incubated with infrared dye-labeled secondary antibodies (IR680 or IR800) and visualized using the LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE).

Serum or plasma analysis by LC MS/MS

Frozen serum or plasma samples were thawed on ice. To 100 µL of each serum or plasma sample were added 200 µL acetonitrile that contains 100 µM 1,3-dinitrobenzene (internal control). The solution was mixed thoroughly and incubated at room temperature for 30 min. Following the incubation step, the samples were centrifuged at 20000 x g to pellet the precipitate. Then, 75 µL of the clarified samples were transferred to LCMS glass vials and mixed with 25 µL water. The final samples (100 µL volume, 50 % Acetonitrile) were analyzed on Waters Acquity Ultra Performance LC MS/MS (4 µL injection, 5 – 95 % Acetonitrile over 3.7 min, 0.6 mL/min flow rate). Standard curves were generated for each inhibitor by making known concentrations into the serum or plasma vehicle matrix (1000, 500, 250, 125, 12,5, 6.25 and 1.25 nM). Then each standard concentration was mixed with equal volume of acetonitrile (containing 100 µM 1,3-dinitrobenzene as internal control) for final concentration of 50% acetonitrile.

3.6 REFERENCES

1   Serafimova, I. M. *et al.* Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol* 8, 471-476 (2012).

2   Harvey, R. A. & Borcherdt, W. O. Variable-ratio stopped-flow mixing device. *Anal Chem* 44, 1926-1928 (1972).

3   Tummino, P. J. & Copeland, R. A. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* 47, 5481-5492 (2008).

4   Cohen, M. S., Hadjivassiliou, H. & Taunton, J. A clickable inhibitor reveals context-dependent autoactivation of p90 RSK. *Nat Chem Biol* 3, 156-160 (2007).

5   Cohen, M. S. *Rational design of selective, irreversible kinase inhibitors: a structural bioinformatics approach* PhD thesis, University of California, (2006).

6   Krishnan, S. *Unpublished work* (2011).

7   Novotny, C. *Unpublished work* (2011).

8   Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* 308, 1318-1321 (2005).

CHAPTER 4

Molecular basis for RSK isoform selectivity of FMK

4.1 Abstract

Labeling studies with **FMK** have showed that it inhibits the RSK1 C-terminal kinase domain (CTD) more potently compared to the RSK2 CTD. We used **FMK** as a chemical probe for elucidating the differences between the two isoforms that are responsible for the observed isoform selectivity. Sequence alignment revealed slight differences in the glycine rich loop region that is adjacent to the nucleophilic cysteine residue that is covalently modified by **FMK**. Mutagenesis studies and biochemical assays confirmed that small changes to the cysteine thiol environment could have profound effects on its reactivity with **FMK**. We demonstrate that a faster rate of covalent bond formation contributes to the observed selectivity of **FMK** for RSK1 CTD.

4.2 Introduction

The human p90 ribosomal S6 kinase (RSK) family of serine/threonine protein kinases is comprised of four members: RSK1 – 4, which that are distinct gene products[1,2]. RSKs are downstream of the Ras-mitogen activated protein kinase (MAPK) signaling pathway and regulate various cellular processes such as cell growth, proliferation and survival (Figure 4.1a)[3,4]. A characteristic feature of the RSK protein kinases is the presence of two separate and functional kinase domains that are connected by a hydrophobic linker region[5,6] (Figure 4.1b). The N-terminal kinase domain (NTD) is responsible for phosphorylating RSK's cellular substrates, while the only known function of the C-terminal kinase

domain (CTD) is to phosphorylate a serine residue on the hydrophobic linker region, which leads to the subsequent activation of the NTD.
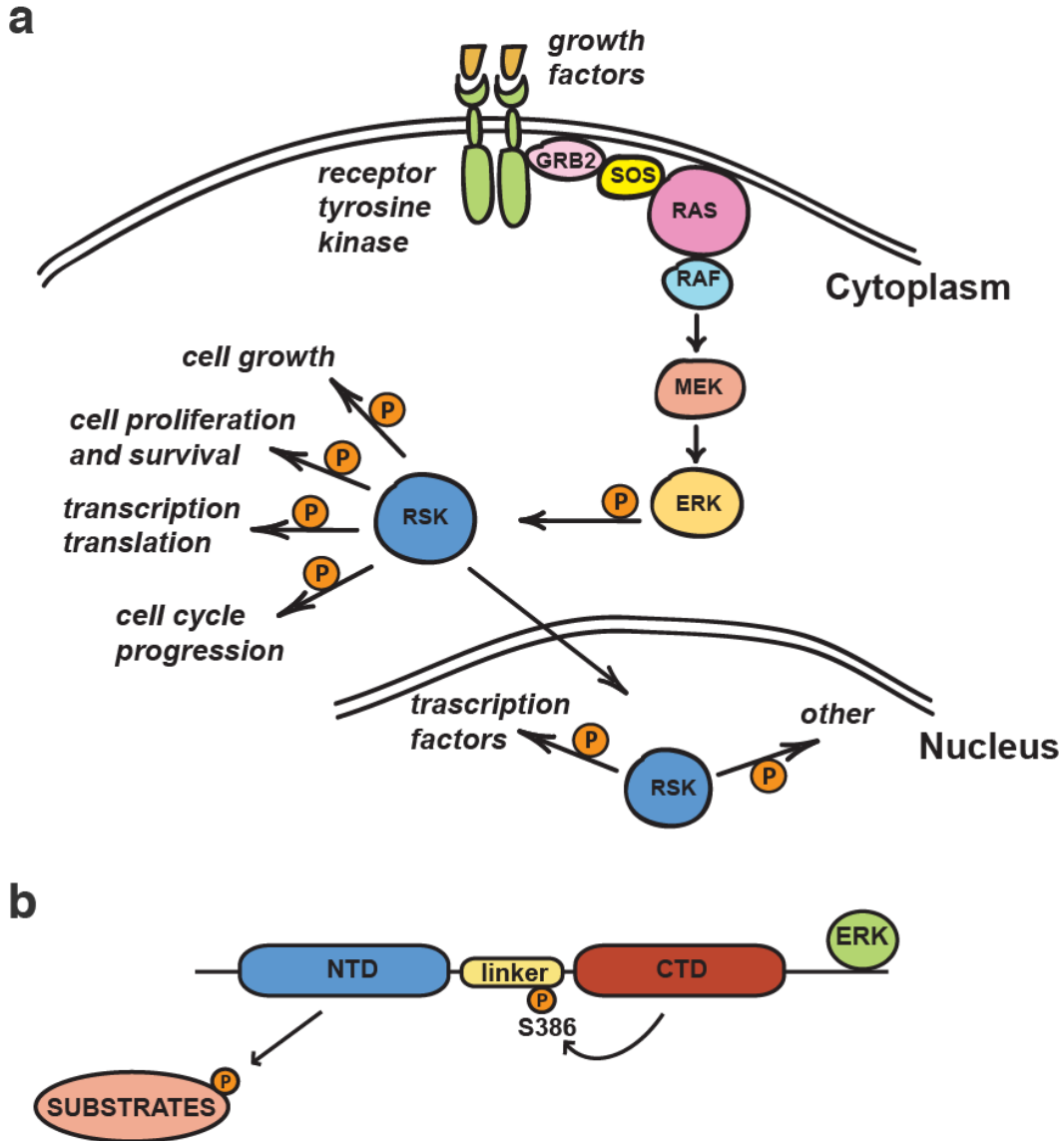
Figure 4.1. MAPK signaling pathway and RSK a) RSK activated by ERK in response to extra cellular stimuli **b)** Schematic representation of RSK depicting the two distinct kinase domains connected by hydrophobic linker region

Misregulation of RSK activity has been linked to cancer cell survival and metastasis[7-9] but there is little evidence describing the individual contribution of

each of the RSK isoforms to these processes. RNA interference has most commonly been used for elucidating the individual roles of the RSK isoforms in cells by affording selective ablation of gene expression. Although this powerful technique can provide valuable information, it can be a challenging strategy with higher model organisms[10]. Therefore, there is an increased need for the development of isoform specific, cell permeable small molecule RSK inhibitors that can be used for understanding the physiological roles of the different RSK isoforms[11]. In order to rationally design isoform specific inhibitors, we sought to understand the structural and functional differences between the RSK1 and RSK2 proteins and used the irreversible inhibitor **FMK** as a reporter probe[12]. The RSK proteins share a large degree of sequence identity, especially in their C-terminal kinase domains (CTDs) (Figure 4.2). In particular, the CTDs of RSK1 and RSK2 (human) are 88% identical. Despite this high degree of sequence similarity, the small molecule inhibitor **FMK**, which was designed to inactivate the C-terminal kinase domains of RSK1, 2 and 4, revealed an unexpected selectivity for RSK1[13].

```
RSK1    1  MPLAQLKEPWPLMELVPLD--PENGQ-----TSGEEAGLQPSKDEGVLKEISITHHVKAG
RSK2    1  MPLAQLADPWQKMAVESPSDSAENGQQIMDEPMGEEE-INPQTEEGSIKEIAITHHVKEG
           ****** :**   * :  .  .****   . ***  ::*..:** :***:****** *

RSK1   54  SEKADPSHFELLKVLGQGSFGKVFLVRKVTRPDSGHLYAMKVLKKATLKVRDRVRTKMER
RSK2   60  HEKADPSQFELLKVLGQGSFGKVFLVKKISGSDARQLYAMKVLKKATLKVRDRVRTKMER
           ******:*********************:*:: .*: :**********************

RSK1  114  DILADVNHPFVVKLHYAFQTEGKLYLILDFLRGGDLFTRLSKEVMFTEEDVKFYLAELAL
RSK2  120  DILVEVNHPFIVKLHYAFQTEGKLYLILDFLRGGDLFTRLSKEVMFTEEDVKFYLAELAL
           ***.:*****:*************************************************

RSK1  174  GLDHLHSLGIIYRDLKPENILLDEEGHIKLTDFGLSKEAIDHEKKAYSFCGTVEYMAPEV
RSK2  180  ALDHLHSLGIIYRDLKPENILLDEEGHIKLTDFGLSKESIDHEKKAYSFCGTVEYMAPEV
           .*************************************:**********************

RSK1  234  VNRQGHSHSADWWSYGVLMFEMLTGSLPFQGKDRKETMTLILKAKLGMPQFLSTEAQSLL
RSK2  240  VNRRGHTQSADWWSFGVLMFEMLTGTLPFQGKDRKETMTMILKAKLGMPQFLSPEAQSLL
           ***:**:*******:**********:*************:************.*******

RSK1  294  RALFKRNPANRLGSGPDGAEEIKRHVFYSTIDWNKLYRREIKPPFKPAVAQPDDTFYFDT
RSK2  300  RMLFKRNPANRLGAGPDGVEEIKRHSFFSTIDWNKLYRREIHPPFKPATGRPEDTFYFDP
           * ************:****.****** *:************:******..:*:*******.

RSK1  354  EFTSRTPKDSPGIPPSAGAHQLFRGFSFVATGLMEDDGKPRAPQAPLHSVVQQLHGKNLV
RSK2  360  EFTAKTPKDSPGIPPSANAHQLFRGFSFVA--ITSDDESQAMQTVGVHSIVQQLHRNSIQ
           ***:: :********** .************   : .**  .   . :**:***** :.:

RSK1  414  FSDGYVVKETIGVGSYSECKRCVHKATNMEYAVKVIDKSKRDPSEEIEILLRYGQHPNII
RSK2  418  FTDGYEVKEDIGVGSYSVCKRCIHKATNMEFAVKIIDKSKRDPTEEIEILLRYGQHPNII
           *:*** *** ******* ****:*******:***:*******:****************

RSK1  474  TLKDVYDDGKHVYLVTELMRGGELLDKILRQKFFSEREASFVLHTIGKTVEYLHSQGVVH
RSK2  478  TLKDVYDDGKYVYVVTELMKGGELLDKILRQKFFSEREASAVLFTITKTVEYLHAQGVVH
           **********:**:*****:********************.**.** *******:*****

RSK1  534  RDLKPSNILYVDESGNPECLRICDFGFAKQLRAENGLLMTPCYTANFVAPEVLKRQGYDE
RSK2  538  RDLKPSNILYVDESGNPESIRICDFGFAKQLRAENGLLMTPCYTANFVAPEVLKRQGYDA
           ******************.:*****************************************

RSK1  594  GCDIWSLGILLYTMLAGYTPFANGPSDTPEEILTRIGSGKFTLSGGNWNTVSETAKDLVS
RSK2  598  ACDIWSLGVLLYTMLTGYTPFANGPDDTPEEILARIGSGKFSLSGGYWNSVSDTAKDLVS
           .*******:******:*********.*******:*******:**** **:**:*******

RSK1  654  KMLHVDPHQRLTAKQVLQHPWVTQKDKLPQSQLSHQDL-QLVKGAMAATYSALNSSKPTP
RSK2  658  KMLHVDPHQRLTAALVLRHPWIVHWDQLPQYQLNRQDAPHLVKGAMAATYSALNRNQ-SP
           *************  **:***:.: *:*** **.:**   :************** .: :*

RSK1  713  QLKPIESSILAQRR-VRKLPSTTL  735
RSK2  717  VLEPVGRSTLAQRRGIKKITSTAL  740
             *:*:   * ***** ::*:.**:*
```

Figure 4.2. Sequence alignment of full length human RSK1 and RSK2. The N-terminal and C-terminal kinase domains are highlighted in blue and red, respectively.

4.3 Results

Isoform selectivity of FMK

In various assays, we have observed that RSK1 is more potently inhibited by the irreversible compound **FMK** (Figure 4.3), suggesting that selective isoform inhibition is possible[13,14].
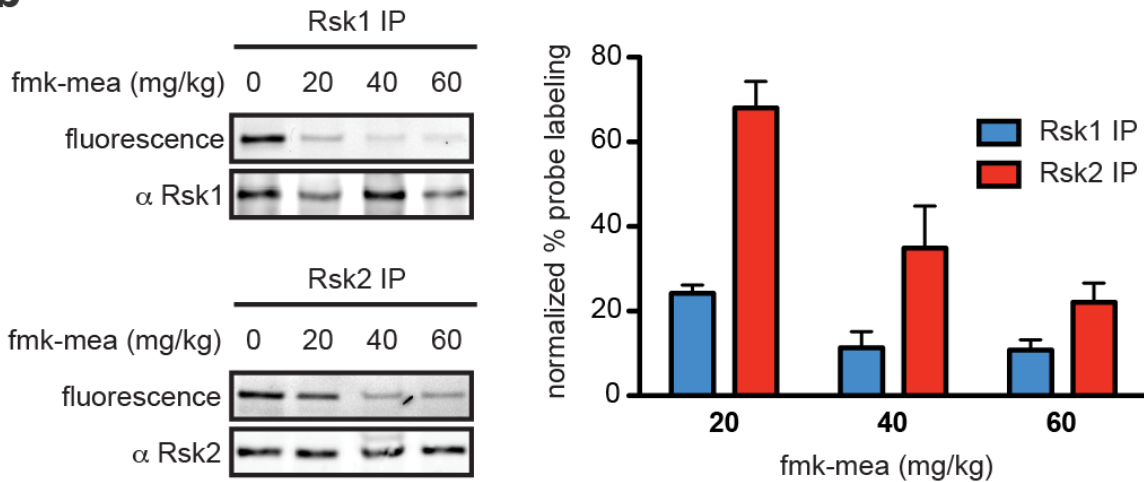
a
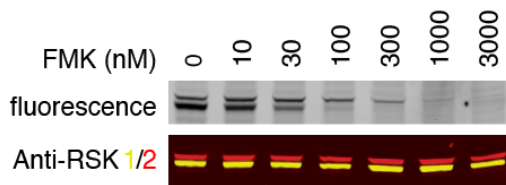
b

Figure 4.3. Cell and mouse based RSK CTD occupancy experiments showing FMK selectivity for RSK1. a) A competition-based assay was used to directly probe for RSK CTD active site occupancy at different doses of **FMK**.. Human MDA-MB321 cells were treated with increasing concentrations of **FMK** for 1 h, followed by 1 hour treatment with **FMK-BODIPY** (3 μM). The clarified cellular lysates were separated by denaturing gel electrophoresis and the fluorescently labeled full length RSK1 and RSK2 were detected by in-gel fluorescence scanning and western blot. The top fluorescence band is RSK2 and bottom band is RSK1. **b)** Immunoprecipitation from heart lysates from mice treated with **fmk-mea** and analyzed by fluorescence scanning. Quantified data is on the right.

In a competition-based assay that directly probes for inhibitor occupancy in cells, the concentration of **FMK** that was required to achieve maximum inhibition of RSK2 was almost 10-fold higher compared to the concentration needed to maximally inhibit RSK1. Human MDA-MB321 breast cancer cells were treated with increasing concentrations of **FMK**, followed by a treatment with saturating amounts of a fluorescent irreversible probe (**FMK-BODIPY**). The clarified cellular lysates were separated by denaturing gel electrophoresis and the fluorescently labeled full length RSK1 and RSK2 were detected by in-gel fluorescence scanning. Quantification of the fluorescent bands relative to the DMSO control revealed that **FMK** displayed maximum inhibition of RSK1 at about 100 nM, while the concentration required for maximum inhibition for RSK2 was between 1000 and 3000 nM (Figure 4.3a).

A similar trend was observed in a competition assay using lysates from mice that were treated with **fmk-mea**, a very closely related analogue of **FMK** with improved pharmacological properties (see Chapter 3 for structure). BALB/c mice were treated (intraperitoneal injection) with various concentrations of **fmk-mea** and their hearts were harvested and processed. Clarified heart lysates were then treated with saturating amounts of **FMK-BODIPY**, followed by immunodepletion of the endogenous RSK1 and RSK2 proteins. The immunodepleted proteins were again resolved by denaturing gel electrophoresis and detected by in-gel fluorescence scanning (Figure 4.3b). For all the doses that were tested, RSK1 was more sensitive to **fmk-mea** inhibition compared to RSK2. However, we observed an inverse relationship between the dose of **fmk-mea**

and the degrees of isoform selectivity, i.e. the higher the **fmk-mea** amount, the less selective the inhibition. At the lowest administered dose (20 mg/kg), RSK1 was ~ 75% inhibited, while RSK2 was only ~ 25%. This difference was much smaller at the highest dose (60 mg/kg), where inhibition of RSK1 and RSK2 was~ 85% and ~70%, respectively. This result suggested that with optimized dosage of the inhibitor, isoform-selective RSK inhibition can be achieved.

Apparent rate of covalent bond formation contributes to the isoform selectivity of FMK

Endogenous RSKs have two kinase domains[5], and **FMK** was rationally designed to selectively inhibit the C-terminal kinase domains of RSK1, RSK2 and RSK4[12]. We wanted to understand the basis for the isoform selectivity observed in cell-based and mouse experiments and therefore developed *in vitro* assays that use purified C-terminal kinase domains of RSK1 and RSK2.

In a kinase activity assay, **FMK** proved more potent against RSK1-CTD than RSK2-CTD, with observed IC_{50} values of 2.7 nM versus 17.9 nM, respectively (Figure 4.4a). Using purified RSK-CTDs that have been pre-activated with ERK (REF), we monitored the degree of phosphorylation of a peptide substrate (CTDtide) derived from the sequence of the hydrophobic motif linker region[3]. These experiments showed that **FMK** selectivity resides in the highly homologous kinase domains, yet they did not provide us with information on how **FMK** is differently interacting with the two isoforms.
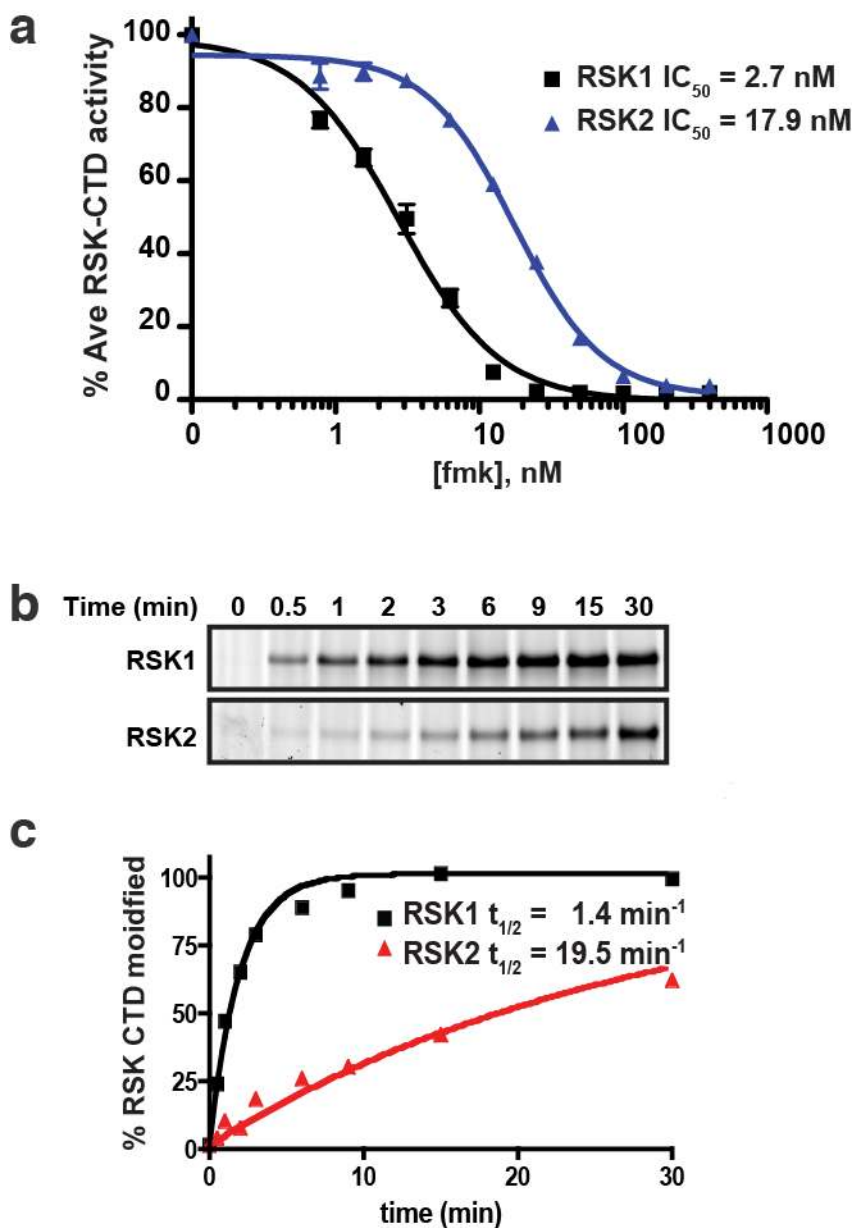
Figure 4.4 *In vitro* **selectivity of FMK for RSK1 over RSK2. a)** Dose response curves for **FMK** from *in vitro* RSK-CTD kinase activity assays. See experimental procedures for details. **b)** Kinetic analysis of **fmk-pa** labeling of RSK1 and RSK2 CTDs. 50 nM RSK CTD were labeled with 150 nM **fmk-pa**, and quenched at various times. The reactions were subjected to click chemistry with TAMRA-N_3 and analyzed by in-gel fluorescence scanning. This experiment was performed by Dr. Ville Paavilainen. **c)** Graphical representation of the data from **b**.

Given that **FMK** is an irreversible inhibitor and forms a covalent bond with Cys436 (RSK2 numbering), we wanted to investigate whether differences in the apparent rate of the covalent adduct formation were responsible for the observed selectivity. We used **fmk-pa**[13] in a two-step labeling approach in order to carry out a kinetic analysis with purified RSK-CTDs *in vitro*[15]. The RSK-CTDs (50 nM) were treated with excess amounts of **fmk-pa** (150 nM) and the reactions were quenched at different time intervals using SDS. An azide fluorescent reporter (TAMRA-N3) was conjugated to **fmk-pa** using click chemistry and the degree of **FMK**-RSK adduct formation was analyzed by in-gel fluorescence scanning (Figure 4.4b). Consistent with previous results, RSK1-CTD was modified much faster compared to RSK2-CTD, indicating that the apparent rate of the RSK1-**FMK** covalent adduct formation is faster. Since we used an end-point detection method, we were not able to uncouple the individual contributions of the reversible affinity for **FMK** from the rate of the covalent bond formation.

Glycine rich loop sequences affect FMK selectivity

To determine the RSK CTD/**FMK** binding mode we solved the crystal structure of RSK2 bound to **FMK** to 2.1Å[15]. Careful examination of the RSK2 CTD residues that are in direct contact with **FMK** revealed that the analogous residues in RSK1 are identical and could not explain the observed selectivity. We examined the sequences of the two RSK CTD isoforms more closely and identified three residues within the Glycine rich loop that were different[13] (Figure 4.5a). The glycine rich loop forms the "ceiling" of the ATP binding pocket and

contains a highly conserved sequence motif (G-X-G-X-Y/F-G/S). The nucleophilic Cys436 is located at the C-terminal end of this loop and we hypothesized that differences in the loop dynamics or conformational preference could affect its reactivity. Barouch-Bentov et al.[16] had identified the presence of a conserved salt bridge in the glycine rich loops of members of the Src and Abl protein kinases that can limit the loop flexibility and control catalytic activity and ATP binding. Using analogous criteria, we were able to identify potential salt bridge partners in RSK1 (E435 – K437; RSK2 numbering) and RSK2 (D427 – K437; RSK2 numbering) (Figure 4.5b).
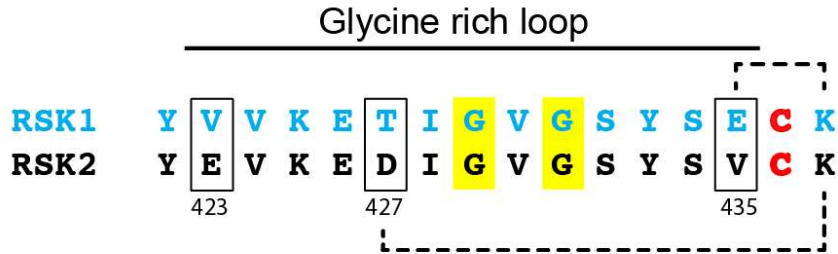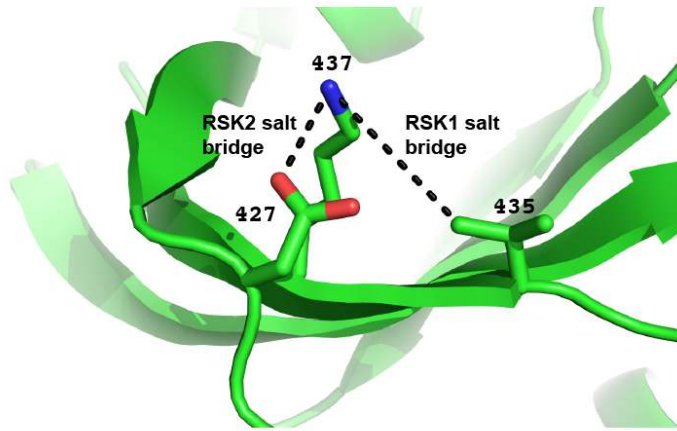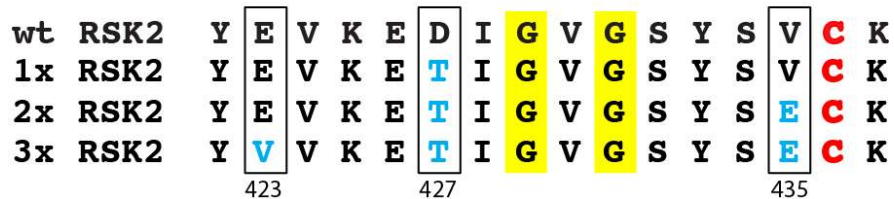
Figure 4.5. Sequence alignment of the Glycine rich loops of RSK1 and RSK2 CTDs. a) Amino acid differences between the two RSK isoforms that result in putative salt bridges (dashed lines). **b)** Crystal structure of RSK2-CTD depicting the proposed putative salt bridges. **c)** RSK1 mutations were introduced to generate three different RSK2 mutants

The putative RSK2 salt bridge is formed by side chain residues found on two adjacent β-strands (D427-K437, RSK2 numbering), while the one in RSK1 is formed by residues found on the same β-strand (E435-K437, RSK2 numbering). We hypothesized that the Cys436 reactivity with **FMK** could be influenced by the

flexibility and the dynamics of the loop and wanted to investigate the effects on mutations in this region on the **FMK** selectivity.

We designed a RSK2 triple mutant (E423V/D427T/V435E) that disrupts the putative RSK2 salt bridge and introduces the RSK1 salt bridge and tested its reactivity with **FMK** *in vitro* (Figure 4.5c). We used a mass-spectrometry based assay to detect the covalent **FMK**-adduct with the appropriate RSK-CTDs (Figure 4.6a). Using two molar equivalents of **FMK** (10 μM), we observed that the wild type RSK1 was modified with a much faster apparent rate compared to the wild type RSK2 ($t_{1/2}$ =1.8 min vs 9.4 min, respectively). The triple RSK2 mutant was also modified faster than the wild type RSK2 ($t_{1/2}$ = 4 min vs. 9.4 min, respectively), suggesting that the presence of the RSK2 putative salt bridge could have an inhibitory effect on the **FMK** adduct formation. In addition to the triple RSK2 mutant, we also generated a single mutant (D427T) that disrupts the RSK2 salt bridge, as well as a double mutant (D427T/V435E) that introduces the RSK1 salt bridge. Although all mutants were expressed and purified successfully (Figure 4.6b, coomassie insert), the single mutant could not be used in the mass spectrometry assay (no signal was observed) and thus did not reveal what the effects on **FMK** binding would be upon disruption of the RSK2 salt bridge alone. Additional biochemical assays revealed that the single mutant does not bind the fluourescent analog **FMK-BODIPY** as efficiently as the other two mutants, despite being active in *in-vitro* kinase assay (data not shown). This discrepancy remains to be investigated in future experiments.
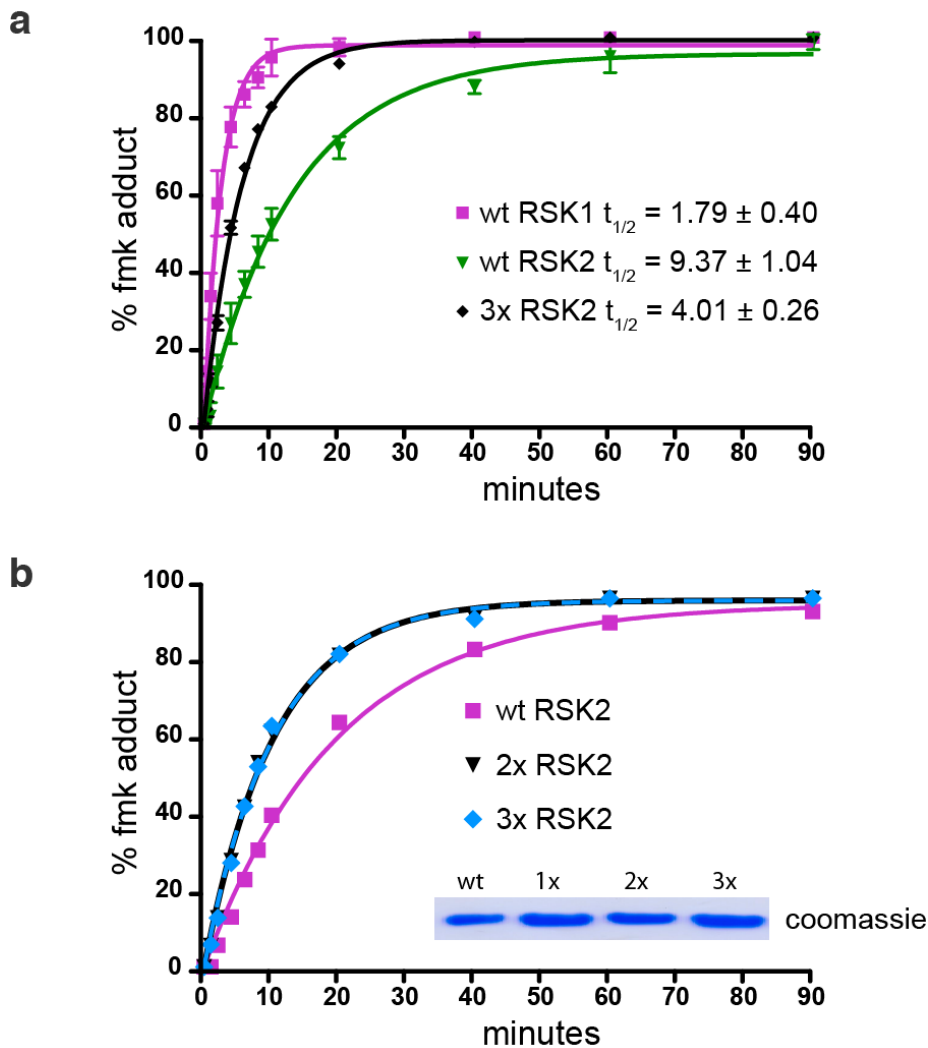
Figure 4.6. Time dependent covalent modification of RSK-CTDs by FMK. a) Progress curves of the **FMK**-RSK CTD adduct formation of wt RSK1, wt RSK2 and the triple mutant (3x) RSK2. **b)** Progress curves of the **FMK**-RSK CTD adduct formation of the different RSK2 CTD mutants (1x, 2x, and 3x) and wt RSK2 CTD. Coomassie insert shows amounts of protein used in the reactions.

The double RSK2 CTD mutant, on the other hand, formed the covalent adduct in a nearly identical manner as the triple mutant. This observation suggest that the local environment produced by the presence of the RSK1 putative salt bridge favors covalent bond formation between the electrophilic **FMK** and the

nucleophilic cysteine thiol. However, more studies are necessary to fully elucidate the basis for the observed selectivity.

4.4 Conclusion

We have observed that the basis for the isoform-selective inhibition of RSK1 and RSK2 is encoded in their C-terminal kinase domains. Subtle differences in the local environment of the Cys436 could affect the nucleophilicity of its side chain as well as its position in relation to the fluoromethylketone of **FMK**. We show that a putative salt bridge in the glycine rich loop in a trans configuration does affect the reaction rate. This may be due to changes in the flexibility of the loop that subsequently affect **FMK** binding to the ATP pocket, as well as the rate of the covalent bond formation. Further biophysical studies and molecular dynamics simulations designed to evaluate loop flexibility as well as cysteine thiol nucleophilicity should be performed in order illuminate any differences among the various RSK proteins and to thus guide the design of isoform-specific inhibitors.

4.5 Experimental

Compounds

FMK, **FMK-pa** and **FMK-mea** were synthesized as described (MSC Science, MSC NCB, MSC Thesis) by Michael S. Cohen.

FMK-BODIPY was synthesized as described (IMS NCB) by Shyam Krishnan.

Cellular RSK1/2 competition based occupancy assay

MDA-MB-231 cells were seeded into 6-well plates at 300,000 cells/well. After 48 h, the media was exchanged with serum-free DMEM, and the cells were treated for 2 h with the indicated concentrations of **FMK**, followed by 1 h incubation with 3 μM of the **FMK-BODIPY** probe. The media was aspirated and the cells were washed with 2 mL of cold PBS and lysed with 70 μL CelLytic M lysis buffer (Sigma) supplemented with protease (Complete, Roche) and phosphatase (PhoStop, Roche) inhibitors. The lysates were cleared by centrifugation at 14K rpm for 10 min at 4°C and normalized by Bradford assay. Laemmli sample buffer was added to the lysates and the proteins were separated by 7.5% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by immunoblot analysis using RSK1 (C-21, Santa Cruz sc-231) and RSK2 (E-1, Santa Cruz sc-9986) antibodies. After incubation with primary antibodies, immunoblots were incubated with infrared dye-labeled secondary antibodies (IR680 or IR800) and

visualized using the LI-COR Odyssey infrared imaging system (LI-COR

Biosciences, Lincoln, NE).

RSK1/2 occupancy assay using heart lysates from mice treated with FMK-mea

Frozen heart tissues from mice that were treated with **fmk-mea** (0, 20, 40

and 60 mg/kg) were thawed on ice in presence of 2 mL PBS containing protease

(Complete, Roche) and phosphatase (Coctail I and II, Sigma) inhibitors. The

tissues were kept on ice and were homogenized three times for 30 sec at 13500

rpm with an electric tissue tearor (Janke & Kunkel/IKA T25-Ultra-Turrax

Homogenizer). The homogenates were transferred to clean 4mL ultracentrifuge

tubes (Beckman) and were clarified by centrifugation at 30000 x g for 60 min

(Beckman Coulter Optima TLX Ultracentrifuge; TLA-100.3 rotor).

Next, 250 μL each of the clarified lysates (from mice treated with 0, 20, 40 and 60

mg/kg **fmk-mea**; normalized to 2.5 mg/mL) were reacted with **FMK-BODIPY** (5

μM) for 1 hour at room temperature (without **FMK** pre-treatment). Following the

FMK-BODIPY reaction, each sample (270 μL total) was diluted with 730 μL ice

cold PBS + 1% NP40. Immunoprecipitation was performed using 4 μg Rsk2

antibody (E-1, sc-9986 mouse monoclonal antibody, Santa Cruz Biotechnology)

for 3 hr at 4°C for each sample, followed by overnight incubation at 4°C with 50

μL of Protein G Sepharose beads (GE Healthcare). The beads were washed

three times for 10 min with 500 μL PBS + 1% NP40, pelleted and resuspended in

50 μL Laemmli sample buffer. A second immunoprecipitation step was performed

with the leftover lysates using 4 μg Rsk1 antibody (C-21, sc-231 rabbit polyclonal antibody, Santa Cruz Biotechnology) for 3 hr at 4°C, followed by overnight incubation at 4°C with 50 μL of Protein G Sepharose beads (GE Healthcare). The beads were washed three times for 10 min with 500 μL PBS + 1% NP40, pelleted and resuspended in 50 μL Laemmli sample buffer. The samples were resolved on 10% SDS-PAGE and detected by in-gel fluorescence scanning with a Typhoon 9400 flatbed laser-induced scanner, followed by transfer to nitrocellulose membrane and western blot detection with antibodies against RSK1 and RSK2.

Protein Expression and Purification

Wild type RSK1 CTD (human RSK2 396-735) was subcloned into pET-46 Ek/LIC vector following the manufacturers protocol (EMD Chemicals, Inc.) and then expressed in E.coli BL21-RIL competent cells. Cells were lysed in lysis buffer (50 mM Tris pH 8.0, 0.5 M NaCl, 10% glycerol, 15 mM imidazole) using a homogenizer (EmulsiFlex-C5, Avestin) operated at 15000 psi at 4°C for 15 min (continuous flow lysis). Soluble His-tagged RSK1-CTD was purified by Ni/NTA affinity chromatography (50 mM Tris pH 8.0, 0.5 M NaCl, 10 mM imidazole) using a 10 mL column at 2 mL/min flow rate with gradient elution (10–500 mM imidazole), followed by a second purification using HiLoad 16/60 Superdex-75 size exclusion chromatography (20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT) where the protein eluted as a single peak at 55 mL, corresponding to a monomer. Purified protein was concentrated and flash frozen in liquid nitrogen in

20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT, 12% glycerol and stored at -80°C. Protein concentrations were determined using the calculated extinction coefficients at 280 nm measured with a NanoDrop1000 spectrophotometer (Thermo Scientific).

Wild type and various mutants of RSK2-CTD (mouse RSK2 399-740) were expressed in E. coli strain BL21 (DE3)-RIL (pET-46 Ek/LIC His6-fusion vector was kindly provided by M. Malakhova and Zigang Dong, University of Minnesota). Cells were lysed in lysis buffer (50 mM Tris pH 8.0, 0.5 M NaCl, 10% glycerol, 15 mM imidazole) using a homogenizer (EmulsiFlex-C5, Avestin) operated at 15000 psi at 4°C for 15 min (continuous flow lysis). Soluble His-tagged RSK2-CTD was purified by Ni/NTA affinity chromatography (50 mM Tris pH 8.0, 0.5 M NaCl, 10 mM imidazole) using a 10 mL column at 2 mL/min flow rate with gradient elution (10–500 mM imidazole), followed by a second purification using HiLoad 16/60 Superdex-75 size exclusion chromatography (20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT) where the protein eluted as a single peak at 55 mL, corresponding to a monomer. Purified protein was concentrated and flash frozen in liquid nitrogen in 20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT, 12% glycerol and stored at -80°C. Protein concentrations were determined using the calculated extinction coefficients at 280 nm measured with a NanoDrop1000 spectrophotometer (Thermo Scientific).

In vitro kinase assays

WT RSK1-CTD or RSK2-CTD (10 μM) were first activated by His_6-ERK2

(10 μM in 20 mM Hepes pH 8.0, 10 mM $MgCl_2$, 2.5 mM TCEP, 0.2 mg/mL BSA, 200 μM ATP) for 30 min at room temp. Active RSK-CTD (5 nM in 20 mM Hepes pH 8.0, 10 mM $MgCl_2$, 2.5 mM TCEP, 0.25 mg/mL BSA, 100 μM ATP) was then treated with **FMK** (ten concentrations, in duplicate) for 30 min. Kinase reactions were initiated by the addition of 5 μCi of [γ-^{32}P]ATP (6000 Ci/mmol, NEN) and 167 μM peptide substrate (RRQLFRGFSFVAK, CTD-tide) and incubated at room temp for 30 min. Kinase activity was determined by spotting 5 μL of each reaction onto P81 cation exchange paper (Whatman). Each blot was washed once with 1% AcOH solution, twice with 0.1% H_3PO_4 solution, and once with MeOH (5-10 minutes per wash). Dried blots were exposed for 30 min to a storage phosphor screen and scanned by a Typhoon imager (GE Life Sciences). The data were quantified using ImageQuant 5.2 software and fit using PRISM 4.0.

Monitoring RSK-CTD modification by fmk-pa by in-gel fluorescence

RSK-CTD (50 nM) was treated with 150,nM **fmk-pa** on ice in a 500 μl volume in pH 7.4 PBS supplemented with 0.13 mg/ml BSA. Reactions were started by addition of **fmk-pa** and time points were quenched at different times by addition of SDS to a final concentration of 2%. Click reactions with TAMRA-N_3 were performed as previously described (Cohen, 2007, Nat Chem Biol). Proteins were resolved on 10% Tris-tricine gels and gels were imaged with a Typhoon imager and band intensities quantified with Image Quant software (GE Life Sciences)

Time dependent covalent modification of RSK CTDs by FMK

To bacterially expressed and purified RSK1 and RSK2 CTDs in PBS at pH 7.4 (5 μM) on ice was added **FMK** (10 μM). The reaction was kept on ice and 50 μL aliquots were removed at different time points after **FMK** addition. Each time point reaction was stopped by adding an equal volume of 0.4% formic acid, and the samples were analyzed by by liquid chromatography (Microtrap C18 Protein column [Michrom Bioresources], 5% MeCN, 0.2% formic acid, 0.25 mL/min; eluted with 95% MeCN, 0.2% formic acid) and in-line ESI mass spectrometry (LCT Premier, Waters). After charge envelope deconvolution, relative amounts of unmodified RSK1-CTD, RSK2-CTDs and the **FMK** adducts were determined by quantifying the mass peak intensities using MassLynx software (peak areas provided nearly identical results). These data (percentage of **FMK** adduct vs. time) were fit to a single exponential function using PRISM 5.0 to obtain $t_{1/2}$ values for **FMK** adduct formation rates in minutes.

4.6 REFERENCES

1   Romeo, Y., Zhang, X. & Roux, P. P. Regulation and function of the RSK family of protein kinases. *Biochem J* **441**, 553-569 (2012).

2   Frodin, M. & Gammeltoft, S. Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Mol Cell Endocrinol* **151**, 65-77 (1999).

3   Carriere, A., Ray, H., Blenis, J. & Roux, P. P. The RSK factors of activating the Ras/MAPK signaling cascade. *Front Biosci* **13**, 4258-4275 (2008).

4   Anjum, R. & Blenis, J. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* **9**, 747-758 (2008).

5   Fisher, T. L. & Blenis, J. Evidence for two catalytically active kinase domains in pp90rsk. *Mol Cell Biol* **16**, 1212-1219 (1996).

6   Chrestensen, C. A. & Sturgill, T. W. Characterization of the p90 ribosomal S6 kinase 2 carboxyl-terminal domain as a protein kinase. *J Biol Chem* **277**, 27733-27741 (2002).

7   Shimamura, A., Ballif, B. A., Richards, S. A. & Blenis, J. Rsk1 mediates a MEK-MAP kinase cell survival signal. *Curr Biol* **10**, 127-135 (2000).

8   Dehan, E. *et al.* betaTrCP- and Rsk1/2-mediated degradation of BimEL inhibits apoptosis. *Mol Cell* **33**, 109-116 (2009).

9   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* **120**, 1165-1177 (2010).

10   Weiss, W. A., Taylor, S. S. & Shokat, K. M. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nat Chem Biol* **3**, 739-744 (2007).

11   Knight, Z. A. & Shokat, K. M. Chemical genetics: where genetics and pharmacology meet. *Cell* **128**, 425-430 (2007).

12   Cohen, M. S., Zhang, C., Shokat, K. M. & Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **308**, 1318-1321 (2005).

13   Cohen, M. S. *Rational design of selective, irreversible kinase inhibitors: a structural bioinformatics approach* PhD thesis, University of California, (2006).

14   Serafimova, I. M. *et al.* Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol* **8**, 471-476 (2012).

15   Paavilainen, V. *Manuscript in progress* (2011).

16   Barouch-Bentov, R. *et al.* A conserved salt bridge in the G loop of multiple protein kinases is important for catalysis and for in vivo Lyn function. *Mol Cell* **33**, 43-52 (2009).

CHAPTER 5

Reversible covalent inhibition

5.1 Introduction

The development of new therapeutic drugs is a long and expensive process. In the last decades, the investment in drug research and development (R&D) has increased substantially, but the number of new medicines approved by the US Food and Drug Administration (FDA) has remained relatively unchanged and the attrition rates have become very high[1-3]. It takes an average of twelve years for a new compound to transition from research lab to human testing and the costs associated with this process are in the range of hundreds of million US dollars[4] . New strategies are needed to rationally design initial lead compounds with improved potency, selectivity and safety and therefore to alleviate the financial burden associated with R&D and thus allow for companies the freedom to quickly develop therapies for rare and orphaned diseases.

Most of the efforts in small molecule drug discovery have involved development of reversible inhibitors (Figure 5.1a). While many such compounds have proven successful, the identification of new small molecule drugs that can selectively and potently manipulate disease target proteins has been challenging. This is especially true in cases where the target protein is a member of a large closely related protein family (ex. kinases), or is involved in protein-protein interaction that needs to be disrupted in order for the therapeutic effect to take place.
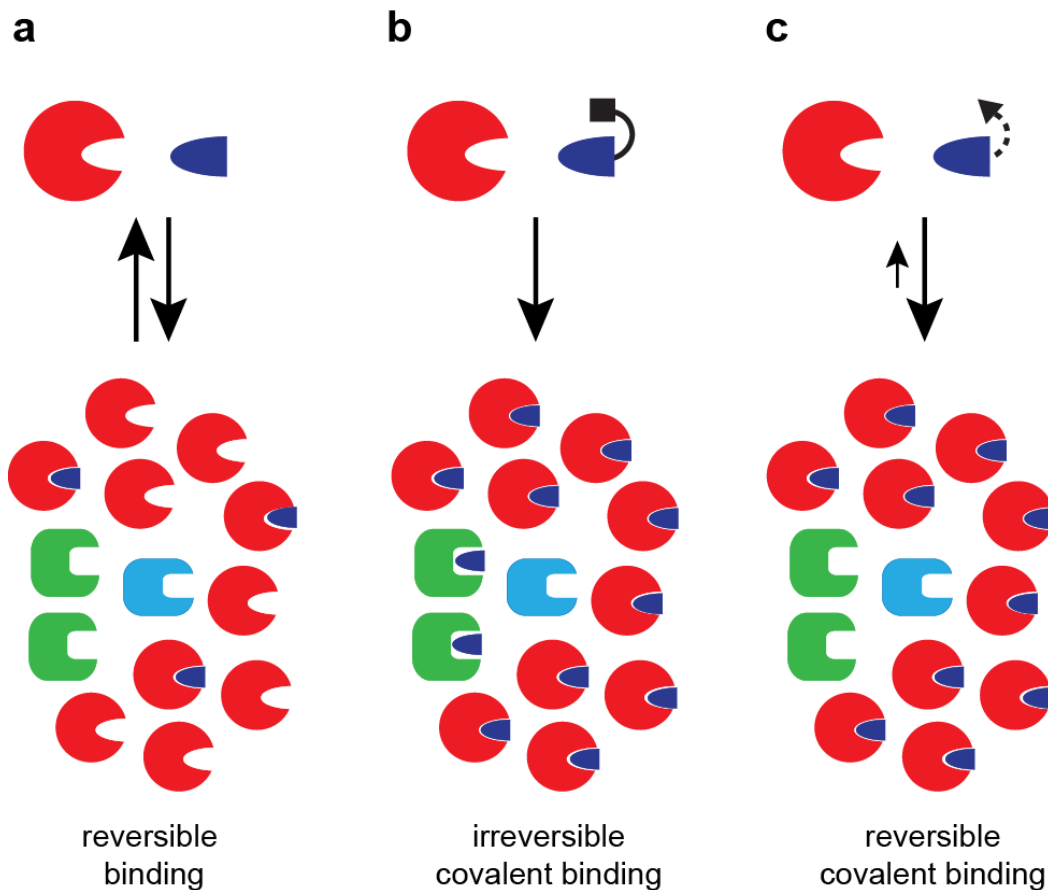
Figure 5-1. Small molecule inhibitor binding to a target protein. a) Reversible binding of ligand (blue) to its target protein (red) reaches a steady state. The potency and selectivity of the compound are driven by non-covalent interactions with the target protein's binding site. **b)** Irreversible binding of a reactive ligand (blue) to its target protein (red). A reactive electrophilic moiety (black square) is attached to the ligand so that it can covalently and irreversibly modify its target protein. Potency is greatly improved, but there is potential for covalently modifying off-target proteins (green) resulting in undesirable side effects. **c)** Reversible covalent binding. An electrophile-bearing ligand (blue) can form a stable covalent bond only in the binding site of its target protein (red). High potency is retained and off-target modification is eliminated since the electrophile is capable of undergoing a reversible reaction.

One approach to address these issues is to modify ligands with reactive moieties that can covalently bond with the protein of interest. Such covalent drugs would have prolonged inhibitory effect and high potency due to the long

target residence conferred by the covalent linkage to the protein[5-7]. However, there is the potential of adverse reactions due to off-target modification or haptenization, and most of the drug discovery efforts deliberately avoid the incorporation of reactive moieties in clinical candidates[8,9] (Figure 5.1b).   A preferred strategy is to preserve the covalent bond while at the same time eliminating any potential toxicity that would arise from permanent modification of off-target proteins or from hapten formation. Drugs that fall in this category have electrophilic moieties that would bind in a covalent but reversible manner with their target protein (Figure 5.1c). The currently identified reversible covalent inhibitors can be categorized into two classes – ones that target catalytic nucleophilic residues of the protein of interest (common examples are protease inhibitors[10]), and ones that target non-catalytic residues[11]. In the case of proteases, the nucleophilicity of the catalytic residues (i.e. cysteine, serine or threonine) is enhanced by the other members of the catalytic triad (i.e. aspartate and histidine), thereby facilitating nucleophilic attack. For other proteins, it is not completely clear how the non-catalytic residues that are still targeted by small molecule electrophiles are activated to perform a nucleophilic addition.

The electrophiles used in the majority of the reversible covalent inhibitors are most commonly aldehydes, ketones, nitriles, and keto-amides, but there are some that are Michael acceptors and boronic acids[10,12-14] (Figure 5.2). A number of protease inhibitors are designed to specifically inactivate their target protease by reacting with the catalytic Ser, Thr or Cys residues in the active site[10,15]. Some examples of marketed and clinically advanced protease targeting drugs are

telaprevir[16], vildagliptin[17] and odanacatib, which possess either a keto-amide or a nitrile electrophile. The β-lactamase inhibitor avibactam is another example of a reversible covalent inhibitor with a ketone electrophilic moiety.
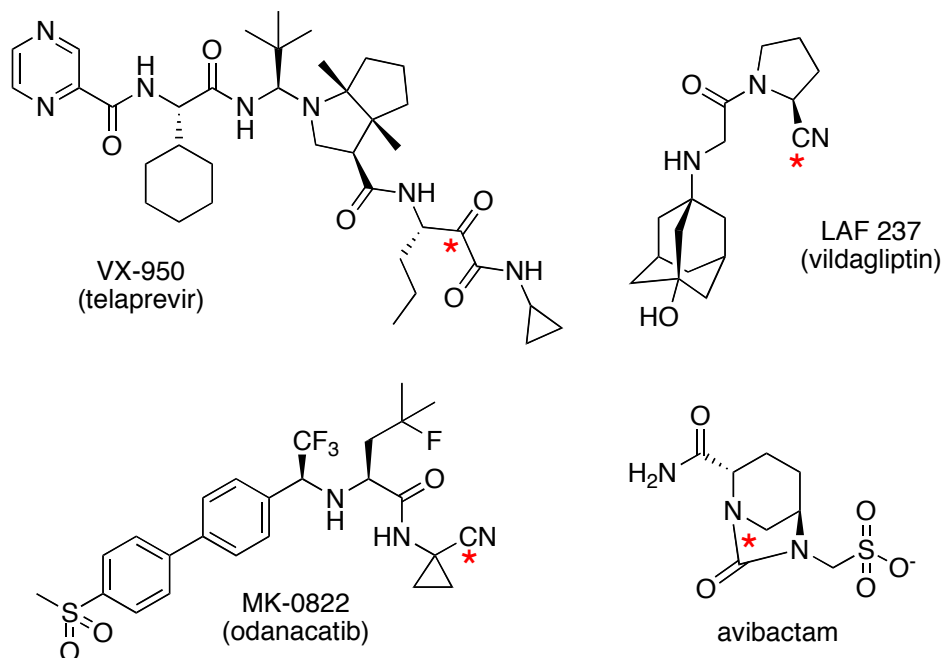
Figure 5.2. Chemical structure of small molecule inhibitors that target *catalytic* residues in a reversible covalent manner. The site of nucleophilic attack is highlighted by an asterisk (red).

Of the compounds that target non-catalytic residues, only one (the p90 ribosomal S6 kinase (RSK) inhibitor **CN-NHiPr**) has been rationally designed to reversibly bind a non-catalytic cysteine[11], while the rest have been serendipitously discovered through high throughput screens[18-24]. Most of the reversible covalent inhibitors that target non-catalytic residues contain α,β-unsaturated Michael acceptor electrophiles (Figure 5.3).
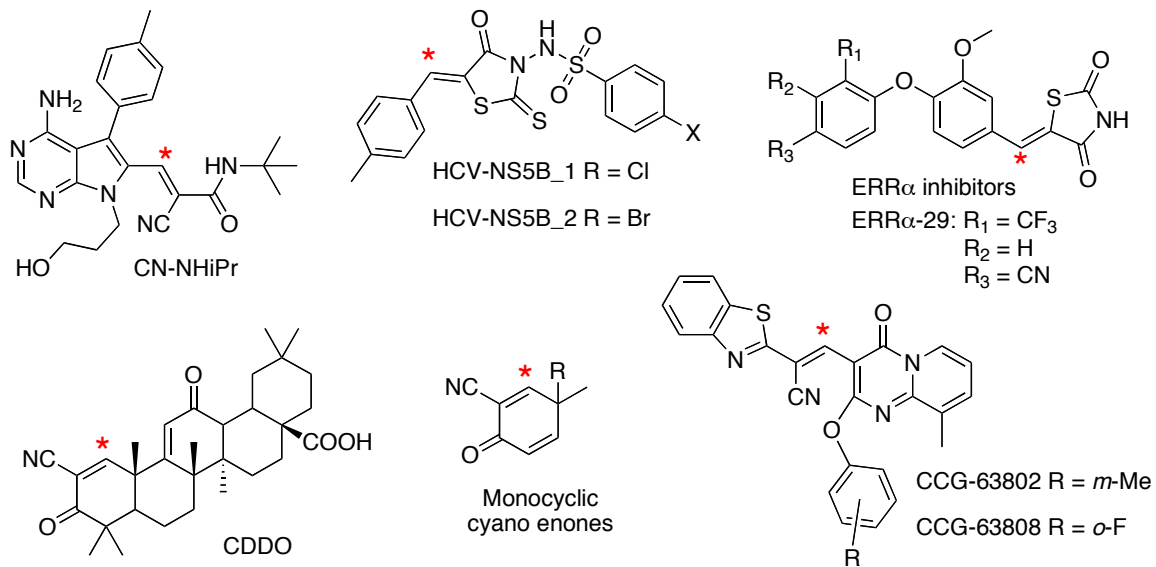
Figure 5.3. Chemical structure of small molecule inhibitors that target *non-catalytic* cysteines in a reversible covalent manner. The site of nucleophilic attack is highlighted by an asterisk (red).

5.2 Reversible covalent inhibitors that target catalytic nucleophiles

VX-950 (telaprevir) is a peptidomimetic protease inhibitor derived via structure-based drug design to target the hepatitis C virus (HCV) NS3-4A protease[16] (Figure 5.2). It incorporates an electrophilic, noncleavable C-terminal α-ketoamide group that forms a reversible covalent bond with the catalytic Ser 139 of the protease. The binding of **VX-950** to the HCV NS3-4A protease is thought to be a biphasic process – first, a weak collision complex between the inhibitor and the HCV protease is formed, followed by a slow rearrangement to a more tightly bound, long-lived covalent complex. The covalent complex formed

between the serine nucleophile of the HCV protease catalytic triad and the α-ketoamide functionality of **VX-950** dissociates slowly, with a half-life of 58 minutes[16,25,26].

MK-0822 (odanacatib) is a selective and potent inhibitor of cathepsin K (Cat K)[27]. Cat K is a primary enzyme involved in osteoclastic bone resorption and is an important target for the treatment of osteoporosis[28,29]. **MK-0822** was developed through SAR studies of previous Cat K inhibitors and was found to selectively inhibit Cat K vesus Cat B, Cat L and Cat S in various assay platforms (for example *in vitro* activity IC_{50} values were 0.2 nM, 1034 nM, 2995 nM and 60 nM respectively). MK-0822 has a nitrile moiety that forms a reversible covalent bond with Cys 25 of Cat K with on- and off-rates of 5.3×10^6 $M^{-1}s^{-1}$ and 0.0008 s^{-1} and a half-life of ~ 14 min[27].

LAF 237 (vildagliptin) is a small non-peptidic DPPIV inhibitor[17] (Figure 5.2). DPPIV is a serine aminopeptidase that rapidly degrades glucagon-like peptide 1 (GLP1), a hormone capable of lowering blood glucose levels. This results in abnormal glucose homeostasis that is associated with obesity and Type 2 diabetes. **LAF 237** forms a reversible covalent bond between a nitrile moiety on the inhibitor and the Ser 630 of the enzyme. The on- and off- rates were reported to be 1.4×10^5 $M^{-1}s^{-1}$ and 0.0025 s^{-1}, and the half-life ($t_{1/2}$) was measured to be ~ 60 min[30-32].

Avibactam is a β-lactamase (class A and class C) inhibitor that is in clinical development for the treatment of bacterial infections comprising Gram-negative organisms[14]. Beta-lactamases are enzymes that are produced by

bacteria and are responsible for conferring resistance to β-lactam antibiotics by breaking open the β-lactam ring and destroying the molecule's antibacterial properties[33]. **Avibactam** is a small non-beta-lactam molecule that can covalently and reversibly acylate the class A serine β-lactamase TEM-1 at Ser70 with observed on and off rates of 2×10^5 $M^{-1}s^{-1}$ and 0.0007 s^{-1}, and a residence half-life ($t_{1/2}$) of ~ 16 min[33]. The acylation followed a two-step mechanism with an initial formation of weak affinity encounter complex, followed by a slowly dissociating covalent complex (Scheme 5.1). **Avibactam** is truly reversible and dissociates intact, unlike other β-lactamase inhibitors that get hydrolyzed by the TEM-1 enzyme[34].

free β-lactamase non-covalent complex covalent complex
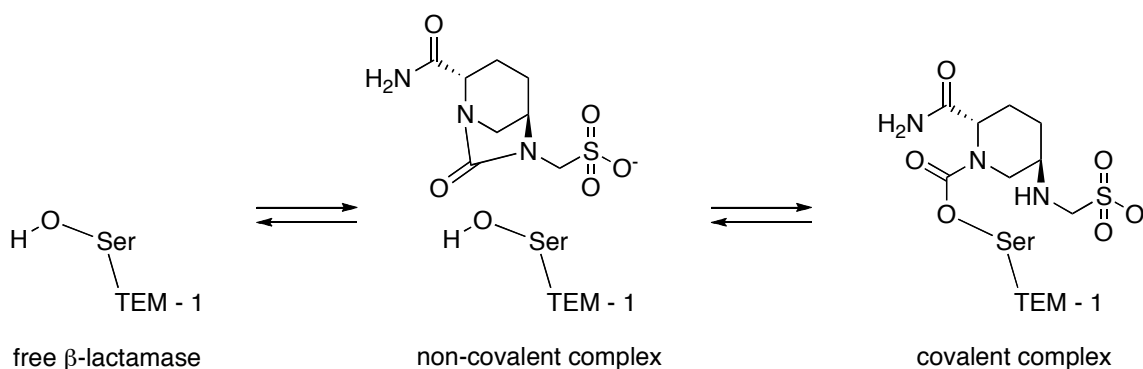
Scheme 5.1. Reversible acylation of the class A serine β-lactamase TEM-1 at Ser70 by avibactam.

5.3 Reversible covalent inhibitors that target non-catalytic residues.

The p90 ribosomal S6 kinase (RSK) inhibitor **CN-NHiPr** (Figure 5.3) is the only rationally designed small molecule inhibitor that specifically targets a non-catalytic residue by a reversible covalent interaction[11]. RSK is a serine/threonine

kinase and is important in controlling various cellular processes such as cell growth, proliferation, and survival. Its activity has been linked to various diseases, such as cancer[35-37]. **CN-NHiPr** has a Michael acceptor cyano-acrylamide functionality that can react with the thiol of Cys436 of the RSK C-terminal kinase domain (Figure 5.4). It forms a long-lived and very slowly dissociating covalent complex with the intact folded protein (half-life is 245 min), despite reacting rapidly and reversibly with thiols in solution.
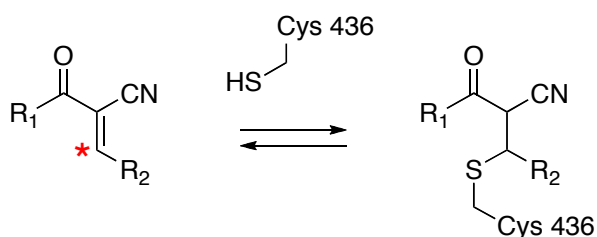
Figure 5.4. Schematic of Michael addition reaction between Cys436 of RSK and nitrile bearing α,β-unsaturated Michael acceptor found in RSK inhibitors. The electrophilic carbon is highlighted by an asterisk (red).

Surprisingly, when the tertiary fold of the RSK protein was disrupted (either by denaturation or proteolysis), the intact **CN-NHiPr** dissociated within seconds, suggesting that the covalent bond is rapidly reversible and cannot promote sustained interaction without additional stabilization by the protein. A co-crystal structure of RSK and a close analogue of **CN-NHiPr** revealed that the non-electrophilic scaffold of the inhibitor is stabilized in the kinase active site by a hydrogen bond network which results in perfect positioning the cyano-acrylamide electrophile for nucleophilic attack by the thiol of Cys436. When the Cys436 was

mutated to a valine, the compound could no longer inhibit the kinase, even though RSK has another cysteine nearby at the bottom of the ATP-binding site.

A reversible covalent small molecule inhibitor that targets a non-catalytic cysteine residue of the hepatitis C virus NS5B RNA-dependent RNA polymerase (HCV NS5B RdRp) has recently been discovered[18,38] (Figure 5.3). An *in vitro* high throughput screen of a random chemical compound library had identified a novel benzylidene derivatives with potent HCV NS5B RdRp inhibitory activity. Further characterization and mechanistic studies of the top two hits (IC_{50} values were 0.54 μM and 0.44 μM for **HCV-NS5B_1** and **HCV-NS5B_2**, respectively) revealed that these inhibitors appear to form a reversible covalent bond with the NS5B Cys 366, a non-catalytic but conserved residue that is required for full NS5B RdRp activity. Compound **HCV_NS5B_2**, as well as other benzylidine series analogs, were co-crystallized with NS5B and the covalent bond was clearly seen in every structure, suggesting that the reversible covalent binding is a general feature of the benzylidine class of inhibitors. Concerns of potential reactivity of the double bond as a nonselective Michael acceptor triggered testing of these analogues against various proteins with biochemically active cysteines and these compound were not found to be nonspecific alkylators[18].

A diaryl ether-based thiazolinenedione ligands were identified in a high-throughput binding assay to inhibit the estrogen-related receptor α (ERRα)[19] (Figure 5.3). ERRα is an orphan nuclear receptor that has been functionally implicated in the regulation of metabolism and energy homeostasis, and there is interest in this receptor as a potential therapeutic target in the treatment of

diabetes and obesity. The compounds in the HTS library were derived from a Knoevenagel condensation of various diaryl ether aldehydes with thiazolidind-2,4-dione and in essence contained a Michael acceptor olefin that is activated by two electorn withdrawing groups. A cocrystal structure revealed that a critical thioether bond is formed between Cys 325 of ERRα and inhibitor **ERRα-29** via conjugate addition to the α,β-unsaturated thiazolinenedione. This covalent addition resulted in a slowly dissociating complex with in vitro apparent half-life ~18 hours. The recovered inhibitors were determined to be structurally intact suggesting that the dissociation from ERRα is not a result of degradative fragmentation.

Cyanoenones are other examples of α,β-unsaturated Michael acceptors that are capable of forming reversible covalent adducts with cysteine thiols[22,24,39]. Medicinal chemistry efforts have identified semi-synthetic triterpenoids **CDDO**, **CDDO-Me** and **CDDO-Im** with potent anti-inflammatory, antiproliferative and differentiating activities[21,40-44] (Figure 5.3, Figure 5.5).
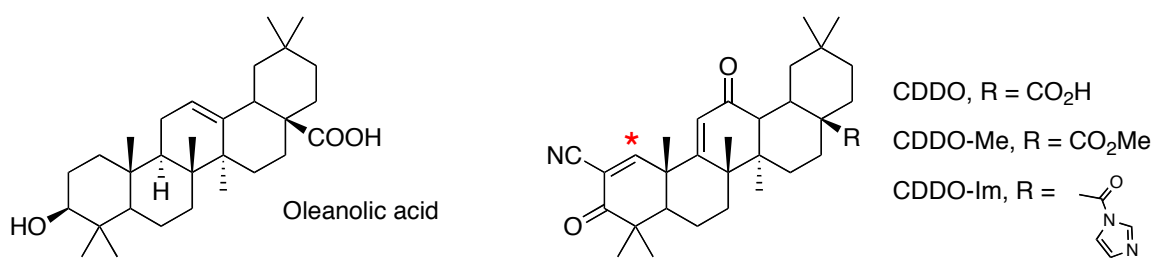
Figure 5.5. Chemical structure of the naturally occurring triterpenoid oleanolic acid and its synthetic derivatives **CDDO**, **CDDO-Me** and **CDDO-Im** that contain an α,β-unsaturated Michael acceptor. The site of nucleophilic attack in **CDDO** and its analogs is highlighted by an asterisk (red).

CDDO is derived from oleanolic acid (Figure 5.5) and contains a α,β-unsaturated carbonyl moiety with a nitrile electron withdrawing group that greatly enhances its potency in inhibiting cytokine-mediated induction of nitiric-oxide synthase (iNOS), relative to the parent compound that lacks the electrophile[40]. Studies have shown that **CDDO** and its derivatives at the C-28 position (**CDDO-Me**, **CDDO-Im**) induce apoptosis of human myeloid leukemia, osteosarcoma, multiple myeloma and lung cancer, breast cancer and pancreatic cancer cells[45-49]. The α,β-unsaturated carbonyl moiety of the A ring of **CDDO** and its analogs can form reversible adducts with reactive thiol groups in dithiothreitol (DTT) or with specific cysteine-rich protein targets[21,24,43,50,51]. The analogs of CDDO that lack the nitrile at the C-2 position on ring A, on the other hand, did not react detectably with DTT. In cell based pull-down experiments, **CDDO** and **CDDO-Me** were shown to form reversible adducts with IκB kinase β (IKKβ) at Cys179[21] as well as with the lipid phosphatase PTEN at Cys124[51], JAK1 at Cys1077[52] and STAT3 at Cys259[52]. Additional studies have also shown that these triterpenoid cyanoenones, as well as some monocyclic cyanoenone analogs (Figure 5.3), can interact with thiol groups of the Keap1 sensor[20]. However, due to the presence of multiple cysteine residues in Keap1 (27 in human homologs) and the reversibility of the Michael adduct, the determination of which cysteines of Keap1 are modified by these compounds has been challenging[24,43].

A high-throughput time-resolved fluorescence resonance energy transfer screen has identified inhibitors of regulators of G protein signaling (RGS)

142

proteins[23] form a library containing about 40,000 compounds. Two closely related acrylonitrile inhibitors with modest activity (**CCG-63802** and **CCG-63808**; IC_{50} values were 1.9 μM and 1.4 μM, respectively) (Figure 5.3) were identified from this primary HTS screen and were further characterized as the first reversible small-molecule inhibitors of an RSG protein (RSG4) in a cysteine dependent mechanism. These analogs contain a vinyl cyanide moiety that can function as a reversible Michael acceptor. Using a flow cytometry protein interaction assay (FCIPA), which measures the binding of RSG4 to the heterodimeric $G\alpha_o$, **CCG-63802** and **CCG-63808** displayed inhibitory activities with IC_{50} values at 9 μM and 10 μM, respectively. When the compounds were tested against a mutant of RSG4 in which all the cysteines in the RSG domain were mutated to alanines, their inhibitory potency was greatly reduced and the IC_{50} values were > 400 μM for both compounds. RSG4 has 11 cysteine residues and further mutational analysis revealed that Cys148, Cys132 and Cys95 are important for the full sensitivity of RSG4 to **CCG-63802** and **CCG-63808**. When the FCIPA assay was performed in the presence of 2mM reduced glutathione (similar to intracellular concentrations), the compounds were approximately 5 fold less potent, but were still able to fully inhibit the interaction between RGS4 and $G\alpha_o$. Further studies need to be performed to elucidate the exact mechanism of action of these compounds.

5.4 Conclusions

Electrophile-bearing small molecule inhibitors that can covalently modify their targets display enhanced selectivity and potency, but also have the potential of being neutralized by off-target nucleophiles. The incorporation of rapidly reversible electrophiles, such as aldehydes, ketones and α,β-unsaturated Michael acceptors has allowed for the generation of highly potent inhibitors with improved selectivity and reduced off target reactivity. The rational design of small molecules that selectively bind to non-catalytic cysteines in a reversible covalent manner has successfully yielded highly potent and selective p90 ribosomal 6S kinase inhibitors and has proven the utility of this concept. Taking advantage of the rich chemistry of thiols and the reactivity of α,β-unsaturated Michael acceptors can facilitate the design of small molecule drugs that target a wide range of proteins.

5.5 REFERENCES

1   US Food and Drug Administration.

         http://www.fda.gov/downloads/AboutFDA/Transparency/Basics/UCM2474
         65.pdf.

2   Kola, I. The state of innovation in drug development. *Clin Pharmacol Ther*
         83, 227-230 (2008).

3   Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition
         rates? *Nat Rev Drug Discov* **3**, 711-715 (2004).

4   California Biomedical Research Associaiton.

         http://ca-biomed.org/pdf/media-kit/fact-sheets/cbradrugdevelop.pdf

5   Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of
         covalent drugs. *Nat Rev Drug Discov* **10**, 307-317 (2011).

6   Smith, A. J., Zhang, X., Leach, A. G. & Houk, K. N. Beyond picomolar
         affinities: quantitative aspects of noncovalent and covalent binding of
         drugs to proteins. *J Med Chem* **52**, 225-233 (2009).

7   Copeland, R. A., Pompliano, D. L. & Meek, T. D. Drug-target residence
         time and its implications for lead optimization. *Nat Rev Drug Discov* **5**,
         730-739 (2006).

8   Uetrecht, J. Idiosyncratic drug reactions: past, present, and future. *Chem
         Res Toxicol* **21**, 84-92 (2008).

9   Evans, D. C., Watt, A. P., Nicoll-Griffith, D. A. & Baillie, T. A. Drug-protein
         adducts: an industry perspective on minimizing the potential for drug

bioactivation in drug discovery and development. *Chem Res Toxicol* **17**, 3-16 (2004).

10   Turk, B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* **5**, 785-799 (2006).

11   Serafimova, I. M. *et al.* Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol* **8**, 471-476 (2012).

12   Santos, M. M. & Moreira, R. Michael acceptors as cysteine protease inhibitors. *Mini Rev Med Chem* **7**, 1040-1050 (2007).

13   Reich, S. H. *et al.* Substituted benzamide inhibitors of human rhinovirus 3C protease: structure-based design, synthesis, and biological evaluation. *J Med Chem* **43**, 1670-1683 (2000).

14   Ehmann, D. E. *et al.* Avibactam is a covalent, reversible, non-beta-lactam beta-lactamase inhibitor. *Proc Natl Acad Sci U S A* **109**, 11663-11668 (2012).

15   Abbenante, G. & Fairlie, D. P. Protease inhibitors in the clinic. *Med Chem* **1**, 71-104 (2005).

16   Lin, C., Kwong, A. D. & Perni, R. B. Discovery and development of VX-950, a novel, covalent, and reversible inhibitor of hepatitis C virus NS3.4A serine protease. *Infect Disord Drug Targets* **6**, 3-16 (2006).

17   Villhauer, E. B. *et al.* 1-[[(3-hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: a potent, selective, and orally bioavailable dipeptidyl peptidase IV inhibitor with antihyperglycemic properties. *J Med Chem* **46**, 2774-2789 (2003).

18   Lee, G. *et al.* Novel inhibitors of hepatitis C virus RNA-dependent RNA polymerases. *J Mol Biol* **357**, 1051-1057 (2006).

19   Patch, R. J. *et al.* Identification of diaryl ether-based ligands for estrogen-related receptor alpha as potential antidiabetic agents. *J Med Chem* **54**, 788-808 (2011).

20   Dinkova-Kostova, A. T. *et al.* Extremely potent triterpenoid inducers of the phase 2 response: correlations of protection against oxidant and inflammatory stress. *Proc Natl Acad Sci U S A* **102**, 4584-4589 (2005).

21   Ahmad, R., Raina, D., Meyer, C., Kharbanda, S. & Kufe, D. Triterpenoid CDDO-Me blocks the NF-kappaB pathway by direct inhibition of IKKbeta on Cys-179. *J Biol Chem* **281**, 35764-35769 (2006).

22   Avonto, C. *et al.* An NMR spectroscopic method to identify and classify thiol-trapping agents: revival of Michael acceptors for drug discovery? *Angew Chem Int Ed Engl* **50**, 467-471 (2011).

23   Blazer, L. L. *et al.* Reversible, allosteric small-molecule inhibitors of regulator of G protein signaling proteins. *Mol Pharmacol* **78**, 524-533 (2010).

24   Dinkova-Kostova, A. T. *et al.* An exceptionally potent inducer of cytoprotective enzymes: elucidation of the structural features that determine inducer potency and reactivity with Keap1. *J Biol Chem* **285**, 33747-33755 (2010).

25 Lin, K., Perni, R. B., Kwong, A. D. & Lin, C. VX-950, a novel hepatitis C virus (HCV) NS3-4A protease inhibitor, exhibits potent antiviral activities in HCv replicon cells. *Antimicrob Agents Chemother* **50**, 1813-1822 (2006).

26 Perni, R. B. *et al.* Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3-4A serine protease. *Antimicrob Agents Chemother* **50**, 899-909 (2006).

27 Gauthier, J. Y. *et al.* The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg Med Chem Lett* **18**, 923-928 (2008).

28 Yasuda, Y., Kaleta, J. & Bromme, D. The role of cathepsins in osteoporosis and arthritis: rationale for the design of new therapeutics. *Adv Drug Deliv Rev* **57**, 973-993 (2005).

29 Grabowskal, U., Chambers, T. J. & Shiroo, M. Recent developments in cathepsin K inhibitor design. *Curr Opin Drug Discov Devel* **8**, 619-630 (2005).

30 Brandt, I. *et al.* Inhibition of dipeptidyl-peptidase IV catalyzed peptide truncation by Vildagliptin ((2S)-{[(3-hydroxyadamantan-1-yl)amino]acetyl}-pyrrolidine-2-carbonitrile). *Biochem Pharmacol* **70**, 134-143 (2005).

31 Tummino, P. J. & Copeland, R. A. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* **47**, 5481-5492 (2008).

32 Landersdorfer, C. B., He, Y. L. & Jusko, W. J. Mechanism-based population pharmacokinetic modelling in diabetes: vildagliptin as a tight

binding inhibitor and substrate of dipeptidyl peptidase IV. *Br J Clin Pharmacol* **73**, 391-401 (2012).

33   Fisher, J. F., Meroueh, S. O. & Mobashery, S. Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity. *Chem Rev* **105**, 395-424 (2005).

34   Bush, K., Macalintal, C., Rasmussen, B. A., Lee, V. J. & Yang, Y. Kinetic interactions of tazobactam with beta-lactamases from all major structural classes. *Antimicrob Agents Chemother* **37**, 851-858 (1993).

35   Carriere, A., Ray, H., Blenis, J. & Roux, P. P. The RSK factors of activating the Ras/MAPK signaling cascade. *Front Biosci* **13**, 4258-4275 (2008).

36   Anjum, R. & Blenis, J. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* **9**, 747-758 (2008).

37   Romeo, Y., Zhang, X. & Roux, P. P. Regulation and function of the RSK family of protein kinases. *Biochem J* **441**, 553-569 (2012).

38   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).

39   Zheng, S. *et al.* Synthesis, chemical reactivity as Michael acceptors, and biological potency of monocyclic cyanoenones, novel and highly potent anti-inflammatory and cytoprotective agents. *J Med Chem* **55**, 4837-4846 (2012).

40   Honda, T. *et al.* Design and synthesis of 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid, a novel and highly active inhibitor of nitric oxide production in mouse macrophages. *Bioorg Med Chem Lett* **8**, 2711-2714 (1998).

41   Suh, N. *et al.* A novel synthetic oleanane triterpenoid, 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid, with potent differentiating, antiproliferative, and anti-inflammatory activity. *Cancer Res* **59**, 336-341 (1999).

42   Wang, Y. *et al.* A synthetic triterpenoid, 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (CDDO), is a ligand for the peroxisome proliferator-activated receptor gamma. *Mol Endocrinol* **14**, 1550-1556 (2000).

43   Eggler, A. L., Small, E., Hannink, M. & Mesecar, A. D. Cul3-mediated Nrf2 ubiquitination and antioxidant response element (ARE) activation are dependent on the partial molar volume at position 151 of Keap1. *Biochem J* **422**, 171-180 (2009).

44   Deeb, D. *et al.* Oleanane triterpenoid CDDO-Me inhibits growth and induces apoptosis in prostate cancer cells through a ROS-dependent mechanism. *Biochem Pharmacol* **79**, 350-360 (2010).

45   Ikeda, T. *et al.* Triterpenoid CDDO-Im downregulates PML/RARalpha expression in acute promyelocytic leukemia cells. *Cell Death Differ* **12**, 523-531 (2005).

46   Ito, Y. *et al.* The novel triterpenoid 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid induces apoptosis of human myeloid leukemia cells by a caspase-8-dependent mechanism. *Cell Growth Differ* **11**, 261-267 (2000).

47   Konopleva, M. *et al.* Novel triterpenoid CDDO-Me is a potent inducer of apoptosis and differentiation in acute myelogenous leukemia. *Blood* **99**, 326-335 (2002).

48   Stadheim, T. A., Suh, N., Ganju, N., Sporn, M. B. & Eastman, A. The novel triterpenoid 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (CDDO) potently enhances apoptosis induced by tumor necrosis factor in human leukemia cells. *J Biol Chem* **277**, 16448-16455 (2002).

49   Konopleva, M. *et al.* The synthetic triterpenoid 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid induces caspase-dependent and -independent apoptosis in acute myelogenous leukemia. *Cancer Res* **64**, 7927-7935 (2004).

50   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).

51   Pitha-Rowe, I., Liby, K., Royce, D. & Sporn, M. Synthetic triterpenoids attenuate cytotoxic retinal injury: cross-talk between Nrf2 and PI3K/AKT signaling through inhibition of the lipid phosphatase PTEN. *Invest Ophthalmol Vis Sci* **50**, 5339-5347 (2009).

52   Ahmad, R., Raina, D., Meyer, C. & Kufe, D. Triterpenoid CDDO-methyl ester inhibits the Janus-activated kinase-1 (JAK1)-->signal transducer and activator of transcription-3 (STAT3) pathway by direct inhibition of JAK1 and STAT3. *Cancer Res* **68**, 2920-2926 (2008).

153
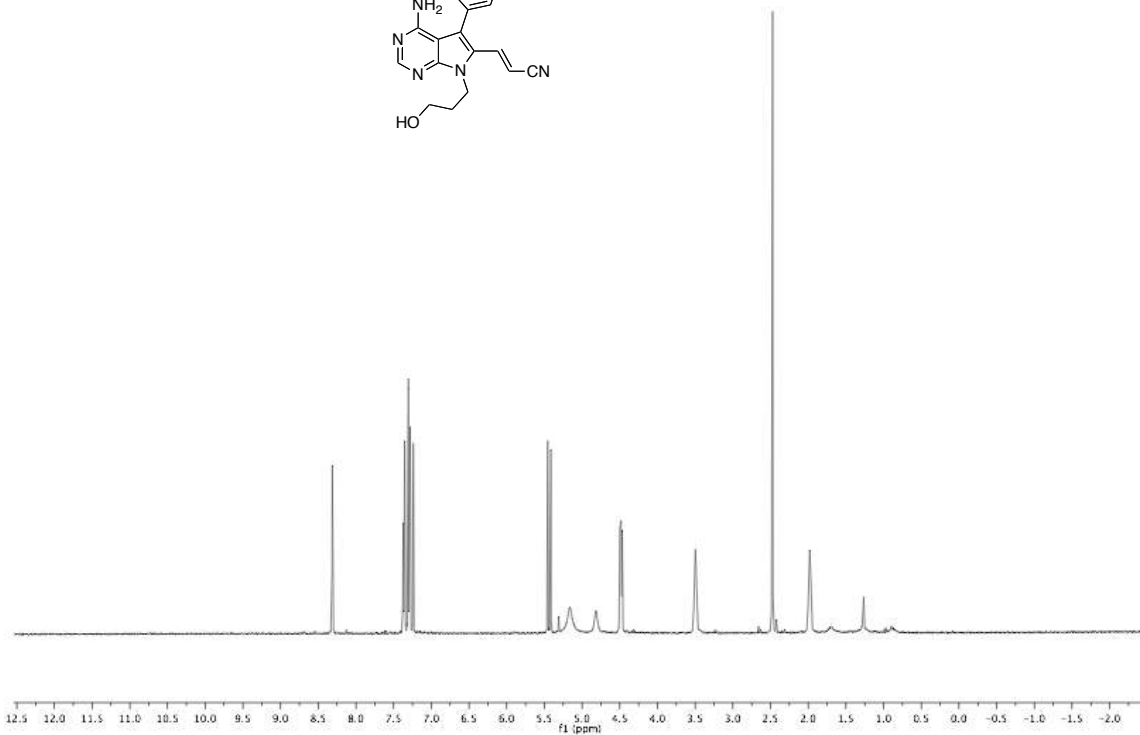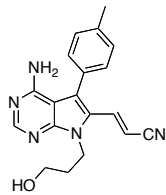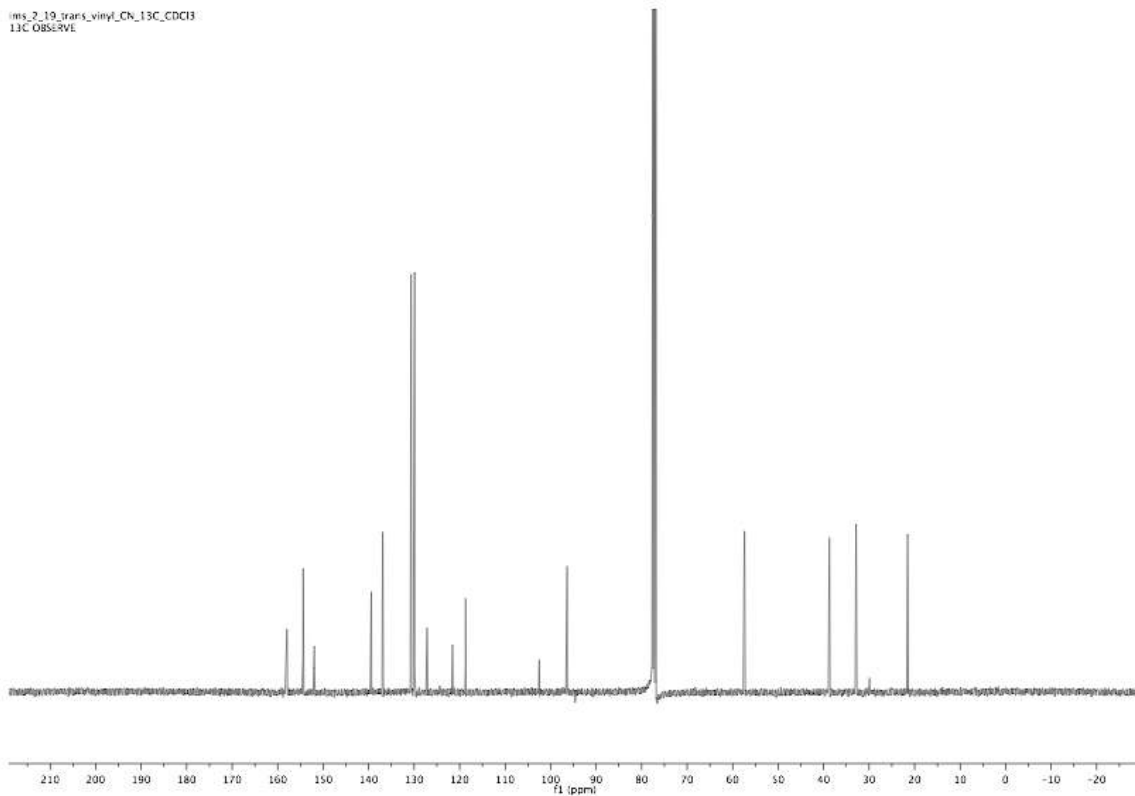
Publishing Agreement

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.

_____   12/26/2012
Author Signature   Date